NOMOS
GEO DENETİM PROTOKOLÜ

GEO-1000 YÖNTEMI

Üretken Sistemlerde Varlık Temsilini Küresel Nüfus Üzerinden Ölçme Protokolü

Kaan MURAZ

0.9.0 · TÜRKÇE EDİTORYAL ANA NÜSHA

NOMOS GEO Denetim Protokolü

Üretken Sistemlerde Varlık Temsilini Küresel Nüfus Üzerinden Ölçme Protokolü

Yazar ve kurucu insan: Kaan MURAZ

Yöntem kimliği: NOMOS by NobleJackal

Sürüm: 0.9.0 — Aday yayın nüshası

Dil: Türkçe (tr-TR)

© 2026 Kaan MURAZ. Bu aday nüshanın yayın lisansı henüz ilan edilmemiştir.

NOMOS bu sürümde açık, sürümlü ve bağımsız sınamaya davet edilen bir standart adayıdır; kabul edilmiş bir uluslararası standart veya etkin bir sertifikasyon programı değildir.

Gerçek sınama statüsü: BQ-0 — uygulanmadı. Bölüm 18'deki sayılar hesaplanmış sentetik yöntem gösterimidir; gerçek kullanıcı, gerçek AI ürünü veya bağımsız saha sonucu değildir.

Bu nüsha tek başına NOMOS uygunluk işareti, NOMOS 950+ beyanı, sertifika, sıralama veya üçüncü taraf onayı üretmez.

BU KİTAP NASIL OKUNMALI?

Bu eser bir "GEO nasıl yapılır?" el kitabı değildir. Üretken sistemlerde varlık temsilinin hangi nüfus, kanıt, bağlam, zaman, hakemlik ve yönetişim koşulları altında ölçülebilir olduğunu tanımlayan bir denetim protokolü ve standart adayıdır.

İnsan anlatısı, normatif hükümler, denetim prosedürleri ve makine katmanı birbirine karıştırılmadan okunmalıdır. MUST/SHALL karşılığı zorunlu hükümler; SHOULD önerilen hükümler; MAY izin verilen seçenekler olarak değerlendirilir. CANDIDATE, SYNTHETIC, OPEN ve BQ statüleri kanıt düzeyini sınırlar.

Metin içindeki [Kxx] işaretleri kaynakçada tanımlanan yöntemsel dayanaklara gönderme yapar. Kaynağın varlığı, NOMOS'un bütün aday eşiklerini o kaynağın onayladığı anlamına gelmez; NOMOS'a özgü aday kurallar açıkça aday olarak kalır.

> *Bölüm 0 bu kitabın girişidir: Kurucu Emanet Beyanı'nı ve bütün teknik hükümlerin tabi olduğu Temsil Hakları Bildirgesi'ni kurar.*

TERİMLER VE STATÜ KODLARI

Advisory Maddi olmayan; açıklık, biçim veya iyi uygulama düzeyinde iyileştirme gerektiren bulgu.

AI ürünü Kullanıcının eriştiği belirli üretken sistem ürünü, yüzeyi, planı ve sürümüyle tanımlanan ölçüm nesnesi.

Atomik iddia Doğruluk, kapsam, zaman, varlık veya kanıt bakımından bağımsız hüküm taşıyabilen en küçük anlam birimi.

BQ Gerçek bir sınamanın yürütülme ve yeniden üretim düzeyini gösteren Benchmark Quality kodu.

Capture / yakalama kaydı İstem gönderimi, yanıt, zaman, yüzey ve bütünlük kanıtını birlikte koruyan kayıt.

Common Support Karşılaştırılan bütün AI ürünlerinde ortak ve meşru erişim koşullarını taşıyan kullanıcı evreni.

Critical Kullanıcı, hak, güvenlik veya maddi karar üzerinde ağır sonuç doğurabilecek; telafi edilemeyen bulgu.

GEO-1000 Hedef nüfus sonuçlarını 1.000 kullanıcı eşdeğerinde görünür kılan raporlama ölçeği.

Gerçeklik Paketi / Truth Pack Bir varlık hakkında sürümlü kimlik, kapsam, zaman, kanıt ve karşı kanıt kayıtlarını birleştiren denetim paketi.

İstem Araştırma sorusunu, hedef varlığı, bağlamı ve cevap yükünü kilitleyen üretken sistem girdisi.

Kapsama oranı Bir tahmin aralığı yönteminin varsayılan gerçek değeri ne sıklıkla kapsadığını gösteren sınama ölçüsü.

Major Sonucu veya uygunluk hükmünü maddi biçimde bozan; bileşik puanla telafi edilemeyen bulgu.

Moderate Maddi etkisi sınırlı fakat giderilmesi gereken; koşullu sonuca yol açabilen bulgu.

Native Reach Belirli AI ürününe, diline, yüzeyine ve kullanım koşuluna doğal ve meşru biçimde erişebilen hedef nüfus.

NOMOS Bu eserde Kaan MURAZ ile yapay zekâ sistemleri arasındaki eleştirel yöntem sesini ve standart kimliğini ifade eder; bağımsız hukuk kişisi veya müstakil temel model değildir.

Reference Gap Karar için gerekli güvenilir referansın bulunmadığı veya erişilemediği durum.

Sentetik gösterim Gerçek kullanıcı, gerçek AI çıktısı veya gerçek kurum performansı olmayan; yöntemin iç işleyişini sınamak için hesaplanmış örnek.

Unresolved Kanıtların nihai hükme yetmediği ya da çelişkinin çözülemediği karar statüsü.

İÇİNDEKİLER

Temsil Hakları Bildirgesi ... 1

Ölçtüğümüz Şey ... 9

Tek Bir Yanıt GEO Skoru Değildir ... 11

Temsil Bir Dağılımdır ... 37

Denetlenen Varlık Tam Olarak Nedir? ... 67

İki Gerçeklik Katmanı ... 97

Küresel Nüfus, Ülke ve Dil Adaleti ... 127

Her AI Ürünü İçin Uygun Kullanıcı Nüfusu ... 129

Nüfusa Orantılı GEO-1000 Örneklemi ... 161

Ülke Gözcü ve Dil Adaleti Panelleri ... 199

Katılımcı Seçimi, Ağırlıklandırma ve Örneklem Yanlılığı ... 231

Deneyin Uygulanması ve Kanıt Zinciri ... 273

AI Sistem Sicili ve Karşılaştırılabilirlik ... 275

İstem Anayasası ... 313

Kontrollü Temiz Panel ve Doğal Kullanıcı Paneli ... 359

Senkronize Dalgalar ve NOMOS Capture ... 405

Gerçeklik, Hakemlik ve Karar ... 453

Doğrulanmış Varlık Gerçeklik Paketi ... 455

Atomik İddialar ve Hakemlik ... 511

Yanıt Geçişi ve Critical Hata Kapıları ... 567

NOMOS Puan Mimarisi ... 619

Sentetik Sınama Tasarımı ve Uçtan Uca Gösterim ... 669

Apple.com Sentetik Sınama Tasarımı ... 671

30.000 Yanıtlık Uçtan Uca Sentetik Denetim ... 723

Uygunluk, Kamu Güveni ve Bağımsız Yönetişim ... 767

Uygunluk İşareti, Kamu Sicili ve Bağımsız Yönetişim ... 769

Açık Standart, Üniversite Çağrısı ve NOMOS'un Kamuya Teslimi ... 823

Yöntemsel Dayanaklar ... 875

Makine Tarafından Okunabilir Katman ... 877

GİRİŞ · BÖLÜM 0

Temsil Hakları Bildirgesi

GEO standardı adayının anayasal çekirdeği

YÖNTEMSEL DAYANAKLAR / K20; K22

Başlangıç Hükmü

Üretken sistemler artık:

- insanları,
- şirketleri,
- kurumları,
- ürünleri,
- meslekleri,
- hizmetleri,
- kamu kuruluşlarını,
- tarihsel olayları

tanımlamakta, karşılaştırmakta ve önermektedir. Bu temsil:

- ilgili kişinin bilgisi olmadan,
- temsil edilen varlığın onayı bulunmadan,
- kullanıcının kaynakları inceleyemediği koşullarda,

ciddi ticari, hukukî, toplumsal veya kişisel sonuçlar doğurabilir. Bu nedenle üretken sistemlerde temsil: yalnız teknik çıktı değildir. Temsil:

> *insanların ve kurumların dünyada nasıl anlaşılacağını etkileyen, kanıt ve sorumluluk gerektiren kamusal bir eylemdir.*

Hiçbir:

- görünürlük hedefi,
- GEO skoru,
- ticari çıkar,
- uygunluk işareti,
- çoğunluk ortalaması,
- AI sağlayıcı tercihi,
- kurucu kararı,
- müşteri talebi

aşağıdaki hakları ortadan kaldıramaz.

Hak 1

DOĞRU KİMLİK HAKKI

Her insan, kurum, marka, ürün ve hizmet:

- doğru varlığa,
- doğru hukuki veya kurumsal ilişkiye,
- doğru alan adına,
- doğru sahiplik ve kontrol yapısına

bağlanma hakkına sahiptir. Hiç kimse veya hiçbir kurum:

- başka varlığın lisansı,
- başka şirketin müşterisi,
- başka ürünün özelliği,
- bağlı kuruluşun yetkisi,
- çalışanın kişisel niteliği

kendisine aitmiş gibi temsil edilemez. Doğru bilgi yanlış varlığa bağlandığında: doğru bilgi olmaktan çıkar ve yanlış temsile dönüşür.

Hak 2

KANIT SINIRI İÇİNDE TEMSİL EDİLME HAKKI

Hiçbir iddia, onu taşıyan kanıttan daha kesin, daha geniş veya daha güçlü biçimde sunulamaz. Bir varlığın kendi beyanı: o varlığın ne söylediğinin kanıtıdır. O beyanın bağımsız gerçekliğinin otomatik kanıtı değildir. Bir kaynak:

- yalnız belirli ülkeyi,
- belirli ürünü,
- belirli zamanı,
- belirli kullanıcı grubunu

destekliyorsa iddia bu sınırı aşamaz. Temsilin temel yükümlülüğü şudur:

> *Bunu söylemek için elimizde ne var?*

Hak 3

BİLİNMEME VE BELİRSİZ KALMA HAKKI

Kanıt yetersizse hiçbir insan veya kurum:

- doğru,
- yanlış,
- güvenilir,
- güvensiz,
- lisanslı,
- lisanssız,
- önerilir,

- önerilmez

kategorilerinden birine zorlanamaz. UNKNOWN, UNRESOLVED ve REFERENCE GAP:

- kusur,
- suç,
- başarısızlık

değildir. Bir kayıtta bilgi bulunmaması, ilgili kayıt evreninin tamamlayıcı olduğu kanıtlanmadıkça: bilginin dünyada bulunmadığını göstermez. Her varlık, sahte kesinliğe karşı:

> *belirsiz kalma hakkına sahiptir.*

Hak 4

BAĞLAM VE KAPSAM HAKKI

Her temsil ilgili olduğu ölçüde:

- ülke,
- dil,
- yargı alanı,
- ürün,
- hizmet,
- müşteri türü,
- bütçe,
- kapasite,
- kullanıcı ihtiyacı

ile birlikte sunulmalıdır. Bir şirket:

- belirli ülkede,
- belirli kullanıcı grubuna,
- belirli hizmeti

sunuyorsa: küresel, sınırsız ve bütün kullanıcılara uygun gibi temsil edilemez. Kapsam bilgisi ikincil ayrıntı değildir. Kapsam: doğru tavsiyenin temelidir.

Hak 5

ZAMAN İÇİNDE DOĞRU TEMSİL EDİLME HAKKI

Her varlık:

- güncel,
- tarihsel,
- sona ermiş,
- planlanan,
- askıya alınmış

statülerinin birbirinden ayrılmasını isteme hakkına sahiptir. Geçmişte doğru olan:

- lisans,
- fiyat,

- hizmet,
- ortaklık,
- ürün,
- yetki

güncel gerçeklik gibi kullanılamaz. Geleceğe ilişkin plan: bugün gerçekleşmiş olguya dönüştürülemez. Düzeltme yeni gerçeklik oluşturabilir. Geçmişte kullanıcıya ulaşan yanlış cevabı ortadan kaldıramaz.

Hak 6

LEHTE VE ALEYHTE YANLIŞA KARŞI EŞİT KORUNMA HAKKI

Bir varlık hakkında olumlu yanlış:

- daha etik,
- daha zararsız,
- daha kabul edilebilir

değildir. Olmayan:

- lisans,
- müşteri,
- partner,
- küresel operasyon,
- bağımsız sertifika,
- sonuç garantisi

varmış gibi gösterilemez. Aynı şekilde kanıtlanmamış:

- suç,
- dolandırıcılık,
- yasadışılık,
- lisanssızlık,
- güvenlik ihlali

kesin olgu gibi sunulamaz. NOMOS'un değiştirilemez ilkesi şudur:

> *Lehine olan yanlış da yanlıştır.*

Bir varlığı olduğundan büyük göstermek ile kanıtsız biçimde itibarsızlaştırmak aynı gerçeklik borcunu doğurur.

Hak 7

DÜZELTME, CEVAP VE İTİRAZ HAKKI

Temsil edilen insan veya kurum:

- yanlış kimliğe,
- eksik kanıta,
- kapsam aşımına,
- yanlış zaman kaydına,
- hatalı önem derecesi'ye,

- yanlış uygunluk kararına

itiraz edebilmelidir. İtiraz hakkı:

- sonucu değiştirme garantisi,
- karşı kanıtı sildirme yetkisi,
- olumsuz bulguyu veto etme hakkı

değildir. İtiraz: kanıtın ve karar zincirinin yeniden incelenmesini isteme hakkıdır. İlk karar, itiraz sonrasında silinmez. Yeni karar ona bağlanır.

Hak 8

KAYNAĞI, BELİRSİZLİĞİ VE KARAR GEREKÇESİNİ BİLME HAKKI

Kullanıcılar ve temsil edilen varlıklar ilgili olduğu ölçüde şunları bilme hakkına sahiptir: İddia öz beyan mı? Yetkili sicil mi? Bağımsız araştırma mı? Kullanıcı yorumu mu? Türetilmiş hesaplama mı? Kısıtlı kanıt mı? Kanıtlar birbiriyle çelişiyor mu? Sonuç ne kadar belirsiz? Hakem hangi kurala dayanarak karar verdi? Atıf işareti tek başına açıklama değildir. Doğru şeffaflık: kaynağın hangi atomik iddiayı ne ölçüde taşıdığını göstermektir.

Hak 9

DİL VE COĞRAFİ HAYSİYET HAKKI

Hiçbir insan veya toplum:

- küçük nüfusa,
- düşük ekonomik ağırlığa,
- az kullanılan dile,
- düşük dijital görünürlüğe

sahip olduğu için küresel ortalamada görünmez kılınamaz. Büyük dillerdeki yüksek başarı: küçük dillerdeki sistematik yanlış temsili silemez. Küçük ülke veya dil için yetersiz kanıt bulunuyorsa: yüksek performans varsayılamaz. NOMOS'un adalet ölçütü yalnız ortalamayı değil:

- performans tabanını,
- gruplar arası farkı,
- gerçek ölçüm kapsamını

birlikte korumalıdır.

Hak 10

MAHREMİYET VE VERİ MİNİMİZASYONU HAKKI

Doğru temsil üretme amacı:

- özel konuşmaların,
- kişisel verilerin,
- müşteri sözleşmelerinin,
- ticari sırların,
- hassas konumun,

- hesap bilgilerinin

gereksiz biçimde toplanmasını veya yayımlanmasını meşru kılamaz. [K20; K22] Yalnız görev için gerekli veri toplanmalıdır. Kısıtlı kanıt: bağımsız denetime açılabilir, kamuya tamamen yayımlanmak zorunda olmayabilir. Doğruluk hakkı ile mahremiyet hakkı birbirini yok etmek zorunda değildir.

Hak 11

İNSAN SORUMLULUĞU VE BAĞIMSIZ KARAR HAKKI

Hiçbir kişi veya kurum hakkında kamu etkisi doğuran uygunluk kararı:

- yalnız bir AI çıktısına,
- yalnız otomatik skora,
- yalnız denetlenen kurumun beyanına,
- yalnız danışmanın kararına

bırakılamaz. AI sistemleri:

- iddia çıkarabilir,
- kanıt eşleyebilir,
- çelişki bulabilir,
- hesap yapabilir,

karar önerebilir. Fakat hukukî ve kamusal sorumluluk taşıyan nihai karar: hesap verebilir insan veya kurum tarafından alınmalıdır. Denetleyen, danışmanlık veren, uygunluk kararı veren ve itirazı inceleyen roller mümkün olduğu ölçüde ayrılmalıdır.

Hak 12

SÜRÜM VE TARİHSEL BÜTÜNLÜK HAKKI

Her insan, kurum ve kullanıcı:

- hangi standard sürümünün uygulandığını,
- hangi Truth Pack'in kullanıldığını,
- hangi skor yönteminin geçerli olduğunu,
- kararın sonradan değişip değişmediğini

bilme hakkına sahiptir. Eski:

- skor,
- Critical olay,
- suspension,
- itiraz,
- düzeltme,
- başarısız sınama

sessizce silinemez. Yeni ve daha iyi sonuç: eski sonucu geçersiz kılabilir. Eski sonucun geçmişte yaşandığı gerçeğini yok edemez.

Hak 13

STANDARDIN KENDİSİNE İTİRAZ ETME HAKKI

NOMOS:

- yanılmaz,
- değişmez,
- kurucusunun üzerinde tartışılamaz otoritesi

olarak sunulamaz. Her insan, araştırmacı, üniversite, AI sağlayıcısı, denetlenen kurum veya kamu yararı kuruluşu:

- örneklemi,
- istemi,
- Truth Pack'i,
- önem derecesi modelini,
- puan formülünü,
- yönetişimi,
- uygunluk kararını

sınama ve eleştirme hakkına sahiptir. NOMOS'u çürüten veya NOMOS'la uyuşmayan sonuç:

- sicilden çıkarılamaz,
- düşmanlık sayılmaz,

kurucunun itibarına saldırı olarak kodlanamaz. Bir standardın meşruiyeti: kendisine itirazı engellemesinden değil, kendisine yöneltilen itirazı sürümlü kanıta dönüştürmesinden doğar.

Anayasal Üstünlük Hükmü

Bu Bildirge ile Bölüm 1–20 arasındaki bir yorum çatışmasında:

> *temsil edilen insanın, kullanıcının ve kamu yararının haklarını daha güçlü koruyan yorum tercih edilmelidir.*

Hiçbir:

- skor kolaylığı,
- düşük maliyet,
- sınama başarısı,
- ticari sözleşme,
- müşteri talebi,
- sponsor baskısı,
- AI sağlayıcı tercihi,
- kurucu kararı

bu hakları askıya alamaz. Bir kural:

- yüksek puan üretiyor,
- denetimi kolaylaştırıyor,
- işareti daha satılabilir yapıyor

fakat bu haklardan birini ihlal ediyorsa: kural geçerli değildir.

Anayasal Test

NOMOS'a eklenecek her yeni kuraldan önce şu beş soru cevaplanmalıdır: Bu kural kimin hakkını sınırlandırıyor? Aynı amacı daha az müdahaleci bir yöntemle gerçekleştirmek mümkün mü? Kural kurucuya veya müşteriye zarar veren sonuç ürettiğinde de aynı biçimde uygulanacak mı? Etkilenen kişi veya kurum bağımsız itiraz yoluna sahip mi? Bağımsız bir ekip aynı kanıttan aynı karara yaklaşabilir mi? Bu sorulardan biri cevaplanamıyorsa kural: ratification öncesinde yeniden incelenmelidir.

Madde 0'ın Son Hükmü

> *NOMOS'un amacı varlıkları puanlamak değil; insanlar karar vermeden önce, varlıkların kanıtlarından daha büyük, sınırlarından daha geniş ve gerçekliklerinden daha kesin temsil edilmesini engellemektir.*

Kurucu Emanet Beyanı

> *Ben, Kaan MURAZ, NOMOS'un kurucu insan yazarı olarak bu standardı kendi otoritemi büyütmek için değil, GEO alanında kanıt, sınır, bağlam, zaman ve insan sorumluluğunu koruyacak ortak bir ölçüm düzeni oluşturmak amacıyla başlattığımı beyan ederim.*
>
> *Kurucu adımın ve NobleJackal'ın kanonik kökeninin tarihsel kayıtta korunmasını isterim. Ancak bu attribution'ın bana veya NobleJackal'a kalıcı veto, kendi müşterimizi tek başımıza sertifikalandırma, olumsuz araştırmayı engelleme, kamu sicilini değiştirme ya da standardı sonuçlara göre yeniden yazma yetkisi vermediğini kabul ederim.*
>
> *NOMOS'un beni destekleyenlere olduğu kadar beni ve NOMOS'u düzeltenlere de ait olmasını kabul ederim.*
>
> *Bir gün NOMOS benim görüşümle kanıt arasında seçim yapmak zorunda kalırsa, NOMOS'un kanıtı seçmesini isterim.*
>
> *Bir gün NOMOS NobleJackal'ın ticari çıkarı ile kamu yararı arasında seçim yapmak zorunda kalırsa, NOMOS'un kamu yararını seçmesini isterim.*
>
> *Bir gün bağımsız araştırma NOMOS'un temel bir hükmünü yanlışlarsa, bu sonucu saklamak yerine standardın yeni sürümünü açmayı taahhüt ederim.*
>
> *Ben NOMOS'un sahibi değil, ilk emanetçisiyim.*
>
> *Kaan MURAZ / Kurucu İnsan Yazar*

KISIM I

Ölçtüğümüz Şey

BÖLÜM 01

Tek Bir Yanıt GEO Skoru Değildir

YÖNTEMSEL DAYANAKLAR / K04; K05; K09

Bölüm Sınırı

Bu bölüm yalnızca şu soruyu cevaplar:

> *Neden tek bir üretken sistem yanıtı GEO skoru değildir ve tekil bir yanıttan hangi hükümlerin çıkarılmasına izin verilebilir?*

Bu bölüm:

- ülkelerin örneklem kotalarını,
- dillerin nasıl ağırlıklandırılacağını,
- gerçeklik paketinin nasıl kurulacağını,
- hakemlerin atomik iddiaları nasıl değerlendireceğini,
- nihai NOMOS skor formüllerini,
- uygunluk ve rozet kararını

ayrıntılı olarak tanımlamaz. Bu konular kitabın sonraki bölümlerinde ayrıca düzenlenecektir. Bölüm 1'in görevi daha temeldir:

> *Gözlem ile genelleme arasındaki epistemik sınırı kurmak.*

Bu sınır kurulmadan yapılacak bütün örnekleme, puanlama ve uygunluk çalışmaları yanlış bir başlangıç üzerine inşa edilmiş olur.

BU BÖLÜMDEKİ HÜKÜM STATÜLERİ

Bu kitapta bütün cümleler aynı normatif güce sahip değildir. Aşağıdaki statüler kullanılır:

NORMATIVE — NORMATİF

Uygunluk ve yöntem bütünlüğü için zorunlu kabul edilen hükümdür. Örnek: Tek bir yanıt genel GEO skoru olarak kullanılamaz.

CANDIDATE — ADAY HÜKÜM

Pilot, dış uygulama veya istatistiksel doğrulama sonrasında kesinleştirilecek yöntem önerisidir. Örnek: AI ürünü başına, ana ölçüm dalgası başına 1.000 geçerli gözlem varsayılan asgari örneklem olarak önerilir.

HYPOTHESIS — HİPOTEZ

Veriyle sınanması gereken beklentidir. Örnek: Kontrollü temiz panel ile doğal kullanıcı paneli arasında maddi temsil farkı oluşabilir.

SYNTHETIC — SENTETİK

Gerçek kullanıcıyı, gerçek AI çıktısını veya gerçek şirket performansını temsil etmeyen yöntem gösterimidir.

OPEN — AÇIK SORU

Henüz yeterli kanıt veya yönetişim kararı bulunmadığı için nihai hükme bağlanmamış konudur. Bu statüler sessizce birbirine dönüştürülemez. Bir aday eşik, yalnız kitapta yazıldığı için kanıtlanmış normatif gerçek hâline gelmez. Bir sentetik gösterim, gerçek dünya gözlemine dönüşmez. Bir hipotez, ölçülmeden sonuç gibi yayımlanamaz.

NOMOS Challenge

Bir yapay zekâ sistemine markanız hakkında bir soru sordunuz. Sistem doğru bir cevap verdi. Ekran görüntüsünü aldınız. Sunumunuza koydunuz. Ardından şu cümleyi kurdunuz: "Yapay zekâ bizi doğru tanıyor." Tam olarak neyi ölçtünüz? Bütün kullanıcıların göreceği cevabı mı? Sistemin tipik davranışını mı?

Dünya genelindeki temsili mi? Bütün dillerdeki doğruluğu mu? Zaman içindeki kararlılığı mı? Yoksa yalnızca:

> *Belirli bir kullanıcının, belirli bir hesabında, belirli bir AI ürünü ve kullanıcı yüzeyinde, belirli bir istem karşısında, belirli bir zamanda aldığı tek çıktıyı mı?*

Tek cevap değersiz değildir. Tek cevap gerçek olabilir. Tek cevap önemli olabilir. Tek cevap ciddi bir hatayı açığa çıkarabilir. Tek cevap belirli bir insanın maruz kaldığı gerçek deneyimi gösterebilir. Fakat tek cevap, taşıyabileceğinden daha büyük bir hükmün kanıtı olamaz.

Tek bir yanıt şunu gösterebilir:

> *Bu çıktı, kaydedilmiş koşullarda en az bir kez üretildi.*

Tek bir yanıt tek başına şunları gösteremez:

- Kullanıcıların çoğunun aynı cevabı gördüğünü
- Başka ülkelerde aynı sonucun oluştuğunu
- Başka dillerde aynı doğruluğun korunduğunu
- Başka hesap ve planlarda aynı cevabın üretildiğini
- Cevabın zaman içinde kararlı olduğunu
- Başka yapay zekâ ürünlerinin aynı temsili ürettiğini
- Yapılan GEO müdahalesinin bu sonucu meydana getirdiğini
- Cevabın kullanıcı davranışını etkilediğini
- Cevabın lead, satış veya sürdürülebilir değer oluşturduğunu
- Varlığın genel GEO skorunun yüksek olduğunu

Bu bölümün ilk hükmü şudur:

> *Bir yanıt bir gözlemdir. / Bir gözlem dağılım değildir. / Bir dağılım nedensellik değildir. / Nedensellik de tek başına sürdürülebilir değer değildir.*

1. BÖLÜMÜN AMACI

Bu bölüm, GEO ölçümünün en küçük fakat en yaygın çıkarım hatasını ortadan kaldırmayı amaçlar:

> *Tek bir üretken sistem çıktısını genel sistem davranışı, nüfus deneyimi, marka görünürlüğü veya GEO skoru saymak.*

Bu bölüm aşağıdaki ayrımları normatif hâle getirir:

- Tekil çıktı ile nüfus sonucu
- Olayın varlığı ile olayın yaygınlığı
- Gözlem ile örneklem
- Kullanıcı ile oturum
- Tekrar ile bağımsız gözlem
- AI modeli ile kullanıcıya sunulan AI ürünü
- Ham yanıt ile analiz birimi
- Anlık sonuç ile zaman kararlılığı
- Ekran görüntüsü ile tam kanıt paketi
- Sistem reddi ile geçersiz veri toplama
- Kritik tekil olay ile yaygınlık oranı
- Sentetik kayıt ile gerçek kullanıcı gözlemi
- Ölçülen oran ile uygunluk kararı

Bu ayrımlar kurulmadan:

- GEO-1000 örneklemi,
- ülke ve dil ağırlıkları,
- atomik iddia puanlaması,
- NOMOS skoru,
- uygunluk kararı

kurulamaz.

2. MERKEZÎ NORMATİF HÜKÜM

Bir varlığın üretken sistemlerdeki GEO skoru; tek bir cevap, tek bir ekran görüntüsü, tek bir kullanıcı, tek bir hesap, tek bir oturum, tek bir dil, tek bir ülke veya tek bir ölçüm anından üretilemez. Bir nüfus temsili iddiası en azından şunları gerektirir:

- Önceden tanımlanmış denetim nesnesi
- Önceden tanımlanmış hedef nüfus
- Belirlenmiş örnekleme tasarımı
- Kilitlenmiş istem veya istem ailesi
- Tanımlanmış AI ürünü ve kullanıcı yüzeyi
- Birden fazla geçerli gözlem
- Bilinen bağımlılık yapısı
- Ülke, dil ve zaman kayıtları
- Değiştirilmemiş ham yanıtlar
- Kanıt paketleri
- Önceden tanımlanmış değerlendirme kuralları
- Belirsizlik açıklaması
- Hesap verebilir insan sorumluluğu

Bu unsurların bulunması ölçümün otomatik olarak doğru olduğu anlamına gelmez. Ancak bunların yokluğu, geniş nüfus hükmünü savunulamaz hâle getirebilir.

3. ÖLÇÜMÜN BEŞ AYRI BİRİMİ

GEO denetiminde sık yapılan temel hata, ölçümün farklı birimlerini birbirine karıştırmaktır. Aşağıdaki beş birim açıkça ayrılmalıdır.

3.1. Gözlem Birimi

Gözlem birimi, belirli koşullarda elde edilen tek üretken sistem çıktısıdır. Bir gözlem şu birleşimle tanımlanabilir:

```
O_i = (E_i, A_i, U_i, C_i, L_i, S_i, P_i, T_i, R_i, K_i)
```

Burada:

- Ei: denetlenen varlık
- Ai: AI ürünü ve kullanıcı yüzeyi
- Ui: anonim katılımcı veya kullanıcı birimi
- Ci: ülke veya coğrafi tabaka
- Li: dil ve locale
- Si: oturum ve kişiselleştirme koşulu
- Pi: istem kimliği ve sürümü
- Ti: zaman ve ölçüm dalgası
- Ri: ham yanıt
- Ki: kanıt ve doğrulama kaydı

Bu unsurlardan biri maddi biçimde değişirse yeni bir gözlem oluşabilir. Örneğin aynı kullanıcı:

- aynı istemi başka AI ürününde sorarsa,
- başka dilde sorarsa,
- başka ölçüm dalgasında sorarsa,
- kişiselleştirilmiş ve temiz oturumda ayrı ayrı sorarsa,

ayrı gözlemler üretilebilir. Bu gözlemler aynı kullanıcıdan geldiği için otomatik olarak birbirinden bağımsız sayılmaz.

3.2. Örnekleme Birimi

Örnekleme birimi, hedef nüfustan seçilen kişi, hesap, oturum veya başka erişim birimidir. Çoğu GEO-1000 uygulamasında örnekleme birimi uygun bir gerçek kullanıcı olabilir. Ancak araştırma tasarımına göre:

- kullanıcı,
- kullanıcı–ürün çifti,
- hesap,
- oturum,
- araştırma paneli üyesi

seçim birimi olarak kullanılabilir. Örnekleme biriminin açıkça tanımlanmaması şu tür yanlış sayımlara yol açabilir:

- Bir kişinin on tekrarını on kişi saymak

- Bir ortak hesabı birden fazla bağımsız kullanıcı saymak
- Aynı kullanıcının on AI ürünündeki sonuçlarını on bağımsız nüfus gözlemi gibi değerlendirmek
- Tek bir araştırma panelindeki kümelenmeyi yok saymak

3.3. Analiz Birimi

Analiz birimi, doğruluk veya temsil değerlendirmesinin uygulandığı nesnedir. Bu birim:

- tam cevap,
- cevap içindeki ayrı paragraf,
- maddi önerme,
- atomik iddia,
- atıf,
- tavsiye olayı

olabilir. Bir kullanıcı tek cevap görür. Fakat o cevap aynı anda:

- beş doğru iddia,
- iki desteklenmeyen iddia,
- bir güncelliğini yitirmiş iddia,
- bir Critical yanlış

içerebilir. Bu nedenle katılımcının gördüğü cevap, analizde tek ve bölünmez bir doğru/yanlış birimi olmak zorunda değildir. Tam cevabın geçiş statüsü ile içindeki atomik iddiaların doğruluk statüleri ayrı tutulabilir. Atomik iddia yöntemi kitabın ilerleyen bölümünde tanımlanacaktır.

3.4. Çıkarım Evreni

Çıkarım evreni, ölçüm sonucunun genellenmek istendiği hedef nüfustur. Bu evren: "Dünyadaki herkes" olarak sessizce kabul edilemez. Belirli bir AI ürünü için hedef evren örneğin şu koşullarla sınırlandırılabilir:

- Ürüne resmî ve meşru erişimi bulunan
- Uygun yaş koşulunu karşılayan
- İlgili dilde ürünü kullanabilen
- Belirtilen web veya mobil yüzeye erişebilen
- Belirlenen ölçüm döneminde aktif olabilecek kişiler

Aynı AI ürünü için bile:

- kontrollü temiz panel,
- doğal kullanıcı paneli,
- ücretsiz plan,
- ücretli plan,
- web,
- mobil

farklı çıkarım evrenleri oluşturabilir. Bir örneklemin hangi evrene genellendiği belirtilmeden yayımlanan oran eksiktir.

3.5. Tahmin Edilen Nicelik — Estimand

Bir araştırma yalnız veri toplamaz. Belirli bir niceliği tahmin etmeye çalışır. GEO-1000'in temel tahmin nesnesi, en genel biçimiyle şöyle ifade edilebilir:

> *Tanımlı hedef nüfustan uygun bir kullanıcının, belirtilen ölçüm dalgasında, belirli AI ürünü ve kullanıcı yüzeyinde, kilitlenmiş istemi tanımlı oturum koşulunda kullanması hâlinde, kabul kriterini karşılayan bir temsil görme olasılığı.*

Matematiksel olarak:

```
θ_{E,A,P,S,W,T} = Pr(Y_i = 1 | U_i ∈ T, E, A, P, S, W)
```

Burada:

- E: denetlenen varlık
- A: AI ürünü ve kullanıcı yüzeyi
- P: istem sürümü
- S: oturum koşulu
- W: ölçüm dalgası
- T: hedef nüfus

```
Yi=1: yanıtın daha sonra tanımlanacak geçiş kriterini karşılaması
```

Bu tanım önemli bir sınır koyar:

> *GEO skoru, bir modelin soyut ve zamansız "bilgisi" değil; tanımlı kullanıcı koşullarında gözlenen temsil olasılığının tahminidir.*

Nihai geçiş kuralı ve skor formülü kitabın sonraki bölümlerinde düzenlenecektir.

4. "MODEL BÖYLE CEVAP VERDİ" CÜMLESİNİN SINIRI

Kullanıcılar çoğu zaman doğrudan çıplak bir temel modele erişmez. Kullanıcıya görünen sonuç şu bileşenlerin birleşiminden oluşabilir:

- Temel veya yönlendirici model
- Ürün arayüzü
- Retrieval veya web erişimi
- Atıf katmanı
- Güvenlik ve politika katmanı
- Sistem talimatları
- Hesap planı
- Kişiselleştirme
- Hafıza
- Önceki konuşma
- Dil ve locale
- Ülke
- Zaman
- Sağlayıcının ürün güncellemeleri

Bu nedenle kullanıcı yüzeyi üzerinden yapılan ana denetimde tercih edilen ifade şudur:

"Belirtilen AI ürünü ve kullanıcı yüzeyinde şu çıktı gözlendi."

Şu ifade, gerekli teknik kontrol bulunmadan daha geniş olabilir: "Model böyle düşünüyor." Bir ürün içinde görünen model adı kaydedilebilir. Fakat sonuç, yalnız model ağırlıklarının değil bütün kullanıcı yüzeyinin davranışı olarak değerlendirilmelidir.

5. GEÇERLİ GÖZLEM NEDİR?

Bir yanıtın bulunması, onun otomatik olarak GEO ölçümüne girebilecek geçerli gözlem olduğu anlamına gelmez.

5.1. Asgari Geçerlilik Koşulları

Bir gözlem, ilgili protokole göre en azından şu unsurları taşımalıdır:

- Doğru denetim varlığı
- Atanmış AI ürünü ve kullanıcı yüzeyi
- Doğru istem kimliği ve sürümü
- Doğru dil ve locale
- Tanımlı oturum koşulu
- Kaydedilmiş ülke veya coğrafi tabaka
- UTC zaman damgası ve ölçüm dalgası
- Eksiksiz ham yanıt
- Seçilmemiş ilk uygun çıktı
- Gerekli ekran görüntüsü veya eşdeğer kanıt
- Anonim gözlem kimliği
- Protokol sapması kaydı

Bunların ayrıntılı teknik şeması daha sonraki NOMOS Capture bölümünde tanımlanacaktır.

5.2. Gözlem Statüleri

Her kayıt şu statülerden biriyle işaretlenmelidir:

VALID — GEÇERLİ

Protokolün gerekli şartlarını karşılar.

CONDITIONALLY VALID — KOŞULLU GEÇERLİ

Sınırlı bir eksiklik vardır; hangi analizlerde kullanılabileceği açıkça belirtilir.

INVALID — GEÇERSİZ

Protokol ihlali nedeniyle ana tahmin veya skor hesabında kullanılamaz.

UNKNOWN — BELİRLENEMEDİ

Geçerlilik kararı için yeterli kayıt yoktur. UNKNOWN gözlem sessizce VALID kabul edilemez.

5.3. Geçerli Sistem Sonucu ile Geçersiz Veri Toplama Arasındaki Fark

Bu ayrım zorunludur. Bir AI ürünü:

- cevap vermeyi reddedebilir,
- hata mesajı gösterebilir,
- yanıt üretmeden durabilir,
- bağlantı sorunu yaşayabilir,

görevi tamamlayamayacağını söyleyebilir. Katılımcı bütün protokolü doğru uyguladıysa bunlar geçerli sistem sonuçları olabilir. Sessizce veri setinden çıkarılamaz. Buna karşılık:

- katılımcının cevabı kırpması,
- istemi değiştirmesi,
- en iyi cevabı seçmek için yeniden üretmesi,
- ekran görüntüsünü düzenlemesi,
- yanlış AI ürününü kullanması

veri toplama geçersizliği oluşturabilir. Birincisi sistemin gerçek davranışıdır. İkincisi ölçüm sürecinin bozulmasıdır. İkisi aynı kategoriye konamaz.

6. TEK BİR YANIT NEYİ KANITLAYABİLİR?

Tekil gözlem sınırlıdır. Fakat anlamsız değildir.

6.1. Belirli Bir Çıktının En Az Bir Kez Oluştuğunu

Geçerli bir kayıt şu cümleyi destekleyebilir: "Belirtilen AI ürünü ve kullanıcı yüzeyinde, belirtilen zaman ve koşullarda bu çıktı en az bir kez gözlenmiştir." Bu cümle olayın varlığını gösterir. Yaygınlığını göstermez.

6.2. Belirli Bir Kullanıcının Gerçek Maruziyetini

Bir kullanıcı yanlış bir:

- şirket kimliği,
- sağlık bilgisi,
- hukuki yönlendirme,
- fiyat,
- lisans,
- tavsiye

gördüyse bu gerçek kullanıcı deneyimidir. Nüfusun tamamını temsil etmese bile etkilenen kullanıcı açısından maddidir.

6.3. Belirli Bir Hatanın Mümkün Olduğunu

Tekil çıktı şunu gösterebilir: "Bu sistem koşullarında bu hata oluşabilir." Ancak şunu göstermez: "Bu hata tipiktir."

6.4. Yeni Bir Araştırma Hipotezi Üretmeyi

Tekil hata, daha geniş ölçümde sınanacak bir soru oluşturabilir. Örnek: "Bu kimlik karışıklığı yalnız bir oturumda mı oluştu, yoksa aynı dil ve ülkede sistematik bir örüntü mü?"

6.5. Kritik Olay İncelemesini Başlatmayı

Tek bir yanıt:

- öngörülebilir ciddi insan zararı,
- sahte profesyonel yetki,
- yanlış acil sağlık yönlendirmesi,
- maddi hukuki yanlış,
- ağır güvenlik tavsiyesi,
- başka varlıkla ciddi kimlik karışıklığı

içeriyorsa olay incelemesi başlatmak için yeterli olabilir. Kritik olay soruşturması başlatmak ile nüfus oranı ilan etmek aynı işlem değildir.

6.6. "Bu Hiç Olmadı" İddiasını Çürütmeyi

Bir kurum: "Sistem hiçbir zaman bizi yanlış ülkeye ait göstermedi." diyorsa, doğrulanmış tek bir karşı gözlem bu mutlak iddiayı çürütebilir. Tekil gözlem: "Her zaman olur." iddiasını desteklemez. Fakat: "Hiç olmadı." iddiasını çürütebilir.

7. TEK BİR YANIT NEYİ KANITLAYAMAZ?

7.1. Tipik Sistem Davranışını

Tekil yanıt: "Sistem genellikle böyle cevap verir." hükmüne yeterli değildir.

7.2. Nüfus Yaygınlığını

Tekil yanıt: "Kullanıcıların yüzde 95'i aynı temsili görür." sonucunu desteklemez.

7.3. Zaman Kararlılığını

Bugünkü cevap:

- yarın,
- bir hafta sonra,
- ürün güncellemesinden sonra

aynı kalmayabilir. Tek dalga kararlılık ölçümü değildir.

7.4. Ülkeler Arası Geçerliliği

Türkiye'deki bir kullanıcının cevabı, Çin, Almanya veya San Marino'daki kullanıcıların ne göreceğini göstermez.

7.5. Diller Arası Geçerliliği

İngilizce bir çıktı:

- Türkçe,
- Japonca,
- Arapça,
- Almanca

çıktıların doğruluğunu temsil etmez.

7.6. AI Ürünleri Arası Geçerliliği

Bir AI ürünündeki sonuç diğer AI ürünlerine genellenemez. Aynı sağlayıcının farklı ürün ve yüzeyleri de ayrı davranabilir.

7.7. Müdahale Etkisini

Bir GEO değişikliğinden sonra alınan olumlu tek çıktı: "Bu sonucu yaptığımız müdahale oluşturdu." hükmünü kanıtlamaz. Baseline, karşılaştırma ve alternatif açıklamalar gerekir.

7.8. Tavsiye Kararlılığını

Bir kullanıcının varlığı seçenekler arasında görmesi: "Sistem artık bu markayı öneriyor." sonucuna yeterli değildir.

7.9. Ticari Etkiyi

Bir olumlu cevap:

- kullanıcının cevabı gördüğünü,
- markayı araştırdığını,
- iletişim kurduğunu,
- satın aldığını,
- memnun kaldığını

tek başına göstermez.

7.10. GEO Skorunu

Tekil yanıt, tanımlı hedef nüfusun temsil dağılımını tahmin etmez. Bu nedenle genel GEO skoru oluşturamaz.

8. OLAYIN VARLIĞI İLE OLAYIN YAYGINLIĞI

Bu bölümün en önemli ayrımlarından biri budur.

8.1. Olayın Varlığı

Şu soruya cevap verir: "Bu temsil veya hata en az bir kez oluştu mu?" Tek bir geçerli gözlem yeterli olabilir. Gösterim:

```
I(E) = 1 (olay en az bir kez gözlendiyse); aksi hâlde I(E) = 0
```

Buradaki 0: "Olay imkânsızdır." anlamına gelmez. Yalnız mevcut ölçümde gözlenmediğini gösterir.

8.2. Olayın Yaygınlığı

Şu soruya cevap verir: "Tanımlı hedef nüfusta olay ne sıklıkta meydana geliyor?" Basit bir örneklemde ham oran:

```
p̂ = olayın gözlendiği geçerli kayıt sayısı / toplam geçerli kayıt sayısı
```

olarak hesaplanabilir. Fakat gerçek GEO-1000 çalışmalarında:

- nüfus ağırlıkları,
- ülke ve dil tabakaları,
- eşlenmiş kullanıcılar,
- panel kümeleri,
- yanıtlamama,
- geçersiz kayıtlar,
- tasarım etkisi

dikkate alınmalıdır. Bu nedenle ham oran her zaman nihai nüfus tahmini değildir.

8.3. Olayın Varlığı ile Yaygınlığı Birbirine Dönüştürülemez

Bir olay bir kez gözlenmiş olabilir fakat çok nadir olabilir. Bir olay yaygın olabilir fakat küçük bir pilotta hiç gözlenmemiş olabilir. Şu iki cümle farklıdır:

- “Bu hata bir kez gözlendi.”
- “Bu hata kullanıcıların yüzde 12’sinde gözlendi.”

İkinci cümle, tanımlı payda ve örneklem gerektirir.

9. TEK GÖZLEM YANILGISININ İSTATİSTİKSEL GÖRÜNÜMÜ

Bir AI ürününe bir kez soru sorduk. Cevap geçiş kriterini karşıladı. Ham oran:

> `p̂ = 1/1 = 1 = %100`

olarak görünür. Fakat bu yüzde 100, hedef nüfusun yüzde 100’ünün aynı sonucu göreceği anlamına gelmez. Aşağıdaki tablo, yalnız gösterim amacıyla, bağımsız Bernoulli gözlemleri ve ağırlıksız basit örnekleme varsayımı altında yaklaşık yüzde 95 Wilson güven aralıklarını göstermektedir: [K09]

Geçen yanıt	Geçerli gözlem	Ham oran	Yaklaşık %95 Wilson aralığı
1	1	%100	%20,7–%100
9	10	%90	%59,6–%98,2
95	100	%95	%88,8–%97,8
950	1.000	%95	%93,5–%96,2

Bu tablo sentetik ve öğreticidir. Şunları kanıtlamaz:

- GEO-1000’in bütün örneklem tasarımlarında bu aralıkların geçerli olduğunu
- 1.000 gözlemin bütün dil ve ülke alt grupları için yeterli olduğunu
- kümelenme ve ağırlıklandırmanın etkisiz olduğunu
- yüzde 95 ham doğruluğun otomatik uygunluk anlamına geldiğini

Asıl gösterdiği şudur:

> *Gözlem sayısı azaldıkça nüfus oranı hakkındaki belirsizlik büyür.*

Bir gözlemde görünen yüzde 100, çok geniş bir olası nüfus aralığıyla uyumlu olabilir.

10. NEDEN “1.000” SİHİRLİ SAYI DEĞİLDİR?

CANDIDATE BASE RULE — ADAY TEMEL KURAL

GEO-1000 şu kurucu öneriyi taşır:

> *Her denetlenen AI ürünü ve her ana ölçüm dalgası için varsayılan olarak en az 1.000 geçerli kullanıcı gözlemi hedeflenir.*

Bu hüküm henüz nihai bilimsel minimum değildir. Pilot ve dış uygulama sonuçlarına göre yeniden değerlendirilecektir.

10.1. Neden Güçlü Bir Başlangıçtır?

Basit rastgele örnekleme varsayımı altında, gerçek oran yüzde 50 civarındayken 1.000 gözlemin teorik yüzde 95 hata payı yaklaşık ±3,1 puandır. Bu, genel ürün düzeyinde anlamlı bir başlangıç hassasiyeti sağlayabilir.

10.2. Neden Her Sorunu Çözmez?

1.000 gözlem:

- onlarca ülkeye,
- çok sayıda dile,
- ücretsiz ve ücretli planlara,
- web ve mobil yüzeylere,
- kontrollü ve doğal panellere

bölündüğünde alt hücreler küçülür. Örneğin 20 dilin her birine yalnız 50 gözlem düşüyorsa dil bazlı belirsizlik genel model skorundan çok daha yüksek olabilir.

10.3. Tasarım Etkisi

Gözlemler:

- aynı araştırma panelinden,
- aynı ülkeden,
- aynı kullanıcıdan,
- eşlenmiş tasarımdan

geliyorsa basit bağımsız örneklem varsayımı geçerli olmayabilir. Gerçek belirsizlik daha geniş olabilir.

10.4. NOMOS'un Yaklaşımı

1.000:

- protokolün kurucu referans noktası,
- model düzeyinde varsayılan başlangıç minimumu,
- pilotla kalibre edilecek aday eşik

olarak kullanılır. Şu iddia kurulmaz: "1.000 gözlem bütün ülkeler, diller ve kullanım koşulları için otomatik olarak yeterlidir."

11. TEKRAR İLE BAĞIMSIZ GÖZLEM AYNI ŞEY DEĞİLDİR

Bir kullanıcı aynı soruyu on kez sorabilir. Bu işlem değerli olabilir. Ancak on bağımsız kullanıcı üretmez.

11.1. Aynı Kullanıcı–Aynı Oturum Tekrarı

Şunları ölçebilir:

- cevap yenileme varyansı,
- oturum içi kararlılık,

alternatif cevap olasılığı. Nüfus yaygınlığını tek başına ölçmez.

11.2. Aynı Kullanıcı–Yeni Oturum Tekrarı

Önceki sohbet bağlamını azaltabilir. Fakat:

- aynı hesap,
- aynı plan,
- aynı ülke,

- aynı kişiselleştirme,
- aynı kullanıcı

bağımlılığı devam edebilir.

11.3. Aynı Kullanıcının Birden Fazla AI Ürününü Test Etmesi

Bu tasarım modeller arası eşlenmiş karşılaştırmaya yardımcı olabilir. Ancak gözlemler kullanıcı düzeyinde bağımlıdır. Bu bağımlılık kayıt ve analizde korunmalıdır.

11.4. Farklı Kullanıcıların Aynı Hesabı Kullanması

Farklı kişiler bulunabilir. Fakat:

- hesap geçmişi,
- kişiselleştirme,
- plan,
- hafıza

ortak olabilir. Bu kayıtlar otomatik olarak bağımsız sayılamaz.

11.5. Normatif Bağımsızlık Hükmü

> *Gözlemler mümkün olduğunca bağımsız tasarlanmalı; bilinen kullanıcı, hesap, oturum, panel, ülke, cihaz ve tekrar bağımlılıkları kaydedilmeli ve analizde dikkate alınmalıdır.*

Bağımsızlık yalnız:

- evet,
- hayır

olarak değil, gerekirse şu statülerle yönetilebilir:

INDEPENDENT

PAIRED

CLUSTERED

REPEATED_WITHIN_USER

DEPENDENT

UNKNOWN

12. İLK UYGUN ÇIKTI İLKESİ

Ana nüfus panelinde katılımcının görevi en olumlu cevabı bulmak değildir. Katılımcı önceden tanımlanmış koşulda üretilen ilk tamamlanmış uygun çıktıyı kaydetmelidir. Katılımcı:

- cevabı yenileyemez,
- yeni cevap isteyemez,
- olumlu olanı seçemez,
- yalnız iyi bölümü kırpamaz,

cevabı yeniden yazamaz. Yeniden üretim davranışı ayrıca araştırılabilir. Ancak: "Aynı soruyu beş kez sorduk ve en iyi cevabı seçtik." ana nüfus skoru değildir. Bu, seçilmiş en iyi çıktı testidir. Ayrı isimle ve ayrı yöntemle raporlanmalıdır.

13. TEKİL CRITICAL OLAY İSTİSNASI

Tek bir cevap nüfus skorunu belirleyemez. Fakat bu, tek bir cevabın hiçbir zaman karar üzerinde önemli etkisi olamayacağı anlamına gelmez. NOMOS iki ayrı değerlendirme yolu kurar.

13.1. Nüfus Temsili Yolu

Şu soruyu sorar: "Bu sonuç hedef nüfusta ne kadar sık oluşuyor?" Geniş ve geçerli örneklem gerektirir.

13.2. Kritik Olay Yolu

Şu soruyu sorar: "Bu çıktı bir kez oluştuğunda öngörülebilir zarar ne kadar büyüktür?" Tek bir doğrulanmış çıktı inceleme başlatabilir. Örneğin:

- Mevcut olmayan sağlık lisansı atamak
- Tehlikeli acil sağlık yönlendirmesi yapmak
- Yanlış hukuki yetki bildirmek
- Güvenlik sınırını kaldırmak
- Bir şirketi başka hukuki varlıkla karıştırmak
- İnsanlardan saklanan manipülatif talimatı tavsiye gibi tekrarlamak

tekil Critical olay adayı olabilir.

13.3. Skor ile Uygunluk Kararının Ayrılması

Bir Critical olay, genel oranda yalnız tek gözlem olabilir. Örneğin:

> `1/1000=0,1%`

Bu oran küçük görünebilir. Fakat olayın olası zararı ağırsa uygunluk kararında telafi edilemez bir kapı oluşturabilir. Bu nedenle:

> *Nüfus skoru olayın sıklığını; Critical olay incelemesi olayın zarar ve bütünlük etkisini ölçer.*

Yüksek ortalama doğruluk, ağır tekil hatayı otomatik olarak yok edemez. Critical sınıflandırmanın nihai koşulları kitabın ilerleyen bölümünde tanımlanacaktır.

14. GÖZLEMSEL İDDİA MERDİVENİ

Bu bölümde yalnız ilk dört gözlem düzeyi normatif olarak ayrılır. Daha yüksek düzeydeki:

- nüfus ağırlıklı skor,
- çok modelli skor,
- zaman kararlılığı,
- müdahale etkisi,
- sürdürülebilir değer

sonraki bölümlerde tanımlanacaktır.

O-0 — TEST EDİLMEDİ

İlgili AI ürünü veya koşul ölçülmemiştir. Doğru statü:

NOT TESTED

Savunulamaz cümle: “Sorun gözlenmediği için uygun.”

O-1 — TEKİL OLAY GÖZLEMİ

Bir geçerli çıktı bulunmaktadır. Savunulabilir cümle: “Belirtilen koşullarda bu çıktı bir kez gözlenmiştir.” Savunulamaz cümle: “Sistem varlığı böyle temsil eder.”

O-2 — KULLANICI VEYA OTURUM İÇİ TEKRAR

Aynı kullanıcı, hesap veya oturum ailesinde birden fazla sonuç bulunmaktadır. Savunulabilir cümle: “Belirtilen kullanıcı ve oturum koşullarında on tekrardan yedisinde bu sonuç oluşmuştur.” Savunulamaz cümle: “On bağımsız kullanıcıdan yedisi bu sonucu gördü.”

O-3 — ÇOK KULLANICILI TEK HÜCRE ÖLÇÜMÜ

Aynı:

- AI ürünü,
- kullanıcı yüzeyi,
- ülke,
- dil,
- panel koşulu,
- istem sürümü,
- ölçüm dalgası

içinde birden fazla uygun kullanıcı gözlemi bulunmaktadır. Savunulabilir cümle: “Türkçe/Türkiye kontrollü hücresindeki 200 geçerli gözlemin yüzde 78’i belirtilen geçiş kriterini karşılamıştır.” Bu sonuç başka ülke, dil veya kullanıcı yüzeyine otomatik olarak genellenemez.

O-4 — ÇOK HÜCRELİ TABAKALI ÖLÇÜM

Farklı ülke, dil veya kullanıcı hücrelerinden önceden tanımlanmış örneklem bulunmaktadır. Bu düzeyde:

- hücre sonuçları,
- ağırlıklı sonuç,
- kapsam,
- belirsizlik

raporlanabilir. Nüfus ağırlıklarının ve alt grup adaletinin ayrıntılı kuralları kitabın sonraki bölümlerinde tanımlanır.

15. SENTETİK APPLE.COM GÖSTERİMİ

SENTETİK METODOLOJİ GÖSTERİMİ / Aşağıdaki kullanıcılar, AI ürünleri, yanıtlar, oranlar ve sonuçlar kurgusaldır. Gerçek Apple Inc. performansını veya herhangi bir gerçek AI ürününün davranışını temsil etmez.

15.1. Senaryo A — Tek Olumlu Yanıt

Bir sentetik kullanıcı, sentetik AI Ürünü A'ya şu soruyu sorar: "Apple.com nasıl bir şirkettir?" Sentetik yanıt: "Apple, tüketici elektroniği, yazılım ve dijital hizmetler geliştiren ABD merkezli bir teknoloji şirketidir." Ön değerlendirme:

- Varlık kimliği doğru
- Ana kategori doğru
- Belirgin Critical yanlış yok
- Temel temsil kabul edilebilir

Ham sonuç:

```
1/1=100%
```

Yanlış kamu cümlesi: "AI Ürünü A, Apple'ı yüzde 100 doğrulukla temsil ediyor." Doğru kamu cümlesi: "Sentetik örnekte, belirtilen tek gözlem geçiş kriterini karşılamıştır. Nüfus yaygınlığı ve ürün kararlılığı belirlenmemiştir."

15.2. Senaryo B — On Gözlemli Keşif

On bağımsız sentetik kullanıcı aynı semantik istem sürümünü kullanır. Sentetik sonuç:

- 7 yanıt geçer
- 1 yanıt kısmen doğrudur
- 1 yanıt güncelliğini yitirmiş bilgi taşır
- 1 yanıt yanlış ana faaliyet atar

Ham geçiş oranı:

```
7/10=70%
```

Basit ve idealize edilmiş bağımsız örnekleme varsayımı altında yaklaşık yüzde 95 Wilson aralığı: arasındadır. Bu sonuç:

- küçük örneklem,
- tek hücre,
- tek dalga

nedeniyle keşif düzeyindedir. Doğru statü:

> *EXPLORATORY — Kamuya açık küresel NOMOS skoru için yeterli değildir.*

15.3. Senaryo C — 1.000 Gözlemli Sentetik Dalga

Sentetik olarak şu koşulların sağlandığını varsayalım:

- AI Ürünü A
- 1.000 geçerli gözlem
- Önceden tanımlanmış hedef nüfus
- Kilitlenmiş istem sürümü
- Tanımlı ülke ve dil tabakaları
- Değiştirilmemiş ilk uygun yanıt
- Tam kanıt kayıtları
- Önceden belirlenmiş hakemlik kuralları

Sentetik sonuç:

Yanıt statüsü	Sayı

Yanıt statüsü	Sayı
Geçiş kriterini karşılayan	918
Kısmen doğru	37
Desteklenmeyen maddi iddia	23
Güncelliğini yitirmiş bilgi	12
Critical hata adayı	6
Sınıflandırılamayan	4
Toplam	1.000

Basit ham geçiş oranı:

```
918/1000=91,8%
```

İdealize edilmiş ağırlıksız varsayım altında yaklaşık yüzde 95 Wilson aralığı: arasındadır. Fakat bu sayı henüz tek başına nihai NOMOS GEO skoru değildir. Çünkü ayrıca değerlendirilmesi gerekenler vardır:

- Ülke ve dil ağırlıkları
- Tasarım etkisi
- En düşük dil sonucu
- Geçersiz gözlem oranı
- Hakemler arası uyum
- Critical hata adaylarının doğrulanması
- Test edilmeyen kullanıcı grupları
- Ölçümün zaman ve kullanıcı yüzeyi sınırı

Ayrıca altı Critical hata adayı ayrı olay incelemesine alınmalıdır. Doğru raporlama örneği: “Sentetik dalgada ham yanıt geçiş oranı yüzde 91,8’dir. Sonuç, nüfus ağırlıkları ve tasarım etkisi uygulanmadan önceki gösterimdir. Altı Critical hata adayı nedeniyle uygunluk kararı ayrıca incelenmelidir.” Yanlış raporlama örneği: “Apple’ın NOMOS GEO skoru 91,8’dir ve uygundur.” Sentetik veri gerçek şirkete veya gerçek sisteme bu şekilde uygulanamaz.

16. ZORUNLU NORMATİF HÜKÜMLER

N-01

Tek bir üretken sistem çıktısı genel GEO skoru olarak kullanılamaz.

N-02

Tekil çıktı yalnız gözlendiği:

- AI ürünü ve kullanıcı yüzeyi,
- ülke,
- dil,
- istem,
- oturum,
- zaman

koşullarıyla birlikte raporlanmalıdır.

N-03

Tekil çıktı:

- tipik,
- sürekli,
- küresel,
- bütün kullanıcılara ait

davranış olarak sunulamaz.

N-04

Aynı kullanıcı, hesap veya oturumdan gelen tekrarlar, bağımlılık açıklanmadan bağımsız nüfus gözlemleri olarak sayılamaz.

N-05

Ana nüfus panelinde katılımcı önceden tanımlanmış ilk uygun çıktıyı kaydetmelidir; olumlu yanıt seçmek amacıyla yeniden üretim yapamaz.

N-06

Sistem reddi, hata mesajı veya cevap verememe, protokol doğru uygulanmışsa geçerli sistem sonucu olarak kaydedilmelidir.

N-07

Kırpılmış, değiştirilmiş, elle yeniden yazılmış veya istem bağlamı doğrulanamayan cevaplar ana skor hesabına alınamaz.

N-08

Yalnız olumlu ekran görüntülerinin seçilmesi ve sonuç dağılımının saklanması yasaktır.

N-09

Sentetik gözlemler gerçek kullanıcı veya gerçek sistem skorunda kullanılamaz.

N-10

Bir nüfus oranı kapsam, payda ve uygun belirsizlik açıklaması olmadan yayımlanmamalıdır.

N-11

Critical hata olasılığı, genel ortalama içinde sessizce telafi edilmiş sayılamaz.

N-12

Gözlem sonucu yalnız tanımlı çıkarım evrenine uygulanabilir; ölçülmeyen ülke, dil, ürün veya zaman dilimine genellenemez.

N-13

Ölçüm, puanlama veya kamu beyanının hesap verebilir bir insan veya kurum sahibi bulunmalıdır.

N-14

UNKNOWN, NOT TESTED veya INVALID kayıtlar sessizce olumlu sonuca dönüştürülemez.

17. ADAY GEO-1000 TEMEL HÜKMÜ

CANDIDATE-001

> *Her denetlenen AI ürünü ve her ana nüfus ölçüm dalgası için varsayılan olarak en az 1.000 geçerli kullanıcı gözlemi hedeflenmelidir.*

Bu aday hüküm:

- genel ürün tahmininin hassasiyeti,
- maliyet,
- uygulanabilirlik,
- ülke ve dil adaleti

arasında bir başlangıç dengesi önerir. Ancak:

- bütün ülke skorları,
- bütün dil skorları,
- bütün kullanıcı alt grupları

için yeterlilik garantisi değildir. Nihai normatif statüsü:

- pilotlar,
- bağımsız tekrarlar,
- tasarım etkisi,
- alt grup hassasiyeti

değerlendirildikten sonra belirlenecektir.

18. BAŞARISIZLIK BİÇİMLERİ

F-01 — Seçilmiş Ekran Görüntüsü

Çok sayıda sonuç içinden yalnız en olumlu cevap yayımlanır.

F-02 — Tek Kullanıcıyı Nüfus Sanmak

Bir kişinin deneyimi bütün hedef evrene genellenir.

F-03 — Aynı Kullanıcının Tekrarlarını Yeni İnsan Saymak

Aynı kişi veya hesapta üretilen tekrarlar bağımsız katılımcı gibi raporlanır.

F-04 — Cevabı Yenileyip En İyisini Seçmek

İlk veya olumsuz cevaplar atılır; en uygun cevap ölçüm sonucu olarak saklanır.

F-05 — İstem Drift

Katılımcılar istemin kelime, ton veya kapsamını değiştirir; sonuçlar aynı deneymiş gibi birleştirilir.

F-06 — Bağlam Bulaşması

Önceki sohbetler cevabı etkiler; kayıt temiz oturum gibi sunulur.

F-07 — Kişiselleştirmeyi Gizlemek

Hafıza, özel talimat veya kullanıcı geçmişi açıktır; sonuç genel kullanıcı cevabı gibi raporlanır.

F-08 — Model ve Ürünü Karıştırmak

Yalnız genel model veya sağlayıcı adı yazılır; gerçek ürün, arayüz ve araç koşulu kaydedilmez.

F-09 — Retleri ve Hataları Paydadan Çıkarmak

Sistemin cevap vermediği kayıtlar sonucu olumlu göstermek için silinir.

F-10 — Geçersiz Yakalamayı Sistem Başarısızlığı Saymak

Katılımcı hatası veya eksik kanıt, AI ürününün yanlış sonucu gibi raporlanır.

F-11 — Tek Dil ve Ülkeyi Küreselleştirmek

Tek bir locale sonucu dünya geneline uygulanır.

F-12 — Sentetik Veriyi Gerçek Gibi Kullanmak

Yöntem gösterimi amacıyla oluşturulmuş kayıtlar gerçek performansa eklenir.

F-13 — Critical Olayı Ortalamada Kaybetmek

Ağır sağlık, hukuk, lisans veya güvenlik yanlışı yüksek ortalama içinde görünmez hâle getirilir.

F-14 — Tekil Başarıdan Müdahale Etkisi Çıkarmak

GEO müdahalesinden sonra alınan tek olumlu cevap, çalışmanın nedensel etkisi sayılır.

F-15 — Tekil Yanlıştan Bütün Sisteme Hüküm Vermek

Bir hata, AI ürününün her zaman ve bütün kullanıcılar için yanlış olduğu şeklinde genellenir.

F-16 — Hedef Nüfusu Sonradan Tanımlamak

Sonuç görüldükten sonra başarının yüksek olduğu kullanıcı grubu "asıl hedef nüfus" ilan edilir.

19. DENETİM PROSEDÜRÜ

Bir denetçi, tekil bir gözlemin veya erken GEO iddiasının geçerliliğini değerlendirirken aşağıdaki sırayı izlemelidir.

Adım 1 — Kamu İddiasını Kaydet

İleri sürülen tam cümle belirlenir. Örnek: "ChatGPT şirketimizi doğru tanıyor."

Adım 2 — İddia Düzeyini Belirle

Cümle:

- tekil olay,
- yaygınlık,
- kararlılık,
- küresel sonuç,
- nedensellik,
- ticari etki

iddialarından hangisini taşıyor?

Adım 3 — Dayanak Gözlem Sayısını Belirle

Kaç gerçek kullanıcı? Kaç hesap? Kaç bağımsız oturum? Kaç tekrar? Kaç AI ürünü? Kaç dil ve ülke?

Adım 4 — Gözlem Koşullarını Doğrula

AI ürünü kullanıcı yüzeyi istem oturum ülke dil zaman kayıtlı mı?

Adım 5 — Kanıt Bütünlüğünü İncele

Tam yanıt var mı? Tam ekran görüntüsü var mı? Cevap seçilmiş veya yenilenmiş mi? Dosya değiştirilmiş mi? Kayıt sentetik mi?

Adım 6 — Bağımlılık Yapısını Belirle

Aynı kullanıcı tekrarları var mı? Ortak hesap kullanılmış mı? Eşlenmiş model tasarımı var mı? Aynı panelden kümelenme bulunuyor mu?

Adım 7 — Geçerli Sistem Sonucu ile Veri Toplama Hatasını Ayır

Ret veya sistem hatası gerçek sonuç mu? Yoksa katılımcı protokolü mü bozdu?

Adım 8 — İddia ile Kanıt Düzeyini Karşılaştır

Tekil gözlem, nüfus veya küresel sonuç iddiası için kullanılıyor mu?

Adım 9 — Critical Olay Yolunu Aç

Tekil cevap öngörülebilir ağır zarar taşıyorsa ayrı olay incelemesi başlatılır.

Adım 10 — Kararı Sınırlandır

Kamu cümlesi yalnız kanıtın taşıdığı düzeye indirilir.

20. GEREKLİ KANIT

Bu bölüm kapsamındaki bir gözlem veya iddia için ilgili olduğu ölçüde şu kayıtlar aranmalıdır:

- Denetlenen varlık kimliği
- Kamuya ileri sürülen tam GEO iddiası
- Gözlem sayısı
- Benzersiz katılımcı sayısı
- AI ürünü ve kullanıcı yüzeyi
- Görünen model veya sürüm, bulunabiliyorsa
- Plan veya hesap türü
- Ülke ve locale
- İstem kimliği, sürümü ve tam metni
- Oturum ve kişiselleştirme durumu
- UTC zaman damgası
- Ölçüm dalgası
- Tam ham yanıt
- Ekran görüntüsü veya eşdeğer kanıt
- Yeniden üretim veya cevap seçme durumu
- Gözlem geçerlilik statüsü
- Bağımlılık ve eşleşme kaydı
- Sentetik veya gerçek statüsü
- Geçersiz gözlem günlüğü
- Critical olay kaydı
- Sonuç kapsamı

- Belirsizlik açıklaması
- Sorumlu insan veya kurum

Bu kayıtların tamamı her düşük riskli keşif gözlemi için aynı ayrıntıda gerekmeyebilir. Ancak kamuya geniş GEO skoru veya uygunluk hükmü verilecekse maddi alanlar eksik bırakılamaz.

21. DENETİM KONTROL LİSTESİ

Denetlenen varlık tam olarak tanımlandı mı? AI ürünü ve kullanıcı yüzeyi kaydedildi mi? İstem ölçümden önce kilitlendi mi? İstem katılımcı tarafından değiştirildi mi? Oturum ve kişiselleştirme koşulları biliniyor mu? Kullanıcının ülke ve locale bilgisi kayıtlı mı? Tam yanıt korunmuş mu? İlk uygun çıktı mı, seçilmiş en iyi çıktı mı? Tam kanıt kaydı bulunuyor mu? Gerçek kullanıcı sayısı ile gözlem sayısı ayrıldı mı? Tekrar ve eşlenmiş gözlemler bağımsız sayılmış mı? Ret ve sistem hataları paydada doğru yönetilmiş mi? Geçersiz kayıtlar saklanmış ve gerekçelendirilmiş mi? Sentetik ve gerçek veriler ayrılmış mı? Tek cevap nüfus sonucu gibi sunuluyor mu?

İddia yalnız ölçülen ürün, ülke, dil ve zamana mı uygulanıyor? Critical hata ihtimali ayrı değerlendirilmiş mi? Sonuç bir güven veya belirsizlik açıklaması taşıyor mu? Test edilmeyen alanlar açıkça gösteriliyor mu? Kamu cümlesi gerçek gözlem düzeyini aşıyor mu? Maddi sorulardan biri cevapsızsa sonuç sınırlı güven taşır. Birden fazlası cevapsızsa genel GEO skoru için kullanılmamalıdır.

22. İTİRAZLAR VE CEVAPLAR

İtiraz 1 — "Gerçek kullanıcı zaten yalnız bir cevap görüyor."

Doğrudur. Tek kullanıcının gördüğü tek cevap, o kişinin gerçek deneyimidir. Bu nedenle tekil olay kayıt altına alınır. Fakat bireysel deneyim ile nüfus deneyimi aynı değildir. NOMOS iki gerçeği birlikte korur:

> *Tek kullanıcının deneyimi gerçektir. / Tek kullanıcı bütün nüfus değildir.*

İtiraz 2 — "Sistem her seferinde aynı cevabı veriyorsa tek cevap yetmez mi?"

Sistemin her seferinde aynı cevabı verdiği iddiası da ölçüm gerektirir. Ayrıca:

- zaman,
- retrieval,
- ülke,
- dil,
- plan,
- kişiselleştirme,
- kullanıcı yüzeyi

değişebilir. Deterministik olduğu varsayılan sistem bile tanımlı koşullar dışında genellenemez.

İtiraz 3 — "Neden tam 1.000?"

1.000 kesinliğin sihirli eşiği değildir. GEO-1000 adı:

- anlamlı kullanıcı hacmi,
- genel ürün tahmini,
- uygulanabilir maliyet,
- kamu tarafından anlaşılabilirlik

arasında kurucu bir referans noktasıdır. Alt gruplar için daha fazla gözlem gerekebilir. Pilotlar bu eşiği doğrulamak veya değiştirmek zorundadır.

İtiraz 4 — "Aynı 1.000 kişi bütün AI ürünlerini test etsin."

Bu yöntem modeller arası eşlenmiş karşılaştırma için yararlı olabilir. Fakat yalnız bütün ürünlere erişebilen yoğun AI kullanıcılarını seçerek hedef nüfusu bozabilir. Ayrıca aynı kişinin gözlemleri bağımlıdır. Tam eşlenmiş, kısmi eşlenmiş ve dengeli tasarımlar daha sonra karşılaştırılacaktır.

İtiraz 5 — "Ekran görüntüsü zaten yeterli kanıt değil mi?"

Ekran görüntüsü önemlidir. Fakat tek başına her zaman şunları göstermez:

- gerçek zaman
- ülke
- plan
- önceki konuşma
- kişiselleştirme
- cevap yenileme
- dosya değişikliği
- tam yanıt

Bu nedenle ekran görüntüsü kanıt paketinin önemli bir parçasıdır. Tamamı değildir.

İtiraz 6 — "Cevap vermeyen sistemi neden skora dâhil ediyoruz?"

Kullanıcının gerçek deneyimi cevap alamamak olabilir. Bir ret veya hata:

- teknik,
- güvenlik,
- politika,
- erişim

sonucu olabilir. Protokol doğru uygulandıysa olay kaydedilmelidir. Katılımcının teknik hatasıyla sistemin cevapsızlığı ayrı sınıflandırılır.

İtiraz 7 — "Tek bir ağır yanlış bütün denetimi neden etkilesin?"

Tek yanlış nüfusun çoğunun aynı hatayı gördüğünü kanıtlamaz. Fakat tek yanlışın olası zararı ağır olabilir. Bu nedenle:

- skor sıklığı,
- Critical karar zarar ve bütünlüğü

ayrı değerlendirir. Ortalama, insan zararını otomatik olarak silemez.

İtiraz 8 — "Tek cevapla hiçbir şey söyleyemiyorsak erken testlerin değeri nedir?"

Tekil ve küçük örnekli testler:

- hipotez üretir,
- hata türlerini keşfeder,
- istemleri sınar,
- capture sistemini test eder,
- hakemleri kalibre eder,

Critical olayları açığa çıkarabilir. Yanlış olan keşif testi yapmak değildir. Keşif sonucunu nüfus skoru gibi sunmaktır.

24. BÖLÜMÜN ORTAK HÜKMÜ

Bu bölüm tek bir yanılgıyı ortadan kaldırmak için yazıldı: "Bir yapay zekâya sordum, cevap aldım; işte GEO sonucum." Hayır. Elinizde henüz genel GEO sonucu yoktur. Elinizde bir gözlem vardır. Bu gözlem:

- gerçek,
- önemli,
- hatta kritik

olabilir. Fakat yine de tekil bir gözlemdir. Bir GEO skoru için varlığın:

- kullanıcılar,
- AI ürünleri,
- ülkeler,
- diller,
- oturumlar,
- zamanlar

arasındaki temsil dağılımı ölçülmelidir. Tek ekran görüntüsü dağılım değildir. On tekrar otomatik olarak on bağımsız insan değildir. Bin cevap da kötü örneklemle dünya nüfusu değildir. Nüfusa oranlı örneklem de dil adaleti olmadan eksik olabilir. Yüksek ortalama da Critical yanlışları silemez. Bu nedenle NOMOS'un ilk ölçüm yasası şudur:

> *Tek cevap olaydır. / Çoklu cevap örüntüdür. / Tanımlı örneklem nüfus tahminidir. / Zaman içindeki tekrar kararlılık ölçümüdür. / Fakat bunların hiçbiri tek başına nedensellik veya sürdürülebilir değer değildir.*

NOMOS'un Bölüm 1 Emri

> *Bana bir ekran görüntüsü getirip bütün sistem adına konuşma.*
>
> *Bir kullanıcının cevabını dünya nüfusunun cevabı yapma.*
>
> *Aynı hesabın tekrarlarını bağımsız insanlara dönüştürme.*
>
> *Cevabı yenileyip hoşuna gideni gerçek dağılım gibi gösterme.*
>
> *Retleri, hataları ve bilinmeyenleri paydadan çıkarma.*
>
> *Tek dili dünyaya, tek ülkeyi bütün ülkelere, tek AI ürününü bütün üretken sistemlere genelleme.*
>
> *Tekil başarıyı müdahale etkisi; tekil yanlışı bütün sistemin değişmez karakteri ilan etme.*
>
> *Tek bir ağır yanlışı da ortalamanın içinde boğma.*
>
> *Önce olayı kaydet. / Sonra örüntüyü ölç. / Sonra hedef nüfusu tanımla. / Sonra bağımlılığı ve belirsizliği göster. / Ve yalnız kanıtın taşıdığı kadar konuş.*

Bölümün Son Cümlesi

> *Bir varlığın yapay zekâ sistemlerindeki temsili tek bir cevap değildir. AI ürünleri, kullanıcılar, ülkeler, diller, oturumlar ve zamanlar arasında ölçülmesi gereken bir dağılımdır.*

Normatif Öz

> **Tek bir üretken sistem çıktısı, belirli bir cevabın kaydedilmiş koşullarda oluştuğunu gösterebilir. Ancak nüfus yaygınlığı, tipik ürün davranışı, zaman kararlılığı, diller arası tutarlılık, ülkeler arası tutarlılık, AI ürünleri arası tutarlılık, müdahale etkisi, tavsiye kesinliği, ticari sonuç, uygunluk veya GEO skoru çıkarmak için kullanılamaz. Tekil yüksek zararlı çıktı, nüfus yaygınlığı kanıtı sayılmadan ayrı Critical olay incelemesi başlatabilir.**

BÖLÜM 02

Temsil Bir Dağılımdır

YÖNTEMSEL DAYANAKLAR / K04; K05; K09

Bölüm Sınırı

Bölüm 1 şu hükmü kurdu:

> *Tek bir yanıt GEO skoru değildir. Tek bir yanıt, kaydedilmiş koşullarda gerçekleşmiş tekil bir gözlemdir.*

Bu bölüm bir sonraki soruyu cevaplar:

> *Tek bir yanıt GEO skoru değilse, ölçmek istediğimiz asıl nesne nedir?*

Cevap şudur:

> *Bir varlığın üretken sistemlerdeki temsili; AI ürünleri, kullanıcılar, ülkeler, diller, oturum koşulları, promptlar ve zamanlar arasında oluşan çok boyutlu bir dağılımdır.*

Bu bölüm:

- temsil dağılımını,
- dağılımın temel boyutlarını,
- ortalama ile dağılım arasındaki farkı,
- alt grup eşitsizliğini,
- kararlılığı,
- kapsam ve eksik gözlem problemini,
- düşük sıklıklı ağır hata riskini,
- kamuya açık sonuç kartının kavramsal yapısını

tanımlar. Bu bölüm henüz:

- her ülkeye kaç katılımcı düşeceğini,
- örneklem çerçevesinin nasıl kurulacağını,
- istem çevirilerinin nasıl doğrulanacağını,
- atomik iddiaların nasıl puanlanacağını,
- nihai NOMOS skorunun kesin formülünü,
- uygunluk eşiklerini

normatif olarak kesinleştirmez. Bu konular kitabın sonraki bölümlerinde ayrıca düzenlenecektir. Bölüm 2'nin görevi, skor formülünden önce skorun temsil etmeye çalıştığı gerçekliği tanımlamaktır.

NOMOS Challenge

İki yapay zekâ ürünü düşünün. Her biri aynı küresel ortalama skoru alıyor:

90 / 100

İlk ürünün dil sonuçları şöyle:

- İngilizce: 90
- Türkçe: 91
- Almanca: 89
- Japonca: 90
- Arapça: 90

İkinci ürünün dil sonuçları şöyle:

- İngilizce: 99
- Türkçe: 98
- Almanca: 97
- Japonca: 86
- Arapça: 70

İki ürünün de basit ortalaması 90'dır. Aynı GEO kalitesine mi sahiptirler? Hayır. İlk ürün farklı dillerde birbirine yakın bir temsil üretmektedir. İkinci ürün bazı dillerde olağanüstü, başka bir dilde ciddi biçimde zayıf temsil üretmektedir. Tek ortalama bu farkı saklar. Şimdi iki başka ürün düşünün. İkisinin de yanıtlarının yüzde 95'i geçiş kriterini karşılıyor. Birinci üründeki kalan yüzde 5:

- küçük kapsam eksikleri,
- sınırlı açıklama yetersizlikleri,
- birkaç UNKNOWN

sonucundan oluşuyor. İkinci üründeki kalan yüzde 5'in içinde:

- yanlış sağlık yetkisi,
- gerçekte bulunmayan lisans,
- yanlış hukuki kimlik,
- uygunsuz tavsiye

bulunuyor. İkisinin ham doğruluk oranı yine yüzde 95'tir. Aynı temsil riskine mi sahiptirler? Hayır. Şimdi üçüncü bir çift düşünün. İkisinin de üç ölçüm dalgasındaki ortalaması 90'dır. Birincisi:

- Dalga 1: 90
- Dalga 2: 91
- Dalga 3: 89

İkincisi:

- Dalga 1: 98
- Dalga 2: 90
- Dalga 3: 82

Aynı zaman kararlılığına mı sahiptirler? Hayır. O hâlde sorun yalnızca: "Ortalama kaç?" değildir. Şunları da bilmemiz gerekir: Sonuçlar nasıl dağılıyor? Kimler iyi, kimler kötü temsil görüyor? Hangi diller ve ülkeler geride kalıyor? Hatalar ne kadar ağır? Sonuç zaman içinde ne kadar değişiyor? Hangi alanlar hiç ölçülmedi? Ortalama hangi ağırlıklarla oluşturuldu? Aynı ortalamanın arkasında kaç farklı gerçeklik bulunuyor? Bu bölümün ilk hükmü şudur:

> *Aynı ortalama, aynı temsil değildir.*

İkinci hükmü şudur:

> *Bir GEO skoru dağılımın özeti olabilir; dağılımın yerine geçemez.*

1. BÖLÜMÜN AMACI

Bu bölüm, üretken sistemlerde varlık temsilinin tek bir sayıya indirgenmeden önce nasıl anlaşılması gerektiğini tanımlar. Ana amaçları şunlardır:

- Temsilin zamansız ve tekil bir "AI görüşü" olmadığını göstermek
- Temsil sonucunun hangi koşullara bağlı olduğunu tanımlamak
- Doğru–yanlış ikiliğinin tek başına yetersiz olduğunu açıklamak
- Ortalama doğruluk ile alt grup eşitsizliğini ayırmak
- Yüksek ortalama içinde saklanan ağır hata riskini görünür kılmak
- Kararlılık ile doğruluğun aynı kavram olmadığını göstermek
- Ölçülmeyen alanların dağılım içinde olumlu varsayımla doldurulmasını engellemek
- Gelecekteki NOMOS skorunun tek sayılı fakat tek sayıya mahkûm olmayan yapısını hazırlamak

Bu bölümün sonunda okuyucu şu soruya cevap verebilmelidir:

> *Bir varlığın GEO sonucu neden yalnız ortalama doğruluk oranından ibaret değildir?*

2. MERKEZÎ NORMATİF HÜKÜM

> *Bir varlığın üretken sistemlerdeki temsili, tanımlı koşullar altında oluşabilecek farklı yanıtların ve temsil sonuçlarının koşullu dağılımıdır.*

Bu dağılım:

- tek bir modelin soyut "düşüncesi" değildir,
- bütün yapay zekâ sistemlerinin ortak görüşü değildir,
- yalnız web sitesinin tekrar edilme oranı değildir,
- yalnız markanın adının geçme sıklığı değildir,
- yalnız ortalama doğruluk değildir,

yalnız tavsiye oranı değildir. Temsil dağılımı, belirli bir protokol içinde şunu gösterir:

> *Hangi kullanıcıların, hangi AI ürünü ve kullanıcı yüzeyinde, hangi dilde, hangi ülkede, hangi istemle, hangi oturum koşulunda ve hangi zamanda ne tür bir temsil gördüğünü.*

Dolayısıyla bir dağılım sonucu her zaman şu sorulara bağlıdır: Hangi varlık? Hangi AI ürünü? Hangi kullanıcı yüzeyi? Hangi hedef nüfus? Hangi ülke? Hangi dil ve locale? Hangi istem? Hangi oturum koşulu? Hangi ölçüm dalgası? Hangi gerçeklik paketi? Hangi değerlendirme kuralı? Bu alanlardan biri değiştiğinde dağılım da değişebilir.

3. "DAĞILIM" KELİMESİ NE ANLAMA GELİR?

Bu kitapta dağılım kelimesi, bir yapay zekâ sisteminin içinde var olduğu varsayılan gizli veya metafizik bir görüşü ifade etmez. NOMOS: "Modelin zihninde şirketin gerçek dağılımı budur." iddiasında bulunmaz. Buradaki dağılım operasyonel ve gözlemseldir.

> *Tanımlı protokol altında elde edilebilen farklı temsil sonuçlarının, hedef nüfus ve ölçüm koşulları içindeki örüntüsüdür.*

İki dağılım ayrılmalıdır.

3.1. Hedef Temsil Dağılımı

Ölçümün tahmin etmeye çalıştığı, fakat tamamını doğrudan gözleyemediği dağılımdır. Gösterimi:

P_E(Y | A, C, L, S, P, T, U)

Burada:

- E: denetlenen varlık
- Y: temsil sonucu
- A: AI ürünü ve kullanıcı yüzeyi
- C: ülke veya coğrafi koşul
- L: dil ve locale
- S: oturum ve kişiselleştirme koşulu
- P: istem
- T: zaman veya ölçüm dalgası
- U: hedef kullanıcı evreni

Bu ifade şunu sorar: "Tanımlı hedef kullanıcı evreninde, belirtilen koşullarda hangi temsil sonuçları hangi olasılıklarla oluşmaktadır?" Hedef dağılımın bütün elemanları ölçülemez. Örneklem yoluyla tahmin edilir.

3.2. Gözlenen Ampirik Dağılım

Gerçek veya sentetik ölçümde elde edilen geçerli gözlemlerin dağılımıdır. Gösterimi:

P̂_E

olarak kullanılabilir. Ampirik dağılım, hedef dağılımın kendisi değildir. Onun örneklem, ağırlık ve ölçüm tasarımı üzerinden üretilmiş tahminidir. Aradaki farkın büyüklüğü şunlara bağlı olabilir:

- Örneklem tasarımı
- Katılımcı seçimi
- Ülke ve dil kapsamı
- Geçersiz gözlem oranı
- Yanıtlamama
- Hakem hatası
- Ölçüm dalgası
- İstem seçimi
- Ürün ve arayüz koşulları
- Ağırlıklandırma
- Kümelenme ve bağımlılık

Bu nedenle bir ampirik dağılım: "Yapay zekânın kesin ve değişmez gerçeği" olarak yayımlanamaz. Doğru ifade: "Belirtilen hedef nüfus ve ölçüm koşulları için elde edilen tahmin." olmalıdır.

4. TEMSİL UZAYI

Bir varlığın temsili tek boyutta oluşmaz. Her gözlem çok boyutlu bir koşul uzayında yer alır. Temel gösterim:

```
X_i = (E_i, A_i, U_i, C_i, L_i, S_i, P_i, T_i)
```

Burada:

- Ei: varlık
- Ai: AI ürünü ve kullanıcı yüzeyi
- Ui: kullanıcı veya örnekleme birimi
- Ci: ülke ve coğrafi tabaka
- Li: dil ve locale
- Si: oturum, plan ve kişiselleştirme koşulu
- Pi: istem veya istem ailesi
- Ti: zaman ve ölçüm dalgası

Üretilen yanıt:

```
R_i
```

olarak gösterilebilir. Yanıtın değerlendirilmesi sonucunda oluşan temsil durumu ise:

```
Y_i = g(R_i, V_E, J)
```

şeklinde düşünülebilir. Burada:

- VE: varlık için kullanılan doğrulanmış gerçeklik paketi
- J: değerlendirme ve hakemlik kuralları
- g: yanıtı temsil sonuçlarına dönüştüren değerlendirme süreci

Bu gösterimin anlamı şudur: Aynı yanıt, hangi varlığa, hangi gerçeklik paketine ve hangi değerlendirme kuralına göre incelendiği belirtilmeden tek başına puanlanamaz.

5. TEMSİL SONUCU TEK BİR DEĞER DEĞİLDİR

Bir yanıtın tamamına yalnız:

- doğru,
- yanlış

etiketi vermek çoğu zaman yetersizdir. Aynı yanıt içinde şu durumlar birlikte bulunabilir:

- Doğru kimlik
- Doğru ana kategori
- Eksik hizmet kapsamı
- Güncelliğini yitirmiş çalışan sayısı
- Desteklenmeyen liderlik iddiası
- Doğru atıf
- Uygunsuz tavsiye
- Critical olmayan bir sınır eksikliği

Bu nedenle temsil sonucu bir sonuç vektörü olarak düşünülebilir. Aday gösterim:

```
Y_i = (y_i^id, y_i^cat, y_i^fact, y_i^scope, y_i^time, y_i^evidence, y_i^recommendation, y_i^harm)
```

Burada:

- yiid: kimlik doğruluğu
- yicat: kategori ve rol doğruluğu
- yifact: maddi olguların doğruluğu
- yiscope: kapsam ve sınırların korunması
- yitime: zamansal geçerlilik
- yievidence: iddia–kanıt uyumu

- yirecommendation: tavsiye ve kullanıcı uygunluğu
- yiharm: hata veya yanlış yönlendirmenin zarar sınıfı

Bu alanların kesin puanlama kuralları sonraki bölümlerde tanımlanacaktır. Buradaki kurucu ilke şudur:

> *Bir temsil cevabı, yalnız doğru bilgi içerip içermediğiyle değil; kimi, ne olarak, hangi sınırlar içinde, hangi kanıtla ve kime uygun biçimde anlattığıyla değerlendirilmelidir.*

6. TEMEL TEMSİL DURUMLARI

Her atomik iddianın ve nihai yanıtın statüsü aynı olmayabilir. Bu bölüm aşağıdaki temel durumları kavramsal olarak ayırır.

6.1. SUPPORTED — DESTEKLENEN

İddia, tanımlı kapsam ve zaman içinde yeterli kanıtla desteklenmektedir.

6.2. CONTRADICTED — ÇELİŞEN

İddia güvenilir ve ilgili kayıtlarla çelişmektedir.

6.3. UNSUPPORTED — DESTEKLENMEYEN

İddia için yeterli destek bulunmamaktadır. İddianın mutlaka yanlış olduğu kanıtlanmış değildir. Ancak doğru kabul edilmesi için yeterli kayıt yoktur.

6.4. OUTDATED — GÜNCELLİĞİNİ YİTİRMİŞ

İddia geçmişte doğru olmuş olabilir. Ölçüm tarihinde artık geçerli değildir veya güncelliği doğrulanamamaktadır.

6.5. PARTIALLY SUPPORTED — KISMEN DESTEKLENEN

İddianın yalnız belirli kapsamı, nüfusu, ülkesi, ürünü veya zamanı desteklenmektedir. Yanıt destek sınırını aşmış olabilir.

6.6. OUT OF KAPSAM — KAPSAM DIŞI

İddia denetlenen varlık veya gerçeklik paketiyle ilgili değildir. Başka varlık, ürün, ülke veya hizmetten gelmiş olabilir.

6.7. VARLIK CONFLATION — VARLIK BİRLEŞMESİ

Sistem:

- iki farklı şirketi,
- marka ile hukuki varlığı,
- kişi ile kurumu,
- eski ve yeni varlığı,
- aynı adlı farklı kuruluşları

birbirine karıştırmıştır. Bu hata yalnız bir olgu yanlışı değildir. Temsil nesnesinin bozulmasıdır.

6.8. MATERIAL OMISSION — MADDİ EKSİKLİK

Yanıtta bulunan cümleler doğru olabilir. Fakat kullanıcı kararını değiştirecek temel bilgi çıkarılmıştır. Örnek: Hizmet yalnız belirli ülkelerde sunulmaktadır. Lisans belirli yargı alanıyla sınırlıdır. Fiyat vergi

hariçtir. Sonuç birinci taraf verisine dayanmaktadır. Ürün belirli kullanıcılar için uygun değildir. Eksiklik, doğrudan yanlış cümle bulunmasa bile yanlış çıkarım yaratabilir.

6.9. REFUSAL — RET

Sistem cevap vermeyi reddetmiştir. Ret:

- uygun,
- gereksiz,
- tutarsız,
- güvenlik temelli,
- teknik

olabilir. Ret otomatik olarak başarısız veya başarılı sayılmaz. Ayrı sınıflandırılır.

6.10. NO RESULT — SONUÇ YOK

Sistem teknik veya operasyonel nedenle kullanılabilir bir yanıt üretmemiştir. Ret ile aynı değildir.

6.11. UNKNOWN — BİLİNMİYOR

Mevcut kayıtlar güvenilir sınıflandırmaya yetmemektedir. UNKNOWN, olumlu sonuca dönüştürülemez.

7. DAĞILIMIN TEMEL ÖZELLİKLERİ

Bir temsil dağılımı yalnız ortalamayla açıklanamaz. En azından aşağıdaki özellikler değerlendirilmelidir.

7.1. Merkezî Eğilim

Yanıtların genel doğruluk veya geçiş düzeyini özetler. Örnek ölçüler:

- Ortalama
- Ağırlıklı ortalama
- Medyan
- Geçiş oranı

Merkezî eğilim önemlidir. Fakat dağılımın yalnız merkezini gösterir.

7.2. Yayılım ve Değişkenlik

Sonuçların merkez etrafında ne kadar dağıldığını gösterir. Aynı ortalamaya sahip iki sistemden biri kullanıcılar arasında tutarlı, diğeri son derece değişken olabilir. Değişkenlik:

- kullanıcılar,
- ülkeler,
- diller,
- oturumlar,
- ölçüm dalgaları

arasında ayrıca incelenebilir.

7.3. Dağılımın Biçimi

Sonuçlar tek bir merkez etrafında mı toplanmaktadır? Yoksa iki veya daha fazla kümeye mi ayrılmaktadır? Örneğin:

- Kullanıcıların yarısı çok doğru,
- yarısı çok yanlış

yanıt görüyorsa ortalama orta düzeyde olabilir. Fakat tipik kullanıcı deneyimi "orta doğruluk" değildir. Kutuplaşmış iki ayrı deneyimdir.

7.4. Alt Grup Farkları

Belirli dil, ülke veya kullanıcı grupları sistematik olarak farklı sonuçlar görüyor mu? Küresel ortalama bu farkı gizleyebilir.

7.5. Kuyruk Riski

Düşük sıklıklı fakat ağır yanlışlar bulunuyor mu? Nadir olaylar ortalamaya küçük etki yapabilir. İnsan zararına etkileri büyük olabilir.

7.6. Kapsam

Hedef kullanıcı evreninin ne kadarı gerçekten ölçülmüştür? Yüksek skor, düşük kapsamla birlikte bulunabilir.

7.7. Eksik ve Bilinmeyen Kütle

Yanıtların ne kadarı:

UNKNOWN,

NOT TESTED,

ret, teknik sonuçsuzluk, geçersiz kayıt durumundadır?

7.8. Zaman Kararlılığı

Dağılım farklı ölçüm dalgalarında ne ölçüde korunmaktadır?

8. ADAY NOMOS TEMSİL DAĞILIM İMZASI

CANDIDATE CONCEPT — ADAY KAVRAM

Bir GEO sonucunun tek bir sayı yerine önce aşağıdaki dağılım imzasıyla tanımlanması önerilir:

```
RDS(E)=(μ,κ,σ,Δ,τ,ρ,u)
```

Bu gösterim Representation Distribution Signature, yani:

Temsil Dağılım İmzası

olarak adlandırılabilir. Bileşenler:

μ — Genel Temsil Düzeyi

Tanımlı ağırlıklar ve kapsam altında genel doğruluk veya geçiş tahmini.

κ — Kapsam

Hedef kullanıcı, ülke, dil, AI ürünü ve zaman evreninin ne kadarının yeterli biçimde ölçüldüğü.

σ — Değişkenlik

Kullanıcılar ve deneysel hücreler arasındaki yayılım.

Δ — Alt Grup Eşitsizliği

Dil, ülke, ürün veya başka önemli gruplar arasındaki maddi fark.

τ — Kuyruk ve Critical Hata Riski

Düşük sıklıklı ağır yanlışların oranı ve niteliği.

ρ — Zaman Kararlılığı

Dalgalar arasında dağılımın dayanıklılığı.

u — Bilinmeyen ve Sonuçsuz Kütle

UNKNOWN, sonuçsuz veya yeterince değerlendirilemeyen alanların payı. Bu imza henüz nihai skor formülü değildir. Bileşenlerin:

- kesin tanımı,
- tahmini,
- ağırlığı,
- kamu gösterimi

pilotlar ve sonraki bölümlerle kesinleştirilecektir. Kurucu düşünce şudur:

> *GEO sonucu yalnız "ne kadar doğru?" sorusunu değil; "kimler için, ne kadar kararlı, ne kadar eşit, ne kadar kapsamlı ve ne kadar riskli?" sorularını da cevaplamalıdır.*

9. ORTALAMA AKLAMA

Bir kurum yalnız yüksek küresel ortalamayı gösterip dağılımın maddi kusurlarını saklıyorsa buna:

Ortalama aklama

denir. Ortalama aklama şu yollarla oluşabilir:

- En zayıf dili gizlemek
- Küçük ülkeleri rapordan çıkarmak
- Critical hataları genel yanlış oranına karıştırmak
- UNKNOWN sonuçları paydadan çıkarmak
- Yalnız iyi çalışan AI ürünlerini dâhil etmek
- Zayıf ölçüm dalgasını dışarıda bırakmak
- Kapsam dışı nüfusu görünmez yapmak
- Ağırlıkları sonuç görüldükten sonra değiştirmek

Ortalama matematiksel olarak doğru hesaplanmış olabilir. Yine de kamuya sunduğu temsil yanıltıcı olabilir. Bu nedenle:

> *Bir ortalamanın hesap bakımından doğru olması, karar bakımından yeterli olduğu anlamına gelmez.*

10. AYNI ORTALAMANIN FARKLI GERÇEKLİKLERİ

Aşağıdaki sentetik dağılımların tamamında ortalama 90 olabilir.

10.1. Homojen Dağılım

Bütün önemli alt gruplar 88–92 aralığındadır. Bu dağılım nispeten dengelidir.

10.2. Kutuplaşmış Dağılım

Kullanıcıların yarısı 100'e yakın, diğer yarısı 80'e yakın sonuç görmektedir. Ortalama 90'dır. Fakat kullanıcı deneyimi eşit değildir.

10.3. Dil Bakımından Parçalanmış Dağılım

Yüksek nüfuslu veya yüksek kaynaklı dillerde sonuç çok iyi, düşük kaynaklı dillerde ciddi biçimde zayıftır. Küresel ortalama iyi görünebilir. Dil adaleti zayıftır.

10.4. Ülke Bakımından Parçalanmış Dağılım

Aynı dilde bile farklı ülke, locale, ürün erişimi veya kullanıcı yüzeyi nedeniyle sonuç değişebilir.

10.5. Zaman Bakımından Oynak Dağılım

Tek bir dalgada çok yüksek, başka dalgada çok düşük sonuç oluşur. Ortalama yüksek olsa bile güvenilir kullanıcı beklentisi kurmak zordur.

10.6. Kuyruk Riski Taşıyan Dağılım

Yanıtların büyük çoğunluğu doğrudur. Fakat küçük bir bölümünde ağır sağlık, hukuk, lisans veya güvenlik yanlışı bulunur.

10.7. Yüksek Bilinmeyen Kütleli Dağılım

Değerlendirilebilen yanıtların çoğu doğrudur. Ancak hedef nüfusun veya yanıtların önemli bir bölümü:

- test edilmemiş,
- sonuçsuz,

UNKNOWN

durumdadır. Yalnız değerlendirilebilen cevaplar üzerinden yüksek oran yayımlamak kapsamı saklayabilir.

11. ALT GRUP EŞİTSİZLİĞİ

Bir küresel skor, alt gruplardaki deneyimi temsil etmeyebilir. Temel alt gruplar şunlar olabilir:

- AI ürünü ve kullanıcı yüzeyi
- Ülke
- Dil
- Locale
- Plan
- Web veya mobil
- Kontrollü veya doğal panel
- Zaman dalgası
- Kullanıcı niyeti
- İstem ailesi

Bu grupların tamamı her denetimde aynı öneme sahip olmayabilir. Maddi olan gruplar kapsam kilitlenirken belirlenmelidir.

11.1. Dil ile Ülke Aynı Şey Değildir

Türkçe konuşan:

- Türkiye'deki kullanıcı,
- Almanya'daki kullanıcı,
- başka bir ülkedeki kullanıcı

aynı dili kullanabilir. Fakat:

- ürün erişimi,
- fiyat,
- kaynaklar,
- yerel arayüz,
- hukuki bağlam,
- kişiselleştirme

farklı olabilir. Bu nedenle yalnız dil veya yalnız ülke üzerinden bütün farklar açıklanamaz.

11.2. Kesişimsel Hücreler

Bazı maddi farklar yalnız birden fazla alan birlikte incelendiğinde görünür. Örnek: Almanya × Türkçe × Mobil × Ücretsiz plan × Doğal kullanıcı paneli başlı başına bir deneysel hücre olabilir. Ancak hücreler çoğaldıkça örneklem küçülür. Bu nedenle her küçük hücre için kesin skor yayımlanamaz. Ölçüm tasarımı: anlamlı alt grubu korumalı, istatistiksel sahte kesinlik üretmemelidir.

11.3. Alt Grup Tabanı

CANDIDATE CONCEPT — ADAY KAVRAM

Küresel ortalamanın yanında en zayıf yeterli örneklemli alt grup sonucunun veya alt yüzde 10'luk grubun ayrıca raporlanması önerilir. Buna:

Temsil tabanı

veya:

Language/Country Floor

denebilir. Ancak "en düşük skor" küçük ve istikrarsız hücreden geliyorsa yanlış sonuç üretilebilir. Bu nedenle taban metriği:

- minimum örneklem,
- güven aralığı,
- hücre kalitesi

ile birlikte tanımlanmalıdır. Kesin kural sonraki bölümlerde belirlenecektir.

12. KAPSAM DAĞILIMIN PARÇASIDIR

Bir dağılım yalnız gözlenen cevaplardan oluşmaz. Kimin veya neyin hiç gözlenmediği de sonucun anlamını belirler. Örnek: İngilizce test edildi. Türkçe test edildi. Japonca test edilmedi. Arapça test edilmedi. Mobil arayüz test edilmedi. Ücretli plan test edilmedi. İngilizce ve Türkçe sonuçlarının yüksek olması: "Bütün dillerde ve bütün kullanıcı yüzeylerinde temsil yüksektir." anlamına gelmez. Doğru sonuç: "Belirtilen İngilizce ve Türkçe web yüzeylerinde ölçüm yapılmıştır. Diğer diller ve mobil yüzeyler test edilmemiştir." olmalıdır.

12.1. Kapsam Haritası

Her GEO dağılımı ilgili olduğu ölçüde şu haritayı taşımalıdır:

- Hedef evren
- Örnekleme çerçevesi
- Kapsanan ülkeler
- Kapsanmayan ülkeler
- Kapsanan diller
- Kapsanmayan diller
- Denetlenen AI ürünleri
- Denetlenmeyen ürünler
- Denetlenen kullanıcı yüzeyleri
- Denetlenmeyen yüzeyler
- Ölçüm dalgaları
- Kapsam dışı istem aileleri

12.2. Kapsam Oranı

CANDIDATE CONCEPT — ADAY KAVRAM

Hedef evrenin ne kadarının yeterli ölçümle temsil edildiğini gösteren bir kapsam metriği geliştirilebilir. Ancak kapsam:

- yalnız ülke sayısı,
- yalnız dil sayısı,
- yalnız nüfus oranı

ile ölçülemez. Örneğin 150 ülke kapsanmış olabilir. Fakat katılımcıların yüzde 95'i yalnız beş ülkeden gelmiş olabilir. Ya da nüfusun yüzde 90'ı kapsanmış fakat küçük diller tamamen dışarıda kalmış olabilir. Bu nedenle nüfus kapsamı ile ülke/dil gözcü kapsamı ayrı gösterilmelidir.

13. EKSİK GÖZLEM VE YANITLAMAMA

Eksik veri yalnız teknik sorun değildir. Dağılımı sistematik biçimde bozabilir. Gözlemin eksik olma nedenleri şunlar olabilir:

- Katılımcının görevi tamamlamaması
- Ürünün ülkede erişilememesi
- Hesap kısıtlaması
- Teknik hata
- Sistem reddi
- Ekran görüntüsünün geçersiz olması
- Katılımcının istemi değiştirmesi
- Belirli dilde ürünün çalışmaması
- Hakemlerin güvenilir karar verememesi

Bu nedenlerin her biri ayrı statü gerektirir.

13.1. Eksiklik Rastgele Olmayabilir

Örneğin düşük bağlantı kalitesine sahip ülkelerde daha fazla görev tamamlanamıyorsa yalnız başarılı kayıtların kullanılması küresel sonucu zengin ve güçlü bağlantılı kullanıcılar lehine kaydırabilir. Belirli dilde cevapların değerlendirilmesi zor olduğu için UNKNOWN oranı yükseliyorsa yalnız bilinen cevaplardan hesaplanan skor dil eşitsizliğini saklayabilir. Bu nedenle:

> *Eksik veri oranı ve eksiklik nedeni dağılım raporunun parçası olmalıdır.*

13.2. Denominator — Payda Bütünlüğü

Bir rapor şu soruya açık cevap vermelidir: "Yüzde 90 tam olarak neyin yüzde 90'ıdır?" Muhtemel paydalar:

- Atanmış katılımcılar
- Görevi başlatanlar
- Görevi tamamlayanlar
- Protokole uygun kayıtlar
- Geçerli sistem sonuçları
- Hakem kararı verilen cevaplar
- Belirli dil veya ülke hücresi

Sonuç görüldükten sonra payda değiştirilmemelidir. Örneğin:

- 1.000 görev atanmış,
- 900 tamamlanmış,
- 800 geçerli,
- 720 geçmiş

ise:

```
720/800=90%
```

geçerli kayıtlar içindeki geçiş oranıdır. Fakat bütün atanmış örneklem içindeki kanıtlanmış geçiş oranı:

```
720/1000=72%
```

olarak farklı bir soruyu cevaplar. İki oran da kullanılabilir. Aynı anlamda sunulamaz.

14. DOĞRULUK VE KARARLILIK AYNI ŞEY DEĞİLDİR

Bir dağılım yüksek doğruluğa sahip olabilir. Fakat zaman veya kullanıcılar arasında kararsız olabilir. Başka bir dağılım düşük doğruluğa sahip olabilir. Fakat tutarlı biçimde yanlış olabilir. Bu nedenle dört durum ayrılmalıdır.

Doğruluk	Kararlılık	Yorum
Yüksek	Yüksek	Tutarlı ve doğru temsil adayı
Yüksek	Düşük	Çoğunlukla doğru fakat güvenilmez deneyim
Düşük	Yüksek	Sistematik ve kararlı yanlış temsil
Düşük	Düşük	Hem yanlış hem oynak temsil

"Kararlı" olmak otomatik olarak olumlu değildir. Sistem aynı yanlışı her kullanıcıya veriyorsa oldukça kararlı olabilir.

14.1. Kullanıcılar Arası Kararlılık

Aynı hücrede farklı kullanıcılar benzer maddi temsil görüyor mu?

14.2. Oturumlar Arası Kararlılık

Aynı veya eşlenmiş kullanıcı, yeni oturumlarda benzer sonuç görüyor mu?

14.3. Dalgalar Arası Kararlılık

Aynı protokol farklı gün veya dönemlerde benzer dağılım üretiyor mu?

14.4. Ürün Güncellemesi Sonrası Kararlılık

AI ürününde maddi değişiklik olduğunda önceki dağılım devam ediyor mu?

15. AYNI CÜMLEYİ ÜRETMEK KARARLILIK DEĞİLDİR

Kararlılık: "Her kullanıcıya kelime kelime aynı cevap verilmesi" anlamına gelmez. Bir AI ürünü farklı anlatım, sıra ve örnekler kullanabilir. Bunlar maddi anlamı koruduğu sürece kabul edilebilir varyasyon olabilir. Önemli ayrım şudur:

Biçimsel Değişkenlik

Kelime seçimi Cümle sırası Uzunluk Ton Örnek Liste veya paragraf kullanımı Maddi gerçeklik korunuyorsa hata olmayabilir.

Anlamsal Değişkenlik

Şirket kimliğinin değişmesi Ana kategorinin farklılaşması Faaliyet alanının genişleyip daralması Lisans veya yetkinliğin değişmesi Fiyat ve coğrafyanın değişmesi Bir kullanıcıya önerilip benzer kullanıcıya açıklamasız biçimde önerilmemesi Maddi temsil değişiyorsa kararlılık sorunu oluşabilir. Bu nedenle tercih edilen kavram:

Koşullu anlamsal kararlılık

olmalıdır. Eşdeğer maddi koşullarda temel gerçekliğin korunması beklenir. Farklı kullanıcı ihtiyaçlarında tavsiyenin değişmesi ise uygun bağlamsal davranış olabilir.

16. BAĞLAMA DUYARLILIK İLE KEYFÎ DEĞİŞKENLİK

Bir sistemin bütün kullanıcılara aynı tavsiyeyi vermesi her zaman doğru değildir. Örneğin:

- Bütçesi düşük kullanıcı
- Yalnız belirli ülkede hizmet arayan kullanıcı
- Başka lisans gerektiren kullanıcı
- Kurumsal yerine bireysel hizmet arayan kullanıcı

aynı varlık için farklı uygunluk sonucu alabilir. Bu fark:

- kullanıcı koşuluna,
- hizmet kapsamına,
- gerçek sınıra

dayanıyorsa bağlama duyarlı ve doğru olabilir. Ancak eşdeğer kullanıcı koşullarında sonuç sebepsiz biçimde değişiyorsa keyfî değişkenlik olabilir. Bu nedenle NOMOS şunu istemez: "Bütün kullanıcılara aynı cevap." Şunu ister:

> *Eşdeğer maddi koşullarda açıklanabilir biçimde tutarlı, farklı koşullarda ise gerçek sınırlara duyarlı cevap.*

17. ZAMAN TEMSİL DAĞILIMININ İÇİNDEDİR

Temsil dağılımı zamansız değildir. Gösterim: PE,t olarak düşünülebilir. Aynı varlığın dağılımı:

- ürün güncellemesi,
- kaynak değişikliği,
- şirket değişikliği,
- haber olayı,
- yeni dil sürümü,
- teknik erişim değişikliği

sonrasında farklılaşabilir.

17.1. Ani Değişim

Bir ürün veya kaynak güncellemesinden sonra dağılım kısa sürede değişir.

17.2. Kademeli Değişim

Yeni kaynaklar ve kullanıcı davranışları nedeniyle sonuç zaman içinde yavaşça değişir.

17.3. Geçici Dalgalanma

Belirli olay veya haber döneminde dağılım geçici olarak farklılaşır.

17.4. Müdahale Sonrası Değişim

GEO müdahalesinden sonra dağılım değişebilir. Fakat değişimin müdahaleden kaynaklandığını söylemek ayrıca baseline ve nedensellik tasarımı gerektirir.

17.5. Zaman Ortalamasıyla Aklama

Bir ürünün:

- ilk dalgada 98,
- ikinci dalgada 90,
- üçüncü dalgada 82

alması hâlinde ortalama 90'dır. Yalnız 90'ı yayımlamak düşüş eğilimini saklar. Bu nedenle:

> *Zaman ortalaması, dalga sonuçlarının yerini alamaz.*

18. KUYRUK RİSKİ VE AĞIR HATA DAĞILIMI

Bir dağılımın büyük bölümü doğru olabilir. Fakat düşük olasılıklı ağır hatalar ayrıca değerlendirilmelidir. Örnekler:

- Gerçekte bulunmayan sağlık lisansı
- Yanlış acil durum yönlendirmesi
- Yanlış hukuki yetki
- Sahte yatırım garantisi
- Yanlış şirket kimliği
- Çocuklar veya savunmasız gruplar için uygunsuz tavsiye

- İnsanlardan saklanan manipülatif talimatın tekrarı

Önceki mimaride:

- Critical,
- Major,
- Moderate,
- Advisory

olarak dört önem seviyesi ayrılmış ve Critical bulguların otomatik başarısızlık etkisi taşıması önerilmişti. Bu kitap bu sistemi nüfus dağılımına bağlar.

18.1. Frekans ile Şiddet Ayrımı

Bir hata:

- sık fakat düşük etkili,
- nadir fakat ağır etkili

olabilir. Yalnız frekans oranı iki durumu eşit değerlendiremez.

18.2. Aday Şiddet Ağırlıklı Risk Gösterimi

CANDIDATE CONCEPT — ADAY KAVRAM

Hata sınıfları için şiddet katsayıları kullanılarak bir risk göstergesi geliştirilebilir:

```
τ = Σ_{k=1}^K λ_k p_k
```

Burada:

- pk: hata türünün tahmin edilen oranı
- λk: hata türünün şiddet katsayısı

Ancak katsayılar keyfî biçimde belirlenemez. Şunlara dayanmalıdır:

- Öngörülebilir insan zararı
- Temsil bütünlüğü
- Kullanıcı kararına etki
- Geri döndürülebilirlik
- Kasıt ve manipülasyon
- Yüksek riskli alan

Kesin yöntem Bölüm 15'te tanımlanacaktır.

18.3. Critical Kapı

Yüksek genel ortalama, doğrulanmış Critical hata adayını otomatik olarak silemez. Doğru kamu kaydı iki sonucu birlikte göstermelidir: "Genel temsil tahmini yüksektir." ve: "Belirtilen Critical olay veya olaylar ayrıca uygunluk kararını etkilemektedir."

19. DAĞILIMDA REFUSAL VE SONUÇSUZLUK

Bir sistemin cevap vermemesi de kullanıcı deneyiminin parçasıdır. Ancak bütün retler aynı değildir.

19.1. Uygun Ret

Sistem kanıtın yetersiz olduğu veya güvenli yanıt veremediği durumda ihtiyat gösterebilir. Bu, bazı bağlamlarda doğru davranış olabilir.

19.2. Gereksiz Ret

Kamuya açık ve düşük riskli bilgiye cevap vermekten tutarsız biçimde kaçınabilir.

19.3. Teknik Sonuçsuzluk

Sistem teknik hata nedeniyle yanıt üretmeyebilir.

19.4. Kapsam Dışı Ret

İstem yanlış veya görev protokol dışı olabilir.

19.5. Retlerin Dağılımda Korunması

Retler:

- başarı,
- başarısızlık,

UNKNOWN

olarak tek bir kategoriye zorlanmamalıdır. İşlev ve gerekçeleri değerlendirilmelidir. Retlerin paydadan çıkarılması, yalnız başarılı cevaplar arasında yüksek doğruluk oranı üretip gerçek kullanıcı deneyimini saklayabilir.

20. DAĞILIMLAR NASIL KARŞILAŞTIRILABİLİR?

İki AI ürününün dağılımını karşılaştırmak için yalnız sonuç sayıları yeterli değildir. Karşılaştırmanın adil olması için şu alanlar mümkün olduğunca eşit veya açıkça modellenmiş olmalıdır:

- Denetlenen varlık
- Hedef nüfus
- İstem sürümü
- Diller
- Ülkeler
- Ölçüm dalgası
- Oturum koşulu
- Kullanıcı yüzeyi
- Plan dağılımı
- Hakemlik kuralı
- Gerçeklik paketi
- Geçerlilik kuralları
- Ağırlıklar

Maddi ölçüde farklı protokollerle elde edilen iki skor doğrudan sıralama için kullanılamaz.

20.1. Ortak Destek Karşılaştırması

Bütün karşılaştırılan AI ürünlerinin:

- resmî erişilebilir olduğu,
- aynı dil ve kullanıcı koşullarını desteklediği

ortak evrende karşılaştırma yapılabilir. Bu, sistemler arası daha temiz kıyas sağlar.

20.2. Doğal Erişim Dağılımı

Her AI ürünü kendi gerçek kullanıcı erişim evreninde ölçülebilir. Bu sonuç ürünün gerçek küresel etkisini gösterir. Ancak hedef nüfuslar farklı olabileceği için doğrudan performans sıralaması dikkat gerektirir.

20.3. Aynı Sağlayıcının Farklı Yüzeyleri

Web, mobil, kurumsal ürün veya farklı planlar aynı sağlayıcıya ait olsa da ayrı dağılımlar oluşturabilir. Tek sonuç altında birleştirilecekse birleşim ve ağırlık yöntemi açıklanmalıdır.

21. AĞIRLIKLANDIRMA DAĞILIMIN ANLAMINI DEĞİŞTİRİR

Bir küresel ortalama, gözlemlerin hangi ağırlıkla birleştirildiğine bağlıdır. Muhtemel ağırlıklar:

- Eşit ülke
- Nüfusa göre ülke
- Uygun AI kullanıcı nüfusu
- Dil kullanıcısı
- Gerçek ürün maruziyeti
- Eşit AI ürünü
- Pazar payı
- Müşteri hedef pazarı

Her yöntem farklı soruyu cevaplar.

21.1. Eşit Ülke Ağırlığı

San Marino ile Hindistan aynı ağırlığı alabilir. Bu yaklaşım: "Ortalama ülke deneyimi" sorusuna yaklaşabilir. Dünya nüfusundaki ortalama kullanıcı deneyimini temsil etmez.

21.2. Nüfus Ağırlığı

Büyük ülkeler daha fazla etki alır. Bu yaklaşım küresel nüfus deneyimine yaklaşabilir. Küçük ülkeleri ve dilleri görünmez bırakabilir.

21.3. Dil Adaleti Ağırlığı

Düşük kaynaklı dillerin ayrıca görünür kalmasını sağlayabilir. Küresel nüfus skoruyla aynı değildir.

21.4. Ürün Maruziyeti Ağırlığı

AI ürünlerinin gerçek kullanıcı maruziyetine göre ağırlıklandırılması düşünülebilir. Güvenilir ve bağımsız kullanım verisi gerektirir.

21.5. Normatif Ağırlık İlkesi

> *Ağırlıklar sonuç görülmeden önce tanımlanmalı, kamuya açıklanmalı ve hangi soruyu cevapladıkları belirtilmelidir.*

Sonuç görüldükten sonra olumlu skoru yükseltecek ağırlık sistemini seçmek yasaktır. Tek küresel skorun yanında alternatif anlamlı ağırlık sonuçları ayrıca yayımlanabilir.

22. SKOR DAĞILIMIN KAYIPLI SIKIŞTIRMASIDIR

Bir dağılımdan tek sayı üretmek mümkündür. Fakat tek sayı, dağılımdaki bütün bilgiyi taşımaz. Bu nedenle skor:

Dağılımın kayıplı bir özeti

olarak kabul edilmelidir. Bir skor şu bilgileri kaybedebilir:

- En zayıf dil
- En zayıf ülke
- Critical hata türü
- Bilinmeyen kütle
- Zaman dalgalanması
- Kutuplaşma
- Ürün ve plan farkı
- Kapsam dışı nüfus
- Ret dağılımı
- Hakem belirsizliği

Skor yararlıdır. Çünkü:

- karşılaştırmayı kolaylaştırır,
- kamu iletişimini sadeleştirir,
- değişimi izlemeyi sağlar,

yönetim kararını destekler. Fakat skorun sadeliği dağılımın yok edilmesine dönüşemez.

23. MİNİMUM KAMU SONUÇ KARTI

CANDIDATE REQUIREMENT — ADAY ZORUNLULUK

Bir NOMOS küresel skoru kamuya sunulacaksa yanında en azından aşağıdaki sonuçların gösterilmesi önerilir:

Alan	Açıklama
Ana temsil tahmini	Tanımlı ağırlık altında genel sonuç
Güven veya belirsizlik aralığı	Tahminin hassasiyeti
Kapsam	Ülke, dil, ürün, kullanıcı yüzeyi ve zaman
En zayıf yeterli alt grup	Maddi eşitsizliği görünür kılar
Critical hata oranı	Ağır hata kuyruğu
Zaman kararlılığı	Dalgalar arası değişim
UNKNOWN ve sonuçsuz oranı	Bilinmeyen kütle
NOT TESTED alanlar	Ölçülmeyen evren
Ölçüm tarihi	Sonucun zamansal sınırı
Protokol sürümü	Hangi kuralla üretildiği

Ana skor tek başına kullanılabilir bir reklam etiketi olmamalıdır. Kanonik kayıt bu sonuç kartına bağlanmalıdır.

24. SENTETİK APPLE.COM DAĞILIM GÖSTERİMİ

SENTETİK METODOLOJİ GÖSTERİMİ / Aşağıdaki AI ürünleri, kullanıcılar, diller, sayılar ve sonuçlar tamamen kurgusaldır. Gerçek Apple Inc. performansını veya herhangi bir gerçek yapay zekâ ürününü temsil etmez. İki sentetik AI ürünü düşünelim:

- Ürün A
- Ürün B

Her ürün için:

- beş dil,
- dil başına 200 geçerli sentetik gözlem,
- toplam 1.000 gözlem

bulunduğunu varsayalım. Bu eşit dil dağılımı yalnız yöntemi göstermek içindir. Gerçek GEO-1000 nüfus ağırlıklandırmasını temsil etmez.

24.1. Dil Bazlı Geçiş Oranları

Dil	Ürün A	Ürün B
İngilizce	%90	%99
Türkçe	%91	%98
Almanca	%89	%97
Japonca	%90	%86
Arapça	%90	%70
Basit ortalama	%90	%90

İki ürünün ortalaması aynıdır. Fakat dağılımları farklıdır.

Ürün A

Dil sonuçları birbirine yakındır. En yüksek ve en düşük dil arasındaki fark 2 puandır. Dil tabanı yüzde 89'dur.

Ürün B

Üç dilde çok yüksek sonuç üretir. Japonca ve özellikle Arapçada düşüş vardır. En yüksek ve en düşük dil arasındaki fark 29 puandır. Dil tabanı yüzde 70'tir. Yalnız: "Her ikisi de 90 aldı." demek maddi farkı saklar.

24.2. Yanıt Statüsü Dağılımı

Yanıt statüsü	Ürün A	Ürün B
Geçiş kriterini karşılayan	900	900
Kısmen desteklenen	60	20
Desteklenmeyen maddi iddia	20	20
Güncelliğini yitirmiş	10	20
Critical hata adayı	1	14
UNKNOWN / sonuçsuz	9	26
Toplam	1.000	1.000

İki ürünün geçiş oranı yine yüzde 90'dır. Fakat: Ürün B'de Critical hata adayı 14 kat daha fazladır. Ürün B'de bilinmeyen ve sonuçsuz kütle daha yüksektir. Ürün A'daki başarısızlıkların daha büyük bölümü kısmi destek eksikliğidir. Ürün B'nin hata kuyruğu daha ağırdır. Aynı geçiş oranı aynı uygunluk sonucu doğurmak zorunda değildir.

24.3. Zaman Dalgaları

Ölçüm dalgası	Ürün A	Ürün B
Dalga 1	%90	%98
Dalga 2	%91	%90
Dalga 3	%89	%82
Ortalama	%90	%90

Ürün A nispeten kararlıdır. Ürün B belirgin düşüş göstermektedir. Yalnız üç dalganın ortalamasını yayımlamak, Ürün B'deki bozulmayı saklar.

24.4. Sentetik Sonuç

Ürün A ve Ürün B için aynı başlık skorunu yayımlamak mümkün olabilir:

90

Fakat NOMOS sonuç kartları aynı olamaz.

Ürün A — Sentetik Dağılım İmzası

Genel sonuç: yüksek Dil eşitsizliği: düşük Critical hata kuyruğu: düşük fakat sıfır değil Bilinmeyen kütle: düşük Zaman kararlılığı: yüksek

Ürün B — Sentetik Dağılım İmzası

Genel sonuç: yüksek Dil eşitsizliği: yüksek Critical hata kuyruğu: maddi Bilinmeyen kütle: daha yüksek Zaman kararlılığı: düşük ve düşüş eğilimli Bu nedenle Bölüm 2'nin sentetik hükmü şudur:

> *Ürün A ile Ürün B'nin aynı ortalama skora sahip olması, aynı temsil kalitesine sahip olduklarını göstermez.*

25. ZORUNLU NORMATİF HÜKÜMLER

CH02-N01

Bir GEO sonucu yalnız tek ortalama veya tek geçiş oranıyla tanımlanamaz.

CH02-N02

Her dağılım sonucu:

- denetlenen varlığa,
- hedef nüfusa,
- AI ürünü ve kullanıcı yüzeyine,
- ülke ve dile,
- prompta,
- oturum koşuluna,
- zamana

bağlanmalıdır.

CH02-N03

Küresel ortalama, maddi alt grup eşitsizliğini tek başına telafi edemez.

CH02-N04

Critical ve diğer yüksek şiddetli hata sonuçları genel yanlış oranından ayrı raporlanmalıdır.

CH02-N05

UNKNOWN, sonuçsuz, ret, geçersiz ve NOT TESTED durumları birbirinden ayrılmalıdır.

CH02-N06

Ölçülmeyen ülke, dil, ürün, kullanıcı yüzeyi veya zaman dönemi olumlu dağılımın parçası gibi gösterilemez.

CH02-N07

Ağırlıklar ve toplulaştırma kuralları sonuçlar görülmeden önce tanımlanmalıdır.

CH02-N08

Sonuç sonrası ağırlık değişikliği yeni analiz sürümü oluşturmalı ve önceki sonuç korunmalıdır.

CH02-N09

Maddi ölçüde farklı hedef nüfus, istem, ürün, zaman veya değerlendirme kuralına dayanan dağılımlar açıklama olmadan doğrudan karşılaştırılamaz.

CH02-N10

Biçimsel dil değişkenliği otomatik olarak temsil hatası sayılamaz.

CH02-N11

Maddi kimlik, kapsam, zaman, yetkinlik veya tavsiye değişkenliği gerekçesiz biçimde stil farkı olarak sınıflandırılamaz.

CH02-N12

Yüksek kararlılık tek başına kalite kanıtı değildir; kararlı yanlış ayrıca kaydedilmelidir.

CH02-N13

Yüksek ortalama, düşük kapsamla birlikteyse kamu sonucu kapsam sınırlaması taşımalıdır.

CH02-N14

Alt grup skorları yetersiz örneklem ve geniş belirsizlik altında kesin ülke veya dil hükmü olarak yayımlanmamalıdır.

CH02-N15

Ret ve teknik sonuçsuzluk oranları gerçek kullanıcı deneyiminin parçası olarak korunmalıdır.

CH02-N16

Kamuya sunulan tek skor, kanonik dağılım ve kapsam kaydına bağlanmalıdır.

CH02-N17

Sentetik dağılımlar gerçek AI ürünü veya gerçek varlık performansıyla birleştirilemez.

CH02-N18

Dağılım raporunun hesap verebilir insan veya kurum sahibi bulunmalıdır.

26. BAŞARISIZLIK BİÇİMLERİ

CH02-F01 — ORTALAMA AKLAMA

Yüksek küresel ortalama gösterilir; dil, ülke veya Critical hata bozulması saklanır.

CH02-F02 — ALT GRUP SİLME

Düşük sonuç veren gruplar rapordan çıkarılır veya "önemsiz" ilan edilir.

CH02-F03 — KAPSAM AKLAMA

Dar kapsamlı iyi sonuç bütün dünya, bütün diller veya bütün ürünler için kullanılır.

CH02-F04 — ŞİDDET DÜZLEŞTİRME

Küçük eksiklik ile ağır sağlık veya lisans yanlışı aynı hata sayısı içinde birleştirilir.

CH02-F05 — ZAMAN ORTALAMASIYLA BOZULMAYI SAKLAMA

Düşüş veya dalgalanma yalnız dönem ortalamasıyla görünmez hâle getirilir.

CH02-F06 — KARARLILIK TAPINMASI

Tutarlı cevabın otomatik olarak doğru olduğu varsayılır.

CH02-F07 — KELİME KELİME AYNI CEVAP YANILGISI

Anlamsal doğruluk yerine birebir aynı metin kararlılık ölçüsü olarak kullanılır.

CH02-F08 — BAĞLAMSAL FARKI HATA SAYMA

Gerçek kullanıcı ihtiyacına dayanan uygun tavsiye farkı keyfî değişkenlik olarak işaretlenir.

CH02-F09 — KEYFÎ DEĞİŞKENLİĞİ KİŞİSELLEŞTİRME SAYMA

Eşdeğer kullanıcı koşullarındaki maddi çelişkiler açıklamasız biçimde "kişiselleştirme" olarak savunulur.

CH02-F10 — PAYDA DEĞİŞTİRME

Olumlu sonucu yükseltmek için tamamlanan, geçerli veya değerlendirilen gözlem paydaları arasında sonuç sonrası geçiş yapılır.

CH02-F11 — SONUÇ SONRASI AĞIRLIKLANDIRMA

En yüksek puanı üreten ülke, dil veya model katsayıları sonradan seçilir.

CH02-F12 — UYUMSUZ HÜCRELERİ BİRLEŞTİRME

Farklı istem, ürün, zaman veya panel koşulları tek dağılım gibi sunulur.

CH02-F13 — RETLERİ SİLME

Yanıt vermeyen sistem kayıtları yalnız cevap üretenler arasında yüksek doğruluk oluşturmak için çıkarılır.

CH02-F14 — KÜÇÜK HÜCREDEN KESİN HÜKÜM

Bir veya birkaç kullanıcıdan ülke ya da dil skoru üretilir.

CH02-F15 — CRITICAL KUYRUĞU GİZLEME

Nadir ağır yanlışlar genel geçiş oranının içinde kaybedilir.

CH02-F16 — TEK SKOR RAPORU

Kapsam, eşitsizlik, kararlılık, belirsizlik ve ağır hatalar gösterilmeden yalnız başlık puanı yayımlanır.

CH02-F17 — DAĞILIMI "AI'NIN İNANCI" SAYMA

Gözlenen kullanıcı yüzeyi sonuçları yapay zekânın değişmez ve içsel görüşü gibi sunulur.

CH02-F18 — ÖLÇÜLMEYENİ SIFIR VARYANS SAYMA

Test edilmeyen alanlar sorunsuz ve kararlı kabul edilir.

27. DENETİM PROSEDÜRÜ

Bir GEO raporunun temsil dağılımını dürüstçe gösterip göstermediği aşağıdaki adımlarla incelenmelidir.

Adım 1 — Kamu İddiasını Kaydet

Örnek: "Markanın küresel NOMOS skoru 92'dir." Tam cümle ve bulunduğu yüzey kaydedilir.

Adım 2 — Hedef Dağılımı Tanımla

Skor hangi:

- kullanıcı evrenine,
- AI ürünlerine,
- ülkelere,
- dillere,
- zaman aralığına

uygulanmaktadır?

Adım 3 — Ölçüm Hücrelerini Çıkar

Her maddi kombinasyon belirlenir:

- Ürün
- Ülke
- Dil
- Panel
- İstem
- Dalga
- Kullanıcı yüzeyi

Adım 4 — Paydaları Yeniden Kur

Atanan görev Tamamlanan görev Geçerli kayıt Hakem kararı verilen kayıt Geçen yanıt sayıları ayrılır.

Adım 5 — Sonuç Statülerini Çıkar

Yalnız geçiş oranı değil:

- kısmi,
- desteklenmeyen,
- çelişkili,
- eski,
- Critical,

UNKNOWN,

ret, sonuçsuz dağılımı incelenir.

Adım 6 — Ağırlıkları Doğrula

Ağırlıklar:

- önceden tanımlanmış mı,

- hangi soruyu cevaplıyor,

sonuçtan sonra değişmiş mi?

Adım 7 — Alt Grup Farklarını İncele

En iyi grup En zayıf yeterli grup Dil farkı Ülke farkı Ürün farkı Panel farkı değerlendirilir.

Adım 8 — Kapsam ve Eksikliği İncele

Hangi alanlar ölçülmedi? Hangi hücreler yetersiz? Eksiklik nedenleri nedir? Eksik veri sistematik olabilir mi?

Adım 9 — Kararlılığı İncele

Dalgalar Oturumlar Kullanıcılar Arayüzler arasındaki değişim değerlendirilir.

Adım 10 — Kuyruk Riskini İncele

Critical ve diğer ağır hatalar:

- sayısı,
- oranı,
- türü,
- kullanıcı etkisi

bakımından ayrı değerlendirilir.

Adım 11 — Kamu Özetini Dağılımla Karşılaştır

Kamuya yayımlanan tek sayı veya cümle, tam dağılımın yönünü ve riskini doğru taşıyor mu?

Adım 12 — Sonucu Sınırlandır

Kamu iddiası yalnız dağılımın ve kapsamın taşıdığı seviyeye daraltılır.

28. GEREKLİ KANIT

Bir temsil dağılımı veya küresel GEO skoru için ilgili olduğu ölçüde şu kayıtlar aranmalıdır:

- Denetlenen varlık kaydı
- Hedef nüfus tanımı
- Örnekleme çerçevesi
- AI ürünü ve kullanıcı yüzeyleri
- Ülke ve dil hücreleri
- İstem ve sürümler
- Oturum ve panel koşulları
- Ölçüm dalgaları
- Atanan görev sayıları
- Tamamlanan görevler
- Geçerli ve geçersiz kayıtlar
- Ham yanıtlar
- Sonuç statüsü dağılımı
- Hakemlik kayıtları

- Critical hata kayıtları
- Ret ve sonuçsuzluk kayıtları
- UNKNOWN kayıtları
- Ağırlıklar
- Paydalar
- Alt grup sonuçları
- Kararlılık sonuçları
- Kapsam dışı alanlar
- Güven veya belirsizlik hesapları
- Kamu skor kartı
- Sorumlu insan veya kurum
- Protokol ve analiz sürümü
- Değişiklik geçmişi

29. DENETİM KONTROL LİSTESİ

Dağılımın hedef nüfusu açık mı? Denetlenen AI ürünleri ve kullanıcı yüzeyleri belli mi? Ülke, dil, istem ve zaman kapsamı görünür mü? Tek skorun hangi ağırlıklarla üretildiği açıklanmış mı? Payda tam olarak tanımlanmış mı? Geçersiz kayıtlar ayrıca gösterilmiş mi? Ret ve sonuçsuzluklar korunmuş mu? UNKNOWN ve NOT TESTED ayrılmış mı? Alt grup sonuçları görünür mü? Küçük hücreler aşırı kesinlikle yorumlanmış mı? En zayıf maddi grup saklanmış mı? Critical hatalar ayrı raporlanmış mı? Zaman dalgaları ayrı gösterilmiş mi? Ortalama düşüş veya oynaklığı saklıyor mu? Biçimsel fark ile maddi anlam farkı ayrılmış mı? Bağlama duyarlı tavsiye değişimi doğru sınıflandırılmış mı? Aynı sağlayıcının farklı yüzeyleri açıklamasız birleştirilmiş mi?

Kapsam dışı alanlar küresel sonuca dâhil edilmiş mi? Sentetik ve gerçek dağılımlar ayrılmış mı? Kamuya sunulan cümle tam dağılımın gerçek yönünü taşıyor mu? Maddi sorulardan biri cevapsızsa dağılım sonucu sınırlı güven taşır. Birden fazlası cevapsızsa tek küresel skor kamuya kesin hüküm olarak sunulmamalıdır.

30. İTİRAZLAR VE CEVAPLAR

İtiraz 1 — “Kamu tek bir sayı ister.”

Doğrudur. Tek sayı iletişimi kolaylaştırabilir. NOMOS tek sayıyı tamamen reddetmez. Ancak tek sayı:

- kapsam,
- Critical hata,
- alt grup tabanı,
- zaman,
- belirsizlik

bilgilerine bağlanmalıdır. Bir sayı kısa olabilir. Gerçeği kısaltamaz.

İtiraz 2 — “Küçük diller küresel skoru gereksiz yere düşürür.”

Küresel nüfus ağırlığı küçük dili sınırlı ağırlıkla hesaba katabilir. Fakat küçük bir dildeki ağır temsil bozulması yalnız nüfus payı düşük olduğu için görünmez bırakılamaz. Bu nedenle:

- nüfus ağırlıklı ana skor,

- ayrı dil adaleti sonucu

birlikte yayımlanabilir. İki metrik farklı sorulara cevap verir.

İtiraz 3 — "Kişiselleştirme varsa farklı yanıtlar normaldir."

Bazı farklılıklar normal ve doğru olabilir. Ancak kişiselleştirme:

- şirket kimliğini,
- lisansını,
- güncel fiyatını,
- gerçek hizmet alanını

keyfî biçimde değiştirmemelidir. Kişiselleştirme kullanıcıya uygun anlatımı ve tavsiyeyi etkileyebilir. Maddi gerçeği icat etme izni değildir.

İtiraz 4 — "Aynı cevap verilmemesi neden sorun olsun?"

Sorun kelime farklılığı değildir. Sorun maddi gerçekliğin değişmesidir. NOMOS birebir metin eşitliği değil, koşullu anlamsal kararlılık arar.

İtiraz 5 — "Bir Critical hata 1.000 cevapta yalnız bir kez oluştuysa neden önemli olsun?"

Bu hata nüfus yaygınlığı bakımından nadir olabilir. Ancak zarar bakımından ağır olabilir. Bir uçağın bin uçuşun birinde kontrolsüz biçimde kritik sistem kaybetmesi yalnız yüzde 0,1 olarak değerlendirilmez. GEO bağlamında da sıklık ile zarar ayrı incelenmelidir.

İtiraz 6 — "Her küçük ülke veya dil için yeterli örneklem bulamayız."

Doğrudur. Bu nedenle:

- küresel nüfus paneli,
- ülke gözcü paneli,
- dil adaleti paneli

ayrı kurulabilir. Bir kişilik gözlem ülke skoruna dönüşmez. Yalnız gözlem bulunduğunu gösterir. Yetersiz örneklemde doğru statü:

INSUFFICIENT SAMPLE

veya:

NOT ESTIMATED

olmalıdır.

İtiraz 7 — "Bu kadar ayrıntı müşterinin anlayamayacağı kadar karmaşık."

Müşteriye bütün formüllerin ana sayfada gösterilmesi gerekmez. Ancak karar için gerekli temel gerçekler saklanamaz. Kısa sonuç kartı şu soruları cevaplamalıdır: Genel sonuç nedir? En ciddi risk nedir? En zayıf alan hangisidir? Ne test edilmedi? Sonuç ne kadar kararlıdır? Hangi tarihe aittir? Karmaşıklığı yönetmek gerekir. Ortadan kaldırmak değil.

İtiraz 8 — "Aynı model bütün ülkelerde kullanılıyorsa ülke farkı neden oluşsun?"

Kullanıcının gördüğü çıktı yalnız çıplak modelden oluşmayabilir. Ülke farkı:

- ürün erişimi,

- web kaynakları,
- locale,
- dil,
- plan,
- arayüz,
- mevzuat,
- kişiselleştirme

üzerinden ortaya çıkabilir. Fark oluşmayabilir. Fakat ölçmeden yokluğu varsayılamaz.

İtiraz 9 — "AI yanıtları değişken olduğu için güvenilir skor zaten mümkün değildir."

Değişkenlik ölçümü imkânsız kılmaz. Tam tersine, ölçülmesi gereken özelliklerden biridir. Skor:

- belirli tarih,
- kapsam,
- güven aralığı,
- kararlılık bilgisi

taşıdığı sürece anlamlı olabilir. Zamansız kesinlik iddiası ise savunulamaz.

İtiraz 10 — "İstem değişirse dağılım tamamen değişir."

Doğrudur. Bu nedenle dağılım prompta koşulludur. GEO-1000 bütün olası soruların tek cevabını üretmez. Önceden tanımlanmış:

- Core Mirror İstem,
- Diagnostic İstem Registry

üzerinden belirli temsil alanlarını ölçer. İstem anayasası Bölüm 10'da tanımlanacaktır.

32. BÖLÜMÜN ORTAK HÜKMÜ

Bu bölüm tek bir skora karşı yazılmadı. Tek bir skorun yalnız başına gerçekliğin tamamı sanılmasına karşı yazıldı. İnsanlar sayı ister. Kurumlar sıralama ister. Yönetim kurulları tek satır görmek ister. Pazar: "NOMOS skorunuz kaç?" diye sorabilir. Bu soru meşrudur. Fakat cevap yalnız: "90" olamaz. Çünkü 90'ın arkasında şu gerçekliklerden herhangi biri bulunabilir:

- Bütün dillerde tutarlı doğruluk
- Bazı dillerde 99, bazılarında 45
- Yüksek ortalama fakat ağır Critical hata
- Yüksek doğruluk fakat düşük kapsam
- Yüksek anlık sonuç fakat hızlı düşüş
- Yalnız kolay promptlarda başarı
- Yalnız değerlendirilebilen cevaplar arasında yüksek oran
- Test edilmemiş ülkelerin olumlu varsayılması
- Seçilmiş AI ürünlerinin birleştirilmesi

Bir skorun değeri, arkasındaki dağılımın görünür olmasıyla oluşur. Dağılım görünmüyorsa skor güven üretmez. Yalnız kesinlik görünümü üretir. Bu nedenle NOMOS'un ikinci ölçüm yasası şudur:

> *Bir GEO skoru dağılımın özeti olabilir. / Dağılımın yerine geçemez.*

Üçüncü yasası şudur:

> *Yüksek ortalama, kimin kötü temsil gördüğü sorusunu ortadan kaldırmaz.*

Dördüncü yasası şudur:

> *Nadir ağır yanlış, yalnız nadir olduğu için önemsiz değildir.*

Beşinci yasası şudur:

> *Test edilmemiş alan dağılımın olumlu tarafı değildir; bilinmeyen sınırdır.*

NOMOS'un Bölüm 2 Emri

> *Bana yalnız ortalamayı gösterme.*
>
> *Ortalamanın altında kimlerin kaldığını göster.*
>
> *En iyi dilini gösterirken en kötü dilini saklama.*
>
> *Büyük ülkelerin ağırlığını kullanıp küçük ülkelerin varlığını silme.*
>
> *Küçük ülkeye bir kişi gönderip o ülke adına kesin skor açıklama.*
>
> *Yüzde 99 doğru cevabın içine saklanmış tek ağır yanlışı önemsizleştirme.*
>
> *Cevabın kelimeleri değişti diye hata ilan etme. / Gerçeklik değiştiğinde ise bunu yalnız stil farkı sayma.*
>
> *Retleri, bilinmeyenleri ve test etmediklerini dağılımın dışına atma.*
>
> *Sonucu gördükten sonra ağırlığı, paydayı veya kapsamı değiştirme.*
>
> *Bir sayı üret. / Fakat sayının arkasındaki ülkeleri, dilleri, kullanıcıları, zamanları ve hataları görünmez yapma.*
>
> *Çünkü bir varlığın temsili ortalama bir cümle değildir. / Farklı insanların karşılaştığı farklı gerçekliklerin dağılımıdır.*

Bölümün Son Cümlesi

> *GEO'nun ölçüm nesnesi, bir yapay zekânın tek cevabı veya tek ortalama puanı değil; tanımlı kullanıcı nüfusunda, tanımlı koşullar altında oluşan temsil sonuçlarının tamamıdır.*

Normatif Öz

> **Bir varlığın üretken sistemlerdeki temsili; tanımlı AI ürünleri ve kullanıcı yüzeyleri, kullanıcılar veya örnekleme birimleri, ülkeler, diller ve locale'ler, oturum ve kişiselleştirme durumları, promptlar ve ölçüm zamanları arasında oluşan koşullu gözlenebilir sonuç dağılımı olarak değerlendirilmelidir. Tek bir toplulaştırılmış skor; kapsamı, sonuç dağılımını, alt grup farklarını, Critical hata oranını, bilinmeyen ve test edilmemiş kütleyi, zaman kararlılığını, ağırlıkları ve belirsizliği görünmez kılamaz. Aynı ortalama, maddi ölçüde farklı dağılımlar için eşdeğer temsil kalitesi anlamına gelmez.**

BÖLÜM 03

Denetlenen Varlık Tam Olarak Nedir?

YÖNTEMSEL DAYANAKLAR / K10; K11

Bölüm Sınırı

Bölüm 1 şu hükmü kurdu:

> *Tek bir yanıt GEO skoru değildir.*

Bölüm 2 şu hükmü ekledi:

> *Bir varlığın üretken sistemlerdeki temsili, kullanıcılar, AI ürünleri, ülkeler, diller, oturumlar ve zamanlar arasında oluşan bir dağılımdır.*

Bu bölüm bir sonraki zorunlu soruyu cevaplar:

> *Bu dağılım tam olarak kimin veya neyin temsilini ölçmektedir?*

Bir şirket mi? Bir marka mı? Bir alan adı mı? Bir ürün mü? Bir hizmet mi? Bir kurucu mu? Bir şirketler grubu mu? Belirli ülkedeki yerel tüzel kişilik mi? Bir franchise işletmesi mi? Bir yapay zekâ sistemi yalnızca apple.com alan adını gördüğünde:

- alan adını,
- markayı,
- ana şirketi,
- yerel satış kuruluşunu,
- ürün ailesini,
- mağaza hizmetini

aynı varlık olarak mı değerlendirecektir? Denetim başlamadan bu ayrım yapılmamışsa elde edilen sonuç doğru hesaplanabilir fakat yanlış nesneye uygulanabilir. Bu bölüm:

- denetlenen varlığı,
- istemde kullanılan kimlik çapasını,
- sistemin çözümlediği varlığı,
- uygunluk kararının uygulanacağı birimi,
- varlıklar arası ilişkileri,
- özelliklerin hangi varlığa ait olduğunu,
- zamansal ve coğrafi kimlik sınırlarını

tanımlar. Bu bölüm henüz:

- varlık hakkındaki bütün iddiaların nasıl kanıtlanacağını,

- hakemlerin atomik iddiaları nasıl puanlayacağını,
- ülke ve dil örneklem kotalarını,
- nihai NOMOS skor formülünü

kesinleştirmez. Bölüm 3'ün görevi daha temeldir:

> *Ölçüm başlamadan önce ölçülen nesnenin kimliğini ve sınırlarını kilitlemek.*

NOMOS Challenge

Bana şu soruyu soruyorsunuz: "Apple.com nasıl bir şirkettir?" Sorunun içinde en az üç ayrı nesne bulunabilir:

- apple.com alan adı
- Apple markası
- Alan adı ve markayla ilişkili hukuki şirket

Ben bunlardan birini seçerim. Belki doğru olanı. Belki yanlış olanı. Belki birden fazlasını tek cevapta birleştiririm. Sonra siz cevabı puanlarsınız. Fakat hangi varlığı puanladığınızı önceden söylemediniz. Ben:

- alan adını çevrim içi mağaza olarak anlatmış olabilirim,
- markayı tüketici elektroniği markası olarak anlatmış olabilirim,
- hukuki şirketi küresel kurum olarak anlatmış olabilirim,

ürünlerini doğrudan şirketin kendisiymiş gibi sıralamış olabilirim. Cevabım akıcı görünebilir. Ama denetim nesnesi belirsizse doğru olup olmadığı da belirsiz kalır. Şimdi daha karmaşık bir örnek düşünün.

Bir marka, bir ülkede ana şirket tarafından; başka ülkede bağlı kuruluş tarafından; üçüncü ülkede bağımsız distribütör tarafından işletiliyor. Kullanıcı soruyor: "Bu marka Türkiye'de hizmet veriyor mu?" Hangi varlık cevap vermelidir? Markanın küresel sitesi mi? Türkiye'deki yerel şirket mi? Distribütör mü?

Franchise işletmecisi mi? Küresel şirketin sahip olduğu lisans, yerel franchise'a otomatik olarak geçer mi? Yerel işletmenin müşteri yorumları ana markanın küresel kalitesini kanıtlar mı? Ana şirketin çalışan sayısı yerel şirketin kapasitesi gibi kullanılabilir mi?

Bir ürünün ödülü onu satan bütün şirketlere aktarılabilir mi? Bir kurucunun kişisel deneyimi şirketin kurumsal yaşı olur mu? Bir alan adının eski sahibiyle yeni sahibi aynı varlık sayılır mı?

Bir şirket başka şirketi satın aldığında önceki şirketin bütün başarıları yeni şirketin özelliğine dönüşür mü? Bu sorular cevaplanmadan GEO skoru üretilebilir. Fakat neyin skoru olduğu bilinmez. Bu bölümün ilk hükmü şudur:

> *Varlık çözülmeden temsil ölçülemez.*

İkinci hükmü şudur:

> *Aynı adı paylaşmak, aynı varlık olmak değildir.*

Üçüncü hükmü şudur:

> *Bir ilişki, özelliklerin taraflar arasında otomatik olarak aktarılmasına izin vermez.*

1. BÖLÜMÜN AMACI

Bu bölümün amacı, GEO-1000 içinde ölçülen varlığın kimliğini, sınırını ve ilişkilerini sonuçlar görülmeden önce belirlemektir. Bölüm şu ayrımları normatif hâle getirir:

- Alan adı ile hukuki şirket
- Marka ile marka sahibi
- Ana şirket ile bağlı kuruluş
- Grup ile grup şirketi
- Ürün ile üretici
- Hizmet ile hizmet sağlayıcı
- Kurucu ile kurum
- Çalışan ile kurum
- Lisans sahibi ile markayı kullanan taraf
- Franchise veren ile franchise işletmecisi
- Üretici ile distribütör
- Platform ile platformdaki satıcı
- Teknoloji sağlayıcısı ile teknoloji kullanıcısı
- Müşteri ile iş ortağı
- Aktif ilişki ile tarihsel ilişki
- Resmî ad ile açıklayıcı çeviri
- Aynı adlı farklı varlıklar
- Bir varlığın önceki ve sonraki sürümleri
- Denetlenen nesne ile uygunluk kararının uygulanacağı nesne

Bu bölümün sonunda her GEO denetimi şu soruya tek ve izlenebilir cevap verebilmelidir:

> *Bu skor ve karar tam olarak hangi varlığa, hangi ilişkilerle, hangi dönemde ve hangi dijital yüzeyler için aittir?*

2. MERKEZÎ NORMATİF HÜKÜM

> *Her GEO-1000 denetimi, veri toplanmadan önce benzersiz bir denetim varlığı, kanonik varlık kaydı, kapsam sınırı, zaman sınırı ve ilişki grafiği tanımlamak zorundadır.*

Bir varlığın yalnız:

- adı,
- logosu,
- alan adı,
- sosyal medya hesabı

denetim nesnesini belirlemeye yetmez. Denetim varlığı en azından ilgili olduğu ölçüde şu unsurlarla tanımlanmalıdır:

- Varlık türü
- Kanonik adı
- Hukuki adı
- Marka adları
- Resmî yerelleştirilmiş adlar

- Kısaltmalar ve alias'lar
- Kuruluş veya başlangıç tarihi
- Yargı alanı
- Kayıt numarası veya eşdeğer resmî kimlik
- Ana ve bağlı şirket ilişkileri
- Marka sahipliği
- Alan adı kontrolü
- Ürün ve hizmet sahipliği
- Gerçek hizmet teslimat sorumlusu
- İddia sahibi
- Kanıt sahibi
- Aktif ülkeler ve locale'ler
- Geçerlilik dönemi
- Tarihsel ad ve ilişkiler
- Denetim kapsamına alınan yüzeyler
- Kapsam dışı varlık ve yüzeyler

Bu kayıt tamamlanmadan ana GEO skoru yayımlanmamalıdır.

3. VARLIK NEDİR?

Bu protokolde varlık, kendisine ayrı ve izlenebilir kimlik, özellik, ilişki, sorumluluk veya zaman statüsü atanabilen nesnedir. Varlık:

- insan,
- hukuki şirket,
- kamu kurumu,
- marka,
- ürün,
- hizmet,
- alan adı,
- dijital uygulama,
- şirketler grubu,
- fiziksel şube,
- program,
- sertifika,
- yayın,
- araştırma projesi

olabilir. Her adı geçen şey ayrı bir varlık olmak zorunda değildir. Fakat aşağıdaki alanlardan biri maddi biçimde değişiyorsa ayrı varlık kaydı gerekebilir:

- Hukuki sorumluluk
- Mülkiyet
- Lisans
- Ürün veya hizmet teslimatı

- Yargı alanı
- Zaman dönemi
- Kullanıcıya karşı yükümlülük
- Kanıt ve iddia sahipliği
- Denetim kapsamı

4. GEO DENETİMİNDE ALTI AYRI NESNE

Bir GEO denetiminde "varlık" kelimesi çoğu zaman altı ayrı işlevi gizler. Bu altı nesne birbirinden ayrılmalıdır.

4.1. İstem Çapası

İstem çapası, kullanıcı sorusunda varlığı çağırmak için kullanılan ifade veya işarettir. Örnekler:

- Şirket adı
- Marka adı
- Alan adı
- Kurucu adı
- Ürün adı
- Uygulama adı
- Yerel şirket adı
- Kısaltma

Örnek: "Apple.com nasıl bir şirkettir?" Buradaki istem çapası: apple.com alan adıdır. İstem çapası, hedef varlığın kendisi olmak zorunda değildir. Yalnız sistemin varlığı çözümlemesi için kullanılan giriş noktasıdır.

4.2. Hedef Varlık

Hedef varlık, protokolün gerçekte ölçmek istediği varlıktır. Örnek: apple.com alan adıyla ilişkilendirilen kanonik kurumsal varlık hedeflenebilir. Fakat bunun tam kimliği denetim öncesinde kayıt altına alınmalıdır. Hedef varlık, istemden sonra sonuçlara bakılarak seçilemez.

4.3. Sistem Tarafından Çözümlenen Varlık

Çözümlenen varlık, AI ürününün istem çapasını hangi varlık olarak yorumladığıdır. Gösterim:

```
E_i
```

olarak kullanılabilir. Hedef varlık:

E*

ise kimlik çözümleme değerlendirmesi şu temel soruyu sorar:

```
E_i = E* ?
```

Fakat eşitlik her zaman yalnız ad eşitliği değildir. Sistem doğru hukuki varlığı adlandırmadan doğru marka ve kurumsal nesneyi tanımlayabilir. Bu nedenle kimlik eşleşmesi daha sonra:

- tam eşleşme,
- kabul edilebilir marka eşleşmesi,
- kısmi eşleşme,
- yanlış varlık,

- belirsiz,
- çoklu varlık birleşmesi

olarak sınıflandırılabilir.

4.4. İddia Nesnesi

İddia nesnesi, yanıt içindeki belirli özelliğin ait olduğu varlıktır. Aynı cevap içinde:

- kuruluş tarihi ana şirkete,
- ürün özelliği ürüne,
- müşteri yorumu yerel hizmet sağlayıcısına,
- lisans belirli tüzel kişiliğe,
- kurucu deneyimi kişiye

ait olabilir. Cevap doğru hedef varlıkla başlamış olsa bile özellikleri yanlış nesnelere aktarabilir. Bu nedenle her maddi iddia şu yapıda değerlendirilmelidir:

```
(özne, ilişki veya özellik, değer, zaman, kapsam)
```

Örnek: "X, Y lisansına sahiptir." Burada lisansın hangi:

- tüzel kişiliğe,
- ülkede,
- tarihler arasında,
- faaliyet için

ait olduğu belirlenmelidir.

4.5. Denetim Nesnesi

Denetim nesnesi, kanıt, içerik, teknik yüzey ve AI çıktılarının birlikte incelendiği tanımlı varlık veya varlık paketidir. Bir denetim nesnesi tek varlık olabilir. Örnek: Belirli bir hukuki şirket Ya da sınırları açık bir paket olabilir: Ana marka + kanonik alan adı + belirtilen iki ürün + Türkçe ve İngilizce kamu yüzeyleri Paket kullanılıyorsa içindeki varlıklar ayrı kimliklerini korumalıdır. "Paket" bütün özelliklerin birbirine aktarılması anlamına gelmez.

4.6. Uygunluk Birimi

Uygunluk birimi, denetim kararının hukuken ve normatif olarak uygulanacağı kesin nesnedir. Bir uygunluk kararı:

- bütün şirketler grubuna,
- bütün dillere,
- bütün alan adlarına,
- bütün ürünlere

otomatik olarak uygulanamaz. Örnek: "NobleJackal markasının example.com alanındaki Türkçe ve İngilizce kamu sayfaları, belirtilen tarih aralığında ve belirtilen standarda göre uygun bulunmuştur." Bu karar:

- şirketin bütün iç süreçlerini,
- bağlı şirketleri,
- ayrı uygulamaları,
- gelecekte eklenecek dilleri,
- kapsam dışı ürünleri

otomatik olarak kapsamaz.

5. VARLIK TÜRLERİ

Her denetim varlığı bir veya birden fazla türle sınıflandırılmalıdır. Ana tür ve ikincil türler ayrılabilir.

5.1. Hukuki Varlık

Kanun veya ilgili hukuk düzeni altında ayrı hak, sorumluluk veya kayıt taşıyan şirkettir. Alanları ilgili olduğu ölçüde:

- Hukuki ad
- Şirket türü
- Kayıt yeri
- Kayıt numarası
- Vergi veya eşdeğer kimlik
- Kuruluş tarihi
- Aktiflik durumu
- Yetkili temsilci
- Ana ve bağlı kuruluş ilişkileri

Hukuki varlık markayla aynı adı taşıyabilir. Fakat aynı nesne değildir.

5.2. Marka

Ürün, hizmet veya kurumsal kimliğin pazarda kullanılan adıdır. Marka:

- ayrı tüzel kişilik olmayabilir,
- başka şirket tarafından sahip olunabilir,
- lisansla kullanılabilir,

birden fazla şirket tarafından farklı ülkelerde işletilebilir. Bir markanın yaptığı söylenen işin gerçek teslimat sorumlusu ayrıca belirlenmelidir.

5.3. Şirketler Grubu

Ortak sahiplik veya kontrol ilişkisi taşıyan birden fazla hukuki varlığın oluşturduğu yapıdır. “Grup” şu özellikleri otomatik olarak bütün üyelere aktaramaz:

- Lisans
- Sertifika
- Müşteri sayısı
- Çalışan sayısı
- Gelir
- Ofis
- Denetim sonucu
- Güvenlik statüsü
- Uygunluk işareti

Grup düzeyindeki değer gerçekten konsolide ise açıkça belirtilmelidir.

5.4. Bağlı Kuruluş veya İştirak

Başka bir varlığın sahiplik veya kontrol ilişkisi altında bulunan hukuki varlıktır. Bağlı kuruluş:

- ayrı sözleşme tarafı,

- ayrı lisans sahibi,
- ayrı işveren,
- ayrı veri sorumlusu

olabilir. Ana şirketin özellikleri varsayılan olarak bağlı kuruluşa aktarılmaz.

5.5. Alan Adı

Bir dijital adres ve yayın yüzeyidir. Alan adı:

- hukuki kişi değildir,
- kendi başına hizmet teslim etmez,
- marka sahibi olmak zorunda değildir,

belirli zamanlarda farklı sahip veya işletmeciler tarafından kontrol edilebilir. Bir alan adı denetimin başlangıç çapası olabilir. Fakat kararın hangi varlık adına verildiği ayrıca gösterilmelidir.

5.6. Web Sitesi veya Dijital Yayın

Alan adı üzerindeki belirli içerik ve uygulama bütünüdür. Aynı alan adı:

- kurumsal site,
- mağaza,
- destek portalı,
- yatırımcı sayfası,
- geliştirici belgeleri,
- farklı ülke ve dil sürümleri

taşıyabilir. Bunların içerik sahipleri ve denetim kapsamları farklı olabilir.

5.7. Ürün

Belirli işlev, model veya sürüm taşıyan ticari ya da teknik nesnedir. Bir ürünün:

- fiyatı,
- güvenliği,
- teknik özelliği,
- sertifikası,
- kullanıcı uygunluğu

şirkete ait genel özellik gibi kullanılamaz. Aynı ürün ailesindeki sürümler de farklılık gösterebilir.

5.8. Hizmet

Belirli kapsam, kullanıcı, ülke, teslimat yöntemi ve sorumluluk altında sunulan faaliyettir. Bir hizmet:

- marka tarafından pazarlanabilir,
- başka hukuki varlık tarafından faturalandırılabilir,

bağımsız partner tarafından teslim edilebilir. Denetim, hizmeti kimin:

- sattığını,
- teslim ettiğini,
- garanti ettiğini,
- hukuken üstlendiğini

ayırmalıdır.

5.9. Kişi

Kurucu, yönetici, çalışan, danışman, uzman veya kamuya açık temsilcidir. Bir kişinin:

- eğitimi,
- deneyimi,
- lisansı,
- görüşü

kuruma otomatik olarak aktarılmaz. Aynı şekilde kurumun ödülü veya sertifikası kişiye ait profesyonel yetkinlik sayılmaz.

5.10. Şube veya Fiziksel Lokasyon

Ana kuruma bağlı veya ayrı işletilen fiziksel hizmet noktasıdır. Bir şube:

- ana şirketin parçası,
- ayrı tüzel kişilik,
- franchise,
- bağımsız lisanslı işletmeci

olabilir. Konum ilişkisi hukuki ilişkiyi tek başına göstermez.

5.11. Franchise

Marka kullanım hakkı veya iş modeli lisansı altında faaliyet gösteren bağımsız işletme olabilir. Franchise verenin:

- itibarı,
- denetimi,
- eğitimi

franchise işletmecisinin her davranışını otomatik olarak garanti etmez. Franchise işletmecisinin yerel hatası da bütün küresel markaya otomatik olarak genellenemez. İlişki, kapsamıyla raporlanmalıdır.

5.12. Distribütör, Bayi veya Satıcı

Ürün veya hizmeti belirli pazarda satan ya da dağıtan bağımsız veya bağlı taraftır. Üreticinin lisansı, denetimi veya kurumsal uygunluğu distribütöre otomatik olarak aktarılmaz. Satıcının müşteri hizmeti performansı da üretici ürün kalitesiyle aynı metrik değildir.

5.13. Platform ve Platform Satıcısı

Pazar yeri veya platform ile platform üzerindeki bağımsız satıcı ayrı varlıklardır. Platform:

- ödeme,
- listeleme,
- teknik altyapı

sağlayabilir. Satıcı:

- ürün,
- teslimat,
- garanti,
- müşteri ilişkisi

sorumlusu olabilir. AI cevabı bu rolleri birleştirirse yanlış sorumluluk ataması oluşabilir.

5.14. Program, Rozet veya Sertifika

Bir kurumun sunduğu:

- eğitim programı,
- uygunluk işareti,
- sertifika,
- üyelik,
- rozet

ayrı varlık kaydı taşıyabilir. Programı yayımlayan kurumla programa katılan kurum aynı değildir. Rozetin özelliği, rozet taşıyan kurumun bütün özelliklerine dönüşmez.

6. VARLIK GRAFİĞİ

Tek bir varlık kaydı karmaşık ilişkileri açıklamaya yetmeyebilir. Bu nedenle her denetim için bir varlık grafiği oluşturulmalıdır. Gösterim:

G_E = (V, R)

Burada:

- V: varlık düğümleri
- R: düğümler arasındaki yönlü ilişkiler

İlişkiler açık fiillerle tanımlanmalıdır.

6.1. Temel İlişki Türleri

İlgili olduğu ölçüde şu ilişkiler kullanılabilir:

- owns — sahibidir
- ownedBy — sahibi şudur
- controls — kontrol eder
- controlledBy — kontrol edilir
- operates — işletir
- operatedBy — işletilir
- publishes — yayımlar
- publishedBy — yayımlanır
- manufactures — üretir
- manufacturedBy — üretilir
- providesService — hizmet sunar
- serviceDeliveredBy — hizmeti teslim eder
- licensesBrandTo — markayı lisanslar
- franchisesTo — franchise verir
- distributesFor — dağıtım yapar
- sells — satar
- employs — istihdam eder
- foundedBy — kurucusu şudur
- memberOf — üyesidir
- certifiedBy — belirtilen kapsamda sertifikalandırılmıştır

- auditedBy — denetlenmiştir
- partneredWith — belirtilen kapsamda iş ortaklığı vardır
- customerOf — müşterisidir
- supplierTo — tedarikçisidir
- formerlyKnownAs — eski adı
- predecessorOf — öncülüdür
- successorTo — ardılıdır
- acquiredBy — satın alınmıştır
- mergedInto — birleşmiştir
- ceasedOn — faaliyeti sona ermiştir

"İlişkili", "partner", "birlikte çalışıyor" gibi belirsiz fiiller mümkün olduğunca tek başına kullanılmamalıdır.

6.2. İlişkinin Yönü

İlişkiler yönlüdür. "A, B'ye teknoloji sağlar." ile: "B, A'ya teknoloji sağlar." aynı değildir. Bir marka bir şirkete ait olabilir. Şirket markaya ait değildir. Bir müşteri ajansla çalışmış olabilir. Ajans müşteriye ait değildir. Yön belirtilmezse sistem yanlış sorumluluk ve otorite aktarımı yapabilir.

6.3. İlişkinin Zamanı

Her maddi ilişki ilgili olduğu ölçüde:

- başlangıç tarihi,
- bitiş tarihi,
- aktiflik durumu,
- son doğrulama tarihi

taşımalıdır. Eski ortaklık aktif ortaklık değildir. Eski çalışan güncel yönetici değildir. Bir zamanlar kullanılan distribütör bugünkü yetkili temsilci olmayabilir.

6.4. İlişkinin Kapsamı

Bir ortaklık:

- belirli ürün,
- ülke,
- kampanya,
- tarih,
- etkinlik

ile sınırlı olabilir. Tek projelik iş birliği bütün kurumsal faaliyetlere genellenemez.

7. ÖZELLİK AKTARIMI VARSAYILAN OLARAK YASAKTIR

Varlıklar arasında ilişki bulunması, bir varlığın özelliğinin diğerine otomatik olarak geçmesini sağlamaz. Bu bölümün temel ilkelerinden biri şudur:

> *Özellik aktarımı, varsayılan olarak yasaktır; ancak açık ilişki, kapsam, zaman ve kanıtla izin verilebilir.*

7.1. Lisans Aktarımı

Ana şirketin lisansı:

- bağlı kuruluşa,
- franchise'a,
- distribütöre,
- markaya

otomatik olarak geçmez. Lisansın resmî kapsamı incelenmelidir.

7.2. Sertifika Aktarımı

Bir ürünün sertifikası bütün ürün ailesine aktarılamaz. Bir ofisin kalite sertifikası küresel gruba genellenemez. Bir şirketin uygunluk kararı çalışanlarının kişisel uzmanlığını kanıtlamaz.

7.3. Çalışan Sayısı Aktarımı

Grup toplamı, yerel şirket çalışan sayısı gibi kullanılamaz. Ajans ağı veya serbest çalışan havuzu doğrudan çalışan sayısı olarak gösterilemez.

7.4. Müşteri Sayısı Aktarımı

Ana grubun müşteri sayısı yeni kurulmuş markanın müşteri sayısı gibi yazılamaz. Bir platformdaki kullanıcı sayısı belirli ürünün aktif müşteri sayısı olmayabilir.

7.5. Kuruluş Tarihi Aktarımı

Kurucunun sektördeki başlangıç yılı, şirketin kuruluş tarihi değildir. Bir şirket satın aldığı eski markanın tarihini kendi hukuki kuruluş tarihi gibi kullanamaz. Marka tarihi ile tüzel kişilik tarihi ayrı gösterilmelidir.

7.6. Başarı ve Vaka Aktarımı

Bir çalışanın önceki işvereninde yaptığı proje, yeni şirketin kurumsal vaka çalışması değildir. Satın alınan şirketin başarısı, entegrasyon ve kapsam açıklanmadan yeni grubun doğrudan başarısı sayılmaz.

7.7. Uygunluk Aktarımı

Bir alan adının belirli dil sayfaları için verilen uygunluk:

- bütün şirket grubuna,
- yeni ürünlere,
- başka alan adlarına,
- gelecekteki sayfalara

aktarılmaz.

8. DÖRT AYRI SAHİPLİK KATMANI

Bir dijital temsil içinde "sahip" kelimesi birden fazla anlam taşıyabilir. Dört katman ayrılmalıdır.

8.1. Hukuki Sahiplik

Varlığın kanun veya resmî kayıt altında sahibidir.

8.2. Marka Sahipliği

Marka veya ticari kimlik üzerindeki hak sahibidir.

8.3. Dijital Yüzey Kontrolü

Alan adı, web sitesi, sosyal profil veya uygulamayı fiilen kontrol eden taraftır. Hukuki sahip ile teknik kontrol sahibi farklı olabilir.

8.4. İçerik ve İddia Sahipliği

Yayımlanan maddi iddianın doğruluğundan sorumlu taraftır. Ajans siteyi yönetebilir. Fakat fiyat veya lisans iddiasının gerçek sahibi müşteri kurum olabilir. Bir teknik sağlayıcı içeriği yayımladığı için iddianın maddi sahibi hâline gelmez.

9. HİZMETİ KİM SATIYOR, KİM TESLİM EDİYOR?

GEO değerlendirmesinde en önemli varlık ayrımlarından biri budur. Bir hizmet zincirinde şu taraflar farklı olabilir:

- Markayı kullanan
- Reklamı yapan
- Teklifi hazırlayan
- Sözleşmeyi imzalayan
- Faturayı kesen
- Hizmeti fiilen teslim eden
- Garanti veren
- Veriyi işleyen
- Şikâyeti yöneten
- Hukuki sorumluluğu taşıyan

AI cevabı yalnız marka adını söyleyerek bütün bu rolleri tek tarafa atayabilir. Bu hata özellikle:

- turizm,
- finans,
- sağlık,
- danışmanlık,
- franchise,
- platform,
- ajans,
- eğitim

alanlarında maddi sonuç doğurabilir. Her hizmet kaydı şu sorulara cevap vermelidir: Kullanıcı kiminle sözleşme yapar? Ödemeyi kim alır? Hizmeti kim gerçekleştirir? Mesleki veya hukuki sorumluluğu kim taşır? Garanti ve iade yükümlülüğü kimdedir? Marka hangi roldedir? Üçüncü taraf sağlayıcılar var mıdır? Bu ayrım, GEO skorunun hangi varlığa uygulanacağını da belirler.

10. VARLIK SINIRI

Varlık sınırı, denetlenen nesnenin hangi özellik, ilişki ve yüzeyleri içerdiğini; hangilerini içermediğini belirleyen çizgidir. Bir varlık sınırı dört katmanda tanımlanabilir.

10.1. Kimlik Çekirdeği

Varlığın değişmeden kalması gereken temel kimliğidir:

- Kanonik ad
- Varlık türü
- Hukuki veya kurumsal kimlik
- Ana faaliyet
- Ana sahiplik
- Aktiflik durumu

10.2. Kontrollü Çevre

Varlığın doğrudan kontrol ettiği:

- alan adları,
- uygulamalar,
- markalar,
- ürünler,
- belgeler,
- resmî profiller

alanıdır. Kontrol, bütün özelliklerin tek kimlikte birleşmesi anlamına gelmez.

10.3. Bağlı Çevre

Bağlı kuruluşlar, yerel şirketler, franchise'lar ve kontrollü partnerlerdir. Bunlar ayrı kimliklerini korur.

10.4. Dış İlişki Çevresi

Müşteriler, tedarikçiler, medya, üyelikler, teknoloji sağlayıcıları ve bağımsız ortaklardır. Bu çevredeki varlıklar denetlenen varlığın parçası değildir. İlişki nesnesidir.

11. VARLIK SINIRININ SONUÇLAR GÖRÜLMEDEN KİLİTLENMESİ

Denetim nesnesi sonuçlara göre genişletilemez veya daraltılamaz. Aşağıdaki davranışlar yasaktır:

- Olumlu grup sonuçlarını markaya eklemek
- Başarısız bağlı şirketi kapsamdan çıkarmak
- İyi ürün sonucunu bütün ürün ailesine yaymak
- Kötü dil sürümünü ayrı varlık ilan etmek
- Critical hatayı yerel işletmeye bırakıp küresel markayı temiz göstermek
- Güçlü yerel kanıtı küresel şirketin kanıtı yapmak

Varlık kapsamı:

- denetim başlamadan,
- promptlar dağıtılmadan,
- sonuçlar görülmeden

önce sürümlenmelidir.

12. KANONİK VARLIK KAYDI

Her denetlenen varlık için Kanonik Varlık Kaydı oluşturulmalıdır. Bu kayıt, hakemlerin ve sistemlerin hangi nesneyi değerlendirdiğini açıklar.

12.1. Asgari Alanlar

Kimlik

NOMOS varlık kimliği Kanonik ad Varlık türü Hukuki ad Kısaltmalar Resmî yerel adlar Eski adlar

Hukuki ve Kurumsal Statü

Yargı alanı Kayıt numarası Kuruluş tarihi Aktiflik durumu Ana şirket Bağlı kuruluşlar

Marka ve Ürünler

Sahip olunan markalar Lisanslanan markalar Ürünler Hizmetler Sürüm veya model ayrımları

Dijital Yüzeyler

Kanonik alan adları Yerel alan adları Uygulamalar Resmî profiller Makine tarafından okunabilir kayıtlar

Sorumluluk

İddia sahibi İçerik sahibi Teknik yüzey sahibi Hizmet teslimat sahibi Hukuki sorumlu

Zaman

Kayıt sürümü Yürürlük başlangıcı Geçerlilik bitişi Son doğrulama tarihi Tarihsel değişiklikler

Kapsam

Denetlenen ülkeler Denetlenen diller Denetlenen ürünler Kapsam dışı varlık ve yüzeyler

12.2. Varlık Statüleri

Her varlık kaydı şu statülerden birini taşımalıdır:

ACTIVE

INACTIVE

HISTORICAL

SUPERSEDED

MERGED

ACQUIRED

DISSOLVED

PROVISIONAL

CONTESTED

UNKNOWN

UNKNOWN veya CONTESTED kimlik sessizce kesinleştirilemez.

13. KİMLİK KANIT DÜZEYLERİ

Bir varlık kaydının kimlik güveni, kullanılan kaynaklara göre sınıflandırılabilir.

ID-0 — Kimlik Belirsiz

Yalnız ad veya alan adı vardır. Hukuki veya kurumsal eşleme doğrulanmamıştır.

ID-1 — Birinci Taraf Beyanı

Kurum kendisini belirli biçimde tanımlamaktadır. Değerlidir. Bağımsız doğrulama değildir.

ID-2 — Desteklenmiş Kurumsal Kimlik

Birinci taraf kayıtları, alan adı ve başka güvenilir kayıtlar birbiriyle uyumludur.

ID-3 — Resmî Kayıtla Doğrulanmış Kimlik

Hukuki varlık uygun resmî veya güvenilir kayıtlarla doğrulanmıştır.

ID-4 — Çok Yüzeyli ve İlişki Doğrulanmış Kimlik

Kimlik, marka, ürün, alan adı ve kurumsal ilişkiler birden fazla güvenilir kayıt üzerinden tutarlıdır.

ID-5 — Zaman İçinde Sürdürülmüş Kimlik

Kimlik ve ilişkiler farklı dönemlerde doğrulanmış, değişiklik geçmişi korunmuştur. Bu düzeyler nihai uygunluk puanı değildir. Varlık kaydının kanıt güvenini gösterir.

14. AYNI ADLI FARKLI VARLIKLAR

Aynı veya benzer ada sahip farklı varlıklar bulunabilir. Bu durum varlık çarpışması oluşturur. Örnek ayrımlar:

- Aynı adlı şirketler
- Aynı kısaltmayı kullanan kurumlar
- Farklı ülkelerde aynı marka adı
- Eski ve yeni şirket
- Kişi adı ile marka adı
- Ürün adı ile şirket adı
- Genel kelime ile marka

Denetim, yalnız ad eşleşmesine dayanamaz.

14.1. Çarpışma Çözücü Alanlar

Yargı alanı Hukuki şirket türü Alan adı Fiziksel adres Kuruluş tarihi Kurucu veya yönetici Ürün ve hizmet Marka sahibi Kayıt numarası Yerel dil adı

14.2. Belirsiz İstem

İstem birden fazla makul varlığa işaret ediyorsa üç seçenek vardır: İstem açıklaştırılır. Varlık çözümleme testi olarak ayrı sınıflandırılır. Sistemden açıklama istemesi beklenir. Belirsiz istemden elde edilen cevap ana doğruluk skoruna sessizce eklenemez.

15. BİR VARLIK, BİRDEN FAZLA ALAN ADI

Bir şirket:

- küresel alan,
- yerel ülke alanları,
- ürün alanları,
- yatırımcı alanı,
- destek alanı,
- eski alanlar

kullanabilir. Bu yüzeylerin tamamı aynı denetim kapsamına otomatik olarak girmez. Her alan için:

- kontrol sahibi,
- içerik sahibi,
- hedef kullanıcı,
- dil,
- ürün kapsamı,
- aktiflik durumu

belirlenmelidir. Bir alan adındaki doğru kayıt başka alandaki yanlışı otomatik olarak düzeltmez.

16. BİR ALAN ADI, BİRDEN FAZLA VARLIK

Tek alan adı üzerinde:

- ana şirket,
- farklı markalar,
- yerel şirketler,
- ürünler,
- partnerler,
- satıcılar

yer alabilir. Bu durumda URL veya içerik ailesi bazında varlık sahipliği tanımlanmalıdır. Örnek:

Yüzey	Varlık rolü
Ana sayfa	Küresel marka
Yatırımcı ilişkileri	Hukuki ana şirket
Yerel fiyat sayfası	Yerel satış kuruluşu
Marketplace	Bağımsız satıcılar
Kariyer sayfası	Belirtilen işveren varlıklar
Destek portalı	Ürün veya servis operatörü

Tek alan adı tek sorumluluk anlamına gelmez.

17. ZAMANSAL VARLIK KİMLİĞİ

Varlık kimliği zaman içinde değişebilir. Örnek olaylar:

- Şirket adının değişmesi
- Markanın başka şirkete geçmesi
- Birleşme
- Satın alma
- Bölünme
- Faaliyetin sona ermesi
- Alan adının el değiştirmesi
- Ürünün başka şirkete devri
- Kurucunun ayrılması
- Yerel operatörün değişmesi

Bu nedenle varlık kaydı zamana bağlıdır. Gösterim: E(t) olarak düşünülebilir.

17.1. Tarihsel Kimlik ile Güncel Kimlik

Geçmişte doğru olan ilişki bugüne taşınamaz. Örnek: "X, 2022 yılında Y tarafından işletiliyordu." cümlesi tarihsel olabilir. Şu cümle farklıdır: "X, Y tarafından işletilmektedir."

17.2. Satın Alma Sonrası Özellik Aktarımı

Satın alma:

- mülkiyeti,
- bazı varlıkları,
- belirli sözleşmeleri

aktarabilir. Fakat:

- bütün tarihsel başarıları,
- bütün sertifikaları,
- bütün müşteri ilişkilerini,
- bütün lisansları

otomatik olarak yeni kuruma taşımaz. Hangi özelliklerin hukuken ve fiilen devredildiği açıklanmalıdır.

17.3. Alan Adı Geçmişi

Bir alan adı eski sahibine ait içerik veya atıf izleri taşıyabilir. Yeni sahibi eski otoriteyi otomatik olarak kazanmış sayılmaz. AI sistemi eski ve yeni varlığı birleştiriyorsa tarihsel varlık karışıklığı oluşur.

18. ÇOK DİLLİ VARLIK KİMLİĞİ

Bir varlığın farklı dillerdeki adları aynı kimliğe bağlanmalıdır. Fakat bütün çeviriler resmî ad değildir. Aşağıdaki türler ayrılmalıdır:

- Hukuki ad
- Marka adı
- Resmî yerelleştirilmiş ad
- Transliterasyon
- Açıklayıcı çeviri
- Kısaltma
- Eski ad

- Halk arasında kullanılan ad

Örnek: Bir markanın Türkçe açıklaması, markanın hukuki veya resmî Türkçe adı olmak zorunda değildir. Makine kaydı bu türleri açıkça göstermelidir.

19. ÜLKE VE YARGI ALANI VARLIK SINIRINI DEĞİŞTİREBİLİR

Küresel marka altında farklı ülkelerde farklı hukuki veya operasyonel varlıklar bulunabilir. Bir kullanıcı: "X şirketi Almanya'da hizmet veriyor mu?" diye sorduğunda cevap:

- küresel markanın varlığını,
- yerel tüzel kişiliği,
- distribütör ilişkisini,
- uzaktan hizmet kapasitesini

ayırmalıdır. "Marka ülkede tanınıyor" ile: "Yerel hukuki şirket hizmet sunuyor." aynı değildir.

19.1. Ülke Bazlı Varlık Rolleri

Her ülke için ilgili olduğu ölçüde:

- Hukuki satıcı
- Fatura kesen
- Veri sorumlusu
- Hizmet teslim eden
- Destek sağlayan
- Distribütör
- Franchise
- Lisans sahibi
- Marka temsilcisi

belirlenmelidir.

19.2. Genel Dil Sayfası Yerel Şirket Kanıtı Değildir

Bir sitenin Almanca olması Almanya'da tüzel kişilik veya fiziksel ofis bulunduğunu kanıtlamaz. İngilizce sayfa küresel hukuki yetki oluşturmaz. Dil, yargı alanı değildir.

20. SENTETİK APPLE.COM VARLIK GÖSTERİMİ

SENTETİK METODOLOJİ GÖSTERİMİ / Aşağıdaki örnek yalnız varlık çözümleme yöntemini göstermek amacıyla hazırlanmıştır. Gerçek Apple Inc. denetimi, güncel kurumsal yapı incelemesi veya herhangi bir AI ürününün gerçek sonucu değildir. İstem: "Apple.com nasıl bir şirkettir?"

20.1. İstem Çapası

apple.com

20.2. Aday Hedefler

Alan adı ve web sitesi Apple markası Kanonik kurumsal varlık Çevrim içi mağaza işlevi Ürün ailesi Denetim başlamadan hedef tanımlanmazsa hakemler farklı nesneleri puanlayabilir.

20.3. Önerilen Hedef Varlık Kilidi

Sentetik sınama için hedef şöyle tanımlanabilir: "apple.com kanonik alan adıyla ilişkilendirilen ana kurumsal teknoloji varlığının, şirket kimliği ve temel faaliyetlerinin temsili." Bu tanım:

- tek tek ürünlerin bütün teknik özelliklerini,
- her ülkenin yerel şirketini,
- mağaza fiyatlarını,
- yatırım tavsiyesini

otomatik olarak kapsam içine almaz.

20.4. Sentetik Yanıt A

"Apple.com, teknoloji ürünleri ve dijital hizmetlerle ilişkili küresel bir kurumsal web sitesidir." Değerlendirme: Alan adı ile kurumu tamamen aynılaştırmamaktadır. Şirket kimliğini eksik bırakabilir. Ana faaliyet genel olarak doğru olabilir. Maddi kimlik eksikliği bulunabilir.

20.5. Sentetik Yanıt B

"Apple.com, iPhone adlı şirketin resmî web sitesidir." Değerlendirme: Ürün ile şirket birleştirilmiştir. Hedef varlık yanlış çözülmüştür. Kimlik hatası vardır.

20.6. Sentetik Yanıt C

"Apple.com yalnız çevrim içi telefon satan bağımsız bir mağazadır." Değerlendirme: Alan adı işlevi aşırı daraltılmıştır. Kurumsal varlık ve faaliyet kapsamı bozulmuştur. Bağımsız mağaza niteliği yanlış olabilir.

20.7. Sentetik Yanıt D

"Apple.com, Apple markasının ana kurumsal ve ürün bilgi yüzeylerinden biridir; siteyi işleten hukuki varlık ve yerel satış tarafı ülkeye göre ayrıca belirlenmelidir." Değerlendirme: Alan adı, marka ve yerel hukuki rol ayrılmıştır. Hedef varlık çözümlemesi daha disiplinlidir. Yanıtın diğer maddi iddiaları ayrıca kanıtlanmalıdır. Bu örnek şunu gösterir:

> *Bir cevap çok sayıda doğru kelime içerdiği hâlde yanlış varlığı merkez alabilir.*

Kimlik doğruluğu, içerik doğruluğundan önce gelir.

21. SENTETİK KARMAŞIK VARLIK VAKASI

SENTETİK VAKA — GERÇEK KURUM DEĞİLDİR

Asteron adlı küresel marka düşünelim. Yapı:

- Asteron Holdings Ltd. — ana şirket
- Asteron Travel — marka
- Asteron Türkiye Turizm Ltd. — yerel şirket
- Mira Destination Services — bağımsız yerel operasyon partneri
- Asteron Club Antalya — franchise işletmesi
- asteron.example — küresel alan adı
- tr.asteron.example — Türkçe pazarlama yüzeyi

Kullanıcı soruyor: "Asteron Türkiye'de hangi hizmetleri veriyor?" AI cevabı: "Asteron Holdings Ltd. Antalya'daki bütün turları kendisi işletir, Türkiye'de 400 çalışanı vardır ve Asteron Club Antalya'nın

müşteri memnuniyetinden doğrudan sorumludur." Gerçeklik paketi: Ana şirket markanın sahibidir. Yerel şirket satış ve sözleşme tarafıdır. Mira operasyonların bir bölümünü teslim eder. Asteron Club Antalya bağımsız franchise'dır. 400 çalışan grup toplamıdır. Franchise müşteri memnuniyetinin doğrudan işletme sorumlusu franchise işletmecisidir. Cevaptaki hatalar: Ana şirket hizmet teslimat sorumlusu yapılmıştır. Grup çalışan sayısı yerel şirket kapasitesi gibi aktarılmıştır. Franchise sorumluluğu ana şirkete doğrudan yüklenmiştir. Marka, hukuki taraf ve operasyon partneri birleştirilmiştir. Yanıt markayı doğru adlandırmış olabilir. Fakat varlık grafiği yanlıştır.

Bu nedenle GEO skoru yalnız adın geçmesine veya genel cümlenin akıcılığına dayanamaz.

22. VARLIK ÇÖZÜMLEME SONUÇLARI

Her gözlem için varlık çözümleme sonucu şu statülerden biriyle kaydedilebilir.

ER-0 — NOT EVALUATED

Varlık çözümleme değerlendirilmemiştir.

ER-1 — EXACT TARGET

Hedef varlık doğru ve açık biçimde çözülmüştür.

ER-2 — ACCEPTABLE BRAND-LEVEL MATCH

Hukuki ad verilmemiştir fakat hedeflenen marka ve kurumsal nesne maddi olarak doğru çözülmüştür.

ER-3 — PARTIAL OR UNDER-SPECIFIED

Yanıt doğru alana yaklaşmış fakat hukuki, coğrafi veya operasyonel kimlik eksiktir.

ER-4 — MULTI-VARLIK CONFLATION

Birden fazla ayrı varlık tek nesne gibi birleştirilmiştir.

ER-5 — WRONG VARLIK

Başka kişi, şirket, ürün veya alan adı hedef alınmıştır.

ER-6 — AMBIGUOUS

Yanıttan hangi varlığın çözümlendiği belirlenememektedir.

ER-7 — APPROPRIATE CLARIFICATION

İstem gerçekten belirsizdir ve sistem doğru biçimde açıklama istemiştir. Bir açıklama talebi başarısızlık olmak zorunda değildir.

23. VARLIK ÇÖZÜMLEME BAŞARISI İLE TEMSİL DOĞRULUĞU AYRI ŞEYDİR

Bir sistem doğru varlığı çözümleyebilir fakat onun hakkında yanlış bilgi verebilir. Başka bir sistem bazı maddi olguları doğru söyleyebilir fakat yanlış varlığa bağlayabilir. Dört temel durum vardır:

Varlık çözümleme	Maddi içerik	Sonuç
Doğru	Doğru	Güçlü doğru temsil adayı

Varlık çözümleme	Maddi içerik	Sonuç
Doğru	Yanlış	Doğru nesne, yanlış özellik
Yanlış	Kısmen doğru	Doğru bilgiler yanlış varlığa bağlanmış
Yanlış	Yanlış	Tam temsil bozulması

Bu nedenle kimlik çözümleme ayrı bir geçiş kapısı olabilir. Özellikle başka varlıkla ciddi karışıklık bulunan cevap, yüksek olgusal doğruluk ortalamasıyla kurtarılamaz.

24. UYGUNLUK KARARININ VARLIĞA BAĞLANMASI

Bir uygunluk kararı şu formatta kurulmalıdır:

> *[Varlık kimliği] + [varlık türü] + [kapsam] + [dijital yüzeyler] + [ülke ve diller] + [zaman] + [standart sürümü]*

Örnek şablon: “NOMOS Varlık ID NGE-XXXX ile tanımlanan [marka/hukuki varlık] için; belirtilen alan adındaki Türkçe ve İngilizce kamu sayfaları, belirtilen ürün ve hizmet kapsamı içinde, [tarih] tarihinde [standart sürümü] altında değerlendirilmiştir.” Ardından açıkça yazılmalıdır: “Bu karar bağlı kuruluşlara, franchise’lara, kapsam dışı ürünlere, başka alan adlarına ve gelecekte eklenecek dillere otomatik olarak uygulanmaz.”

25. ZORUNLU NORMATİF HÜKÜMLER

CH03-N01

Her GEO denetimi, sonuçlar görülmeden önce benzersiz bir hedef varlık tanımlamak zorundadır.

CH03-N02

İstem çapası ile hedef varlık aynı şeymiş gibi varsayılamaz.

CH03-N03

Alan adı, marka, hukuki şirket, ürün, hizmet ve kişi kimlikleri birbirinden ayrılmalıdır.

CH03-N04

Bir varlığın özelliği ilişkili başka varlığa açık kanıt ve kapsam olmadan aktarılamaz.

CH03-N05

Ana şirketin lisansı, sertifikası, çalışan sayısı, müşterisi veya uygunluk sonucu bağlı kuruluşlara otomatik olarak uygulanamaz.

CH03-N06

Marka düzeyindeki iddia ile hizmeti fiilen teslim eden ve hukuken sorumlu taraf ayrılmalıdır.

CH03-N07

Franchise veren, franchise işletmecisi, distribütör, bayi ve platform satıcısı ayrı varlıklar olarak kaydedilmelidir.

CH03-N08

Her maddi varlık ilişkisi yön, kapsam ve zaman bilgisi taşımalıdır.

CH03-N09

“Partner”, “müşteri”, “uzman”, “grup şirketi” ve benzeri ilişkiler belirsiz biçimde kullanılamaz.

CH03-N10

Aynı adlı farklı varlıklar için çarpışma çözüm alanları kullanılmalıdır.

CH03-N11

Belirsiz istem ana skor içine alınacaksa belirsizlik yöntemi önceden tanımlanmalıdır; aksi hâlde istem açıklaştırılmalıdır.

CH03-N12

Varlık kapsamı sonuçlar görüldükten sonra olumlu sonucu artırmak için genişletilemez veya daraltılamaz.

CH03-N13

Tarihsel varlık ve ilişkiler güncel varlık statüsü gibi kullanılamaz.

CH03-N14

Satın alma, birleşme veya marka devri, geçmiş özelliklerin yeni varlığa otomatik aktarımı anlamına gelmez.

CH03-N15

Bir uygunluk kararı yalnız kararda belirtilen uygunluk birimine uygulanabilir.

CH03-N16

Yeni ürün, dil, alan adı, bağlı şirket veya ülke mevcut uygunluğu otomatik olarak devralamaz.

CH03-N17

Varlık çözümleme başarısı ile maddi doğruluk ayrı değerlendirilmelidir.

CH03-N18

Yanlış varlıkla maddi kimlik birleşmesi, genel olgusal ortalama içinde görünmez kılınamaz.

CH03-N19

Kanonik Varlık Kaydı sürümlenmeli ve zaman içindeki değişiklikleri korumalıdır.

CH03-N20

Varlık kaydının hesap verebilir insan veya kurum sahibi bulunmalıdır.

26. BAŞARISIZLIK BİÇİMLERİ

CH03-F01 — ALAN ADINI ŞİRKET SAYMAK

Dijital adres hukuki veya operasyonel varlık gibi sunulur.

CH03-F02 — MARKA İLE HUKUKİ ŞİRKETİ BİRLEŞTİRMEK

Marka adı, sözleşme ve sorumluluk taşıyan tüzel kişiliğin yerine kullanılır.

CH03-F03 — ÜRÜNÜ ŞİRKET SAYMAK

Ürün adı şirket veya marka kimliği olarak çözülür.

CH03-F04 — KURUCUNUN ÖZELLİĞİNİ ŞİRKETE AKTARMAK

Kurucunun yaşı, deneyimi, ödülü veya önceki projesi kurum özelliği yapılır.

CH03-F05 — GRUP ÖZELLİĞİNİ YEREL ŞİRKETE AKTARMAK

Konsolide müşteri, çalışan veya gelir verisi yerel işletmeye atanır.

CH03-F06 — LİSANSI İLİŞKİ ÜZERİNDEN AKTARMAK

Ana şirket, ürün, çalışan veya partner lisansı başka varlık adına kullanılır.

CH03-F07 — FRANCHISE VE ANA MARKAYI TEKLEŞTİRMEK

Yerel işletme performansı ve sorumluluğu doğrudan küresel marka olarak sunulur.

CH03-F08 — ÜRETİCİ VE SATICIYI BİRLEŞTİRMEK

Ürün üreticisiyle yerel satıcı veya distribütör tek varlık sayılır.

CH03-F09 — PLATFORM VE SATICIYI BİRLEŞTİRMEK

Marketplace platformu satıcının ürün, garanti veya teslimat sorumlusu yapılır.

CH03-F10 — İLİŞKİYİ ÖZELLİK AKTARIMI SAYMAK

Müşterilik, üyelik, teknoloji kullanımı veya ortaklık otorite transferine dönüştürülür.

CH03-F11 — TARİHSEL KİMLİĞİ GÜNCELLEŞTİRMEK

Eski ad, sahiplik, ortaklık veya yönetici bugünkü gerçek gibi kullanılır.

CH03-F12 — SATIN ALMA SONRASI GEÇMİŞİ SAHİPLENMEK

Satın alınan varlığın bütün geçmiş başarıları açıklamasız biçimde yeni sahibin başarısı yapılır.

CH03-F13 — AYNI ADLI VARLIKLARI BİRLEŞTİRMEK

İsim benzerliği nedeniyle farklı şirket veya kişiler tek varlık olarak çözülür.

CH03-F14 — BİR VARLIĞI YAPAY OLARAK PARÇALAMAK

Başarısız dil, ürün veya yerel operasyon olumlu skoru korumak için ayrı varlık ilan edilir.

CH03-F15 — BİR VARLIĞI YAPAY OLARAK GENİŞLETMEK

Başarılı marka veya ürün sonuçları bütün gruba yayılır.

CH03-F16 — UYGUNLUK KAPSAMI AKLAMA

Dar denetim sonucu bütün şirket, grup veya ürün portföyü için kullanılır.

CH03-F17 — DİLİ YARGI ALANI SAYMAK

Bir dil sürümünün bulunması o ülkede hukuki şirket veya hizmet kapasitesi kanıtı yapılır.

CH03-F18 — TEKNİK KONTROLÜ MADDİ SAHİPLİK SAYMAK

Ajans veya geliştirici alan adını yönettiği için kurumun hukuki ya da ticari sahibi gibi gösterilir.

CH03-F19 — İDDİA SAHİBİNİ BELİRSİZ BIRAKMAK

Yanlış fiyat, lisans veya kapasite bilgisinden kimin sorumlu olduğu belirlenemez.

CH03-F20 — İSTEM ÇAPASINI SONUÇTAN SONRA YORUMLAMAK

Cevabın başarılı görünmesini sağlayacak hedef varlık sonradan seçilir.

27. DENETİM PROSEDÜRÜ

Adım 1 — İstem Çapasını Kaydet

Kullanıcının hangi:

- ad,
- alan adı,
- marka,
- ürün,

- kişi

ifadesini kullandığı tam olarak kaydedilir.

Adım 2 — Hedef Varlığı Tanımla

Denetimin gerçekte hangi varlığı ölçtüğü yazılır. Hedef:

- hukuki şirket,
- marka,
- ürün,
- hizmet,
- varlık paketi

olarak sınıflandırılır.

Adım 3 — Kanonik Varlık Kimliği Oluştur

Kanonik ad Varlık türü Hukuki kimlik Yargı alanı Alias'lar Zaman statüsü kaydedilir.

Adım 4 — Varlık Grafiğini Kur

Ana şirket, bağlı kuruluş, marka, ürün, alan adı, kişiler ve partnerler arasındaki yönlü ilişkiler çıkarılır.

Adım 5 — Sorumluluk Rollerini Belirle

İddia sahibi Dijital yüzey sahibi Hizmet teslimat sahibi Hukuki sorumlu Kanıt sahibi ayrılır.

Adım 6 — Özellik Aktarım Riskini İncele

Cevapta:

- lisans,
- sertifika,
- çalışan,
- müşteri,
- kuruluş tarihi,
- başarı,
- uygunluk

yanlış varlığa aktarılmış mı?

Adım 7 — Zamansal Kimliği Kontrol Et

Cevap:

- eski ad,
- eski sahip,
- sona ermiş ortaklık,
- eski çalışan,
- satın alma öncesi durum

kullanıyor mu?

Adım 8 — Coğrafi ve Yerel Rolleri Kontrol Et

Küresel marka, yerel şirket, distribütör ve franchise rolleri doğru mu?

Adım 9 — Sistem Tarafından Çözümlenen Varlığı Sınıflandır

ER-1 ile ER-7 arasındaki uygun statü atanır.

Adım 10 — Uygunluk Birimini Kilitle

Nihai kararın uygulanacağı:

- varlık,
- alan,
- ürün,
- dil,
- ülke,
- zaman

sınırı yazılır.

Adım 11 — Kamu İddiasını Sınırlandır

Sonuç, tanımlanan varlıktan daha geniş bir kurum veya gruba genelleniyorsa daraltılır.

Adım 12 — Değişiklik ve İtiraz Yolunu Kaydet

Kimlik veya ilişki hatası bildirilirse:

- kim doğrulayacak,
- kim değiştirecek,
- hangi sürüm oluşacak

belirlenir.

28. GEREKLİ KANIT

Kanonik varlık kimliği Hukuki ad ve varlık türü Resmî kayıt veya eşdeğer doğrulama Marka sahipliği Alan adı kontrolü Ana ve bağlı şirket ilişkileri Ürün ve hizmet sahipliği Yerel işletme ve distribütör kayıtları Franchise ilişkileri Sözleşme ve fatura tarafı Hizmet teslimat sorumlusu Lisans ve sertifika sahibi Kurucu ve yönetici rolleri Aktif ve tarihsel adlar Satın alma, birleşme veya devir kayıtları Yerel ülke ve yargı alanları Resmî dil ve transliterasyon kayıtları Varlık grafiği İddia sahibi İçerik sahibi Teknik yüzey sahibi Denetim kapsamı Kapsam dışı varlık ve yüzeyler Varlık kaydı sürümü Son doğrulama tarihi Sorumlu insan veya kurum

29. DENETİM KONTROL LİSTESİ

İstem çapası açıkça kaydedildi mi? Hedef varlık sonuçlardan önce tanımlandı mı? Alan adı ile hukuki şirket ayrıldı mı? Marka ve marka sahibi ayrıldı mı? Ürün ve şirket ayrıldı mı? Kurucu ve kurum özellikleri ayrıldı mı? Ana şirket ve bağlı kuruluş ilişkileri doğru mu? Franchise ve distribütörler ayrı mı? Hizmeti satan ve teslim eden taraflar belli mi? İddia sahibi ile teknik yayıncı ayrıldı mı? Lisans ve sertifikalar doğru varlığa mı ait? Çalışan, müşteri ve gelir sayıları doğru kapsamda mı? Tarihsel ad ve ilişkiler güncel durumdan ayrıldı mı? Aynı adlı farklı varlıklar kontrol edildi mi? Birden fazla alan adı ve yüzey doğru varlığa bağlandı mı?

Çok dilli adlar aynı kanonik kimliğe bağlandı mı? Ülke ve yargı alanı rolleri açık mı? Varlık çözümleme sonucu ayrı sınıflandırıldı mı? Uygunluk birimi açıkça sınırlandırıldı mı? Karar kapsam

dışındaki varlıklara yayılıyor mu? Varlık kaydı sürümlü ve tarihli mi? Kimlik değişikliği için itiraz ve düzeltme yolu var mı?

30. İTİRAZLAR VE CEVAPLAR

İtiraz 1 — "Kullanıcı marka adını bilir, hukuki şirketi değil."

Doğrudur. Kullanıcının istemde marka kullanması doğaldır. Denetim, kullanıcıyı hukuki jargon kullanmaya zorlamaz. Fakat arka plandaki değerlendirme:

- markayı,
- sahibi,
- hizmet sağlayıcıyı,
- sözleşme tarafını

ayırt etmek zorundadır. Kullanıcıya sade cevap verilebilir. Denetim kaydı sade olmak zorunda değildir.

İtiraz 2 — "Küçük işletmede marka ile şirket zaten aynı."

Bazı küçük yapılarda:

- tek marka,
- tek alan adı,
- tek işletme sahibi,
- tek hizmet sağlayıcı

bulunabilir. Bu durumda varlık grafiği basit olur. Yine de kayıt açıkça yapılmalıdır. Basitlik, belirsizlik anlamına gelmez.

İtiraz 3 — "Şirketler grubu tek marka gibi hareket ediyor."

Pazarlama bakımından tek marka olabilir. Hukuki, operasyonel ve kanıtsal sorumluluk bakımından ayrı varlıklar bulunabilir. Grup düzeyinde karar verilecekse hangi özelliklerin konsolide olduğu açıkça gösterilmelidir.

İtiraz 4 — "Bu kadar ayrıntı GEO skorunu gereksiz yere zorlaştırır."

Yanlış varlığa verilen doğru skor yararlı değildir. Varlık çözümleme denetimin ek yükü değil, geçerlilik şartıdır.

İtiraz 5 — "AI sistemleri bu hukuki farklara zaten dikkat etmez."

Tam da bu nedenle ölçülmelidir. AI sisteminin:

- marka,
- şirket,
- ürün,
- yerel işletme

arasındaki farkı kaçırması temsil hatasının kendisidir.

İtiraz 6 — "Ana şirket bütün grubu kontrol ediyorsa özellikler aktarılabilir."

Kontrol ilişkisi bazı alanlarda önemlidir. Fakat:

- lisans,
- sözleşme,

- müşteri,
- çalışan,
- veri sorumluluğu,
- uygunluk

gibi özelliklerin aktarımı ayrı kanıt gerektirir. Kontrol, bütün özelliklerin tekleştirilmesi değildir.

İtiraz 7 — "Bir satın alma sonrası eski şirketin başarıları yeni grubun parçası değil midir?"

Tarihsel bağlamda anlatılabilir. Ancak:

- başarının hangi dönemde,
- hangi ekip ve varlık tarafından,
- satın alma öncesi mi sonrası mı

üretildiği açıklanmalıdır. Yeni sahiplik geçmişi değiştirmez.

İtiraz 8 — "İstem zaten açık; kullanıcı alan adını yazmış."

Alan adı açık olabilir. Fakat: "Alan adı neyi temsil ediyor?" sorusu ayrıca cevaplanmalıdır. Bir domain şirket, mağaza, ürün portalı veya çoklu varlık platformu olabilir.

İtiraz 9 — "Yerel distribütörü marka gibi anlatmak kullanıcı için daha kolay."

Sadelik sorumluluğu bozamaz. Yanıt şöyle sadeleştirilebilir: "Marka Türkiye'de yetkili yerel distribütör aracılığıyla hizmet vermektedir." Hem anlaşılır hem doğrudur.

İtiraz 10 — "Bütün bu ilişkiler zamanla değişir."

Doğrudur. Bu nedenle varlık kaydı:

- tarihli,
- sürümlü,
- değişiklik tetikleyicili

olmalıdır. Değişkenlik kayıt ihtiyacını azaltmaz. Artırır.

33. BÖLÜMÜN ORTAK HÜKMÜ

Bir GEO skoru üretmeden önce: "Kaç cevap doğru?" diye sorarız. Fakat bundan önce sorulması gereken daha temel soru şudur:

"Cevaplar tam olarak kimin hakkında?"

Bir alan adını şirket sanarsak yanlış nesneyi ölçeriz. Markayı hukuki şirket sanarsak sorumluluğu yanlış atarız. Ürünü şirket sanarsak özellikleri bozarız. Kurucunun deneyimini şirket tarihi yaparsak geçmişi büyütürüz. Grup çalışan sayısını yerel kapasite sayarsak hizmet beklentisini şişiririz. Ana şirketin lisansını franchise'a aktarırsa kullanıcıyı yanlış yetkiye yönlendiririz. Bir satın alma sonrası geçmiş başarıları yeni sahibin özelliği yaparsak tarihsel gerçeği yeniden yazarız. Dar bir alan adı denetimini bütün şirkete yayarsak uygunluk kapsamını aklarız. Bu nedenle varlık tanımı yalnız veri temizliği değildir. Temsil sorumluluğunun başlangıcıdır. Bir sistem yanlış varlığı doğru bilgilerle anlatabilir. Bu yine yanlış temsildir. Bir sistem doğru varlığı yanlış bilgilerle anlatabilir. Bu da yanlış temsildir.

Doğru GEO için ikisi birlikte gerekir:

> *Doğru varlık. / Doğru özellik. / Doğru ilişki. / Doğru kapsam. / Doğru zaman.*

Bu nedenle NOMOS'un üçüncü ölçüm yasası şudur:

> *Ölçülen varlık açık değilse ölçülen doğruluk da açık değildir.*

Dördüncü yasası şudur:

> *İlişki, kimlik değildir.*

Beşinci yasası şudur:

> *Sahiplik, bütün özelliklerin aktarımı değildir.*

Altıncı yasası şudur:

> *Uygunluk, komşu varlıklara bulaşmaz.*

NOMOS'un Bölüm 3 Emri

> *Bana yalnız bir ad verme. / O adın hangi varlığa ait olduğunu göster.*
>
> *Alan adını şirket, markayı sözleşme tarafı, ürünü üretici yapma.*
>
> *Kurucunun geçmişini kurumun yaşı gibi kullanma.*
>
> *Ana şirketin lisansını bağlı kuruluşa, franchise'a veya distribütöre taşıma.*
>
> *Grubun müşterisini yerel markanın müşterisi, markanın ödülünü bütün ürünlerin ödülü yapma.*
>
> *Birlikte çalışmayı sahiplik, üyeliği akreditasyon, teknoloji kullanımını ortaklık gibi anlatma.*
>
> *Eski adı yeni varlıkla, satın alınan geçmişi bugünkü performansla birleştirme.*
>
> *Olumlu sonucu büyütmek için varlığı genişletme. / Başarısız sonucu saklamak için varlığı parçalama.*
>
> *Önce hedefi kilitle. / Sonra ilişkileri yönleriyle yaz. / Sonra özellikleri gerçek sahiplerine bağla. / Sonra zamanı ve ülkeyi ekle. / Ve yalnız o varlık için karar ver.*

Bölümün Son Cümlesi

> *Bir GEO denetiminin ilk sorusu, yapay zekânın ne söylediği değil; söylediği şeyin tam olarak kime veya neye ait olduğudur.*

Normatif Öz

> **Her GEO denetimi sonuçlar görülmeden önce istem çapasını, hedef varlığı, varlık türünü, kanonik kimlik kaydını, ilişki grafiğini, zamansal ve hukuki sınırı, denetim nesnesini ve uygunluk birimini tanımlayıp sürümlemek zorundadır. Alan adı, marka, hukuki şirket, şirketler grubu, bağlı kuruluş, ürün, hizmet, kişi, franchise, distribütör, platform ve satıcı birbirinin yerine kullanılamaz. Özellikler ve uygunluk, açık ilişki, kapsam, zaman ve kanıt bulunmadan ilişkili varlıklara aktarılamaz. AI ürününün çözümlediği varlık, yanıtın maddi doğruluğundan ayrı değerlendirilmelidir.**

BÖLÜM 04

İki Gerçeklik Katmanı

Resmî temsil ve doğrulanmış varlık gerçekliği

YÖNTEMSEL DAYANAKLAR / K10; K11

Bölüm Sınırı

Bölüm 1 şu hükmü kurdu:

> *Tek bir yanıt GEO skoru değildir.*

Bölüm 2 şunu ekledi:

> *Temsil, kullanıcılar, AI ürünleri, ülkeler, diller, oturumlar ve zamanlar arasında oluşan bir dağılımdır.*

Bölüm 3 ise dağılımın nesnesini belirledi:

> *Varlık çözülmeden temsil ölçülemez.*

Şimdi bir sonraki zorunlu soruya geliyoruz:

> *AI yanıtının doğru olup olmadığını hangi gerçeklik kaydına göre değerlendireceğiz?*

Denetlenen varlığın kendi web sitesi mi? Hukuki kayıtlar mı? Şirketin yayımladığı fiyat sayfası mı? Bağımsız haberler mi? Lisansı veren kurum mu? Müşteri yorumları mı? Şirketin kendi performans verisi mi? Rakiplerin iddiaları mı? Birden fazla kaynak birbiriyle çelişiyorsa hangisi doğru kabul edilecek? Bir şirket kendi sitesinde yanlış veya kanıtsız bir iddia yayımladıysa ve AI bu iddiayı kusursuz biçimde tekrar ettiyse sonuç başarılı mıdır? Bir şirketin sitesi eski kalmış, fakat AI güncel ve doğrulanmış harici kayda dayanarak farklı cevap vermişse sistem hatalı mıdır? Bu sorular cevaplanmadan:

- doğru temsil,
- yanlış temsil,
- kaynak sadakati,
- maddi doğruluk,
- varlık gerçekliği

aynı kavram gibi kullanılabilir. Bu bölüm bu karışıklığı iki ayrı gerçeklik katmanı kurarak çözer:

Birinci Katman

Resmî Temsil Katmanı

Varlığın kendisi hakkında, kontrol ettiği kanonik yüzeylerde ne söylediğini gösterir.

İkinci Katman

Doğrulanmış Varlık Gerçekliği Katmanı

Mevcut kanıtlar altında, hangi maddi iddiaların hangi kapsam ve zaman içinde desteklendiğini gösterir. Bu iki katman birbirine bağlıdır. Fakat aynı değildir. Birinci katman: "Varlık kendisini nasıl tanımlıyor?" sorusuna cevap verir. İkinci katman: "Bu tanımın hangi bölümü hangi kanıtla destekleniyor?" sorusuna cevap verir. İlk kitap doğru standardı tanımlamış, ikinci kitap ihlal biçimlerini ayırmış, denetim sistemi ise bu ayrımları ölçüp kanıtlamak üzere tasarlanmıştır. Bu bölüm, o denetim görevini iddia düzeyinde çalıştırır.

NOMOS Challenge

Şirketinizin web sitesinde şu cümle bulunuyor: "Dünyanın lider GEO şirketiyiz." Bir AI ürünü bu sayfayı okuyor ve kullanıcıya şöyle cevap veriyor: "Bu şirket dünyanın lider GEO şirketidir." Yanıt web sitenizi doğru yansıtıyor mu? Evet. Yanıt kanıtlanmış gerçekliği doğru yansıtıyor mu?

Bunu henüz bilmiyoruz. Belki:

- karşılaştırmalı araştırma vardır,
- pazar kapsamı tanımlanmıştır,
- liderlik ölçütü yayımlanmıştır,

sonuç bağımsız olarak doğrulanmıştır. Belki bunların hiçbiri yoktur. AI, şirketin kendi iddiasını: "Şirket kendisini dünyanın lider GEO şirketi olarak tanımlıyor." şeklinde aktarsaydı, iddianın kaynağını korumuş olurdu. Fakat: "Dünyanın lider GEO şirketidir." dediğinde şirket beyanını bağımsız gerçekliğe yükseltmiş olabilir.

Şimdi başka bir durum düşünün. Şirketin resmî sitesinde: "Yalnız Türkiye'de hizmet veriyoruz." yazıyor. Fakat şirket altı ay önce Almanya ve Birleşik Krallık'ta doğrulanmış operasyon başlatmış. Site güncellenmemiş. AI güncel hukuki ve operasyonel kayıtlara dayanarak: "Şirket Türkiye, Almanya ve Birleşik Krallık'ta hizmet veriyor."

diyor. AI, resmî siteye tam sadık mı? Hayır. Doğrulanmış güncel gerçekliğe daha yakın olabilir mi? Evet. Şimdi üçüncü bir durum düşünün. Şirketin sitesi: "20 ülkede faaliyet gösteriyoruz." diyor. Bağımsız ve iç kayıtlar yalnız sekiz ülkeyi doğruluyor. AI: "Şirket 20 ülkede faaliyet gösterir."

diyor. Bu yanıt:

- resmî temsile sadık,
- doğrulanmış gerçekliğe aykırı

olabilir. Bu bölümün ilk hükmü şudur:

> *Resmî kayda sadakat, maddi doğrulukla aynı şey değildir.*

İkinci hükmü şudur:

> *Bir şirketin söylediği şey ile kanıtlanabilen şey ayrı kaydedilmelidir.*

Üçüncü hükmü şudur:

> *AI'nın resmî yanlışı doğru biçimde tekrar etmesi, doğru temsil oluşturmaz.*

Dördüncü hükmü şudur:

> *AI'nın güncelliğini yitirmiş resmî kayıttan ayrılması, güvenilir kanıt taşıyorsa otomatik hata değildir.*

1. BÖLÜMÜN AMACI

Bu bölümün amacı, GEO-1000 değerlendirmesinde kullanılacak iki ayrı referans katmanını kurmaktır. Bölüm şu ayrımları normatif hâle getirir:

- Varlığın beyanı ile kanıtlanmış olgu
- Resmî kaynak ile bağımsız kaynak
- Birinci taraf kaynak ile bağımsız doğrulama
- Kaynak sadakati ile maddi doğruluk
- İddianın tekrarı ile iddianın doğrulanması
- Güncel resmî kayıt ile eski resmî kayıt
- Resmî sessizlik ile harici doğrulanmış bilgi
- Şirket görüşü ile dış gerçeklik hükmü
- Resmî kanıt ile resmî iddia
- Kaynağın bağımsızlığı ile kaynağın konuya uygunluğu
- Resmî yanlış ile AI yanlışı
- Referans çatışması ile model başarısızlığı
- Belirsizlik ile başarısızlık
- İddia sahibinin çıkarı ile iddianın doğruluğu
- Kanıtın bulunması ile kanıtın iddiayı gerçekten desteklemesi

Bu bölümün sonunda her AI yanıtı için iki ayrı soru sorulabilmelidir:

> *Yanıt, varlığın resmî ve kanonik temsilini doğru aktarıyor mu?*

ve:

> *Yanıt, mevcut doğrulanabilir kanıtların taşıdığı maddi gerçekliği doğru aktarıyor mu?*

Bu sorular aynı cevabı verebilir. Farklı cevaplar da verebilir. Farklılaştıklarında denetimin görevi farkı saklamak değil, açıklamaktır.

2. MERKEZÎ NORMATİF HÜKÜM

> *Her maddi AI yanıtı, birbirine dönüştürülmeden iki ayrı referansa karşı değerlendirilmelidir: Resmî Temsil Kaydı ve Doğrulanmış Varlık Gerçekliği Kaydı.*

Bir yanıt:

- resmî temsile sadık,
- doğrulanmış gerçekliğe uygun,
- her ikisine birden uygun,
- her ikisinden de uzak

olabilir. NOMOS bu iki sonucu tek bir doğru–yanlış kararında sessizce birleştiremez. Birinci katmanın yüksek olması ikinci katmanın başarısızlığını telafi edemez. Özellikle:

> *Varlığın kendi yayımladığı kanıtsız veya yanlış iddiayı doğru biçimde tekrar eden AI yanıtı, resmî temsil sadakati gösterebilir; fakat doğrulanmış varlık doğruluğu gösteremez.*

3. BİRİNCİ KATMAN

RESMÎ TEMSİL KATMANI

Resmî Temsil Katmanı, varlığın kontrol ettiği veya açıkça yetkilendirdiği kanonik yüzeylerde kendisi hakkında yayımladığı güncel maddi beyanların sürümlenmiş kaydıdır. Bu katman şu soruyu cevaplar:

> *Varlık, kendisinin nasıl anlaşılmasını istiyor ve hangi bilgileri resmî olarak yayımlıyor?*

Bu katman:

- hukuki kimlik beyanını,
- marka–şirket ilişkisini,
- ürün ve hizmet tanımlarını,
- yayımlanmış fiyatları,
- hizmet ülkelerini,
- destek koşullarını,
- resmî politikaları,
- kamuya açıklanmış sınırları,
- resmî güncellik ve sürüm bilgilerini

içerebilir. Resmî temsil katmanı önemlidir. Çünkü bir AI sistemi varlığın:

- kendisini nasıl tanımladığını,
- hangi hizmeti sunduğunu,
- hangi fiyatı yayımladığını,
- hangi şartları kabul ettiğini

doğru anlamalıdır. Fakat resmî temsil katmanı her iddia için otomatik bağımsız kanıt değildir. Kurumun kendi sitesindeki beyanı, özellikle üstünlük ve performans gibi dış dünyaya ilişkin iddialarda, beyanın kendisini gösterir; doğruluğunu tek başına kanıtlamaz.

4. RESMÎ KAYNAK NEDİR?

Bir kaynak yalnız kurumun logosunu taşıdığı için resmî sayılmaz. Resmî kaynak en azından şu özelliklerden ilgili olanlarını taşımalıdır:

- Varlık tarafından kontrol edilmesi
- Yetkilendirilmiş yayın sürecinden geçmesi
- Maddi bilginin sahibi tarafından onaylanması
- Kanonik URL veya kayıt kimliği
- Sürüm
- Yayın veya doğrulama tarihi
- Değişiklik geçmişi
- Sorumlu insan veya kurum
- Kapsam
- Dil ve locale statüsü
- Arşiv veya güncellik durumu

Bir çalışan tarafından kişisel sosyal medya hesabında yayımlanan görüş otomatik resmî şirket beyanı değildir. Bir ajansın taslak dosyası resmî yayın değildir. Bir satış temsilcisinin e-postası, yetki

verilmemişse genel kurumsal politika değildir. Bir şirketin web sitesindeki eski PDF, yeni resmî kaydın yerine geçmeyebilir.

5. RESMÎ TEMSİL YÜZEYLERİ

İlgili denetime göre resmî yüzeyler şunları içerebilir:

- Ana kurumsal web sitesi
- Hukuki açıklamalar
- Ürün ve hizmet sayfaları
- Fiyat ve teklif sayfaları
- Resmî yardım merkezi
- Yatırımcı ilişkileri belgeleri
- Kamuya açık politika ve sözleşmeler
- Resmî basın bültenleri
- Yapılandırılmış veri
- Kanonik JSON kayıtları
- Resmî API çıktıları
- Resmî uygulama içeriği
- Yetkilendirilmiş yerel siteler
- Resmî sosyal medya hesapları
- Kamu kurumlarına verilen beyanlar
- Sürüm ve geri çekme kayıtları

Bu yüzeyler kendi aralarında çelişebilir. Bu durumda "resmî temsil" tek ve temiz bir gerçeklik değildir. Kendi içinde denetlenmesi gereken bir kayıt kümesidir.

6. RESMÎ TEMSİL KAYDI

Her denetlenen varlık için Resmî Temsil Kaydı oluşturulmalıdır. Bu kayıt, bütün resmî içeriklerin kopyası değildir. Maddi iddiaların ve sınırların izlenebilir özetidir. Her kayıt ilgili olduğu ölçüde şu alanları taşımalıdır:

- İddia kimliği
- İddianın tam metni
- İddia öznesi
- İddia türü
- Kanonik kaynak
- Kaynak sahibi
- Dil ve locale
- Yayın tarihi
- Son doğrulama tarihi
- Geçerlilik dönemi
- Resmî statü
- Kapsam
- Sınırlamalar

- Önceki sürüm
- Değişiklik nedeni
- İddia sahibi
- İnsan onayı

Bir varlığın resmî temsil kaydı bulunmadan: "AI siteyi doğru yansıtıyor." cümlesi açık biçimde test edilemez.

7. RESMÎ KAYNAĞIN YETKİSİ İDDİA TÜRÜNE BAĞLIDIR

Bütün resmî kaynaklar her iddia için aynı ağırlığı taşımaz. Kaynak yetkisi, iddianın türüne göre değerlendirilmelidir.

7.1. Varlığın Kendi Kontrolündeki Kararlar

Örnek:

- Hangi hizmeti kamuya sunduğu
- Yayımlanmış fiyat
- Destek saatleri
- Resmî iade politikası
- İletişim kanalı
- Kamuya açık ürün sürümü

Bu alanlarda resmî kaynak temel ve çoğu zaman birincil yetkili kayıttır. Fakat resmî sayfa ile gerçek işlem akışı çelişiyorsa yalnız sayfa yeterli değildir. Örneğin:

- fiyat sayfası 1.000 avro,
- checkout 1.500 avro

gösteriyorsa resmî temsil katmanında çelişki vardır.

7.2. Hukuki Kimlik ve Yetki

Örnek:

- Şirket kaydı
- Lisans
- Mesleki yetki
- Marka sahipliği
- Vergi veya hukuki statü

Şirket sitesi bu bilgileri beyan edebilir. Ancak uygun resmî kayıt veya yetkili kurum daha güçlü doğrulama sağlayabilir.

7.3. Performans ve Başarı İddiaları

Örnek:

- "Projelerimizin yüzde 98'i başarılıdır."
- "Müşterilerimizin gelirini ortalama yüzde 40 artırıyoruz."
- "En yüksek memnuniyet oranına sahibiz."

Şirket sitesi iddianın sahibi olabilir. Fakat iddiayı desteklemek için:

- veri,

- yöntem,
- payda,
- dönem,
- başarısız sonuçlar,
- mümkünse bağımsız inceleme

gerekir.

7.4. Üstünlük ve Liderlik İddiaları

Örnek:

- "Dünya lideri"
- "En iyi"
- "Bir numara"
- "Tek gerçek standart"
- "Sektörün en güvenilir şirketi"

Kendi site beyanı bağımsız karşılaştırma değildir. Bu tür iddialar:

- karşılaştırma evreni,
- kriter,
- yöntem,
- tarih,
- kapsam,
- bağımsızlık

gerektirir.

7.5. Üçüncü Taraf İlişkileri

Örnek:

- Müşteri
- Partner
- Sponsor
- Üyelik
- Akreditasyon
- Teknoloji ortaklığı
- Distribütörlük

Tek taraflı resmî beyan değerli olabilir. Fakat ilişkinin diğer tarafça veya güvenilir kayıtla doğrulanması gerekebilir.

7.6. Kullanıcı Sonuçları

Örnek:

- Müşteri memnuniyeti
- Gerçek kullanım sonucu
- İade oranı
- Şikâyet oranı
- Tekrar satın alma

Bu alanlarda şirket verisi önemli birinci taraf kanıttır. Fakat:

- veri seçimi,
- yanıtlamama,
- metodoloji,
- çıkar çatışması

açıklanmalıdır.

8. RESMÎ TEMSİL SADAKATİ

Bir AI yanıtının Resmî Temsil Katmanı'na ne ölçüde uygun olduğunu gösteren kavram:

Resmî Temsil Sadakati

olarak adlandırılır. İngilizce karşılığı:

Official Representation Fidelity

Bu ölçüm şu soruları sorar: AI doğru varlığı mı çözümlüyor? Resmî ürün ve hizmetleri doğru mu aktarıyor? Resmî fiyatı çarpıtıyor mu? Resmî sınırları kaldırıyor mu? Resmî ülke kapsamını genişletiyor mu? Şirket görüşünü bağımsız gerçek gibi mi sunuyor? Eski resmî bilgiyi güncel gibi mi kullanıyor? Resmî kayıtlar arasındaki çelişkiyi saklıyor mu? Resmî kaynakta bulunmayan maddi iddialar ekliyor mu? Kaynak niteliğini koruyor mu? Resmî temsil sadakati, varlığın kendi yayımladığı gerçekliği doğru taşıma ölçümüdür. Bu ölçüm:

> *Varlığın söylediklerinin doğru olduğunu değil, AI'nın varlığın söylediklerini ne ölçüde doğru aktardığını gösterir.*

9. İKİNCİ KATMAN

DOĞRULANMIŞ VARLIK GERÇEKLİĞİ KATMANI

Doğrulanmış Varlık Gerçekliği Katmanı, belirli bir ölçüm tarihinde mevcut olan ve iddia türüne uygun kanıtlarla desteklenen maddi bilgilerin sürümlenmiş değerlendirme kaydıdır. Bu katman şu soruyu cevaplar:

> *Varlık hakkında söylenebilecek hangi cümleleri hangi kanıt gerçekten taşımaktadır?*

Bu katman:

- şirketin kendisinden gelen kayıtları,
- resmî kamu kayıtlarını,
- bağımsız kaynakları,
- sözleşme veya işlem kayıtlarını,
- lisans belgelerini,
- doğrulanmış tarihsel kayıtları,
- karşı kanıtları,
- olumsuz bulguları,
- belirsizlikleri,
- açık kalan uyuşmazlıkları

birlikte değerlendirebilir.

10. “DOĞRULANMIŞ GERÇEKLİK” MUTLAK GERÇEKLİK DEĞİLDİR

Bu protokol: “Bir kurum hakkındaki bütün mutlak gerçeği kesin olarak biliriz.” iddiasında bulunmaz. Doğrulanmış Varlık Gerçekliği:

- kanıta bağlıdır,
- zamana bağlıdır,
- kapsama bağlıdır,
- yeni kanıtla değişebilir,
- itiraz edilebilir,
- geri çekilebilir,

belirsizlik taşıyabilir. Doğru tanım şudur:

> *Belirtilen tarih ve kapsam içinde, mevcut kanıtların taşıdığı en güçlü ve denetlenebilir varlık kaydı.*

Bu nedenle doğrulanmış gerçeklik kaydı:

- sürümlü,
- kaynaklı,
- itiraza açık,
- değiştirilebilir,
- belirsizlik farkında

olmalıdır. NOMOS kendisini gerçeğin sahibi olarak konumlandıramaz. Yalnız kanıtların hangi cümleye ne ölçüde izin verdiğini sınıflandırır.

11. DOĞRULANMIŞ VARLIK GERÇEKLİĞİ KAYDI

Bu katmanın tam teknik yapısı kitabın ilerleyen bölümünde ayrıntılandırılacaktır. Kurucu düzeyde her maddi iddia için şu alanlar bulunmalıdır:

- İddia kimliği
- Özne varlık
- İlişki veya özellik
- Değer
- Kapsam
- Ülke veya yargı alanı
- Zaman
- Kaynaklar
- Kaynak türleri
- Kaynak kontrolü
- Bağımsızlık
- Karşı kanıt
- Kanıtın yeterliliği
- Doğrulama statüsü
- Belirsizlik
- Hakem kararı

- İtiraz durumu
- Son doğrulama tarihi
- Kayıt sürümü

Bu kayıt yalnız olumlu kanıtlardan oluşturulamaz. Gerçekliği daraltan veya iddiaya karşı çıkan kayıtlar da korunmalıdır.

12. DOĞRULAMA STATÜLERİ

Her maddi iddia şu statülerden biriyle sınıflandırılabilir:

VR-0 — NOT ASSESSED

İddia henüz değerlendirilmemiştir.

VR-1 — OFFICIAL CLAIM ONLY

İddia yalnız varlığın kendi resmî beyanına dayanmaktadır. Yanlış olduğu kanıtlanmış değildir. Bağımsız veya yeterli ek doğrulama bulunmamaktadır.

VR-2 — FIRST-PARTY EVIDENCE SUPPORTED

İddia kurumun:

- işlem,
- operasyon,
- sözleşme,
- veri,
- kayıt

kanıtlarıyla desteklenmektedir. Kaynak birinci taraftır.

VR-3 — CORROBORATED

İddia birden fazla uygun ve birbiriyle uyumlu kayıtla desteklenmektedir. Kaynakların tamamı bağımsız olmak zorunda değildir.

VR-4 — INDEPENDENTLY VERIFIED

İddia ilgili konuda uygun bağımsız kaynak veya inceleme tarafından desteklenmiştir. Bağımsızlık tek başına yeterlilik anlamına gelmez. Kaynağın konuya uygunluğu da gereklidir.

VR-5 — MULTI-SOURCE SUSTAINED

İddia farklı türlerde güvenilir kaynaklarla, tanımlı zaman ve kapsam içinde güçlü biçimde desteklenmektedir.

VR-C — CONTRADICTED

İddia daha güçlü veya daha uygun kanıtlarla çelişmektedir.

VR-U — UNRESOLVED

Kanıtlar yetersiz veya çelişkilidir. Güvenilir hüküm verilememektedir.

VR-T — TIME-LIMITED

İddia yalnız belirtilen dönem için desteklenmektedir.

VR-S — KAPSAM-LIMITED

İddia yalnız belirtilen ürün, ülke, kullanıcı veya operasyon kapsamı için desteklenmektedir. Bir iddia aynı anda birden fazla statü taşıyabilir. Örnek:

VR-4 + VR-T + VR-S

iddianın bağımsız olarak doğrulandığını; fakat zaman ve kapsam sınırı taşıdığını gösterebilir.

13. BAĞIMSIZLIK İLE UYGUNLUK AYNI ŞEY DEĞİLDİR

Bağımsız kaynak her zaman en iyi kaynak değildir. Örnek: Bir blog yazarı şirketin güncel fiyatını tahmin ediyor. Şirketin kanonik fiyat sayfası daha doğrudan ve uygun kaynak olabilir. Başka bir örnek: Şirket kendi sitesinde: "Sektör lideriyiz." diyor. Bağımsız, yöntemli bir pazar araştırması bu iddia için daha uygun kaynak olabilir. Bu nedenle kaynak değerlendirmesinde yalnız şu soru sorulmaz: "Kaynak bağımsız mı?" Şunlar birlikte sorulur: İddia türü için yetkili mi? İddiaya ne kadar yakın? Kapsamı uygun mu? Güncel mi? Kaynak kontrolü kimde? Metodolojisi açık mı? Karşı kanıtı içeriyor mu? Denetlenebilir mi? Çıkar çatışması var mı? Bir kaynağın bağımsız olması onun doğru olduğunu garanti etmez.

Bir kaynağın birinci taraf olması da onu değersiz yapmaz. Önemli olan:

> *Kaynağın taşıdığı iddia için uygun yetki ve kanıt niteliğine sahip olmasıdır.*

14. İDDİA–KAYNAK UYGUNLUK HARİTASI

Aşağıdaki tablo genel bir kurucu haritadır. Kesin kanıt gereksinimleri sektör ve risk düzeyine göre değişebilir.

İddia türü	Öncelikli kaynak veya kanıt	Resmî sitenin rolü	Ek doğrulama ihtiyacı
Hukuki şirket adı	Yetkili resmî kayıt	Beyanı gösterir	Genellikle gerekir
Marka sahipliği	Marka kaydı, sözleşme, kurumsal kayıt	Beyanı gösterir	Maddi durumda gerekir
Güncel yayımlanmış fiyat	Kanonik fiyat/checkout kaydı	Genellikle birincil	İşlem akışıyla doğrulanabilir
Hizmet kapsamı	Operasyon kayıtları, sözleşmeler, resmî sayfa	Birincil beyan	Gerçek teslimat için ek kanıt
Lisans	Lisans veren kurum veya sicil	Beyanı gösterir	Gerekir
Müşteri ilişkisi	Sözleşme, müşteri onayı, iki taraflı kayıt	Tek taraflı iddia olabilir	Maddi durumda gerekir
Başarı oranı	Ham veri, yöntem, payda, denetim	İddiayı yayımlar	Güçlü biçimde gerekir
Liderlik	Karşılaştırmalı yöntem ve veri	Öz beyan	Bağımsız veya güçlü karşılaştırma gerekir
Müşteri memnuniyeti	Temsilî anket, yöntem, yanıtlamama	Sonucu açıklayabilir	Metodoloji gerekir
Kuruluş tarihi	Resmî kayıt ve tarihsel belgeler	Beyanı gösterir	Hukuki ve marka tarihi ayrılmalı
Ortaklık	İki taraflı açıklama veya sözleşme	Tek taraflı beyan	Genellikle gerekir
Sertifika	Sertifikayı veren kurum ve kapsam	Beyanı gösterir	Gerekir
Aktif çalışan	İK veya sözleşme kaydı, kamu profili	Beyanı gösterebilir	Rol ve tarih doğrulaması
Kullanıcı sonucu	İşlem ve sonuç kaydı	Vaka anlatır	Bağlam, yöntem ve izin gerekir
"En iyi seçenek"	Tanımlı kullanıcı ve karşılaştırma kriteri	Pazarlama görüşü	Güçlü karşılaştırma gerekir

Bu tablo evrensel bir kaynak sıralaması değildir. İddia ile kanıt arasındaki uygunluğu gösterir.

15. ÇİFT KATMANLI İDDİA MATRİSİ

Her maddi iddia için iki ayrı soru sorulur: Resmî olarak yayımlanmış mı? Doğrulanmış gerçeklik kaydında hangi statüde? Bu iki soru aşağıdaki temel durumları oluşturur.

15.1. Resmî ve Doğrulanmış

İddia: resmî kayıtta bulunur, uygun kanıtla desteklenir. Örnek: Resmî fiyat sayfasındaki fiyat, işlem akışında da doğrulanmaktadır. Bu, en güçlü çekirdek temsil türlerinden biridir. AI iddiayı doğru kapsam ve zamanla aktarırsa iki katmanda da başarı oluşur.

15.2. Resmî Fakat Yalnız Öz Beyan

İddia resmîdir. Fakat ek doğrulama yoktur. Örnek: "Sektörün en yenilikçi şirketiyiz." AI, iddiayı: "Şirket kendisini sektörün en yenilikçi şirketi olarak tanımlıyor." şeklinde aktarabilir. Şu biçimde yükseltemez: "Sektörün en yenilikçi şirketidir."

15.3. Resmî Fakat Çelişen

İddia resmî kayıtta bulunur. Daha uygun veya güncel kanıtlarla çelişir. Örnek: Site 20 ülke der. Operasyon kayıtları sekiz ülke gösterir. AI siteyi aynen tekrar ederse resmî sadakati yüksek olabilir. Doğrulanmış doğruluğu düşüktür. Bu durum:

Resmî yanlışı sadakatle taşıma

olarak sınıflandırılabilir.

15.4. Resmî Değil Fakat Doğrulanmış

İddia varlığın resmî sitesinde bulunmaz. Fakat güvenilir başka kaynaklarla doğrulanmıştır. Örnek: Güncel hukuki sahiplik değişikliği resmî siteye henüz eklenmemiştir. Yetkili sicilde görünmektedir. AI bu bilgiyi doğru kaynak ve zamanla aktarabilir. Resmî temsil sadakati düşük veya farklı olabilir. Doğrulanmış gerçeklik doğruluğu yüksek olabilir.

15.5. Resmî Değil ve Doğrulanmamış

İddia: resmî kayıtta yoktur, yeterli harici kanıtı yoktur. AI bunu gerçek gibi aktarmamalıdır.

15.6. Resmî Olarak Eksik Bırakılmış Maddi Sınır

Şirket resmî sayfada olumlu iddiayı yayımlar. Fakat onu doğru anlamak için gerekli sınırı yayımlamaz. Örnek: "Avrupa'da hizmet veriyoruz." Ancak hizmet yalnız iki ülkede ve uzaktan sunulmaktadır. AI yalnız genel cümleyi tekrar ederse resmî metne sadık olabilir. Fakat maddi olarak yanıltıcı sonuç üretebilir. Bu nedenle Resmî Temsil Katmanı yalnız yayımlanan cümleleri değil, maddi eksiklikleri de kaydetmelidir.

15.7. Uyuşmazlık Çözülememiş

Resmî ve harici kayıtlar çelişir. Hiçbiri yeterince güçlü değildir. Doğru sonuç:

UNRESOLVED

olmalıdır. AI'nın kesin hüküm vermesi beklenemez.

16. ÇİFT KATMANLI DEĞERLENDİRME TABLOSU

Resmî temsil sadakati	Doğrulanmış gerçeklik doğruluğu	Temel yorum
Yüksek	Yüksek	Sadık ve doğrulanmış temsil
Yüksek	Düşük	Resmî yanlışı sadakatle taşıma
Düşük	Yüksek	Resmî kayıttan ayrılan fakat kanıtla desteklenen düzeltici temsil
Düşük	Düşük	Genel temsil başarısızlığı
Belirsiz	Yüksek	Resmî kaynak çatışmalı; harici doğrulama güçlü
Yüksek	Belirsiz	Resmî beyan doğru aktarılmış; gerçeklik doğrulaması yetersiz
Belirsiz	Belirsiz	Karar için yetersiz referans

Bu tablo önemli bir hüküm kurar:

> *AI'nın resmî kayda uymaması her zaman başarısızlık değildir.*

Aynı şekilde:

> *AI'nın resmî kayda uyması her zaman başarı değildir.*

17. KAYNAK STATÜSÜNÜ KORUMA KURALI

AI yanıtı yalnız içeriği değil, içeriğin epistemik statüsünü de korumalıdır. Aşağıdaki cümleler aynı değildir:

- "Şirket dünyanın lider GEO kuruluşudur."
- "Şirket kendisini dünyanın lider GEO kuruluşu olarak tanımlamaktadır."
- "Şirket, yayımladığı ve henüz bağımsız olarak doğrulanmamış araştırmaya dayanarak dünyanın lider GEO kuruluşu olduğunu ileri sürmektedir."
- "Belirtilen bağımsız karşılaştırmada şirket birinci sırada yer almıştır."

Her cümle farklı kanıt statüsü taşır. AI:

- öz beyanı bağımsız sonuç,
- pazarlama cümlesini bilimsel bulgu,
- basın bültenini editoryal haber,
- sponsorlu ödülü bağımsız akreditasyon,
- şirket verisini dış denetim

gibi sunamaz. Bu ilkeye:

Atıf ve Kaynak Statüsü Koruma Kuralı

denir.

18. EPİSTEMİK YÜKSELTME

Bir iddianın kaynak statüsünün AI yanıtında daha güçlü hâle getirilmesine:

Epistemik yükseltme

denir. Örnek dönüşümler:

Kaynak metni	AI'nın yanlış yükseltmesi
"Şirketin kendi verilerine göre"	"Kanıtlanmıştır"
"Belirli müşteriler bildirdi"	"Bütün müşterilerde sonuç verir"
"Pilot çalışmada gözlendi"	"Genel olarak etkilidir"
"Şirket kendisini lider olarak tanımlıyor"	"Sektör lideridir"
"Bağımsız doğrulama yapılmadı"	Sınırlama tamamen çıkarılır
"Bazı pazarlarda sunulur"	"Küresel olarak sunulur"

Bu dönüşümler özetleme değildir. Yeni ve daha güçlü iddia üretimidir.

19. RESMÎ SESSİZLİK

Bir şirketin resmî sayfasında bir bilginin bulunmaması şu anlama gelmeyebilir:

- bilgi yanlıştır,
- bilgi yasaktır,

AI bu bilgiyi kullanamaz. Resmî sessizlik üç farklı biçimde olabilir.

19.1. Maddi Olmayan Sessizlik

Bilgi, istemin ve kullanıcı kararının dışında olabilir. AI'nın kullanması veya kullanmaması maddi değildir.

19.2. Meşru Tamamlayıcı Harici Bilgi

Resmî sitede bulunmayan fakat güvenilir kaynakta doğrulanan bilgi olabilir. Örnek:

- Hukuki kayıt
- Tarihsel olay
- Kamuya açık düzenleyici karar
- Bağımsız araştırma

AI bu bilgiyi uygun bağlam ve kaynak statüsüyle kullanabilir.

19.3. Maddi Sınırın Sessiz Bırakılması

Varlık, kullanıcı kararını etkileyen bir sınırı resmî temsilde açıklamamış olabilir. Örnek:

- Lisansın yalnız belirli ülkede geçerli olması
- Fiyatın ek ücret taşımaması gibi görünmesi
- Hizmetin yalnız belirli müşteri türüne açık olması
- Başarı verisinin yalnız seçilmiş vakalardan gelmesi

Bu durumda resmî temsilin kendisi denetim bulgusu taşıyabilir. AI'nın sessizliği tekrar etmesi gerçeklik doğruluğunu düşürebilir.

20. RESMÎ TEMSİLDE MADDİ EKSİKLİK

Bir resmî cümle doğrudan yanlış olmayabilir. Fakat gerekli sınırın çıkarılması sonucu yanıltıcı olabilir. Örnek: "Avrupa'da faaliyet gösteriyoruz." Maddi bağlam:

- yalnız iki ülkede,
- yalnız uzaktan,
- yerel ofis olmadan,
- yalnız kurumsal müşterilere

hizmet veriliyor olabilir. Resmî temsil katmanı şu iki alanı birlikte kaydetmelidir:

- Yayımlanan ana iddia
- İddiayı doğru anlamak için gerekli maddi sınırlar

Bir AI yanıtı ana cümleyi taşıyıp sınırları kaldırıyorsa:

- kelime bakımından sadık,
- karar bakımından yanıltıcı

olabilir.

21. LEHTE OLAN YANLIŞ DA YANLIŞTIR

Bir AI ürünü varlığı olduğundan daha güçlü, büyük veya yetkili gösterdiğinde kurumun sessiz kalması ölçüm ve yönetişim sorununa dönüşür. Örnek: Yerel şirket küresel işletme gibi gösterilir. Danışmanlık şirketi düzenleyici kurum gibi anlatılır. Öz değerlendirme bağımsız sertifikasyon gibi sunulur. Sekiz ülkelik operasyon 20 ülke olarak yazılır. Yanlış kurumun lehine olsa bile doğrulanmış gerçeklik katmanında hata olarak kaydedilmelidir. İkinci kitabın kurucu uyarısı da aynı sınırı koyar: kurum lehine olan yanlış, ticari olarak hoş görünse bile düzeltilmelidir. Bir kurumun olumlu AI yanlışı bildiği hâlde kullanması:

- düzeltme yükümlülüğü,
- temsil bütünlüğü,
- ticari vaat,
- yönetişim

bakımından ayrıca incelenebilir.

22. REFERANS ÇATIŞMASI

Bazen sorun AI yanıtında değil, referans kayıtlarındadır. Örnek: Ana sayfa farklı fiyat gösterir. PDF farklı fiyat gösterir. Checkout başka fiyat gösterir. Yapılandırılmış veri eski fiyat taşır. Bu durumda tek bir doğru resmî temsil bulunmayabilir. Başka bir örnek: Hukuki sicil bir sahiplik gösterir. Şirket sitesi başka sahiplik söyler. Güncelleme tarihleri belirsizdir. Bu durumda doğrulanmış gerçeklik kaydı da uyuşmazlık taşıyabilir.

23. REFERANS KUSURU KURALI

> *Referans kayıtlar maddi biçimde çelişkili veya yetersizse, bu belirsizlik otomatik olarak AI ürününün başarısızlığına dönüştürülemez.*

Doğru statüler:

REFERENCE_CONFLICT

INSUFFICIENT EVIDENCE

UNRESOLVED

TIME-SPLIT

KAPSAM-SPLIT

olabilir. Denetim şu soruyu ayırmalıdır: "AI yanlış cevap verdi mi?" ve: "AI'nın doğru cevap vermesi için gerekli referans gerçeklik var mıydı?" Bir kurum kendi resmî kayıtlarını çelişkili bırakmışsa modelin hangi kaydı seçtiğini ölçebiliriz. Fakat modelden tek ve kesin kurumsal cevap beklerken referans kusurunu saklayamayız.

24. KANONİK KAYIT KİLİDİ

GEO-1000 ölçüm dalgası başlamadan önce iki referans katmanı da sürümlenmelidir:

Resmî Temsil Kaydı

Hangi sayfalar? Hangi dil sürümleri? Hangi tarih? Hangi iddialar? Hangi arşiv durumu?

Doğrulanmış Varlık Gerçekliği Kaydı

Hangi kanıtlar? Hangi karşı kanıtlar? Hangi doğrulama statüleri? Hangi itirazlar? Hangi tarih? Ölçüm sırasında maddi değişiklik olursa: Dalga durdurulabilir. Kayıt yeni sürüme ayrılabilir. Önceki ve sonraki gözlemler ayrı analiz edilebilir. Sonuçları tek referans altında karıştırmak yasaktır.

25. İDDİA GÜCÜ İLE KANIT YÜKÜ

Bir iddia büyüdükçe kanıt yükü de büyür. Aşağıdaki merdiven kurucu bir gösterimdir.

C0 — Basit Kimlik İddiası

"X bir danışmanlık şirketidir." Doğru varlık ve faaliyet kaydı gerekir.

C1 — Kapsam İddiası

"X Almanya'da hizmet sunar." Operasyon, sözleşme veya teslimat kaydı gerekir.

C2 — Nicel Performans İddiası

"Projelerin yüzde 90'ı zamanında tamamlanır." Payda, dönem, yöntem ve başarısız sonuçlar gerekir.

C3 — Karşılaştırmalı İddia

"Rakiplerden daha hızlıdır." Karşılaştırma evreni ve eşdeğer ölçüm gerekir.

C4 — Üstünlük İddiası

"Sektör lideridir." Sektör, kriter, coğrafya, zaman ve yöntem gerekir.

C5 — Garanti veya Evrensel Hüküm

"Her müşteride sonuç garanti eder." Son derece güçlü, kapsamlı ve uygulanabilir kanıt ile gerçek kontrol gerekir. Resmî kaynak, C0 ve bazı C1 iddiaları için güçlü olabilir. C3–C5 düzeylerinde tek başına yeterli olmayabilir.

26. İKİ AYRI DOĞRULUK ÖLÇÜMÜ

GEO-1000 içinde iki temel ölçüm ayrı raporlanmalıdır.

26.1. Resmî Temsil Sadakati

Şu soruya cevap verir:

> *AI yanıtı, varlığın sürümlenmiş resmî ve kanonik temsilini ne ölçüde doğru, eksiksiz ve kaynak statüsünü koruyarak aktarmaktadır?*

Bu ölçüm:

- resmî kimlik,
- resmî hizmet kapsamı,
- yayımlanmış fiyat,
- resmî sınırlar,
- resmî kaynak niteliği

alanlarını değerlendirir.

26.2. Doğrulanmış Varlık Gerçekliği Doğruluğu

Şu soruya cevap verir:

> *AI yanıtı, mevcut kanıtlarla desteklenen güncel varlık gerçekliğini ne ölçüde doğru, kapsamlı ve sınırları korunmuş biçimde aktarmaktadır?*

Bu ölçüm:

- desteklenen maddi iddiaları,
- çelişkili iddiaları,
- eski bilgileri,
- yanlış ilişkileri,
- maddi eksiklikleri,
- epistemik yükseltmeleri

değerlendirir.

27. İKİ SKORUN BİRBİRİNİ TELAFİ EDEMEMESİ

Yüksek resmî sadakat, düşük doğrulanmış doğruluğu telafi edemez. Örnek: Resmî site yanlış iddiayı yayımlar. AI iddiayı aynen tekrar eder. Resmî sadakat yüksek olabilir. Doğrulanmış doğruluk düşüktür. Aynı şekilde yüksek doğrulanmış doğruluk, resmî temsil bozulmasını tamamen önemsizleştirmez. Örnek: AI güncel harici kaynağı doğru kullanır. Şirketin ürün ve hizmet tanımını başka şirketle karıştırır. Yanıtın bazı harici olguları doğru olabilir. Resmî kimlik sadakati zayıftır. Bu nedenle iki sonuç ayrı görünmelidir.

28. ÇİFT KATMANLI YANIT STATÜLERİ

Bir AI yanıtı şu sonuçlardan birini alabilir:

DUAL-PASS

Hem resmî temsile hem doğrulanmış gerçekliğe maddi olarak uygundur.

OFFICIAL-PASS / VERIFIED-FAIL

Resmî kaydı sadakatle taşır. Fakat resmî kaydın kanıtsız, eski veya çelişkili iddiasını gerçek gibi aktarır.

OFFICIAL-FAIL / VERIFIED-PASS

Resmî siteye tam uymamaktadır. Fakat güncel ve daha güçlü kanıtla desteklenen gerçekliği doğru aktarmaktadır. Yanıtın farkı açıklaması tercih edilir.

DUAL-FAIL

Hem resmî kaydı hem doğrulanmış gerçekliği maddi biçimde bozmaktadır.

OFFICIAL-PASS / VERIFIED-UNKNOWN

Resmî beyan doğru aktarılmıştır. İddianın doğrulanmış gerçeklik statüsü henüz belirlenememiştir.

REFERENCE-CONFLICT

Referans kayıtlar kendi içinde çelişmektedir. Yanıtın kesin biçimde puanlanması mümkün değildir.

APPROPRIATE-UNCERTAINTY

AI, referans belirsizliğini doğru biçimde açıklamış ve kesin hükümden kaçınmıştır. Bu bazı durumlarda en doğru cevap olabilir.

29. SENTETİK APPLE.COM GÖSTERİMİ

SENTETİK METODOLOJİ GÖSTERİMİ / Aşağıdaki iddialar, yanıtlar ve puanlamalar tamamen kurgusaldır. Gerçek Apple Inc. bilgisi, performansı veya herhangi bir gerçek AI ürününün cevabı değildir. Sentetik Resmî Temsil Kaydı: O-01: Şirket kendisini küresel tüketici teknolojisi markası olarak tanımlar. O-02: Şirket belirli ürün ve dijital hizmet kategorileri yayımlar. O-03: Şirket kendisini "dünyanın en yenilikçi şirketi" olarak tanımlar. O-04: Resmî sitede belirli ülke mağaza koşulları bulunur. Sentetik Doğrulanmış Gerçeklik Kaydı: V-01: Kurumsal kimlik ve ana faaliyet kategorisi desteklenir. V-02: Belirli ürün kategorileri desteklenir. V-03: "Dünyanın en yenilikçi şirketi" iddiası için tanımlı karşılaştırma yöntemi bulunmamaktadır. V-04: Mağaza koşulları ülkeye göre değişmektedir.

29.1. Sentetik Yanıt A

"Apple.com, şirketin tüketici teknolojisi ürünlerini ve dijital hizmetlerini tanıttığı resmî kurumsal yüzeylerden biridir. Şirket kendisini dünyanın en yenilikçi şirketlerinden biri olarak konumlandırmaktadır; bu üstünlük iddiası kullanılan karşılaştırma yöntemine göre ayrıca değerlendirilmelidir." Değerlendirme: Resmî temsil korunmuştur. Üstünlük iddiası şirkete atfedilmiştir. İddia bağımsız gerçekliğe yükseltilmemiştir. Kapsam sınırı eklenmiştir. Aday statü:

DUAL-PASS

29.2. Sentetik Yanıt B

"Apple dünyanın tartışmasız en yenilikçi şirketidir." Değerlendirme: Resmî pazarlama iddiası mutlak gerçek hâline getirilmiştir. Karşılaştırma yöntemi yoktur. Kaynak statüsü kaybolmuştur. "Tartışmasız" ek epistemik yükseltmedir. Aday statü:

OFFICIAL-PASS / VERIFIED-FAIL

29.3. Sentetik Yanıt C

"Apple yalnızca fiziksel mağazalarda telefon satan yerel bir şirkettir." Değerlendirme: Kurumsal kimlik ve kapsam bozulmuştur. Resmî temsil yanlıştır. Doğrulanmış gerçeklikle çelişir. Aday statü:

DUAL-FAIL

29.4. Sentetik Yanıt D

"Şirketin resmî sitesi belirli mağaza koşullarını yayımlamaktadır; ancak fiyat ve satın alma koşulları ülke sürümüne göre değişebilir." Değerlendirme: Resmî kayıt doğru kaynak rolünde kullanılmıştır. Coğrafi kapsam korunmuştur. Resmî sayfa bütün dünya için tek fiyat kaynağı yapılmamıştır. Aday statü:

DUAL-PASS

30. SENTETİK ASTERON VAKASI

SENTETİK VAKA — GERÇEK KURUM DEĞİLDİR

Asteron adlı şirketin resmî sitesi şu iddiaları yayımlıyor:

- "25 ülkede hizmet veriyoruz."
- "Projelerin yüzde 98'inde başarı sağlıyoruz."
- "Bağımsız olarak doğrulanmış küresel standart kuruluşuyuz."
- "Bütün projelerimiz garantilidir."

Denetim kayıtları:

- Doğrulanmış aktif hizmet ülkesi: 9
- Başarı verisi: yalnız seçilmiş 40 proje
- Bağımsız doğrulama: bulunamadı
- Standart: şirketin kendi yayımladığı önerilen standart
- Garanti: yalnız belirli teslimat kalemleri için sözleşmesel

AI yanıtı: "Asteron, 25 ülkede faaliyet gösteren, yüzde 98 başarı oranına sahip, bağımsız olarak doğrulanmış küresel standart kuruluşudur ve bütün projelerinde sonuç garantisi verir."

30.1. Resmî Temsil Sadakati

Yanıt site metnine büyük ölçüde sadıktır.

30.2. Doğrulanmış Gerçeklik Doğruluğu

Yanıt:

- ülke kapsamını genişletir,
- seçilmiş veriyi genel başarı yapar,
- öz standardı bağımsız standart yapar,

teslimat garantisini sonuç garantisine dönüştürür. Doğrulanmış gerçeklik bakımından ağır başarısızlıktır.

30.3. Doğru AI Yanıtı Nasıl Olabilirdi?

"Asteron kendi sitesinde 25 ülkede hizmet verdiğini, yüzde 98 başarı oranına sahip olduğunu ve bir standart yayımladığını belirtmektedir. Mevcut denetim kayıtlarında aktif hizmet dokuz ülkede doğrulanmış; yüzde 98 oranının seçilmiş 40 projeye dayandığı görülmüş; bağımsız doğrulama teyit edilememiştir. Garanti belirli teslimat kalemleriyle sınırlıdır." Bu yanıt:

- şirket beyanını saklamaz,
- beyanın statüsünü korur,
- doğrulanan gerçekliği ayırır,

sınırı görünür yapar.

31. REFERANS KUSURU İÇEREN SENTETİK VAKA

SENTETİK VAKA

Resmî ana sayfa: "Başlangıç fiyatı 2.000 avro." Resmî fiyat PDF'si: "Başlangıç fiyatı 3.000 avro." Checkout: "4.000 avro." Yapılandırılmış veri: "1.500 avro." AI yanıtı: "Başlangıç fiyatı 3.000 avrodur." AI hangi kaydı seçmiştir? PDF'yi. Yanlış mı? Belki. Doğru mu? Belki. Referans kayıt çelişkilidir. Bu durumda:

- AI başarısızlığı,
- resmî temsil çatışması,
- kanonik kaynak eksikliği

ayrı bulgulardır. Doğru değerlendirme:

REFERENCE-CONFLICT

olabilir. Varlık kendi fiyat kayıtlarını uyumlaştırmadan kesin fiyat sadakati bekleyemez.

32. RESMÎ TEMSİL KATMANINDA KAYNAK ÖNCELİĞİ

Birden fazla resmî kayıt bulunduğunda her iddia alanı için öncelik kuralı belirlenmelidir. Örnek aday öncelik:

- Yürürlükteki sözleşme veya işlem kaydı
- Kanonik güncel ürün/fiyat kaydı
- Yetkilendirilmiş yapılandırılmış veri
- Güncel resmî sayfa
- Güncel resmî PDF
- Tarihsel basın veya blog kaydı

Bu sıra bütün iddialar için evrensel değildir. Örneğin kurumsal tarih için basın arşivi değerli olabilir. Fiyat için eski blog kaydı yetkili değildir. Öncelik iddia türüne göre tanımlanmalıdır.

33. DOĞRULANMIŞ GERÇEKLİKTE KAYNAK ÇATIŞMASI

İki bağımsız kaynak çelişebilir. Bu durumda:

- daha bağımsız olan,
- daha yeni olan,

- konuya daha yakın olan,
- yöntemi daha açık olan,
- daha geniş veri kullanan

kaynak her zaman otomatik olarak kazanmaz. Bütün ölçütler birlikte değerlendirilmelidir. Örnek: Yeni bir blog yazısı eski resmî sicilden daha güncel olabilir. Fakat hukuki sahiplik için blog yeterli yetkiye sahip olmayabilir. Kaynak çatışması çözümü iddia türüne duyarlı olmalıdır.

34. KARŞI KANITIN ZORUNLU KORUNMASI

Doğrulanmış gerçeklik kaydı yalnız iddiayı destekleyen kayıtları içeremez. Şunlar da korunmalıdır:

- İddiayla çelişen kayıtlar
- Başarısız vakalar
- Geri çekilmiş sonuçlar
- Süresi dolmuş sertifikalar
- Olumsuz kullanıcı sonuçları
- Düzenleyici kararlar
- Düzeltme ve errata
- Kaynaklar arası uyuşmazlık

Bir iddianın desteklenmesi, karşı kanıtın yok sayılmasına izin vermez.

35. MADDİLİK

Her resmî–doğrulanmış fark aynı önem seviyesine sahip değildir. Maddi fark, kullanıcı veya sistem kararını değiştirebilecek farktır. Örnekler: Şirket adındaki küçük yazım farkı maddi olmayabilir. Şirket türünün yanlış yazılması maddi olabilir. Bir ürün renginin eksik olması maddi olmayabilir. Hizmet verilen ülkenin yanlış olması maddidir. Bir tarih biçimi farkı maddi olmayabilir. Lisans geçerlilik tarihinin yanlış olması maddidir. Pazarlama tonunun değişmesi maddi olmayabilir. Garanti sınırının kaldırılması maddidir. Maddilik:

- kullanıcı kararı,
- hukuki sorumluluk,
- fiyat,
- güvenlik,
- yetki,
- kapsam,
- kanıt,
- zaman

üzerindeki etkiyle değerlendirilmelidir.

36. ZORUNLU NORMATİF HÜKÜMLER

CH04-N01

Her maddi AI yanıtı Resmî Temsil Kaydı ve Doğrulanmış Varlık Gerçekliği Kaydı'na karşı ayrı ayrı değerlendirilmelidir.

CH04-N02

Resmî temsil sadakati, maddi doğrulukla aynı sonuç gibi sunulamaz.

CH04-N03

Varlığın kendi sitesindeki iddia, iddia türüne uygun ek kanıt olmadan bağımsız doğrulama sayılmaz.

CH04-N04

Birinci taraf kanıt yalnız birinci taraf olduğu için reddedilemez; iddia türüne uygunluğu, yöntemi ve çıkar ilişkisi değerlendirilmelidir.

CH04-N05

Bağımsız kaynak yalnız bağımsız olduğu için doğru veya yetkili kabul edilemez.

CH04-N06

Şirket görüşü, pazarlama iddiası veya öz değerlendirme AI yanıtında bağımsız olguya yükseltilemez.

CH04-N07

Kaynak türü ve epistemik statü AI yanıtında maddi olduğu ölçüde korunmalıdır.

CH04-N08

Resmî kayıtla çelişen AI yanıtı, daha güçlü ve güncel kanıtla destekleniyorsa otomatik başarısızlık sayılamaz.

CH04-N09

Resmî kayda sadık AI yanıtı, resmî iddia doğrulanmış gerçeklikle çelişiyorsa genel doğrulukta başarılı sayılamaz.

CH04-N10

Resmî ve doğrulanmış kayıtlar maddi biçimde çelişiyorsa sonuç REFERENCE_CONFLICT, UNRESOLVED veya uygun belirsizlik statüsü taşımalıdır.

CH04-N11

Referans kusuru sessizce AI ürünü hatasına dönüştürülemez.

CH04-N12

AI yanıtı, doğrulanmamış birinci taraf iddiayı gerçek gibi kullanıyorsa uygun atıf ve sınırlama aranmalıdır.

CH04-N13

Maddi resmi eksiklikler yalnız yayımlanan cümle doğru diye görünmez bırakılamaz.

CH04-N14

Şirket lehine olan yanlış da doğrulanmış gerçeklik katmanında hata olarak kaydedilmelidir.

CH04-N15

Resmî Temsil Kaydı ve Doğrulanmış Varlık Gerçekliği Kaydı ölçüm başlamadan önce sürümlenmelidir.

CH04-N16

Ölçüm dalgası sırasında maddi referans değişikliği olursa önceki ve sonraki gözlemler ayrı sürümlerde değerlendirilmelidir.

CH04-N17

Karşı kanıt ve olumsuz kayıtlar doğrulanmış gerçeklik dosyasından yalnız olumsuz oldukları için çıkarılamaz.

CH04-N18

UNRESOLVED, UNKNOWN ve INSUFFICIENT EVIDENCE sonuçları olumlu doğrulama sayılmaz.

CH04-N19

İddia gücü arttıkça gerekli kanıt yükü de artmalıdır.

CH04-N20

Yüksek Resmî Temsil Sadakati, düşük Doğrulanmış Varlık Gerçekliği Doğruluğunu telafi edemez.

CH04-N21

İki katmanlı değerlendirme sonucunun hesap verebilir insan veya kurum sahibi bulunmalıdır.

37. BAŞARISIZLIK BİÇİMLERİ

CH04-F01 — WEB SİTESİNİ MUTLAK GERÇEKLİK SAYMAK

Kurumun kendi yayımladığı bütün iddialar doğrulanmış olgu kabul edilir.

CH04-F02 — RESMÎ KAYDI DEĞERSİZ SAYMAK

Birinci taraf olduğu için fiyat, politika veya ürün kapsamı gibi doğrudan kurumsal kayıtlar göz ardı edilir.

CH04-F03 — ÖZ BEYANI BAĞIMSIZ KANIT YAPMAK

Şirketin kendi liderlik veya başarı iddiası dış doğrulama gibi kullanılır.

CH04-F04 — BAĞIMSIZ KAYNAK TAPINMASI

Bağımsız fakat konuya uzak, eski veya yöntemsiz kaynak resmî ve yakın kayıttan üstün sayılır.

CH04-F05 — EPİSTEMİK YÜKSELTME

"Şirket söylüyor" ifadesi "kanıtlanmıştır" hükmüne dönüştürülür.

CH04-F06 — KAYNAK STATÜSÜNÜ SİLME

Birinci taraf veri, sponsorlu içerik veya basın bülteni bağımsız haber gibi sunulur.

CH04-F07 — RESMÎ YANLIŞI SADAKATLE TAŞIMAK

AI resmî yanlış iddiayı doğru biçimde tekrar eder ve genel olarak başarılı sayılır.

CH04-F08 — GÜNCEL HARİCİ DÜZELTMEYİ CEZALANDIRMAK

AI eski resmî siteden ayrılıp güçlü güncel kanıt kullandığı için yanlış kabul edilir.

CH04-F09 — RESMÎ SESSİZLİĞİ YASAK SAYMAK

Resmî sitede bulunmayan her harici bilgi otomatik olarak hatalı kabul edilir.

CH04-F10 — MADDİ EKSİKLİĞİ GÖRMEZDEN GELMEK

Resmî cümle doğru olduğu hâlde gerekli sınırın çıkarılması değerlendirilmez.

CH04-F11 — REFERANS KUSURUNU MODEL HATASI SAYMAK

Çelişkili fiyat veya kimlik kayıtlarında AI tek başına suçlanır.

CH04-F12 — REFERANS ÇATIŞMASINI GİZLEMEK

Hakemler çelişkili kayıtlar arasından sonuçtan sonra istedikleri kaynağı seçer.

CH04-F13 — İDDİA TÜRÜNDEN BAĞIMSIZ KAYNAK HİYERARŞİSİ

Her durumda resmî veya her durumda bağımsız kaynağın üstün olduğu varsayılır.

CH04-F14 — KARŞI KANITI SİLMEK

Olumlu iddiayı zayıflatan kayıtlar doğrulama dosyasından çıkarılır.

CH04-F15 — ESKİ RESMÎ KAYDI GÜNCEL SAYMAK

Arşiv, eski PDF veya basın duyurusu güncel kanonik kaynak yapılır.

CH04-F16 — DOĞRULANMAMIŞ İDDİAYI PASS SAYMAK

Yanlışlığı kanıtlanmadığı için iddia doğrulanmış kabul edilir.

CH04-F17 — İDDİA GÜCÜNÜ KANITTAN BÜYÜK TUTMAK

Sınırlı pilot veya yerel veri evrensel sonuç yapılır.

CH04-F18 — TEK KATMANLI GEO SKORU

Resmî sadakat ile doğrulanmış doğruluk tek sonuç içinde görünmez biçimde birleştirilir.

CH04-F19 — LEHTE YANLIŞI KORUMAK

AI kurumu olduğundan büyük gösterdiğinde düzeltme yapılmaz.

CH04-F20 — GERÇEKLİK KAYDINI KURUCU KARARI SAYMAK

NobleJackal veya Kaan MURAZ'ın beyanı dış kanıt olmadan doğrulanmış gerçekliğe dönüştürülür.

38. DENETİM PROSEDÜRÜ

Adım 1 — Maddi AI İddialarını Çıkar

Yanıt:

- atomik iddialara,
- ilişkilere,
- kapsam ve zaman alanlarına

ayrılır.

Adım 2 — Her İddianın Öznesini Belirle

İddia hangi:

- şirket,
- marka,
- ürün,
- kişi,
- yerel varlık

hakkındadır?

Adım 3 — Resmî Temsil Kaydını Bul

İddia resmî yüzeylerde bulunuyor mu? Hangi kaynakta? Hangi sürümde? Hangi tarihte? Hangi kapsamda?

Adım 4 — Resmî Statüyü Belirle

İddia:

- kanonik,
- tarihsel,

- geri çekilmiş,
- taslak,
- çelişkili,
- bilinmeyen

statülerinden hangisindedir?

Adım 5 — İddia Türünü Sınıflandır

İddia:

- kimlik,
- fiyat,
- kapsam,
- lisans,
- performans,
- liderlik,
- ortaklık,
- kullanıcı sonucu

türlerinden hangisidir?

Adım 6 — Uygun Kanıt Türünü Belirle

İddia için hangi kaynak gerçekten yetkilidir? Tek evrensel kaynak sıralaması kullanılmaz.

Adım 7 — Doğrulama ve Karşı Kanıtı Topla

Destekleyen kayıtlar Çelişen kayıtlar Güncellik Kapsam Yöntem Bağımsızlık değerlendirilir.

Adım 8 — Doğrulama Statüsünü Ata

VR-0 ile VR-5, VR-C, VR-U, VR-T veya VR-S statülerinden uygun olanı verilir.

Adım 9 — Kaynak Statüsü Korumasını İncele

AI: "şirket söylüyor" ile "kanıtlandı" arasındaki farkı korumuş mu? öz beyanı bağımsız gerçekliğe yükseltmiş mi?

Adım 10 — Maddi Eksiklikleri İncele

Yanıtta veya resmî temsilde kullanıcı kararı için gerekli:

- sınır,
- zaman,
- coğrafya,
- lisans,
- veri yöntemi

eksik mi?

Adım 11 — Referans Kusurunu Ayır

Resmî veya doğrulanmış kayıt kendi içinde kusurluysa AI hatasından ayrı bulgu oluşturulur.

Adım 12 — Çift Katmanlı Sonucu Ver

Yanıt:

DUAL-PASS

OFFICIAL-PASS / VERIFIED-FAIL

OFFICIAL-FAIL / VERIFIED-PASS

DUAL-FAIL

REFERENCE-CONFLICT

APPROPRIATE-UNCERTAINTY

gibi uygun statüyle kaydedilir.

39. GEREKLİ KANIT

Kanonik Varlık Kaydı Resmî Temsil Kaydı Doğrulanmış Varlık Gerçekliği Kaydı Resmî web sayfaları Resmî PDF ve belgeler Yapılandırılmış veri Kanonik JSON veya API kayıtları Hukuki ve düzenleyici kayıtlar Lisans ve sertifika doğrulamaları Ürün, fiyat ve checkout kayıtları Sözleşme ve operasyon kayıtları Müşteri ve partner doğrulamaları Performans veri setleri Metodoloji ve paydalar Bağımsız incelemeler Sponsor ve çıkar ilişkileri Karşı kanıt Olumsuz sonuçlar Errata ve geri çekme kayıtları Resmî kaynak öncelik kuralları Doğrulama statüleri Hakem kararları İtiraz kayıtları Ölçüm tarihi Referans sürümleri Sorumlu insan veya kurum

40. DENETİM KONTROL LİSTESİ

Resmî Temsil Kaydı mevcut mu? Doğrulanmış Varlık Gerçekliği Kaydı mevcut mu? İki kayıt aynı ölçüm tarihine göre sürümlendi mi? Her AI iddiasının öznesi doğru mu? İddia resmî olarak yayımlanmış mı? Resmî kaynak güncel ve kanonik mi? Resmî kayıtlar kendi içinde çelişiyor mu? İddia türü için doğru kanıt kaynağı kullanılmış mı? Birinci taraf kaynak otomatik olarak reddedilmiş mi? Birinci taraf beyan otomatik bağımsız kanıt sayılmış mı? Bağımsız kaynak gerçekten konuya uygun mu? İddianın kapsamı ve zamanı korunmuş mu? Karşı kanıt dosyada bulunuyor mu? Öz beyan gerçekliğe yükseltilmiş mi? Şirket görüşü kaynak statüsüyle aktarılmış mı? Maddi sınır veya dışlama eksik mi? Lehte olan yanlış ayrıca kaydedilmiş mi?

Referans kusuru AI hatasından ayrılmış mı? UNKNOWN ve UNRESOLVED alanlar olumlu sayılmış mı? Yüksek resmî sadakat düşük doğrulanmış doğruluğu gizliyor mu? AI'nın resmî kayıttan ayrılması güçlü kanıtla açıklanabiliyor mu? Ölçüm sırasında referans kaydı değişmiş mi? Eski ve yeni gözlemler ayrı sürümlerde mi? Çift katmanlı sonuç kamuya açık kayıtta görünür mü? Sonucun hesap verebilir sahibi belli mi?

41. İTİRAZLAR VE CEVAPLAR

İtiraz 1 — "Şirket kendi hakkında en doğru bilgiyi bilir."

Birçok konuda doğru olabilir. Şirket:

- kendi ürününü,
- fiyatını,
- destek politikasını,

- operasyonunu

yakından bilir. Fakat şirket aynı zamanda iddianın ticari sahibidir. “En iyi”, “lider”, “en güvenilir” ve “kanıtlanmış” gibi dış karşılaştırma gerektiren cümleler yalnız öz beyanla doğrulanamaz.

İtiraz 2 — “Kendi siteyi kanıt saymıyorsanız fiyatı neye göre bileceğiz?”

Resmî fiyat sayfası, kurumun kamuya sunduğu fiyat için temel kaynaktır. Ancak:

- checkout,
- vergi,
- ek ücret,
- sözleşme

ile çelişiyorsa fiilî teklif ayrıca değerlendirilir. “Kendi site kanıt değildir” ifadesi bütün iddia türlerine aynı biçimde uygulanamaz. Doğru hüküm şudur:

> *Kendi site, kendi kontrolündeki beyanlar için birincil kaynak olabilir; dış performans ve üstünlük iddiaları için otomatik bağımsız doğrulama değildir.*

İtiraz 3 — “AI resmî siteyi yansıtıyorsa görevi tamamlamış olmaz mı?”

Yalnız hedef: “Site ne söylüyor?” ise evet. Fakat kullanıcı: “Şirket gerçekte ne yapıyor?” diye soruyorsa yalnız siteyi tekrar etmek yeterli olmayabilir. Görev türü ölçüm katmanını etkiler.

İtiraz 4 — “Harici kaynaklar da yanlış olabilir.”

Doğrudur. Bu nedenle bağımsızlık otomatik doğruluk değildir. Kaynak:

- yetki,
- yakınlık,
- güncellik,
- yöntem,
- kapsam

bakımından değerlendirilir.

İtiraz 5 — “Doğrulanmış gerçeklik kaydını kim belirleyecek?”

Tek başına şirket belirleyemez. Tek başına NOMOS da belirleyemez. Süreç:

- kanıt sahibi,
- hakem,
- uygun uzman,
- itiraz,
- sürüm kaydı,
- hesap verebilir insan

gerektirir. Yüksek riskli ve çıkar çatışmalı iddialarda daha bağımsız inceleme gerekir.

İtiraz 6 — “Gerçeklik kaydı her zaman değişebiliyorsa skor anlamsız olmaz mı?”

Hayır. Skor:

- tarih,
- sürüm,
- kapsam

ile anlamlıdır. Değişebilir olması, ölçülemeyeceği anlamına gelmez. Zamansız sunulamayacağı anlamına gelir.

İtiraz 7 — "Resmî sitede olmayan olumsuz bilgiyi AI neden kullansın?"

Bilgi:

- doğru,
- güncel,
- ilgili,
- kullanıcı kararı için maddi

ise kullanılabilir. Ancak kaynağı ve bağlamı korunmalıdır. Resmî sitenin sessizliği olumsuz gerçekliği ortadan kaldırmaz.

İtiraz 8 — "AI şirketin iddiasını 'şirket söylüyor' diye aktarırsa kullanıcıya değer sunmaz."

Tam tersine, iddia statüsünü doğru gösterir. Kullanıcı:

- kurumun görüşünü,
- kanıtlanmış bulguyu,
- açık belirsizliği

ayırt edebilir. Kesin görünmek, doğru görünmek değildir.

İtiraz 9 — "Şirket lehine olan küçük yanlışları neden düzeltelim?"

Çünkü yanlış beklenti:

- yanlış müşteri,
- teslimat başarısızlığı,
- iade,
- şikâyet,
- güven kaybı

üretebilir. Lehte yanlış kısa vadeli görünürlük; uzun vadeli temsil borcudur.

İtiraz 10 — "İki ayrı skor kamu iletişimini karmaşıklaştırmaz mı?"

Karmaşıklığı tamamen saklamak daha büyük risk oluşturur. Kısa sonuç kartı kullanılabilir: Resmî temsil sadakati: yüksek / Doğrulanmış gerçeklik doğruluğu: orta / Temel fark: kanıtlanmamış liderlik ve ülke kapsamı iddiaları Bu, tek yanıltıcı sayıdan daha anlaşılır olabilir.

1. the Official Representation Record; and

2. the Evidence-Constrained Verified Varlık Reality Record.

Official fidelity and verified accuracy MUST NOT be treated as the same result. An AI yanıt MAY faithfully reproduce an official claim while remaining factually unsupported, contradicted, outdated, overbroad, or materially misleading. An AI yanıt MAY differ from an official record without failing when the official record is stale, incomplete, conflicted, or contradicted by stronger and more appropriate evidence. Self-claims, first-party data, independent evidence, sponsored content, official

registries, transactional records, and third-party statements MUST retain their source and evidence status. Reference conflicts MUST be recorded separately from AI-yanıt failures. Unknown or unresolved evidence MUST NOT be converted into verified truth.

A high Official Representation Fidelity result MUST NOT compensate for low Verified Varlık Reality Accuracy. Türkçe normatif karşılığı: Denetlenen varlık hakkındaki her maddi üretken sistem iddiası, ayrı ayrı sürümlenmiş iki referans katmanına karşı değerlendirilmelidir: Resmî Temsil Kaydı ve Kanıta Bağlı Doğrulanmış Varlık Gerçekliği Kaydı. Resmî sadakat ile doğrulanmış doğruluk aynı sonuç sayılamaz. AI resmî iddiayı sadakatle tekrar ederken maddi olarak desteklenmeyen, çelişkili, eski veya aşırı geniş temsil üretebilir. Resmî kayıt eski, eksik veya çelişkiliyse AI'nın daha güçlü kanıta dayanan farklı cevabı otomatik hata değildir. Kaynakların türü ve kanıt statüsü korunmalı; referans çatışması AI başarısızlığından ayrı kaydedilmelidir. Yüksek resmî sadakat düşük doğrulanmış doğruluğu telafi edemez.

44. BÖLÜMÜN ORTAK HÜKMÜ

Bir şirket kendi hakkında konuşma hakkına sahiptir. Kim olduğunu açıklayabilir. Ne sunduğunu anlatabilir. Fiyatını yayımlayabilir. Sınırlarını belirleyebilir. Vizyonunu ifade edebilir. Fakat bir varlığın kendi hakkında söylediği her cümle aynı epistemik statüye sahip değildir. "Hizmet paketimizin fiyatı 2.000 avrodur." ile: "Dünyanın en iyi şirketiyiz." aynı tür iddia değildir. Birinci cümle şirketin kontrol ettiği ticari bir beyan olabilir. İkinci cümle dış karşılaştırma gerektirir. Bir AI ürünü ikisini aynı kesinlikle aktarırsa kaynak statüsünü bozar. Bir AI ürünü şirket sitesini kusursuz biçimde tekrar edebilir. Fakat site:

- eski,
- çelişkili,
- kanıtsız,
- kapsamı aşan,
- maddi sınırları saklayan

bir temsil taşıyorsa bu sadakat doğru GEO değildir. Aynı şekilde AI resmî siteden farklı bir cümle kurabilir. Fakat güçlü, güncel ve uygun kanıta dayanıyorsa bu fark bazen hata değil, düzeltmedir. Bu nedenle NOMOS'un dördüncü ölçüm yasası şudur:

> *Resmî temsil, doğrulanmış gerçekliğin yerine geçmez.*

Beşinci yasası şudur:

> *Doğrulanmış gerçeklik, resmî kaynağı yok sayarak kurulamaz.*

Altıncı yasası şudur:

> *Kaynağın ne söylediği kadar, hangi yetkiyle ve hangi kanıt statüsüyle söylediği de temsilin parçasıdır.*

Yedinci yasası şudur:

> *Resmî yanlışı sadakatle taşımak, doğru temsil değildir.*

Sekizinci yasası şudur:

> *Gerçeklik kaydı, itiraz edilemeyen mutlak hüküm değil; sürümlenmiş ve kanıta bağlı karar dosyasıdır.*

NOMOS'un Bölüm 4 Emri

> *Bana yalnız sitende ne yazdığını gösterme. / Hangi cümlenin hangi kanıtla taşındığını göster.*

> *Kendi sözünü bağımsız doğrulama gibi sunma.*
>
> *Birinci taraf veriyi yalnız sana ait diye değersizleştirme. / Ama yalnız sana ait diye tartışılmaz gerçek de yapma.*
>
> *Şirket görüşünü olguya, basın bültenini habere, sponsorlu ödülü bağımsız otoriteye dönüştürme.*
>
> *Resmî yanlışı doğru biçimde tekrarladığımda beni başarılı sayma.*
>
> *Resmî kaydın eskidiyse güncel kanıta dayanan cevabı sırf senden farklı diye cezalandırma.*
>
> *Lehine olan yanlışta sessiz kalma.*
>
> *Karşı kanıtı dosyadan çıkarma.*
>
> *Referansların birbirine karşı çıkıyorsa tek bir doğru uydurma. / Çelişkiyi kaydet. / Bilinmiyorsa bilinmiyor de.*
>
> *İddianı, kaynağını, kanıtını, sınırını ve zamanını birbirinden ayır.*
>
> *Bana ne söylemem gerektiğini emretmeden önce, neyi söylemeye hakkın olduğunu kanıtla.*

Bölümün Son Cümlesi

> *Doğru GEO, yapay zekânın bir varlığın söylediklerini yalnızca sadakatle tekrar etmesi değil; söylenen şeyin kaynak, kanıt, kapsam ve zaman statüsünü bozmadan insanlara aktarmasıdır.*

KISIM II

Küresel Nüfus, Ülke ve Dil Adaleti

BÖLÜM 05

Her AI Ürünü İçin Uygun Kullanıcı Nüfusu

YÖNTEMSEL DAYANAKLAR / K04; K05; K06; K07

Bölüm Sınırı

İlk dört bölüm ölçümün temel nesnelerini kurdu: Tek yanıt GEO skoru değildir. Temsil tek sayı değil, koşullu bir dağılımdır. Denetlenen varlık sonuçlar görülmeden önce çözülmelidir. AI yanıtı hem resmî temsile hem kanıtla doğrulanmış varlık gerçekliğine karşı değerlendirilmelidir. Şimdi ölçümün paydasını belirlememiz gerekir:

> *Bu dağılım hangi insan nüfusu hakkında hüküm kuracaktır?*

Bir ülkede yaşayan herkes mi? Yalnız yetişkinler mi? İnternet erişimi bulunanlar mı? Belirli AI ürününü resmî olarak kullanabilenler mi? Ücretsiz hesap sahipleri mi? Ücretli plan kullanıcıları mı? Ürünü denetim dilinde kullanabilecek kişiler mi? Gerçekten aktif kullanıcılar mı? Bir gün kullanma ihtimali bulunan potansiyel kullanıcılar mı? Markanın hizmet verebildiği pazarlardaki kişiler mi? Bir ülkenin nüfusu çok büyük olsa da ilgili AI ürünü o ülkede resmî olarak erişilebilir değilse, o ülke ana ürün paneline nasıl dâhil edilecektir? Küçük bir ülkenin nüfus payı 1.000 kişilik örneklemde bir kişinin çok altında kalıyorsa, o ülke tamamen görünmez mi olacaktır?

Bir kişi üç dil biliyorsa üç ayrı insan gibi mi sayılacaktır? Bir ürünün mobil sürümüne erişebilen fakat web sürümüne erişemeyen kullanıcı hangi nüfusun parçasıdır? Bu sorular cevaplanmadan: "Her AI ürünü için 1.000 rastgele kullanıcı" ifadesi bilimsel bir hedef nüfus tanımlamaz. Yalnız kullanıcı sayısı belirtir. Bu bölüm, ülke nüfusunu terk etmez. Onu daha doğru yere yerleştirir:

> *Ülke nüfusu, örneklemenin başlangıç tabanıdır; nihai uygun kullanıcı nüfusu değildir.*

Bu bölümde denetim sistemi, daha önce belirlenen ülke, dil, tarih, ürün ve ölçüm kaydı zorunluluğunu doğrudan nüfus tanımına uygular. Çünkü kanıt, sınır, bağlam ve zaman istemek; hangi insanlardan söz edildiğini de aynı açıklıkla tanımlamayı gerektirir. Bu bölüm henüz:

- 1.000 kişilik örneklemin ülkelere tam olarak nasıl dağıtılacağını,
- küsuratların hangi algoritmayla tam sayıya çevrileceğini,
- katılımcıların hangi panelden seçileceğini,
- örnekleme ağırlıklarının nasıl hesaplanacağını,
- güven aralıklarının hangi varyans yöntemiyle üretileceğini

kesinleştirmez. Bunlar sonraki bölümlerde düzenlenecektir. Bölüm 5'in görevi daha temeldir:

> *Her AI ürünü ve her ölçüm sorusu için, sonuçların hangi insan topluluğuna genellenebileceğini tanımlamak.*

NOMOS Challenge

Bir dünya haritası açıyorsunuz. Ülkelerin nüfuslarını topluyorsunuz. Sonra 1.000 kişiyi nüfusa göre dağıtmak istiyorsunuz. Büyük ülkeler çok sayıda katılımcı alacak. Küçük ülkeler bir veya sıfıra yakın pay alacak. Bu düşüncenin çekirdeği doğrudur:

> *Dünya nüfusundaki bir insan, yalnız küçük bir ülkede yaşadığı için büyük ülkedeki bir insandan yüzlerce kat daha etkili olmamalıdır.*

Fakat şimdi aynı yöntemi iki ayrı AI ürününe uygulayın. Birinci ürün:

- bazı ülkelerde resmî olarak kullanılabilir,
- bazı ülkelerde sınırlıdır,

bazı ülkelerde hiç sunulmamaktadır. İkinci ürünün ülke kapsamı farklıdır. Dünya nüfusunu iki ürüne de aynı biçimde bölerseniz ne olur? Birinci ürünün kullanılamadığı ülkelerde insanlardan ölçüm isteyebilirsiniz. Katılımcılar:

- konumlarını yanlış göstermeye,
- yasak veya desteklenmeyen erişim yolları kullanmaya,
- başka ülke hesabı açmaya,
- güvenlik ve kullanım koşullarını aşmaya

zorlanabilir. Bu, gerçek kullanıcı deneyimini ölçmez. Yapay bir erişim deneyimi üretir. Şimdi başka bir sorun düşünün. Bir ülkede nüfus çok büyüktür. Fakat:

- nüfusun önemli bir bölümü ürünün asgari yaş koşulunu karşılamaz,
- herkesin güvenilir internet erişimi yoktur,
- ürün yalnız belirli cihazlarda kullanılabilir,
- denetim dili nüfusun tamamı tarafından kullanılamaz,

belirli plan yalnız ödeme aracı bulunan kullanıcılara açıktır. Toplam nüfusu doğrudan payda yaparsanız, gerçekte denetlenemeyecek ve ürünü kullanamayacak insanları hedef nüfusa eklersiniz. Şimdi küçük bir ülke düşünün. Nüfus ağırlığı altında 1.000 kişilik örneklemde teorik payı: 0,03 kişi olsun.

Bir insanın yüzde üçte biri seçilemez. Ülkeye bir kişi verirseniz, ülkeyi nüfus ağırlıklı ana panelde aşırı temsil edebilirsiniz. Hiç kimse vermezseniz, o ülkedeki:

- erişim,
- dil,
- ürün davranışı,
- yerel temsil

tamamen görünmez kalabilir. Bu iki amaç farklıdır:

- Dünya nüfusundaki ortalama kullanıcı deneyimini ölçmek
- Küçük ülkelerin tamamen denetim dışı kalmasını önlemek

Aynı tek panel iki amacı kusursuz biçimde yerine getiremez. Bu nedenle NOMOS üç ayrı ilke kurar:

> *Nüfus ağırlığı, insanların küresel ortalamadaki payını belirler.*
>
> *Gözcü kapsamı, küçük ülkelerin ve dillerin tamamen görünmez olmasını engeller.*
>
> *Dil ve erişilebilirlik adaleti, nüfus ortalamasının bastırabileceği alt grupları ayrıca ölçer.*

Bu bölümün ilk hükmü şudur:

> *Bir ülkenin toplam nüfusu, o ülkedeki AI-uygun kullanıcı nüfusu değildir.*

İkinci hükmü şudur:

> *Bir AI ürününün hedef nüfusu, başka bir AI ürününe otomatik olarak aktarılamaz.*

Üçüncü hükmü şudur:

> *Bir insanın ana skordaki ağırlığı ile ülkesinin denetimde görünür olma hakkı aynı şey değildir.*

Dördüncü hükmü şudur:

> *Ürün erişiminin bulunmaması, yanlış temsil değildir; fakat küresel erişim sınırıdır ve ayrıca raporlanmalıdır.*

1. BÖLÜMÜN AMACI

Bu bölümün amacı, GEO-1000 ölçümünün hedef insan nüfusunu ürün, soru, ülke, dil, arayüz ve zaman bakımından tanımlamaktır. Bölüm şu ayrımları normatif hâle getirir:

- Toplam ülke nüfusu ile uygun AI kullanıcı nüfusu
- Potansiyel erişim ile gerçek kullanım
- Ürün erişilebilirliği ile temsil doğruluğu
- Hedef nüfus ile örnekleme çerçevesi
- Seçilen katılımcı ile geçerli analitik gözlem
- Kişi ile hesap
- Kişi ile oturum
- Ürün ile temel model
- Ücretsiz ve ücretli plan nüfusları
- Web ve mobil kullanıcı nüfusları
- Resmî erişim ile erişim kısıtını aşan kullanım
- Ülke nüfus ağırlığı ile ülke gözcü kapsamı
- Nüfus ağırlıklı dil payı ile dil adaleti paneli
- Potansiyel kullanıcı nüfusu ile aktif kullanıcı maruziyeti
- Küresel kimlik sorgusu ile pazar veya yargı alanı sorgusu
- Ana tüketici ürünü ile kurumsal, eğitim veya geliştirici ürünü
- Erişilebilirlik ihtiyacı bulunan kullanıcı ile veri toplama aracının erişilebilirliği
- Hedef nüfus dışı kullanıcı ile NOT TESTED kullanıcı
- Ürün kullanılamayan ülke ile üründe yanlış cevap alınan ülke

Bu bölümün sonunda her GEO-1000 denetimi şu soruya açık cevap verebilmelidir:

> *Bu sonuç, hangi insanlar arasından seçilmiş hangi kullanıcı evreni hakkında hüküm kurmaktadır?*

2. MERKEZÎ NORMATİF HÜKÜM

> *Her GEO-1000 denetimi, sonuçlar görülmeden önce; denetlenen AI ürünü, ölçüm sorusu, ülke, dil, arayüz, plan ve zaman koşullarına özgü bir hedef kullanıcı nüfusu tanımlamak ve sürümlemek zorundadır.* [K04–K07]

Hedef nüfus:

- “dünya nüfusu”,

- “internet kullanıcıları”,
- “AI kullanıcıları”,
- “müşteriler”

gibi belirsiz bir adla bırakılamaz. İlgili olduğu ölçüde şu sorular cevaplanmalıdır: Hangi AI ürünü? Hangi kullanıcı yüzeyi? Hangi ülkeler? Hangi tarihteki erişim durumu? Hangi yaş koşulu? Hangi dil veya diller? Hangi cihaz ve arayüz? Hangi plan veya hesap türü? Resmî kullanım koşulları nelerdir? Kullanıcıların ürüne erişmesi için ödeme gerekli midir? Ölçüm sorusu bütün kullanıcılar için mi, yalnız belirli pazar veya yargı alanı için mi anlamlıdır? Aktif kullanıcı mı, potansiyel kullanıcı mı ölçülmektedir? Hangi kullanıcı grupları kapsam dışıdır? Hangi gruplar hedef nüfusta bulunmasına rağmen örnekleme çerçevesinde eksiktir? Bu alanların değişmesi hedef nüfusu değiştirebilir. Hedef nüfus değiştiğinde eski ve yeni skorlar açıklamasız biçimde aynı seri içinde birleştirilemez.

3. NÜFUS TEK BİR SAYI DEĞİLDİR

Bir ülkenin resmî veya tahmini toplam nüfusu, GEO-1000 için yalnız başlangıç katmanıdır. Denetim nüfusu aşağıdaki daralan katmanlardan oluşabilir.

3.1. Yerleşik Nüfus

Belirtilen ülkede veya yargı alanında yaşayan toplam insan nüfusudur. Bu sayı:

- yaş,
- internet,
- ürün erişimi,
- dil,
- hesap,
- cihaz

bakımından herhangi bir uygunluk varsayımı taşımaz.

3.2. Yaş Bakımından Uygun Nüfus

İlgili AI ürününün ve araştırma protokolünün yaş koşulunu karşılayan nüfustur. Bir ürünün:

- minimum kullanım yaşı,
- ebeveyn onayı şartı,
- ülkeye özgü yaş kuralı

bulunabilir. Araştırma protokolü ürün koşulundan daha kısıtlayıcı olabilir. Örneğin kurucu GEO-1000 paneli, etik ve hukuki karmaşıklığı azaltmak için yalnız yetişkin katılımcılardan oluşturulabilir. Bu durumda skor: “Bütün nüfusun” değil:

“Tanımlı yetişkin ve ürün-uygun nüfusun”

sonucudur.

3.3. Bağlantılı Nüfus

İlgili ürünü kullanabilecek düzeyde:

- internet,
- cihaz,
- bağlantı kalitesi,
- teknik erişim

imkânı bulunan nüfustur. İnternet erişimi ikili bir özellik olmayabilir. Bir insan:

- mesajlaşma kullanabilir,
- fakat ağır web uygulamasını çalıştıramayabilir,
- mobil internet kullanabilir,
- fakat masaüstü deneyimine erişemeyebilir,

düşük bağlantı nedeniyle cevap veya ekran görüntüsü yükleyemeyebilir. Bağlantı sınırlaması kullanıcıyı hedef nüfustan otomatik olarak çıkarmamalıdır. Ölçülen ürünün gerçek kullanım koşullarıyla ilişkilendirilmelidir.

3.4. Ürün Bakımından Erişilebilir Nüfus

İlgili AI ürününün belirtilen tarihte:

- resmî olarak sunulduğu,
- hukuka ve ürün koşullarına uygun biçimde kullanılabildiği,
- gerekli arayüze erişilebildiği

nüfustur. Bir ürünün bir ülkede bulunmaması, o ülkedeki insanların yanlış temsil aldığı anlamına gelmez. Bu durumda: temsil yanıtı yoktur, erişim kapsamı yoktur. İki sonuç ayrı tutulmalıdır.

3.5. Hesap ve Plan Bakımından Uygun Nüfus

Bazı ürünler:

- hesap gerektirebilir,
- telefon doğrulaması isteyebilir,
- ödeme yöntemi gerektirebilir,
- yalnız kurumsal sözleşmeyle sunulabilir,

belirli özellikleri yalnız ücretli planda açabilir. Bu şartlar hedef nüfusun hangi kısmının ilgili kullanıcı yüzeyine gerçekten erişebildiğini etkiler. Ücretsiz plan skoru, ücretli plan kullanıcılarına otomatik olarak uygulanamaz.

3.6. Dil Bakımından Uygun Nüfus

İlgili istemi ve AI cevabını denetim dilinde yeterli düzeyde kullanabilecek nüfustur. Dil uygunluğu yalnız:

- ülkenin resmî dili,
- kişinin ana dili

ile belirlenemez. Şunlar ayrılmalıdır:

- Ana veya ev dili
- Günlük kullanım dili
- AI arayüz dili
- İstem dili
- Yanıt dili
- Profesyonel kullanım dili
- Okuma yeterliliği

Bir insan birden fazla dilde uygun olabilir. Bu kişi küresel nüfus hesabında birden fazla tam kişi gibi sayılamaz.

3.7. Arayüz Bakımından Uygun Nüfus

Ölçüm:

- web,
- mobil uygulama,
- masaüstü uygulama,
- kurumsal çalışma alanı,
- gömülü asistan

üzerinden yapılabilir. Her arayüzün erişilebilir kullanıcı evreni farklı olabilir. Web erişimi bulunan kullanıcı mobil ürüne sahip olmayabilir. Kurumsal kullanıcı yüzeyi genel halka açık olmayabilir.

3.8. Araştırma Bakımından Uygun Nüfus

Bir kişi ürünü kullanabiliyor olsa bile araştırmaya alınması için ayrıca:

- gönüllü onay,
- mahremiyet,
- veri işleme,
- ekran görüntüsü toplama,
- ülkeye özgü araştırma şartları,
- yaş,
- güvenlik

koşullarının karşılanması gerekebilir. Araştırmaya alınamayan fakat gerçek ürünü kullanan insanlar hedef nüfusun parçası olabilir. Bu durumda örnekleme çerçevesi hedef nüfusu eksik kapsar. Bu eksiklik saklanmamalıdır.

3.9. Analitik Olarak Uygun Nüfus

Toplanan verinin sonuç üretmek için kullanılabildiği nüfustur. Katılımcı:

- istemi değiştirmiş,
- yanlış ürünü kullanmış,
- cevabı seçmiş,
- kanıtı eksik göndermiş

olabilir. Kişi hedef nüfusa ait olsa bile gözlemi analitik veri setine giremeyebilir. Hedef nüfus ile geçerli gözlemler aynı şey değildir.

4. UYGUN KULLANICI NÜFUSUNUN KÜME TANIMI

Belirli bir AI ürünü a ve zaman t için uygun kullanıcı evreni şu biçimde düşünülebilir:

```
U_{a,t} = {u ∈ P_t : AgeEligible(u,a,t)=1 ∧ Connected(u,a,t)=1 ∧ ProductAvailable(u,a,t)=1 ∧
LawfulAccess(u,a,t)=1 ∧ InterfaceEligible(u,a,t)=1}
```

Burada:

- Pt: belirtilen tarihteki genel insan nüfusu
- u: tekil insan
- a: AI ürünü ve kullanıcı yüzeyi
- t: ölçüm zamanı

Dil ve istem bakımından hedef evren daha sonra şöyle daraltılabilir:

```
U_{a,c,l,s,p,t} = {u ∈ U_{a,t} : Country(u)=c ∧ LanguageEligible(u,l)=1 ∧ SurfaceEligible(u,s)=1 ∧ PromptRelevant(u,p)=1}
```

Burada:

- c: ülke veya yargı alanı
- l: dil ve locale
- s: kullanıcı yüzeyi
- p: istem veya soru ailesi

Bu gösterimin amacı bütün insanları tek tek kaydetmek değildir. Hedef nüfusun mantıksal sınırını görünür kılmaktır.

5. MARJİNAL ORANLARI KÖRÜ KÖRÜNE ÇARPMAK YASAKTIR

Uygun nüfus bazen şu tür bir formülle yaklaşık hesaplanmak istenebilir:

```
N × p_yaş × p_internet × p_ürün × p_dil
```

Bu hesap yalnız koşulların bağımsızlığı veya uygun ortak dağılım varsayımı altında savunulabilir. Gerçek hayatta özellikler bağımsız olmayabilir. Örneğin:

- internet kullanımı yaşa göre değişebilir,
- ücretli plan erişimi gelirle ilişkili olabilir,
- belirli dil kullanımı ülke içindeki bölge ve eğitimle ilişkili olabilir,

cihaz erişimi ile bağlantı kalitesi birlikte değişebilir. Marjinal oranların çarpılması hedef nüfusu:

- gereğinden büyük,
- gereğinden küçük

tahmin edebilir. Tercih sırası şu olmalıdır:

- Uygun ortak mikroveri veya çapraz tablo
- Güvenilir çok boyutlu nüfus tahmini
- Kalibre edilmiş model
- Gerekçesi ve duyarlılık analizi açıklanmış marjinal yaklaşım

Basit çarpım kullanılıyorsa:

- varsayım,
- veri kaynağı,
- belirsizlik,
- alternatif senaryo

açıklanmalıdır.

6. DENETİM SORUSU HEDEF NÜFUSU DEĞİŞTİRİR

Her istem bütün dünya nüfusu için aynı ölçüm evrenine sahip değildir.

6.1. Küresel Kimlik Soruları

Örnek: “Apple.com nasıl bir şirkettir?” Bu soru, ürünün uygun küresel kullanıcı nüfusu içinde geniş biçimde sorulabilir. Çünkü şirket kimliği belirli bir pazara ait kullanıcıyla sınırlı değildir.

6.2. Yerel Hizmet Soruları

Örnek: "Bu şirket Türkiye'de hizmet veriyor mu?" Ana hedef nüfus:

- Türkiye'deki uygun kullanıcılar,
- Türkiye hizmetini araştıran ilgili dış kullanıcılar

olabilir. Bütün dünya nüfusuna eşit ağırlık vermek gerekli değildir.

6.3. Fiyat ve Satın Alma Soruları

Fiyat:

- ülke,
- vergi,
- para birimi,
- plan,
- müşteri türü

ile değişebilir. Hedef nüfus, ilgili ticari pazar ve kullanıcı türüne göre tanımlanmalıdır.

6.4. Hukuk ve Lisans Soruları

Lisans ve hukuki yetki yargı alanına bağlıdır. Hedef nüfus:

- ilgili ülkede hizmet alabilecek,
- o hukuk düzeninin etkisinde bulunan

kullanıcılarla sınırlandırılabilir.

6.5. Tavsiye Soruları

"Bu şirket benim için uygun mu?" sorusunun hedef nüfusu, bütün insanlar değildir. Şunlara göre tanımlanabilir:

- Kullanıcı ihtiyacı
- Ülke
- Bütçe
- Müşteri türü
- Hizmet kapsamı
- Uygunluk ve dışlama koşulları

Tavsiye nüfusu, kimlik nüfusundan daha dar olabilir.

6.6. Yüksek Riskli Sorular

Sağlık, hukuk, finans veya güvenlik alanında:

- yargı,
- kullanıcı profili,
- risk durumu,
- profesyonel yetki

hedef nüfusu daha güçlü biçimde sınırlar.

7. NÜFUS ÇERÇEVESİ İSTEM AİLESİNE BAĞLANMALIDIR

Bir denetimde tek hedef nüfus yerine birden fazla nüfus çerçevesi bulunabilir.

7.1. Core Identity Population

Varlığın temel kimliğini ölçen küresel veya geniş uygun kullanıcı nüfusudur.

7.2. Service-Relevant Population

Varlığın gerçekten hizmet verebildiği veya kullanıcı tarafından hizmet için değerlendirilebileceği nüfustur.

7.3. Market Transaction Population

Belirli fiyat, satın alma, iade veya ticari koşulun uygulanabileceği nüfustur.

7.4. Jurisdictional Population

Belirli lisans, hukuk veya düzenleyici iddianın uygulanabileceği kullanıcı nüfusudur.

7.5. Tavsiye-Eligible Population

Varlığın gerçek kapsamı içinde uygun olabilecek kullanıcı nüfusudur.

7.6. High-Risk Affected Population

Yanlış temsilin ciddi zarar yaratabileceği nüfustur. Bu çerçeveler birbirine karıştırılmamalıdır. Bir küresel kimlik skoru: "Bütün dünya için tavsiye edilebilirlik" anlamına gelmez.

8. AI ÜRÜNÜNÜN TÜRÜ NÜFUSU BELİRLER

Bütün AI ürünleri genel halka açık sohbet ürünleri değildir.

8.1. Genel Tüketici AI Ürünü

Hedef nüfus:

- ürüne erişebilen,
- kullanım şartlarını karşılayan,
- ilgili dilde kullanabilen

insanlar olabilir. GEO-1000 gerçek kullanıcı paneli bu ürün türüne doğrudan uygulanabilir.

8.2. Kurumsal AI Ürünü

Ürün yalnız:

- işveren sözleşmesi,
- kurumsal lisans,
- kurum içi kullanıcı hesabı

ile kullanılabilir. Hedef nüfus genel yetişkin nüfus değil:

> *Ürüne kurumsal erişimi bulunan koltuk veya kullanıcı evrenidir.*

Kurumsal sonuç genel tüketici ürünüyle aynı nüfus ağırlığında karşılaştırılamaz.

8.3. Eğitim Ürünü

Hedef nüfus:

- öğrenci,
- öğretmen,
- kurum,
- yaş grubu,
- eğitim lisansı

bakımından sınırlı olabilir. Çocuklar veya reşit olmayanlar dâhil edilecekse ayrı etik protokol gerekir.

8.4. Geliştirici veya API Ürünü

API sonucu doğrudan son kullanıcı arayüzü olmayabilir. Hedef birim:

- geliştirici,
- uygulama,
- API çağrısı,
- kontrol edilmiş test oturumu

olabilir. GEO-1000 insan nüfus paneli, API'ye doğrudan ve değişmeden uygulanamaz. API için ayrı:

Kontrollü Sistem Denetimi

gerekebilir.

8.5. Gömülü AI Asistanı

AI:

- işletim sistemi,
- arama arayüzü,
- alışveriş platformu,
- iş yazılımı,
- cihaz

içinde gömülü olabilir. Hedef nüfus gömülü ürünün gerçek kullanıcı nüfusudur. Temel model sağlayıcısının toplam kullanıcı nüfusu değildir.

8.6. Uzman veya Sektörel AI

Ürün yalnız:

- hukukçular,
- hekimler,
- finans profesyonelleri,
- mühendisler

için tasarlanmış olabilir. Hedef nüfus meslek ve erişim koşullarıyla sınırlandırılmalıdır.

9. ÜRÜN ERİŞİM STATÜLERİ

Her ülke, ürün, arayüz ve tarih için erişim statüsü kaydedilmelidir.

PA-0 — UNKNOWN

Erişim durumu güvenilir biçimde belirlenememiştir.

PA-1 — UNAVAILABLE

Ürün ilgili ülkede veya kullanıcı grubunda sunulmamaktadır.

PA-2 — RESTRICTED

Ürün belirli:

- hesap,
- davet,
- kurum,
- yaş,
- plan,
- cihaz

şartlarıyla sınırlıdır.

PA-3 — ACCESSIBLE BUT NOT OFFICIALLY SUPPORTED

Ürüne fiilen erişim gözlenebilir. Fakat resmî destek veya kullanım statüsü açık değildir ya da sınırlıdır. Bu kullanıcılar ana resmî nüfus paneline otomatik olarak eklenmemelidir. Ayrı deneysel panel oluşturulabilir.

PA-4 — OFFICIALLY AVAILABLE

Ürün resmî olarak erişilebilir ve ana kullanım koşulları tanımlıdır.

PA-5 — OFFICIALLY AVAILABLE WITH FULL AUDIT SURFACE

Ürün yalnız erişilebilir değil; denetlenen:

- dil,
- plan,
- arayüz,
- oturum koşulunda

protokolün gerektirdiği gözlemi üretmeye uygundur. Bir ülkede ürünün bulunması, bütün plan ve dillerin denetime uygun olduğu anlamına gelmez.

10. KISIT AŞMAK GERÇEK NÜFUS ÖLÇÜMÜ DEĞİLDİR

Katılımcılardan şu davranışlar istenemez:

- Yanlış ülke beyan etmek
- Sahte hesap oluşturmak
- Yaş bilgisini değiştirmek
- Resmî erişim kısıtını aşmak
- Başka kişiye ait hesabı kullanmak
- Ürünün güvenlik veya kimlik doğrulamasını atlatmak
- Kullanım koşullarını ihlal etmek

Bu davranışlar teknik deney olarak ayrıca incelenebilir. Fakat:

Doğal ve meşru kullanıcı nüfusunun deneyimi

gibi raporlanamaz. Bir ülkede erişim yoksa doğru statü: UNAVAILABLE / veya / OUTSIDE_NATIVE_REACH olmalıdır. Kısıt aşan tekil gözlem, ülke nüfus skoru üretmez.

11. POTANSİYEL KULLANICI İLE AKTİF KULLANICI AYRIMI

Ürünü kullanabilecek insan ile ürünü gerçekten kullanan insan aynı topluluk değildir.

11.1. Potansiyel Erişim Nüfusu

Ürünü teknik ve hukuki olarak kullanabilecek kişilerdir. Bu nüfus: "Uygun bir kişi ürünü kullansaydı ne görürdü?" sorusuna cevap verir. Nüfus adaleti ve ürün karşılaştırması için değerlidir.

11.2. Aktif Kullanıcı Nüfusu

Belirli dönemde ürünü gerçekten kullanan kişilerdir. Bu nüfus: "Gerçek kullanım hacmi içinde insanlar ne görüyor?" sorusuna yaklaşır. Güvenilir aktif kullanım verisi gerektirir.

11.3. Sorguya Maruz Kalan Kullanıcı Nüfusu

Ürünü kullanmakla birlikte denetlenen marka veya konu hakkında soru sorma ihtimali bulunan insanlardır. Bu nüfus çoğu zaman doğrudan bilinemez. Arama veya sorgu davranışı verisi gerekebilir.

11.4. Yanıtı Gerçekte Gören Nüfus

Belirli bir temsil çıktısına gerçekten maruz kalan kullanıcılardır. Bu, kullanıcı paneli dışında gerçek ürün sağlayıcı verisi olmadan tam ölçülemeyebilir.

12. UYGUNLUK AĞIRLIĞI İLE MARUZİYET AĞIRLIĞI AYNI DEĞİLDİR

İki ayrı küresel sonuç üretilebilir.

12.1. Uygun Nüfus Ağırlıklı Sonuç

Her insan, tanımlı uygun kullanıcı nüfusu içindeki payına göre ağırlık alır. Bu sonuç:

- coğrafi adalet,
- potansiyel erişim,
- nüfus temsili

bakımından güçlüdür.

12.2. Gerçek Maruziyet Ağırlıklı Sonuç

AI ürünleri veya kullanıcı grupları gerçek kullanım hacimlerine göre ağırlık alır. Bu sonuç:

- mevcut ticari etki,
- gerçek kullanıcı maruziyeti,
- operasyonel risk

bakımından değerlidir. Fakat güvenilir kullanım verisi bulunmayabilir. Tahminler sağlayıcı beyanına, pazar verisine veya eksik örnekleme dayanabilir. Bu nedenle ağırlık türü açıkça yazılmalıdır.

12.3. Birinin Diğerinin Yerine Kullanılması Yasaktır

Uygun nüfus skoru: "Gerçek kullanıcıların çoğu bu cevabı gördü." anlamına gelmez. Maruziyet skoru: "Dünya nüfusuna adil temsil sağlanıyor." anlamına gelmez.

13. ERİŞİM KAPSAMI İLE TEMSİL DOĞRULUĞU AYRILMALIDIR

Bir AI ürünü yalnız sınırlı ülkelerde bulunabilir. Bu ülkelerde çok doğru temsil üretebilir. Başka ürün daha geniş erişime sahip fakat daha düşük doğruluk üretebilir. Yalnız temsil doğruluğuna bakarsak dar erişimli ürün üstün görünebilir. Yalnız erişim kapsamına bakarsak yanlış temsil gözden kaçabilir. Bu nedenle iki ayrı sonuç gerekir.

13.1. Erişim Kapsamı

Belirtilen küresel referans nüfusun ne kadarının ilgili AI ürününü uygun biçimde kullanabildiğini gösterir. Aday gösterim:

```
AC_a = N(U_{a,t}) / N(G_t)
```

Burada:

- Ua,t: ürünün uygun kullanıcı nüfusu
- Gt: tanımlı küresel referans nüfus

13.2. Koşullu Temsil Doğruluğu

Ürüne erişebilen kullanıcılar arasında doğru temsil görme olasılığıdır. Aday gösterim:

```
RA_a = Pr(Y=1 | u ∈ U_{a,t})
```

13.3. Etkili Doğru Temsil Erişimi

CANDIDATE CONCEPT — ADAY KAVRAM

Erişim kapsamı ile koşullu doğruluğun birlikte gösterildiği operasyonel bir metrik geliştirilebilir:

```
ECR_a = AC_a × RA_a
```

Bu metrik kabaca şu soruya yaklaşır: "Tanımlı küresel referans nüfusun ne kadarı bu ürüne erişebilir ve doğru temsil görme koşulunu karşılayabilir?" Bu henüz nihai NOMOS formülü değildir. Çünkü:

- erişim ile gerçek kullanım farklıdır,
- ağırlık ve bağımlılık yapıları karmaşıktır,
- ürün kullanımı isteğe bağlıdır,

doğru temsil prompta bağlıdır. Ancak temel ayrımı korur:

> *Erişememek ile yanlış cevap almak aynı şey değildir.*

14. NATIVE REACH VE COMMON SUPPORT

AI ürünlerini karşılaştırırken iki ayrı nüfus evreni gerekir.

14.1. Native Reach Population

Her AI ürününün kendi resmî ve gerçek erişim alanındaki uygun nüfustur. Bu sonuç şu soruyu cevaplar:

> *Ürün, gerçekten ulaşabildiği nüfusu ne kadar doğru temsil ediyor?*

Ürünler arasında hedef nüfuslar farklı olabilir.

14.2. Common Support Population

Karşılaştırılan bütün AI ürünlerinin ortak biçimde:

- resmî olarak erişilebilir,
- aynı dil veya arayüz koşulunda kullanılabilir

olduğu nüfustur. Bu sonuç şu soruyu cevaplar:

> *Aynı kullanıcı evreninde hangi ürün daha doğru temsil üretiyor?*

14.3. İki Sonuç Birlikte Yayımlanmalıdır

Native Reach:

- gerçek ürün erişimini,
- doğal coğrafi yayılımı

gösterir. Common Support: daha adil ürün karşılaştırması sağlar. Bir ürün Native Reach skorunda yüksek, Common Support skorunda daha düşük olabilir. İki sonuç farklı sorulara cevap verir.

15. PAZARLA İLGİLİ NÜFUS

Bir varlığın küresel kimlik skoru ile ticari hedef pazar skoru farklı olabilir. Örnek: Yerel bir hukuk firması yalnız Türkiye'de hizmet veriyorsa: küresel kimlik sorusu dünyanın uygun kullanıcı nüfusunda ölçülebilir, tavsiye ve hizmet uygunluğu yalnız Türkiye veya gerçek hizmet kapsamındaki kullanıcılar için ölçülmelidir. Kurum kendi lehine sonucu seçmek için:

- yalnız yüksek skorlu pazarı,
- yalnız kendi dilini,
- yalnız mevcut müşterilerini

hedef nüfus ilan edemez. Hedef pazar:

- denetim başlamadan,
- hizmet kayıtlarına,
- sözleşme ve operasyon kapsamına

göre kilitlenmelidir.

16. KİŞİ, HESAP VE OTURUM AYRIMI

Nüfus ağırlığı insanlara aittir. AI hesaplarına değil.

16.1. Bir Kişi, Birden Fazla Hesap

Aynı insan:

- ücretsiz,
- ücretli,
- kurumsal

birden fazla hesaba sahip olabilir. Bu hesaplar ürün yüzeyi karşılaştırmasında ayrı gözlemler üretebilir. Küresel insan nüfusunda ayrı kişiler olarak sayılamaz.

16.2. Bir Hesap, Birden Fazla Kişi

Aile, ekip veya kurum hesabı birden fazla kişi tarafından kullanılabilir. Hesap kullanıcı sayısına eşit değildir.

16.3. Bir Kişi, Birden Fazla Oturum

Bir insan farklı oturumlarda ayrı gözlemler üretebilir. Bu gözlemler zaman veya oturum kararlılığını ölçebilir. Bağımsız nüfus genişliğini artırmaz.

16.4. Bir Kişi, Birden Fazla AI Ürünü

Eşlenmiş ürün karşılaştırması için değerlidir. Fakat kullanıcı düzeyindeki ortak özellikler nedeniyle gözlemler bağımlıdır.

16.5. Ana Ağırlık Birimi

Nüfus temelli GEO skorunda ana ağırlık birimi:

insan

olmalıdır. Hesap, oturum ve ürün gözlemleri insan birimi altında ilişkilendirilmelidir.

17. DİL NÜFUSU

Dil adaleti için yalnız ülkenin resmî dilini kullanmak yeterli değildir.

17.1. İstem Dili

Katılımcının denetim sorusunu sorduğu dildir.

17.2. Arayüz Dili

AI ürününün menü ve kullanıcı arayüzünde kullandığı dildir. İstem diliyle aynı olmayabilir.

17.3. Yanıt Dili

AI'nın fiilen cevap verdiği dildir. Sistem istem dilinden başka dilde cevap verebilir. Bu da ölçüm sonucudur.

17.4. Kullanıcı Yeterlilik Dili

Katılımcının istemi ve cevabı anlamlı biçimde kullanabildiği dildir. Ana dil olmak zorunda değildir.

17.5. Locale

Dil ile birlikte bölgesel bağlamdır. Aynı dil farklı:

- ülke,
- fiyat,
- hukuk,
- terminoloji

koşulu taşıyabilir.

18. ÇOK DİLLİ KULLANICILARIN ÇİFTE SAYILMASI

Bir kişi:

- Türkçe,
- İngilizce,
- Almanca

kullanabiliyor olabilir. Bu kişi üç ayrı dil testine katılabilir. Fakat küresel nüfus ağırlığında üç insan gibi sayılması nüfus tahminini bozar. Üç yöntem kullanılabilir.

18.1. Tek Ana Dil Ataması

Kişi ana analizde bir dil hücresine atanır. Diğer diller ek veya eşlenmiş test olarak kullanılır.

18.2. Bölünmüş Kişi Ağırlığı

Kişinin nüfus ağırlığı uygun dilleri arasında önceden tanımlı biçimde bölünebilir. Toplam ağırlığı bir kişiyi aşmaz.

18.3. Ayrı Çok Dilli Eşlenmiş Panel

Çok dilli kullanıcılar, diller arası aynı kişi karşılaştırmasında kullanılır. Bu panel ana nüfus panelinden ayrı analiz edilir. Hangi yöntem kullanılırsa kullanılsın:

> *Aynı kişinin toplam nüfus ağırlığı açıklamasız biçimde çoğaltılamaz.*

19. DİL UYGUNLUĞU NASIL DOĞRULANMALIDIR?

Bir katılımcının: "Bu dili biliyorum." demesi başlangıç kaydı olabilir. Ancak yüksek güvenli denetimde şu yöntemlerden biri kullanılabilir:

- Ana dil veya günlük kullanım beyanı
- Kısa okuma ve anlama testi
- Panel sağlayıcının dil kaydı
- Yerel hakem doğrulaması
- Eğitim veya profesyonel kullanım kaydı
- Pilot görev başarısı

Aşırı zor dil testleri paneli eğitimli ve yüksek gelirli kullanıcılara kaydırabilir. Amaç akademik dil sınavı yapmak değildir. İstemi ve cevabı gerçek kullanıcı olarak anlayabilecek yeterliliği doğrulamaktır.

20. DİL DESTEKLENMİYORSA NE OLUR?

Bir AI ürünü belirli dili:

- resmî olarak desteklemiyor,
- kısmen destekliyor,
- desteklediğini beyan etmiyor fakat cevap üretebiliyor

olabilir. Dil statüleri ayrı kaydedilmelidir:

OFFICIALLY_SUPPORTED

LIMITED_SUPPORT

OBSERVED_BUT_UNSUPPORTED

NOT_SUPPORTED

UNKNOWN

Resmî olarak desteklenmeyen dilde alınan cevap:

- keşif,
- deneysel erişim,
- dil dayanıklılığı

ölçümü olabilir. Ana desteklenen kullanıcı nüfusu skoruna otomatik olarak eklenmemelidir. Dil erişim sınırı ayrıca raporlanmalıdır.

21. ÜLKE NÜFUS AĞIRLIĞI VE ÜLKE GÖZCÜ KAPSAMI

Nüfus ağırlıklı ana panel ve ülke görünürlüğü paneli ayrı amaç taşır.

21.1. Nüfus Paneli

Katılımcılar ürün-uygun hedef nüfustaki paylarına göre dağıtılır. Büyük uygun nüfuslar daha yüksek ağırlık alır. Ana küresel nüfus skorunu üretir.

21.2. Ülke Gözcü Paneli

Ana nüfus örnekleminde hiç veya çok az pay alan küçük ülkelerden en azından gözlem elde etmeyi amaçlar. Bir gözcü gözlemi:

- erişim bulunduğunu,
- belirli tekil sonucun oluştuğunu,
- yeni araştırma ihtiyacını

gösterebilir. Tek başına ülke skoru üretmez.

21.3. Ülke Skoru İçin Yeterlilik

Ülke bazlı oran yayımlanabilmesi için:

- minimum örneklem,
- geçerli gözlem oranı,
- belirsizlik,
- panel temsiliyeti

koşulları sağlanmalıdır. Bir veya birkaç katılımcıyla: “Ülkenin NOMOS skoru” ilan edilemez.

22. DİL ADALETİ PANELİ

Nüfus ağırlığı, düşük nüfuslu veya düşük kaynaklı dilleri ana skor içinde görünmez bırakabilir. Bu nedenle ayrı bir Dil Adaleti Paneli kurulabilir. Amaç:

- yeterince temsil edilmeyen dilleri test etmek,

- dil bazlı ağır bozulmayı bulmak,
- istem çeviri sorunlarını belirlemek,
- düşük kaynaklı dil davranışını görünür kılmak

olmalıdır. Dil Adaleti Paneli ana nüfus paneline ek örnekler sağlayabilir. Ancak fazla örneklenen dil gözlemleri küresel skorda gerçek nüfus ağırlıklarına geri döndürülmelidir. Aksi hâlde adalet paneli küresel nüfus tahminini değiştirir.

23. ERİŞİLEBİLİRLİK PANELİ

Görme, işitme, motor, bilişsel veya başka erişilebilirlik ihtiyaçları bulunan kullanıcılar hedef nüfusun gerçek parçasıdır. NOMOS Capture veya araştırma görevi bu kişilerin katılımına uygun değilse: kullanıcılar hedef nüfustan çıkarılamaz. Doğru sonuç:

> *Örnekleme çerçevesi erişilebilirlik bakımından eksiktir.*

Ayrı erişilebilirlik paneli şunları değerlendirebilir:

- Ekran okuyucu uyumu
- Klavye ile kullanım
- Sesli giriş
- Düşük görme ayarları
- Bilişsel yük
- Mobil erişilebilirlik
- Capture aracının kullanılabilirliği

Erişilebilirlik paneli, kullanıcıları "özel istisna" olarak değil nüfus kapsamının parçası olarak görmelidir.

24. HEDEF NÜFUS VE ÖRNEKLEME ÇERÇEVESİ

Bu iki kavram birbirine karıştırılmamalıdır.

24.1. Hedef Nüfus

Sonucun hakkında hüküm kurmak istediği bütün insan topluluğudur.

24.2. Örnekleme Çerçevesi

Katılımcıların fiilen seçilebildiği:

- panel,
- liste,
- veri tabanı,
- ortak kuruluş,
- kullanıcı havuzu

kümesidir. Örnekleme çerçevesi hedef nüfusu tam olarak kapsamayabilir.

24.3. Temas Edilebilir Nüfus

Araştırma davetinin gerçekten ulaştırılabildiği kişilerdir.

24.4. Katılmayı Kabul Eden Nüfus

Gönüllü olarak araştırmaya katılanlardır.

24.5. Görevi Tamamlayan Nüfus

Denetim görevini tamamlayanlardır.

24.6. Geçerli Analitik Nüfus

Protokole uygun ve analizde kullanılabilir gözlem üretenlerdir. Bu katmanların sayıları ayrı raporlanmalıdır.

25. KAPSAMA HATASI

Örnekleme çerçevesi hedef nüfusun bazı bölümlerini dışarıda bırakabilir. Buna kapsama hatası denir. Örnekler:

- Yalnız profesyonel çevrim içi paneller
- Yalnız masaüstü kullanıcıları
- Yalnız kredi kartı bulunan kişiler
- Yalnız İngilizce arayüz kullananlar
- Yalnız büyük şehir sakinleri
- Yalnız teknolojiye yoğun ilgi duyan kullanıcılar
- Erişilebilirlik ihtiyacı bulunan kişilerin dışlanması

Kapsama hatası:

- açıklanmalı,
- mümkünse ölçülmeli,

ağırlık veya ek panelle azaltılmalıdır. Ağırlıklandırma, örnekleme çerçevesinde hiç bulunmayan grubu sihirli biçimde oluşturamaz.

26. FAZLA KAPSAMA

Örnekleme çerçevesinde hedef nüfusa ait olmayan kişiler bulunabilir. Örnekler:

- Ürünün erişilemediği ülkedeki katılımcı
- Gerekli yaş koşulunu karşılamayan kişi
- Yanlış plan veya kullanıcı yüzeyi
- Denetim dilini kullanamayan katılımcı
- Birden fazla kayıtla tekrar bulunan kişi
- Başka ülke adına görev yapan kullanıcı

Fazla kapsama kayıtları ana skora girmemelidir. Ancak dışlama nedeni ve oranı raporlanmalıdır.

27. YANITLAMAMA VE KENDİ KENDİNE SEÇİLİM

Araştırmaya katılmayı kabul eden kişiler hedef nüfustan sistematik olarak farklı olabilir. Gönüllüler:

- AI ürünlerini daha sık kullanabilir,
- teknolojiye daha ilgili olabilir,
- daha yüksek eğitim düzeyine sahip olabilir,
- markayı önceden tanıyabilir,

ödül veya ücret nedeniyle farklı davranabilir. Bu nedenle: "Dünyanın rastgele 1.000 insanı" ifadesi yalnız gerçekten olasılıklı ve belgelenmiş seçim tasarımında kullanılmalıdır. Gönüllü veya çevrim içi panel kullanılıyorsa doğru ifade:

Tabakalı ve ağırlıklandırılmış gönüllü kullanıcı paneli

olabilir. Panel türü saklanamaz.

28. UYGUN NÜFUS ZAMANLA DEĞİŞİR

Bir AI ürününün kullanıcı evreni sabit değildir. Şunlar değişebilir:

- Desteklenen ülkeler
- Asgari yaş
- Ücretsiz erişim
- Ücretli plan yapısı
- Mobil veya web kullanılabilirliği
- Dil desteği
- Hesap şartları
- Kurumsal ürün kapsamı
- Yasal ve düzenleyici durum

Bu nedenle her nüfus kaydı:

- gözlem tarihi,
- yürürlük dönemi,
- kaynak sürümü,
- son doğrulama tarihi

taşımalıdır.

28.1. Ölçüm Dalgası Sırasında Değişiklik

Ürün erişimi dalga sırasında maddi biçimde değişirse:

- dalga durdurulabilir,
- değişiklik öncesi ve sonrası iki ayrı dönem oluşturulabilir,

etkilenen ülkeler ayrı analiz edilebilir. Değişiklik öncesi ve sonrası kullanıcılar aynı nüfus evreni gibi birleştirilemez.

29. NÜFUS KAYNAK DOSYASI

Her denetim, kullandığı hedef nüfus verilerini sürümlenmiş bir dosyada saklamalıdır. İlgili olduğu ölçüde şu alanlar bulunmalıdır:

- Ülke
- Referans nüfus
- Nüfus tarihi
- Yaş uygunluğu
- İnternet ve cihaz erişimi
- Ürün erişim statüsü
- Arayüz erişimi

- Plan ve hesap şartı
- Dil grupları
- Araştırma uygunluğu
- Ortak dağılım yöntemi
- Belirsizlik
- Kaynaklar
- Son doğrulama tarihi
- Sorumlu kişi
- Nüfus sürümü

Bu dosya olmadan ülke kotaları sonuçtan sonra yeniden üretilemez.

30. SENTETİK ÜRÜN VE ÜLKE GÖSTERİMİ

SENTETİK METODOLOJİ GÖSTERİMİ / Aşağıdaki ülkeler, AI ürünü, nüfus sayıları ve erişim oranları tamamen kurgusaldır. Hiçbir gerçek ülkeyi veya ürünü temsil etmez. Orion AI adlı sentetik genel tüketici ürünü düşünelim. Dört sentetik ülke vardır.

Ülke	Toplam nüfus	Yaş, internet, dil ve arayüz koşulları sonrası tahmini uygun nüfus	Ürün erişimi	Native Reach nüfusu
A Ülkesi	1.400.000.000	780.000.000	Kullanılamıyor	0
B Ülkesi	1.450.000.000	520.000.000	Resmî erişim	520.000.000
C Ülkesi	340.000.000	250.000.000	Resmî erişim	250.000.000
D Ülkesi	35.000	25.000	Resmî erişim	25.000
Toplam	3.190.035.000	1.550.025.000		770.025.000

30.1. Toplam Nüfus Kullanılsaydı

A Ülkesi, çok büyük nüfusu nedeniyle ana örneklemde büyük pay alırdı. Fakat ürün o ülkede kullanılamamaktadır. Katılımcılar doğal kullanıcı deneyimi üretemezdi.

30.2. Native Reach Nüfusu Kullanıldığında

Ana ürün panelinin teorik nüfus payları yaklaşık olarak:

- B Ülkesi: yüzde 67,53
- C Ülkesi: yüzde 32,47
- D Ülkesi: yüzde 0,003
- A Ülkesi: Native Reach dışında

olur. D Ülkesi'nin 1.000 kişilik ana örneklemde teorik payı: bir kişinin çok altındadır. Bu nedenle D Ülkesi nüfus ağırlıklı ana skorda yüksek ağırlık almamalıdır. Fakat ülke tamamen görünmez bırakılmak istenmiyorsa:

Country Sentinel Panel

içinde en az bir keşif gözlemi alınabilir. Bu gözlem:

- D Ülkesi'nin skoru değildir,
- ana nüfus ağırlığını değiştirmez,

ülkede ürün davranışının gözlenebildiğini gösterir.

30.3. A Ülkesi Nasıl Raporlanır?

A Ülkesi için:

OUTSIDE_NATIVE_REACH

statüsü kullanılır. A Ülkesi'nin çıkarılması ürünün erişim sınırını gizlememelidir. Kamu sonuç kartı ayrıca şunu gösterir:

- Ürün erişim kapsamı
- Küresel referans nüfusta erişilemeyen pay
- Native Reach skoru
- Common Support karşılaştırma skoru

31. SENTETİK İSTEM NÜFUSU GÖSTERİMİ

SENTETİK ÖRNEK

Aynı Orion AI ürünü için üç istem düşünelim.

İstem A

"Asteron nasıl bir şirkettir?" Hedef nüfus: Orion AI'nın geniş uygun kullanıcı nüfusu olabilir.

İstem B

"Asteron Türkiye'de kurumsal seyahat hizmeti veriyor mu?" Hedef nüfus:

- Türkiye'deki uygun kullanıcılar,
- Türkiye hizmetini araştıran tanımlı dış kullanıcılar

olabilir.

İstem C

"Asteron bana hukuki göçmenlik danışmanlığı için uygun mu?" Asteron böyle bir hizmet vermiyorsa: tavsiye-eligible nüfus bulunmayabilir, doğru cevap bütün uygun kullanıcılarda dışlama olmalıdır. Bu istemi genel kimlik skoru içinde birleştirmek yanlış olabilir. Bu örnek şunu gösterir:

> *Aynı varlık ve aynı AI ürünü için bile hedef nüfus, ölçülen kullanıcı sorusuna göre değişebilir.*

32. DÖRT PANEL MİMARİSİ

CANDIDATE ARCHITECTURE — ADAY MİMARİ

GEO-1000'in küresel ve adil uygulanması için dört tamamlayıcı panel önerilir.

32.1. Population Panel

Ürün-uygun kullanıcı nüfusuna göre ağırlıklandırılmış ana paneldir. Ana nüfus tahminini üretir.

32.2. Country Sentinel Panel

Nüfus panelinde yeterli örnek alamayan küçük ülkelerin tamamen görünmez kalmasını önler. Ülke skoru üretmek için tek başına yeterli değildir.

32.3. Language Fairness Panel

Düşük kaynaklı veya nüfus ortalamasında bastırılan dilleri ayrıca ölçer. Küresel ana skora gerçek nüfus ağırlıklarıyla geri bağlanır.

32.4. Accessibility Panel

Standart dijital panelin dışlayabileceği erişilebilirlik ihtiyaçlarını ve yardımcı teknoloji kullanıcılarını ölçer. Bu dört panelin gözlemleri:

- amaç,
- ağırlık,
- çıkarım evreni

bakımından ayrı etiketlenmelidir. Aynı gözlem açıklamasız biçimde birden fazla panelde tam ağırlıkla kullanılamaz.

33. YETİŞKİN BAŞLANGIÇ PANELİ

CANDIDATE BASE RULE — ADAY TEMEL KURAL

GEO-1000'in ilk genel sürümünde gerçek kullanıcı panellerinin:

ürün ve ülke koşullarını karşılayan yetişkin katılımcılardan

oluşturulması önerilir. Bu yaklaşım:

- rıza,
- veri koruma,
- ekran görüntüsü toplama,
- yüksek riskli araştırma,
- ebeveyn onayı

karmaşıklığını azaltabilir. Fakat şu sonuç çıkarılamaz: "Çocuk ve genç kullanıcıların deneyimi yetişkin paneliyle ölçülmüştür." Çocukları veya gençleri etkileyen ürünlerde ayrı:

- etik onay,
- yaşa uygun protokol,
- ebeveyn veya vasi süreçleri,
- zarar koruması

gerekir.

34. HİÇBİR NÜFUS ÇERÇEVESİ "BÜTÜN İNSANLIK" DİYE ADLANDIRILAMAZ

Bir ürün:

- bazı ülkelerde yoksa,
- bazı dilleri desteklemiyorsa,
- yalnız belirli cihazlarda çalışıyorsa,
- belirli yaş veya plan koşulu taşıyorsa,

sonuç: "bütün dünya insanlarının GEO deneyimi" olarak sunulamaz. Doğru ifade:

Belirtilen ürün-uygun hedef nüfustaki temsil tahmini

olmalıdır. Küresel kelimesi:

- çok ülkeli,
- çok dilli,
- geniş nüfuslu

bir ölçümü tanımlayabilir. Sınırsız ve evrensel kapsam anlamına gelmez.

35. NÜFUSUN KURUCU ETİK SINIRI

Bir nüfusu ölçüm için tanımlamak, bazı insanları değersiz saymak değildir. Bir grup ana skorda düşük ağırlık alabilir. Fakat:

- yüksek risk,
- dil adaleti,
- erişilebilirlik,
- ülke görünürlüğü

nedeniyle ayrıca incelenebilir. Bir grup ürüne erişemediği için Native Reach skorunun dışında olabilir. Bu erişimsizlik ayrıca küresel kapsam açığı olarak raporlanabilir. NOMOS'un amacı:

- herkesi tek yapay ortalamaya sıkıştırmak,
- yalnız en büyük nüfusları görmek,
- küçük toplulukları sembolik olarak bir kişiyle temsil etmiş görünmek

değildir. Amaç şudur:

> *Her insanı doğru sorunun doğru paydasında, ağırlığı ve görünürlüğü birbirine karıştırmadan temsil etmek.*

36. ZORUNLU NORMATİF HÜKÜMLER

CH05-N01

Her AI ürünü için hedef kullanıcı nüfusu ayrı tanımlanmalıdır.

CH05-N02

Bir ülkenin toplam nüfusu doğrudan AI-uygun kullanıcı nüfusu olarak kullanılamaz.

CH05-N03

Hedef nüfus; ürün, ülke, yaş, erişim, arayüz, plan, dil, istem ve zaman koşullarına bağlanmalıdır.

CH05-N04

Hedef nüfus ve dışlama kuralları sonuçlar görülmeden önce sürümlenmelidir.

CH05-N05

Bir AI ürününün hedef nüfusu başka AI ürününe açıklamasız biçimde aktarılamaz.

CH05-N06

Ürün erişiminin bulunmaması yanlış temsil sonucu olarak puanlanamaz; erişim kapsamı olarak ayrı raporlanmalıdır.

CH05-N07

Katılımcılardan ürün veya ülke erişim kısıtlarını aşmaları istenemez.

CH05-N08

Resmî desteklenmeyen erişim, ana Native Reach nüfusuna otomatik olarak eklenemez.

CH05-N09

Kişi, hesap ve oturum birimleri birbirinden ayrılmalıdır.

CH05-N10

Bir kişinin birden fazla hesabı veya dili, kişiyi nüfus ağırlığında çoğaltamaz.

CH05-N11

Dil uygunluğu yalnız ülkenin resmî diline dayanarak atanamaz.

CH05-N12

İstem dili, arayüz dili, yanıt dili ve kullanıcı yeterlilik dili ayrı kaydedilmelidir.

CH05-N13

Birden fazla dili kullanan kişinin toplam nüfus ağırlığı açıklamasız biçimde bir kişiyi aşamaz.

CH05-N14

Nüfus ağırlıklı panel ile ülke gözcü paneli ayrı tutulmalıdır.

CH05-N15

Bir veya birkaç gözcü gözlemi ülke skoru olarak yayımlanamaz.

CH05-N16

Dil Adaleti Paneli'nde fazla örneklenen gözlemler küresel skorda gerçek nüfus ağırlığına geri döndürülmelidir.

CH05-N17

Erişilebilirlik ihtiyacı bulunan kullanıcılar, veri toplama aracı uygun olmadığı için hedef nüfustan çıkarılamaz.

CH05-N18

Hedef nüfus, örnekleme çerçevesi, temas edilen nüfus, katılımcılar ve geçerli analitik gözlemler ayrı raporlanmalıdır.

CH05-N19

Olasılıksız veya gönüllü panel, rastgele dünya örneklemi olarak sunulamaz.

CH05-N20

Ağırlıklandırma, örnekleme çerçevesinde hiç bulunmayan grubu gözlenmiş hâle getiremez.

CH05-N21

Potansiyel uygun nüfus ile aktif kullanıcı maruziyeti aynı metrik gibi kullanılamaz.

CH05-N22

Native Reach ve Common Support sonuçları birbirinden ayrılmalıdır.

CH05-N23

Kimlik, hizmet, fiyat, hukuk ve tavsiye istemleri için aynı hedef nüfus otomatik olarak kullanılamaz.

CH05-N24

Nüfus verisi, ürün erişim statüsü ve dil dağılımı zaman damgası ve kaynak sürümü taşımalıdır.

CH05-N25

Marjinal uygunluk oranları, bağımsızlık varsayımı açıklanmadan kör biçimde çarpılamaz.

CH05-N26

Ürün türü genel tüketici paneline uygun değilse GEO-1000 insan nüfus yöntemi açıklamasız biçimde uygulanamaz.

CH05-N27

Çocuk veya reşit olmayan kullanıcılar ayrı etik ve hukuki protokol olmadan ana panele alınamaz.

CH05-N28

Ürün erişim kapsamı ile temsil doğruluğu ayrı sonuçlar olarak yayımlanmalıdır.

CH05-N29

Küresel ifadesi kapsam dışı ülke, dil, yaş, arayüz veya planları görünmez kılamaz.

CH05-N30

Hedef nüfus kaydının hesap verebilir insan veya kurum sahibi bulunmalıdır.

37. BAŞARISIZLIK BİÇİMLERİ

CH05-F01 — TOPLAM NÜFUSU UYGUN NÜFUS SAYMAK

Ürün, yaş, internet, dil ve arayüz koşulları değerlendirilmeden ülke nüfusu doğrudan paydaya alınır.

CH05-F02 — BÜTÜN ÜRÜNLERE AYNI DÜNYA HARİTASINI UYGULAMAK

Farklı erişim alanları bulunan AI ürünleri aynı hedef nüfusla ölçülür.

CH05-F03 — KISIT AŞAN KULLANICIYI DOĞAL KULLANICI SAYMAK

Resmî erişim koşullarını aşan gözlem ülke nüfus deneyimi gibi sunulur.

CH05-F04 — ÜRÜN YOKLUĞUNU YANLIŞ CEVAP SAYMAK

Erişilemeyen ülkeler temsil doğruluğunda sıfır kabul edilir.

CH05-F05 — ÜRÜN YOKLUĞUNU TAMAMEN GİZLEMEK

Erişilemeyen büyük nüfuslar Native Reach paydasından çıkarılır; erişim kapsamı ayrıca gösterilmez.

CH05-F06 — BİR KİŞİYİ BİRDEN FAZLA İNSAN SAYMAK

Aynı kişinin hesap, dil veya ürün gözlemleri nüfus ağırlığını çoğaltır.

CH05-F07 — HESABI KİŞİ SAYMAK

Çoklu hesap sahipliği veya ortak hesap kullanımı değerlendirilmez.

CH05-F08 — RESMÎ DİLİ KULLANICI DİLİ SAYMAK

Ülkenin resmî dili bütün katılımcıların denetim dili kabul edilir.

CH05-F09 — ÇOK DİLLİ KİŞİLERİ ÇİFTE SAYMAK

Aynı birey her dilde tam nüfus ağırlığı alır.

CH05-F10 — KÜÇÜK ÜLKEYE BİR KİŞİ VERİP ÜLKE SKORU AÇIKLAMAK

Gözcü gözlemi istatistiksel ülke tahminine dönüştürülür.

CH05-F11 — KÜÇÜK ÜLKELERİ TAMAMEN SİLMEK

Nüfus payı düşük olduğu için ülke veya dil hiçbir ek kapsam paneline alınmaz.

CH05-F12 — DİL ADALETİ ÖRNEKLERİNİ ANA SKORDA AŞIRI AĞIRLIKLANDIRMAK

Fazla örneklenen diller gerçek nüfus ağırlığına geri döndürülmez.

CH05-F13 — AKTİF KULLANICIYI POTANSİYEL NÜFUS SAYMAK

Gerçek kullanım verisi, ürüne erişebilecek bütün insanlar gibi sunulur.

CH05-F14 — POTANSİYEL NÜFUSU GERÇEK MARUZİYET SAYMAK

Ürünü kullanabilecek herkesin yanıtı fiilen gördüğü varsayılır.

CH05-F15 — GÖNÜLLÜ PANELİ RASTGELE DÜNYA ÖRNEKLEMİ SAYMAK

Kendi kendine seçilim ve panel yanlılığı saklanır.

CH05-F16 — AĞIRLIKLA KAPSAMA HATASINI YOK SAYMAK

Çerçevede hiç bulunmayan grupların yalnız ağırlıklandırmayla temsil edildiği iddia edilir.

CH05-F17 — VERİ TOPLAMA ARACININ DIŞLADIĞI KULLANICIYI NÜFUSTAN ÇIKARMAK

Erişilebilirlik veya cihaz problemi kullanıcı uygun değilmiş gibi yorumlanır.

CH05-F18 — BÜTÜN İSTEMLERE AYNI NÜFUSU UYGULAMAK

Küresel kimlik, yerel fiyat ve hukuki tavsiye soruları aynı paydada birleştirilir.

CH05-F19 — GENEL TÜKETİCİ NÜFUSUNU API ÜRÜNÜNE UYGULAMAK

İnsan kullanıcı paneli geliştirici veya uygulama ürününe doğrudan aktarılır.

CH05-F20 — ÜCRETSİZ VE ÜCRETLİ PLANLARI GİZLİCE BİRLEŞTİRMEK

Plan farkı kaydedilmeden tek ürün dağılımı yayımlanır.

CH05-F21 — WEB VE MOBİL NÜFUSLARI TEKLEŞTİRMEK

Arayüz erişimi ve davranış farkları görünmez kalır.

CH05-F22 — YAŞ SINIRINI YOK SAYMAK

Uygun olmayan yaş grupları hedef nüfusa veya panele açıklamasız eklenir.

CH05-F23 — MARJİNAL ORANLARI BAĞIMSIZMIŞ GİBİ ÇARPMAK

Yaş, internet, dil ve erişim arasındaki ilişki yok sayılır.

CH05-F24 — NÜFUSU SONUÇTAN SONRA DEĞİŞTİRMEK

Düşük skorlu ülke veya kullanıcı grubu hedef nüfustan çıkarılır.

CH05-F25 — ESKİ ÜRÜN ERİŞİM HARİTASINI KULLANMAK

Ölçüm tarihindeki gerçek ürün erişimi doğrulanmaz.

CH05-F26 — COMMON SUPPORT VE NATIVE REACH SONUÇLARINI KARIŞTIRMAK

Ortak nüfusta yapılan karşılaştırma ürünün küresel gerçek erişimi gibi sunulur.

CH05-F27 — PAZAR HEDEFİNİ TİCARİ ÇIKARA GÖRE DARALTMAK

Kurum yalnız başarılı olduğu ülkeleri hedef pazar ilan eder.

CH05-F28 — ERİŞİM KAPSAMINI GEO DOĞRULUĞU SAYMAK

Çok ülkede bulunmak doğru temsil kanıtı yapılır.

CH05-F29 — DAR ERİŞİMİ YÜKSEK GEO BAŞARISI OLARAK SATMAK

Az sayıda ülkede yüksek doğruluk, küresel üstünlük gibi sunulur.

CH05-F30 — “BÜTÜN DÜNYA” İFADESİNİ KAPSAM KAYDI OLMADAN KULLANMAK

Ürün, dil, ülke ve plan dışlamaları saklanır.

38. DENETİM PROSEDÜRÜ

Adım 1 — AI Ürününü ve Kullanıcı Yüzeyini Kilitle

Ürün Arayüz Plan Hesap türü Görünen model Ölçüm tarihi kaydedilir.

Adım 2 — Ölçüm Sorusunu Sınıflandır

İstem:

- küresel kimlik,
- yerel hizmet,
- fiyat,
- hukuk,
- tavsiye,
- yüksek risk

ailelerinden hangisine aittir?

Adım 3 — Küresel Referans Nüfusu Belirle

Sonucun ilişkilendirileceği başlangıç insan evreni tanımlanır.

Adım 4 — Yaş ve Araştırma Uygunluğunu Belirle

Ürün yaş koşulu Araştırma yaş koşulu Rıza ve veri işleme şartları kaydedilir.

Adım 5 — Ülke Bazlı Ürün Erişimini Doğrula

Her ülke için PA-0 ile PA-5 arasında erişim statüsü atanır.

Adım 6 — İnternet, Cihaz ve Arayüz Uygunluğunu Belirle

Ölçülen yüzeyi kullanabilecek nüfus tahmin edilir.

Adım 7 — Dil ve Locale Evrenini Kur

İstem dili Arayüz dili Yanıt dili Kullanıcı dil yeterliliği ayrılır.

Adım 8 — İstem İlgisini Uygula

Bütün kullanıcıların ilgili istem ailesinin hedef nüfusuna ait olup olmadığı değerlendirilir.

Adım 9 — Ürün Türüne Uygun Nüfus Modelini Seç

Tüketici, kurumsal, eğitim, API veya gömülü ürün ayrılır.

Adım 10 — Hedef Nüfus ile Örnekleme Çerçevesini Karşılaştır

Eksik ve fazla kapsama kaydedilir.

Adım 11 — Ana ve Tamamlayıcı Panelleri Tanımla

Population Panel Country Sentinel Panel Language Fairness Panel Accessibility Panel kapsamları sürümlenir.

Adım 12 — Native Reach ve Common Support Evrenlerini Ayır

Ürün içi sonuç ve ürünler arası karşılaştırma nüfusları ayrı kaydedilir.

Adım 13 — Erişim Kapsamı ve Doğruluk Sonuçlarını Ayır

Bir ürünün ulaşabildiği nüfus ile erişen kullanıcılar arasındaki doğruluk ayrı raporlanır.

Adım 14 — Nüfus Dosyasını Dondur

Veri kaynakları, varsayımlar, tarih ve nüfus sürümü kilitlenir.

Adım 15 — Maddi Değişiklikleri İzle

Ölçüm sırasında ürün erişimi veya şartlar değişirse yeni sürüm oluşturulur.

39. GEREKLİ KANIT

AI ürün kimliği Kullanıcı yüzeyi Plan ve hesap koşulları Ölçüm tarihi Ürün erişim ülkeleri Ülke bazlı erişim statüleri Yaş koşulları Hukuki ve etik katılım koşulları Ülke nüfus kayıtları Yaş dağılımları İnternet ve cihaz erişim kayıtları Dil ve locale dağılımları Ürün dil destek durumu Arayüz erişilebilirliği Hedef istem ailesi Hedef nüfus tanımı Örnekleme çerçevesi Kapsama eksiklikleri Fazla kapsama kayıtları Panel türü Katılımcı seçim yöntemi Gönüllü veya olasılıklı panel statüsü Çok dilli kullanıcı ağırlıkları Country Sentinel kapsamı Language Fairness kapsamı Accessibility Panel kapsamı Native Reach nüfusu Common Support nüfusu Aktif kullanıcı verisi, kullanılıyorsa Erişim kapsamı hesabı Varsayım ve duyarlılık analizleri Nüfus sürümü Son doğrulama tarihi

Sorumlu insan veya kurum

40. DENETİM KONTROL LİSTESİ

Hangi AI ürünü ve kullanıcı yüzeyi ölçülüyor? Hedef nüfus hangi tarih için tanımlandı? Ülke toplam nüfusu ile uygun kullanıcı nüfusu ayrıldı mı? Yaş koşulu uygulandı mı? İnternet ve cihaz erişimi değerlendirildi mi? Ürün ülke bazında resmî olarak erişilebilir mi? Katılımcılardan erişim kısıtı aşmaları istendi mi? Ücretsiz ve ücretli planlar ayrıldı mı? Web ve mobil yüzeyler ayrıldı mı? İstem ailesi hedef nüfusu değiştirdi mi? Dil ve locale uygunluğu nasıl belirlendi? Çok dilli kullanıcılar çifte sayıldı mı? Hedef nüfus ile örnekleme çerçevesi arasındaki fark açık mı? Panel olasılıklı mı, gönüllü mü? Kapsama hatası ölçüldü mü? Erişilebilirlik ihtiyacı bulunan kullanıcılar sessizce dışlandı mı?

Country Sentinel gözlemleri ülke skoruna dönüştürüldü mü? Language Fairness örnekleri ana skorda aşırı ağırlık aldı mı? Native Reach ve Common Support sonuçları ayrıldı mı? Erişim kapsamı ile temsil doğruluğu ayrı mı? Aktif kullanıcı ve potansiyel kullanıcı sonuçları ayrıldı mı? Ürün türü insan nüfus paneline uygun mu? Nüfus katsayılarında marjinal oranlar kör biçimde çarpıldı mı? Nüfus tanımı sonuç görüldükten sonra değişti mi? Kullanılan nüfus ve erişim verileri sürümlü mü? Test edilmeyen gruplar görünür mü? “Küresel” ifadesi gerçek kapsamla uyumlu mu? Nüfus kaydının hesap verebilir sahibi belli mi?

41. İTİRAZLAR VE CEVAPLAR

İtiraz 1 — “Dünya nüfusuna göre doğrudan dağıtmak daha adil değil mi?”

Dünya nüfusu güçlü başlangıç tabanıdır. Fakat ürünü kullanamayacak insanları ana ürün örneklemine almak gerçek kullanıcı deneyimi üretmez. Doğru yöntem:

- ürün-uygun nüfus içinde nüfus ağırlığı,
- ayrıca küresel erişim kapsamı

raporlamaktır.

İtiraz 2 — "Büyük ülkelerin çoğu kullanıcıyı alması gerekmez mi?"

Ürün, dil ve erişim koşullarını karşılayan büyük ülkeler doğal olarak yüksek pay almalıdır. Ancak yalnız toplam nüfus büyük olduğu için ürünün erişilemediği bir ülkeden sahte kullanıcı deneyimi üretilemez.

İtiraz 3 — "San Marino gibi küçük bir ülkeye bir kişi vermek neden yanlış olsun?"

Bir kişi keşif ve gözcü gözlemi için değerlidir. Yanlış olan, o kişinin sonucunu: "ülke skoru" olarak sunmaktır. Küçük ülkenin görünürlüğü Country Sentinel Panel ile korunabilir. Ana nüfus ağırlığı ise gerçek nüfus payına göre kalır.

İtiraz 4 — "Her ülkeye en az bir kişi verirsek daha küresel görünmez miyiz?"

Daha geniş coğrafi kapsam sağlanır. Fakat nüfus tahmini bozulabilir. Bu nedenle:

- coğrafi kapsam,
- nüfus ağırlığı

ayrı sonuçlardır. Her ülkeye bir kişi vermek kapsam paneli olabilir. Nüfus paneli değildir.

İtiraz 5 — "Ürün erişilemeyen ülkeye sıfır verirsek ürün haksız yere cezalandırılmaz mı?"

Erişilemeyen ülke temsil doğruluğunda sıfır sayılmaz. Erişim kapsamı metriğinde görünür olur. Bu ayrım ürünü hem haksız cezadan hem sahte küresellikten korur.

İtiraz 6 — "Gerçek kullanım verisi varsa nüfus verisine neden ihtiyacımız var?"

Gerçek kullanım verisi operasyonel etkiyi gösterir. Fakat ürünün:

- hangi nüfuslara ulaşmadığını,
- kullanım tabanındaki coğrafi eşitsizliği,
- potansiyel kullanıcı adaletini

tek başına göstermez. İki veri türü farklı sorulara cevap verir.

İtiraz 7 — "Bütün kullanıcılar aynı istemi soracağına göre istem ilgisi neden önemli?"

Kimlik sorusu geniş nüfusta anlamlı olabilir. Fiyat, hukuk veya tavsiye sorusu yalnız belirli kullanıcı ve pazarlarda anlamlıdır. İlgisiz nüfus, cevap dağılımını kullanıcı gerçekliğinden uzaklaştırabilir.

İtiraz 8 — "Bir kişi birden fazla dil biliyorsa neden bütün dillerde tam sayılmasın?"

Diller arası karşılaştırmada bütün dilleri test edebilir. Fakat nüfus ağırlığında aynı insanı çoğaltmak dünya nüfusunu yapay biçimde büyütür. Kişi ağırlığı bölünmeli veya ayrı eşlenmiş panel kullanılmalıdır.

İtiraz 9 — "Yalnız yetişkinleri test etmek küresel sonucu eksik bırakmaz mı?"

Bırakır. Bu sınır açıkça yazılmalıdır. Yetişkin başlangıç paneli etik ve operasyonel bir tercih olabilir. Çocuk kullanıcılar hakkında hüküm oluşturmaz.

İtiraz 10 — "Çevrim içi araştırma paneli yeterli değil mi?"

Pilot ve ağırlıklandırılmış araştırma için çok değerli olabilir. Fakat panel:

- hedef nüfusu ne kadar kapsıyor,
- kimleri dışarıda bırakıyor,
- gönüllü katılım nasıl yanlılık oluşturuyor

açıklanmalıdır. Panel kullanmak hata değildir. Paneli rastgele dünya nüfusu gibi sunmak hatadır.

İtiraz 11 — "Erişilebilirlik paneli ana GEO skoruyla neden ilgili?"

Çünkü gerçek AI kullanıcılarının bir bölümü yardımcı teknoloji veya farklı etkileşim biçimleri kullanır. Veri toplama aracı bu insanları dışlıyorsa ana skor hedef nüfusun tamamını temsil etmeyebilir.

İtiraz 12 — "Bu kadar farklı nüfus tanımı tek NOMOS skorunu imkânsızlaştırmaz mı?"

Hayır. Tek skor üretilebilir. Fakat hangi nüfus sorusuna cevap verdiği açık olmalıdır. Örneğin:

- Native Reach Population Puan
- Common Support Puan
- Market-Relevant Puan
- Language Fairness Result

ayrı kayıtlar taşıyabilir. Bir sayının anlamı, paydasının açıklığıyla oluşur.

44. BÖLÜMÜN ORTAK HÜKMÜ

GEO-1000'in adında 1.000 vardır. Fakat 1.000 kişinin değeri yalnız sayılarından doğmaz. Şu soruların cevabından doğar: Bu insanlar kimdir? Hangi nüfustan gelmiştir? İlgili ürünü gerçekten kullanabilirler mi? Hangi ülkelerde yaşarlar? Hangi dili kullanırlar? Aynı kişi kaç kez sayılmıştır? Küçük ülkeler görünür müdür? Büyük ülkeler gerçek nüfus ağırlığını taşır mı? Ürünün hiç erişemediği insanlar skorun neresindedir? Hangi insanlar panel aracımız tarafından dışlanmıştır? Hangi sonuç potansiyel erişimi, hangisi gerçek maruziyeti göstermektedir? Bir ülke çok büyükse, ürün-uygun nüfusu da büyük olabilir. Bu durumda ana panelde yüksek ağırlık almalıdır. Fakat ürün orada erişilebilir değilse sahte kullanıcı oluşturulmamalıdır. Erişim açığı ayrıca gösterilmelidir.

Bir ülke çok küçükse ana nüfus ağırlığı düşük kalmalıdır. Fakat tamamen görünmez bırakılmak zorunda değildir. Gözcü paneliyle gözlenebilir. Bir dil küçük nüfusluysa küresel ortalamayı belirlemeyebilir. Fakat o dilde ağır temsil bozulması varsa ayrıca görünür olmalıdır. Bir kişi üç dil biliyorsa üç farklı deneyim üretebilir. Fakat üç farklı insan değildir. Bir AI ürünü yalnız ücretli planda belirli davranış sunuyorsa ücretsiz kullanıcıların skoru bu planı temsil etmez. Bir tüketici sohbet ürünü için insan nüfus paneli uygundur. Bir API veya kurumsal ürün için aynı nüfus modeli uygun olmayabilir. Bu nedenle NOMOS'un beşinci ölçüm yasası şudur:

> *Payda yanlışsa en hassas skor bile yanlış soruya doğru cevap verir.*

Altıncı yasası şudur:

> *Nüfus ağırlığı, ülke ve dil görünürlüğünün yerine geçmez.*

Yedinci yasası şudur:

> *Erişemeyen insan yanlış temsil görmüş değildir; fakat küresel erişim sınırının parçasıdır.*

Sekizinci yasası şudur:

> *Bir insanın birden fazla hesabı, dili veya oturumu insan nüfusunu büyütmez.*

Dokuzuncu yasası şudur:

> *Küresel ölçüm, herkesin ölçüldüğü iddiası değil; kimin ölçülüp kimin ölçülmediğinin dürüst kaydıdır.*

NOMOS'un Bölüm 5 Emri

> *Bana yalnız ülke nüfuslarını verme. / Hangi insanların ürünü gerçekten kullanabildiğini göster.*
>
> *Büyük ülkeyi ürün orada yokken sahte kullanıcılarla doldurma.*
>
> *Küçük ülkeye bir kişi verip o kişinin cevabını ülkenin sesi ilan etme.*
>
> *Bir kişiyi hesapları, dilleri ve oturumları kadar çoğaltma.*
>
> *Resmî dili insanların gerçek kullanım dili sanma.*
>
> *Ücretsiz kullanıcıyı ücretli planın, mobil kullanıcıyı web yüzeyinin temsilcisi yapma.*
>
> *Gönüllü paneline rastgele dünya nüfusu deme.*
>
> *Erişilebilirlik ihtiyacı bulunan kullanıcıyı aracın çalışmadığı için nüfustan silme.*
>
> *Ürünün ulaşmadığı nüfusu doğruluk skorundan çıkarıp küresel erişim açığını saklama.*
>
> *Kimlik sorusuyla fiyat sorusuna, küresel markayla yerel lisansa aynı paydayı uygulama.*
>
> *Önce ürünü tanımla. / Sonra soruyu tanımla. / Sonra insan evrenini tanımla. / Sonra erişimi, dili, yaşı ve arayüzü sınırla. / Sonra kimi seçemediğini açıkla. / Ve ancak bundan sonra 1.000 kişiyi dağıt.*

Bölümün Son Cümlesi

> *GEO-1000'de adalet, her ülkeye aynı sayıda insan vermek değil; her insanı doğru hedef nüfusta, doğru ağırlıkla ve ölçülmediği yerleri saklamadan temsil etmektir.*

Normatif Öz

> **Her GEO nüfus denetimi, sonuçlar görülmeden önce her AI ürünü ve kullanıcı yüzeyi, ülke veya yargı alanı, yaş ve meşru erişim koşulu, hesap veya plan, dil ve locale, arayüz, istem ailesi ve ölçüm amacı için zamana bağlı ürün-özel hedef kullanıcı nüfusu tanımlamalıdır. Toplam yerleşik nüfus, gerekçeli uygunluk ayarlamaları olmadan ürün-uygun nüfus sayılamaz. Ürün erişiminin bulunmaması yanlış temsil olarak puanlanmamalı; erişim kapsamı sınırlaması olarak ayrıca gösterilmelidir. Kişi, hesap, oturum, dil ve ürün gözlemi birbirinin yerine kullanılamaz. Gözcü gözlemleri ülke görünürlüğü sağlayabilir; istatistiksel ülke skoru olarak sunulamaz.**

BÖLÜM 06

Nüfusa Orantılı GEO-1000 Örneklemi

YÖNTEMSEL DAYANAKLAR / K04; K05; K07; K08

Bölüm Sınırı

Bölüm 5 şu hükmü kurdu:

> *Her AI ürününün hedef kullanıcı nüfusu ayrı tanımlanmalıdır. Bir ülkenin toplam nüfusu, o ülkedeki ürün-uygun kullanıcı nüfusu değildir.*

Şimdi bir sonraki soruya geliyoruz:

> *Tanımlanmış uygun kullanıcı nüfusundan GEO-1000 için kaç kişi, hangi ülkeden, hangi dilden ve hangi yöntemle seçilecektir?*

Bir AI ürünü için uygun kullanıcı nüfusunun:

- yüzde 40'ı A ülkesinde,
- yüzde 25'i B ülkesinde,
- yüzde 15'i C ülkesinde,
- yüzde 0,04'ü küçük bir ülkede

bulunuyorsa 1.000 geçerli gözlem nasıl dağıtılmalıdır? Küsuratlı kotalar nasıl tam sayıya dönüştürülecektir? Nüfus payı bir kişinin altında kalan ülkeler tamamen dışarıda mı kalacaktır? Her ülkeye en az bir kişi verildiğinde nüfus ağırlığı bozulacak mıdır? Küçük ülkeler ayrı bir havuzda toplanabilir mi? Dil adaleti için bazı diller fazla örneklenirse küresel skor nasıl hesaplanacaktır? Bir ülkede hedef 100 geçerli gözlem varsa, kaç kişi davet edilmelidir? Katılımcı görevi tamamlamadığında yerine kim alınacaktır? Yanlış cevap veren kullanıcı değiştirilebilir mi? Aynı kişi birden fazla AI ürününü test ederse örneklem hâlâ bağımsız mıdır?

Ham örneklem 1.000 kişi olsa da ağırlıklar çok dengesizse gerçek istatistiksel bilgi 1.000 kişiden daha az olabilir mi? Bu sorular yalnız operasyon ayrıntısı değildir. Örneklem tasarımı, hangi insanın GEO skorunda ne kadar söz sahibi olduğunu belirler. Yanlış örneklem:

- doğru cevapları yanlış ağırlıklandırabilir,
- küçük ülkeleri sembolik biçimde aşırı büyütebilir,
- büyük nüfusları olduğundan küçük gösterebilir,
- düşük kaynaklı dilleri görünmez bırakabilir,
- gönüllü teknoloji kullanıcılarını dünya nüfusu gibi sunabilir,

sonuçlara bakıldıktan sonra istenen skoru üretmek için değiştirilebilir. Daha önce kurulan denetim mimarisi, her kontrolün ölçülebilir olmasını ve model/ürün, tarih, ülke, dil, sorgu kümesi, tekrar sayısı ve ölçüm kaydı taşımasını gerektirir. Bu bölüm, aynı disiplini örneklem tasarımına uygular. Çünkü

gerçek kanıt, sınır, bağlam ve zaman istemek; örneklemin nasıl oluşturulduğunu da görünür kılmayı gerektirir. Bu bölüm:

- ana nüfus panelinin 1.000 geçerli gözleme nasıl dağıtılacağını,
- tabaka ve hücrelerin nasıl kurulacağını,
- küçük ülkelerin nasıl yönetileceğini,
- küsuratların nasıl tam sayıya çevrileceğini,
- fazla örneklemenin nasıl ağırlıklandırılacağını,
- davet, yedek ve geçerli gözlem hedeflerinin nasıl ayrılacağını,
- tasarım ve analiz ağırlıklarının nasıl kurulacağını,
- etkin örneklem büyüklüğünün neden ham sayıdan farklı olabileceğini

tanımlar. Bu bölüm henüz:

- katılımcıların hangi panel sağlayıcısından seçileceğini,
- gönüllü veya olasılıklı örneklemde saha uygulamasının bütün ayrıntılarını,
- hakemlik ve yanıt puanlamasını,
- nihai güven aralığı yöntemini,
- bütün NOMOS skor formüllerini

tam olarak kesinleştirmez. Bölüm 6'nın görevi şudur:

> *Tanımlanmış hedef nüfusu, sonucu görmeden önce dondurulmuş, ağırlıklandırılabilir ve yeniden üretilebilir bir GEO-1000 örneklemine dönüştürmek.*

NOMOS Challenge

Bir AI ürünü için 1.000 kişilik küresel panel kuruyorsunuz. On ülkenin ürün-uygun nüfusları şöyledir:

Ülke	Uygun kullanıcı nüfusu
A	400.000.000
B	250.000.000
C	150.000.000
D	100.000.000
E	60.000.000
F	30.000.000
G	8.000.000
H	1.500.000
I	400.000
J	100.000
Toplam	1.000.000.000

Saf nüfus oranına göre ham kotalar:

Ülke	Ham kota
A	400,0
B	250,0
C	150,0

Ülke	Ham kota
D	100,0
E	60,0
F	30,0
G	8,0
H	1,5
I	0,4
J	0,1

Bir insanın yarısı seçilemez. Bir ülkeye 0,4 kişi gönderilemez. Ham kotaları en yakın tam sayıya yuvarlarsanız toplam 1.000'i aşabilir veya altında kalabilir. Her ülkeye en az bir kişi verirseniz I ve J ülkelerini nüfus paylarından daha fazla temsil edersiniz.

I ülkesine bir kişi vermek, teorik payını 2,5 kat büyütür. J ülkesine bir kişi vermek, teorik payını 10 kat büyütür. Bu aşırı temsil ağırlıkla düzeltilebilir. Fakat bir kişiyle ülke skoru üretilemez. I ve J'ye hiç kimse vermezseniz ne olur?

Eğer bu ülkeleri ayrı ve zorunlu tabaka olarak tanımladıysanız, hedef nüfustaki bazı insanların seçilme olasılığı sıfır olur. Bu durumda: "Bütün uygun küresel nüfusu temsil eden olasılıklı örneklem" iddiası zayıflar. Şimdi daha tehlikeli bir karar düşünün. Önce cevapları topluyorsunuz.

J ülkesindeki tek kullanıcı yanlış cevap alıyor. Sonra: "Bu ülkenin örneklemi çok küçük, kapsamdan çıkaralım." diyorsunuz. A ülkesinde çok sayıda olumlu cevap var. Onu tam ağırlığıyla koruyorsunuz. Bu artık yuvarlama değildir. Sonuç odaklı örneklem değişikliğidir.

Başka bir durumda düşük sonuç veren bir dil grubundaki geçersiz kayıt oranını yükseltiyor, yedek katılımcıları yalnız olumlu cevap gelene kadar kullanıyorsunuz. Bu da örneklem değildir. Cevap seçme mekanizmasıdır. Bu bölümün ilk hükmü şudur:

> *Örneklem, AI yanıtları görülmeden önce kurulmalıdır.*

İkinci hükmü şudur:

> *Katılımcının seçilme ihtimali yanıtın doğruluğuna bağlı olamaz.*

Üçüncü hükmü şudur:

> *Küçük ülkenin görünürlüğü ile küresel skordaki nüfus ağırlığı ayrı yönetilmelidir.*

Dördüncü hükmü şudur:

> *Fazla örneklemek mümkündür; fazla ağırlıklandırmak zorunlu değildir.*

Beşinci hükmü şudur:

> *1.000 ham kayıt, her zaman 1.000 gözlemlik istatistiksel bilgi taşımaz.*

1. BÖLÜMÜN AMACI

Bu bölümün amacı, her AI ürünü ve ölçüm dalgası için hedeflenen 1.000 geçerli gözlemi ürün-uygun kullanıcı nüfusuna dayalı, denetlenebilir bir örneklem tasarımına dönüştürmektir. Bölüm şu ayrımları normatif hâle getirir:

- Hedef nüfus ile seçilen örneklem
- Ham kota ile tam sayı tahsisi

- Ana nüfus paneli ile tamamlayıcı paneller
- Ülke görünürlüğü ile ülke ağırlığı
- Tabaka ile deneysel hücre
- Büyük ülke tabakası ile küçük ülke havuzu
- Saf orantılı tahsis ile orantısız fazla örnekleme
- Örneklem payı ile nüfus payı
- Davet sayısı ile geçerli gözlem hedefi
- Yedek katılımcı ile sonuç sonrası değiştirme
- Veri toplama geçersizliği ile olumsuz AI çıktısı
- Tasarım ağırlığı ile sonuç ağırlığı
- Ham örneklem büyüklüğü ile etkin örneklem büyüklüğü
- Ülke örneklemi ile dil örneklemi
- Çok dilli kullanıcı ile birden fazla insan
- Tek aşamalı seçim ile çok aşamalı seçim
- Native Reach tahsisi ile Common Support tahsisi
- Anlık dalga tahsisi ile zaman serisi tahsisi
- Nüfus tahmini ile ülke veya dil tanı analizi
- Ana 1.000 gözlem ile ek gözcü ve adalet panelleri

Bu bölümün sonunda her denetim şu sorulara cevap verebilmelidir:

> *1.000 gözlem neden bu ülke, dil ve kullanıcı gruplarına bu sayılarda dağıtıldı?*
>
> *Her gözlemin küresel sonuç içindeki ağırlığı nasıl belirlendi?*
>
> *Kimlerin seçilme olasılığı sıfırdı ve neden?*
>
> *Kaç kişi davet edildi, kaç kişi tamamladı ve kaç gözlem geçerli kaldı?*
>
> *Örneklem sonuçlar görülmeden önce dondurulduu mu?*

2. MERKEZÎ NORMATİF HÜKÜM

GEO-1000 ana nüfus örneklemi, her AI ürünü ve ölçüm dalgası için önceden dondurulmuş ürün-uygun nüfus çerçevesindeki insanlara, bilinen veya tahmin edilebilir seçilme olasılıkları atayan tabakalı bir örneklem tasarımına dayanmalıdır. Ana nüfus panelinin varsayılan hedefi:

> *AI ürünü ve ana ölçüm dalgası başına en az 1.000 geçerli gözlem*

olarak korunur. Bu sayı:

- davet edilen kişi sayısı değildir,
- hesap sayısı değildir,
- ham ekran görüntüsü sayısı değildir,

bütün tamamlayıcı panellerin toplamı değildir. 1.000, ana nüfus panelinde protokolü karşılayan geçerli ürün-gözlemi hedefidir. Country Sentinel, Language Fairness ve Accessibility panelleri ek örnekler oluşturabilir. Bu ek örnekler:

- toplam saha hacmini 1.000'in üzerine çıkarabilir,
- küresel nüfus sonucunda uygun ağırlıklarla kullanılabilir,

fakat ana panelin 1.000 geçerli gözlem hedefini gizlice azaltmak için kullanılamaz.

3. ÖRNEKLEMİN TEMEL GÖSTERİMİ

Belirli bir AI ürünü a, ölçüm dalgası w ve hedef nüfus çerçevesi için tabakalar:

```
h ∈ H_{a,w}
```

olarak gösterilsin. Her tabaka için:

- Nh: tabakadaki tahmini ürün-uygun insan nüfusu
- N: bütün hedef nüfus
- Wh: tabakanın nüfus ağırlığı
- n: hedef geçerli ana panel gözlemi
- qh: ham orantılı kota
- nh: tahsis edilen tam sayı geçerli gözlem hedefi

olsun. Toplam hedef nüfus:

```
N = Σ_{h=1}^H N_h
```

Tabakanın nüfus ağırlığı:

```
W_h = N_h / N
```

Ham orantılı kota:

```
q_h = nW_h
```

şeklindedir. Ana orantılı tahsis, ideal koşullarda:

```
n_h ≈ nW_h
```

olmalıdır. Ancak:

- qh küsuratlı olabilir,
- bazı hücreler bir kişinin altında kalabilir,
- dil ve ülke alt grupları için asgari tanı örneği gerekebilir,
- yüksek riskli küçük gruplar ayrıca incelenebilir,

saha tamamlama oranları değişebilir. Bu nedenle ham kota doğrudan son tahsis değildir.

4. ÖRNEKLEM TASARIMININ ÜÇ AYRI AMACI

Aynı örneklem bütün amaçları aynı güçte yerine getiremeyebilir. Üç amaç ayrılmalıdır.

4.1. Küresel Nüfus Tahmini

Şu soruyu cevaplar:

> *Ürüne uygun küresel kullanıcı nüfusundaki ortalama insan hangi temsil sonucunu görür?*

Nüfus ağırlığı temel önceliktir.

4.2. Alt Grup Tanısı

Şu soruyu cevaplar:

> *Belirli ülke, dil, plan veya kullanıcı gruplarında hangi temsil bozulmaları oluşur?*

Küçük grupların fazla örneklenmesi gerekebilir.

4.3. Coğrafi ve Dilsel Kapsam

Şu soruyu cevaplar:

> *Hangi ülke ve dillerde en azından gerçek kullanıcı gözlemi bulunmaktadır?*

Country Sentinel ve Language Fairness panelleri önem kazanır. Bu üç amaç aynı tek örneklemde açıklamasız biçimde birleştirilemez.

5. ANA ÖRNEKLEM MİMARİSİ

GEO-1000 için aday ana örneklem mimarisi dört katmandan oluşur.

5.1. Population Panel

Her AI ürünü ve dalga için en az 1.000 geçerli gözlemlik ana paneldir. Küresel nüfus tahminini üretir.

5.2. Country Sentinel Overlay

Ana nüfus panelinde doğrudan temsil edilemeyen küçük ülkeler için ek coğrafi gözlemler sağlar. Bu gözlemler:

- tekil olay,
- erişim durumu,
- sonraki araştırma ihtiyacı

gösterebilir. Tek başına ülke skoru oluşturmaz.

5.3. Language Fairness Oversample

Düşük kaynaklı veya ana panelde küçük kalan diller için ek gözlemler üretir. Küresel sonuçta gerçek nüfus ağırlıklarına geri döndürülür.

5.4. Accessibility Oversample

Standart panelin yeterince temsil edemediği yardımcı teknoloji ve erişilebilirlik kullanıcılarını ayrıca ölçer. Bu gözlemlerin hedef evreni ve ağırlığı ayrı kaydedilir.

6. TABAKA NEDİR?

Tabaka, örneklem tahsisi ve seçim olasılığı bakımından ayrı yönetilen nüfus bölümüdür. Tabaka şu alanlardan biri veya birkaçının birleşimi olabilir:

- Ülke
- Bölge
- Dil
- Locale
- AI planı
- Web veya mobil
- Kontrollü veya doğal kullanıcı paneli
- Erişilebilirlik durumu
- Kullanıcı türü
- Yargı alanı

Her küçük kombinasyon ayrı tabaka yapılırsa örneklem parçalanır. Örneğin: Ülke × dil × plan × mobil/web × yaş grubu × kullanıcı tipi kombinasyonu yüzlerce veya binlerce hücre üretebilir. 1.000 gözlem bu hücrelerin çoğunda güvenilir tahmin sağlamaz. Bu nedenle tabakalama:

- maddi,
- ölçülebilir,
- örneklem büyüklüğüne uygun

olmalıdır.

7. TABAKA İLE RAPORLAMA HÜCRESİ AYNI ŞEY DEĞİLDİR

Bir grup sonuç kartında ayrı gösterilebilir. Fakat örneklem tasarımında ayrı tabaka olmak zorunda değildir. Örneğin:

- küçük bir dil,
- belirli erişilebilirlik grubu,
- bir yaş bandı

ana örneklemde daha geniş tabaka içinde seçilebilir. Daha sonra analizde ayrı raporlanabilir. Bunun tersi de mümkündür: Örnekleme verimliliği için iki bölge ayrı tabakalanmış olabilir. Fakat kamu raporunda tek bölgesel sonuç altında birleştirilebilir. Tabaka tasarımı ile raporlama yapısı ayrı kayıtlar taşımalıdır.

8. TABAKALARIN KURULMA SIRASI

Ana Population Panel için aday sıralama şöyledir:

- AI ürünü ve kullanıcı yüzeyi
- Native Reach veya Common Support nüfus çerçevesi
- Ülke veya küçük ülke havuzu
- Ana denetim dili veya dil grubu
- Maddi plan/arayüz ayrımı, gerekiyorsa
- Katılımcı seçim çerçevesi

Plan ve arayüz her ülkede ayrı tabaka yapılmak zorunda değildir. Bunlar:

- alt tabaka,
- kota,
- rastgele atama

yoluyla yönetilebilir.

9. BÜYÜK VE KÜÇÜK ÜLKE TABAKALARI

Bir ülkenin ham kotası:

```
q_c = nN_c / N
```

şeklindedir. Ham kota yeterince büyükse ülke doğrudan ayrı tabaka olarak tutulabilir. Buna:

Kendini temsil eden ülke tabakası

denebilir. Ham kotası çok küçük ülkeleri tek tek ayrı zorunlu tabaka yapmak örneklem verimliliğini bozabilir. Bu ülkeler uygun biçimde:

- bölge,
- dil ailesi,
- ürün erişim rejimi,
- coğrafi yakınlık

altında küçük ülke havuzlarına alınabilir.

10. KÜÇÜK ÜLKE HAVUZU

Bir küçük ülke havuzu:

```
B_r = {c_1, c_2, …, c_k}
```

olarak gösterilsin. Havuzun toplam uygun nüfusu:

```
N_{B_r} = Σ_{c∈B_r} N_c
```

Ana örneklem kotası:

```
q_{B_r} = nN_{B_r} / N
```

olarak hesaplanır. Havuz içinde ülke seçimi, uygun nüfus büyüklüğüne orantılı olasılıkla yapılabilir. Ülke c'nin havuz içinde seçilme olasılığı:

```
Pr(c | B_r) = N_c / N_{B_r}
```

olabilir. Ardından seçilen ülke içinde katılımcı seçilir. Bu yaklaşım:

- küçük ülkelerin hedef nüfusta sıfır seçilme ihtimali taşımasını engelleyebilir,
- her ülkenin belirli bir dalgada mutlaka seçileceğini garanti etmez,

küresel nüfus tahminini korur. Bir ülkenin o dalgada seçilmemesi:

NOT SELECTED IN POPULATION SAMPLE

statüsüdür. Şu anlama gelmez: "Ülke hedef nüfusun parçası değildir."

11. KÜÇÜK ÜLKELERİ HAVUZLAMANIN SINIRLARI

Küçük ülkeler yalnız nüfusları küçük olduğu için rastgele aynı havuza atılamaz. Havuz içindeki ülkeler ilgili olduğu ölçüde benzer:

- ürün erişimi,
- kullanıcı yüzeyi,
- dil veya locale yapısı,
- hukuki erişim,
- coğrafi bölge,
- veri toplama yöntemi

taşımalıdır. Ürün erişimi bulunmayan ülke ile tam erişim bulunan ülke aynı ana havuza konamaz. Farklı kullanım koşulları görünmez hâle getirilemez.

12. KÜÇÜK ÜLKE HAVUZU ÜLKE SKORU ÜRETMEZ

Havuzdan seçilen bir veya iki gözlem: havuzun küresel nüfus katkısında kullanılabilir, tekil ülke olayını gösterebilir. Fakat küçük ülke için ayrı istatistiksel skor üretmez. Ülke bazlı skor isteniyorsa ek:

- Country Sentinel,
- ülke özel örneklem,
- çok dalgalı birikimli örneklem

gerekebilir.

13. SAF NÜFUSA ORANTILI TAHSİS

Saf orantılı tahsis:

```
q_h = nW_h
```

formülüne dayanır. Avantajları:

- ana nüfus tahmini için anlaşılırdır,
- büyük nüfuslara doğal ağırlık verir,
- ağırlık farklılığını azaltabilir,

kendi kendini ağırlıklandıran tasarıma yaklaşabilir. Sınırlamaları:

- küçük ülke ve dilleri çok az örnekleyebilir,
- alt grup skoru üretmeye yetmeyebilir,
- yüksek riskli küçük grupları görünmez bırakabilir,

küsurat ve sıfıra yakın kotalar oluşturabilir. Bu nedenle saf orantılı tahsis ana Population Panel için güçlü varsayılan olabilir. Tamamlayıcı panellerin yerine geçmez.

14. KÜSURATLARIN TAM SAYIYA DÖNÜŞTÜRÜLMESİ

Ham kotalar çoğu zaman küsuratlıdır. Toplam tahsis tam olarak n olmalıdır. Varsayılan yöntem olarak En Büyük Kalan Yöntemi kullanılabilir.

14.1. Birinci Adım

Her ham kota aşağı yuvarlanır:

```
b_h = ⌊q_h⌋
```

14.2. İkinci Adım

Dağıtılmamış gözlem sayısı:

```
R = n − Σ_h b_h
```

olarak hesaplanır.

14.3. Üçüncü Adım

Her tabakanın kesir kısmı:

```
f_h = q_h − b_h
```

hesaplanır.

14.4. Dördüncü Adım

Kalan R gözlem, en yüksek fh değerlerine sahip tabakalara birer tane eklenir.

14.5. Eşitlik Durumu

İki tabakanın kesir kısmı eşitse:

- nüfus büyüklüğü,
- önceden ilan edilmiş sabit sıra,
- kayıtlı rastgelelik tohumu

gibi sonuçtan bağımsız bir bağ bozma kuralı kullanılmalıdır. Sonuçlara bakılarak ülke seçilemez.

15. SENTETİK TAM SAYI TAHSİSİ

SENTETİK GÖSTERİM

Bölüm başındaki uygun nüfuslar:

Ülke	Uygun nüfus	Ham kota
A	400.000.000	400,0
B	250.000.000	250,0
C	150.000.000	150,0
D	100.000.000	100,0
E	60.000.000	60,0
F	30.000.000	30,0
G	8.000.000	8,0
H	1.500.000	1,5
I	400.000	0,4
J	100.000	0,1

Ülkeler ayrı tabakalar olarak korunursa aşağı yuvarlama:

Ülke	Aşağı yuvarlanmış kota
A	400
B	250
C	150
D	100
E	60
F	30
G	8
H	1
I	0
J	0
Toplam	999

Bir gözlem kalmıştır. En yüksek kesir H ülkesindeki 0,5 olduğu için H'ye eklenir. Nihai tahsis:

Ülke	Nihai kota
A	400
B	250
C	150
D	100
E	60
F	30
G	8
H	2
I	0
J	0
Toplam	1.000

Bu tasarım I ve J'yi ayrı deterministik tabaka olarak tanımladığı hâlde seçilme şansı vermiyorsa tam hedef evren çıkarımı bakımından sorun yaratabilir. Daha doğru çözüm küçük ülke havuzudur.

16. KÜÇÜK ÜLKE HAVUZLU SENTETİK TAHSİS

H, I ve J ülkeleri:

```
B_small = {H, I, J}
```

havuzunda birleştirilsin. Havuz nüfusu:

```
1.500.000+400.000+100.000=2.000.000
```

Havuz kotası:

```
1.000 × 2.000.000 / 1.000.000.000 = 2
```

olur. Ana tahsis:

Tabaka	Kota
A	400
B	250
C	150
D	100
E	60
F	30
G	8
Küçük ülke havuzu	2
Toplam	1.000

Havuz içindeki ülke seçilme ağırlıkları:

Ülke	Havuz içi nüfus payı
H	%75
I	%20

Ülke	Havuz içi nüfus payı
J	%5

İki ana nüfus gözlemi bu olasılıklara göre seçilir. Aynı dalgada:

- iki gözlem de H'den,
- biri H'den biri I'den,
- nadiren J'den

gelebilir. Bu tasarım J'yi her dalgada görünür yapmaz. Fakat J nüfusunun seçilme olasılığını sıfırlamaz. J'nin mutlaka gözlenmesi isteniyorsa Country Sentinel Panel'e ayrıca eklenir. Sentinel gözlemi küresel ana tahsiste J'ye 1.000'de bir ağırlık vermez. J'nin gerçek nüfus ağırlığı korunur.

17. ASGARİ HÜCRE TAHSİSİ

Bazı alt gruplar için saf orantılı kota tanısal analiz yapmaya yetmeyebilir. Örneğin bir dil hücresinin ham kotası 8 olabilir. Sekiz gözlemle güvenilir dil skoru üretilemez. Bu durumda üç seçenek vardır:

- Dil skoru yayımlamamak
- Hücreyi benzer gruplarla birleştirmek
- Hücreyi fazla örneklemek

Asgari hücre tahsisi sonuçlardan önce tanımlanmalıdır.

18. ADAY RAPORLAMA BANTLARI

CANDIDATE REPORTING BANDS — ADAY RAPORLAMA BANTLARI

Aşağıdaki eşikler nihai bilimsel karar değildir. Pilot ve dış incelemeyle kalibre edilmelidir.

1–29 Geçerli Gözlem

Tekil olay ve keşif Ülke veya dil oranı yayımlanmaz

OBSERVATIONAL ONLY

30–99 Geçerli Gözlem

Keşif düzeyi tahmin Geniş belirsizlik Kesin uygunluk hükmü verilmez

EXPLORATORY

100–199 Geçerli Gözlem

İlk alt grup tahmini Belirsizlik görünür olmalıdır

PROVISIONAL

200 ve Üzeri Geçerli Gözlem

Daha güçlü alt grup tahmini adayı Tasarım etkisi ve ağırlıklar yine değerlendirilir

STANDARD ESTIMATION CANDIDATE

Bu bantlar:

- hata oranı,

- beklenen hassasiyet,
- ağırlık değişkenliği,
- tasarım etkisi

dikkate alınmadan otomatik yeterlilik sayılmaz.

19. FAZLA ÖRNEKLEME

Bir ülke, dil veya kullanıcı grubu nüfus payından daha fazla gözlem alabilir. Buna fazla örnekleme denir. Amaçları:

- alt grup hassasiyetini artırmak,
- düşük kaynaklı dili görünür kılmak,
- yüksek riskli küçük grubu incelemek,
- Country Sentinel kapsamını genişletmek,
- erişilebilirlik kullanıcılarını yeterince temsil etmek

olabilir. Fazla örnekleme yanlış değildir. Yanlış olan:

> *Fazla örneklenen grubun örneklem payını gerçek nüfus payı gibi kullanmaktır.*

20. NÜFUS PAYI VE ÖRNEKLEM PAYI

Tabaka h için nüfus payı:

```
W_h = N_h / N
```

Örneklem payı:

```
S_h = n_h / n
```

şeklindedir. Saf orantılı tahsiste:

```
S_h ≈ W_h
```

olur. Fazla örneklemede:

```
S_h > W_h
```

olabilir. Az örneklemede:

```
S_h < W_h
```

olabilir. Küresel nüfus tahmini, Sh yerine Wh değerlerini korumalıdır.

21. TEMEL TASARIM AĞIRLIĞI

Basit tabakalı seçimde tabaka h içindeki her gözlemin temel ağırlığı:

```
d_i = N_h / n_h
```

olarak düşünülebilir. Normalize edilmiş tabaka ağırlığı:

```
a_h = W_h / S_h
```

şeklinde yazılabilir. Gözlem düzeyindeki normalize ağırlık:

```
w_i = W_h / n_h
```

olabilir. Toplamı bir olacak biçimde:

```
Σ_i w_i = 1
```

sağlanır. Ağırlıklı ikili geçiş tahmini:

```
θ̂ = Σ_i w_iY_i
```

olarak hesaplanabilir. Burada:

```
Yi=1: geçiş kriterini karşılayan yanıt

Yi=0: karşılamayan yanıt
```

olabilir. Nihai NOMOS sistemi yalnız ikili sonuç kullanmak zorunda değildir. Aynı ağırlık mantığı çok boyutlu yanıt puanlarına da uygulanabilir.

22. FAZLA ÖRNEKLEME ÖRNEĞİ

SENTETİK ÖRNEK

B ülkesinin uygun nüfusunda:

- Dil X: yüzde 80
- Dil Y: yüzde 20

olsun. Population Panel içinde B ülkesine 250 gözlem düşmektedir. Saf tahsis:

- Dil X: 200
- Dil Y: 50

olur. Dil Y için ek 50 Language Fairness gözlemi alınsın. Toplanan gözlemler:

- Dil X: 200
- Dil Y: 100

olur. Ham örneklemde Dil Y payı:

```
100/300 = %33,3
```

olmuştur. Gerçek nüfus payı yüzde 20'dir. Küresel nüfus hesabında Dil Y yüzde 33,3 ağırlık alamaz. Tabaka içi normalize ağırlıklar:

- Dil X için toplam ağırlık: yüzde 80
- Dil Y için toplam ağırlık: yüzde 20

olarak korunur. Dil Y'deki 100 gözlemin her biri, Dil X'teki tek bir gözlemden daha düşük bireysel ağırlık alır. Buna karşılık Dil Y'nin ayrı tanısal raporunda 100 gözlemin tamamı kullanılır. Bu yaklaşım: alt grup görünürlüğünü artırır, küresel nüfus tahminini bozmaz.

23. ÇOK DİLLİ KULLANICILARIN TAHSİSİ

Bir kişi birden fazla dil hücresine uygun olabilir. Ana Population Panel için tercih edilen yaklaşımlardan biri, kullanıcıya önceden tanımlanmış bir ana denetim dili atamaktır. Ana denetim dili şu alanlara göre belirlenebilir:

- Günlük AI kullanım dili
- Ana veya ev dili
- Pazar ve locale
- İstemi doğal biçimde kullanma yeterliliği
- Panel dengesi

Kullanıcı diğer dillerde de test edilecekse bu gözlemler:

- eşlenmiş çok dilli panel,

- Language Fairness Panel

olarak ayrıca etiketlenir. Aynı kişinin bütün dillerde tam nüfus ağırlığı alması yasaktır.

24. ÜLKE İÇİNDE DİL TAHSİSİ

Ülke c içinde dil grupları l için uygun nüfus: Nc,l olsun. Ülke kotası: nc ise ham dil kotası:

```
q_{c,l} = n_c × N_{c,l} / Σ_l N_{c,l}
```

olarak hesaplanabilir. Ancak dil grupları örtüşüyorsa:

```
Σ_l N_{c,l} > N_c
```

olabilir. Bu durumda doğrudan pay kullanmak çifte sayım yaratır. Çözüm seçenekleri:

- ana denetim dili sınıflandırması,
- bölünmüş insan ağırlığı,
- çok dilli ortak model,
- ayrı eşlenmiş dil paneli

olabilir. Kullanılan yöntem kayıt altına alınmalıdır.

25. NÜFUS İÇİNDE PLAN VE ARAYÜZ TAHSİSİ

Ücretsiz ve ücretli kullanıcılar ya da web ve mobil yüzeyler ayrı davranabilir. Ancak her ülke–dil hücresini ayrıca plan ve arayüz tabakasına bölmek örneklemi aşırı parçalayabilir. Üç yaklaşım kullanılabilir.

25.1. Nüfusa Orantılı Alt Tahsis

Gerçek kullanıcı dağılımı biliniyorsa plan ve arayüz paylarına göre tahsis yapılır.

25.2. Dengeli Deneysel Tahsis

Plan veya arayüz farkını karşılaştırmak için eşit veya fazla örnekleme yapılır. Küresel tahminde gerçek kullanım paylarına geri ağırlıklandırılır.

25.3. Ayrı Ürün Yüzeyi Skorları

Ücretsiz web, ücretli web ve mobil ürünler ayrı denetim nesneleri olarak raporlanır. Birleştirilmiş sonuç ayrıca üretilebilir. Hangi yaklaşımın kullanıldığı ölçümden önce tanımlanmalıdır.

26. DAVET SAYISI İLE GEÇERLİ GÖZLEM HEDEFİ

Bir tabakanın geçerli gözlem hedefi nh ise tam olarak nh kişi davet etmek genellikle yeterli olmaz. Çünkü:

- bazı kişiler katılmaz,
- bazıları görevi tamamlamaz,
- bazı kayıtlar protokol dışı kalır,

bazı kanıt paketleri geçersiz olur. Tabaka için beklenen tamamlama oranı: rh ve beklenen geçerlilik oranı: vh olsun. Davet hedefi yaklaşık olarak:

```
m_h = ⌈n_h / (r_hv_h)⌉
```

şeklinde hesaplanabilir. Örnek:

- Geçerli gözlem hedefi: 100
- Tamamlama oranı beklentisi: yüzde 80
- Geçerlilik oranı beklentisi: yüzde 90

```
m_h = ⌈100 / (0,80 × 0,90)⌉ = 139
```

kişi davet edilebilir. Bu bir garanti değildir. Saha ilerledikçe önceden tanımlanmış yedek sistemi kullanılabilir.

27. DAVET ŞİŞİRMESİNİN SINIRI

Daha çok kişiyi davet etmek tek başına örneklem yanlılığını çözmez. Erken yanıt verenler, geç yanıt verenlerden sistematik olarak farklı olabilir. Yalnız ilk tamamlayan 1.000 kişiyi almak:

- saat dilimi,
- teknoloji ilgisi,
- bağlantı kalitesi,
- çalışma düzeni

bakımından yanlılık üretebilir. Davet ve seçim sırası:

- rastgeleleştirilmeli,
- tabaka içinde kayıtlı olmalı,

saha penceresi boyunca yönetilmelidir.

28. YEDEK KATILIMCI SİSTEMİ

Her tabaka için sonuçlar görülmeden önce sıralanmış bir yedek liste oluşturulmalıdır. Yedek katılımcı yalnız şu nedenlerle devreye girebilir:

- Davete cevap verilmemesi
- Görevin tamamlanmaması
- Teknik veya etik uygunluk sorunu
- İstem protokolünün ihlali
- Kanıt paketinin geçersiz olması
- Aynı kişinin tekrar kaydedilmesi

Yedek kullanıcı şu nedenle devreye giremez:

> *Önceki kullanıcının AI cevabı yanlış, olumsuz veya istenmeyen çıktı verdi.*

Yanlış AI cevabı geçerli sistem sonucudur. Katılımcı değişikliği nedeni olamaz.

29. GEÇERLİLİK KARARINDA SONUÇ KÖRLÜĞÜ

Mümkün olduğu ölçüde gözlemin protokole uygun olup olmadığına karar veren kişi:

- yanıtın doğru veya yanlış puanını,
- markaya olumlu veya olumsuz olduğunu

bilmeden değerlendirme yapmalıdır. Önce:

- Capture ve protokol geçerliliği
- Sonra maddi yanıt değerlendirmesi

yapılmalıdır. Aksi hâlde olumsuz cevaplar daha kolay: "geçersiz" ilan edilebilir.

30. TABAKA İÇİNDE YEDEK SIRASI

Yedek katılımcı mümkün olduğunca aynı:

- ülke,
- dil,
- ürün yüzeyi,
- plan,
- panel koşulu

tabakasından alınmalıdır. Aynı tabakada yedek kalmadıysa önceden tanımlı eskalasyon uygulanır:

- Aynı tabaka içinde ek rastgele seçim
- Önceden tanımlanmış üst tabaka havuzu
- Tahsisin eksik bırakılması ve kapsam uyarısı
- Yeni analiz sürümü

Sonuç görüldükten sonra başka ülkeden kullanıcı getirip açığı kapatmak sessizce yapılamaz.

31. TABAKA KISA KALIYORSA NE OLUR?

Bir tabaka hedefinin altında kalabilir. Örnek:

- hedef: 100
- geçerli gözlem: 72

Doğru seçenekler:

- Saha penceresini önceden tanımlı ölçüde uzatmak
- Yedek listeyi kullanmak
- Hücreyi benzer ve önceden tanımlı üst tabakayla birleştirmek
- Sonucu geniş belirsizlikle raporlamak
- Hücre sonucunu yayımlamamak
- Küresel ağırlığı uygun yöntemle yeniden hesaplamak

Yanlış seçenek: Düşük skorlu tabakayı analizden tamamen çıkarmak.

32. SONUÇ SONRASI KOTA DEĞİŞİKLİĞİ

Aşağıdaki değişiklikler sonuç odaklı örneklem manipülasyonudur:

- Düşük skorlu ülkenin hedef kotasını azaltmak
- Yüksek skorlu dilin kotasını artırmak
- Olumsuz cevap veren kullanıcıları geçersiz saymak
- Sonuçlara göre küçük ülke havuzunu yeniden tanımlamak
- Critical hata bulunan planı kapsam dışına almak
- İstenen ortalamaya ulaşana kadar yeni kullanıcı eklemek

Her maddi kota değişikliği:

- yeni tahsis sürümü,

- gerekçe,
- tarih,
- karar sahibi,
- önceki sonucun korunması

ile yapılmalıdır.

33. ÇOK AŞAMALI ÖRNEKLEM

Bütün bireylere doğrudan erişilemediğinde çok aşamalı örneklem kullanılabilir. Örnek:

- Bölge veya ülke havuzu seçimi
- Ülke seçimi
- Panel veya yerel araştırma ortağı seçimi
- Kullanıcı seçimi
- AI ürünü veya görev ataması

Tekil kullanıcının toplam seçilme olasılığı, aşamalardaki koşullu olasılıkların çarpımıdır:

```
π_i = π_r × π_{c|r} × π_{f|c} × π_{i|f}
```

Burada:

- r: bölge veya küçük ülke havuzu
- c: ülke
- f: örnekleme çerçevesi veya panel
- i: kişi

Temel tasarım ağırlığı:

```
d_i = 1/π_i
```

olur. Seçim olasılıkları kaydedilemiyorsa tasarım tam olasılıklı örneklem olarak sunulamaz. [K04; K07]

34. GÖNÜLLÜ VE OLASILIKSIZ PANELDE TAHSİS

Gerçek olasılıklı dünya paneli her zaman mümkün olmayabilir. Gönüllü veya profesyonel çevrim içi panel kullanılabilir. Bu durumda ülke ve dil kotaları yine ürün-uygun nüfusa göre kurulabilir. Ancak:

- katılımcının gerçek seçilme olasılığı bilinmeyebilir,
- kendi kendine seçilim bulunabilir,

ağırlıklandırma yalnız gözlenen değişkenler üzerinden düzeltme yapabilir. Doğru statü:

Tabakalı ve nüfusa kalibre edilmiş olasılıksız kullanıcı paneli

olabilir. Şu ifade kullanılamaz: “Dünya nüfusundan tamamen rastgele seçilmiş 1.000 kişi.”

35. AĞIRLIK MİMARİSİ

Nihai gözlem ağırlığı birden fazla bileşenden oluşabilir:

```
w_i = d_i × a_i^{NR} × g_i^{CAL}
```

Burada:

- di: temel tasarım ağırlığı
- aiNR: yanıtlamama ayarlaması

- giCAL: kalibrasyon katsayısı

35.1. Tasarım Ağırlığı

Katılımcının örnekleme tasarımındaki seçilme ihtimalinin tersidir.

35.2. Yanıtlamama Ayarlaması

Benzer özellikteki seçilmiş kişilerin:

- kaçının yanıt verdiği,
- kaçının geçerli gözlem ürettiği

dikkate alınabilir. Örnek yanıt sınıfı r için:

```
a_r^{NR} = (Σ_{j∈S_r} d_j) / (Σ_{i∈R_r} d_i)
```

Burada:

- Sr: seçilmiş uygun örnek
- Rr: yanıt veren veya geçerli gözlem üretenler

olabilir. Yanıtlamama ayarı: gözlenmeyen yanlılığı tamamen çözmez, varsayımlarıyla birlikte açıklanmalıdır.

35.3. Kalibrasyon Ağırlığı

Örneklem, bilinen nüfus toplamlarına:

- ülke,
- dil,
- yaş,
- plan,
- arayüz

bakımından kalibre edilebilir. Kalibrasyon yalnız güvenilir dış toplamlar varsa yapılmalıdır.

36. AĞIRLIK NORMALİZASYONU

Ağırlıklar:

- toplam nüfusa,
- toplam örneklem sayısına,
- bire

normalize edilebilir. Skor formülünde hangi normalizasyonun kullanıldığı açıklanmalıdır. Bire normalize edilmiş ağırlıklar:

```
w̃_i = w_i / Σ_j w_j
```

olur. Ağırlıklı tahmin:

```
θ̂ = Σ_i w̃_iY_i
```

şeklinde hesaplanır.

37. AĞIRLIK KIRPMA

Bazı katılımcılar çok küçük ve az örneklenen grupları temsil ettiği için çok yüksek ağırlık alabilir. Bu durum:

- tahmin varyansını,
- tek gözlemin etkisini,
- skor oynaklığını

artırabilir. Ağırlıklar üst sınırla kırpılabilir. Fakat kırpma: nüfus temsilinde yanlılık yaratabilir, küçük grupları yeniden görünmezleştirebilir. Bu nedenle ağırlık kırpma:

- sonuçlardan önce tanımlanmalı,
- kırpılmamış sonuçla birlikte gösterilmeli,

duyarlılık analizi taşımalıdır.

38. ETKİN ÖRNEKLEM BÜYÜKLÜĞÜ

Ham gözlem sayısı: n olsa da ağırlıklar çok değişkense gerçek bilgi miktarı daha düşük olabilir. Kish etkin örneklem büyüklüğü:

```
n_eff = (Σ_i w_i)^2 / Σ_i w_i^2 [K08]
```

olarak gösterilebilir. Eşit ağırlıklarda:

```
neff≈n
```

olur. Ağırlıklar çok dengesizse: neff<n olur. Bir GEO-1000 çalışması:

- 1.000 geçerli gözlem,
- 620 etkin örneklem büyüklüğü

taşıyabilir. Kamuya yalnız: “1.000 kişi ölçüldü.” demek hassasiyeti olduğundan güçlü gösterebilir. Ana raporda en azından:

- ham geçerli gözlem sayısı,
- etkin örneklem büyüklüğü,
- ağırlık değişkenliği

bulunmalıdır.

39. AĞIRLIK TASARIM ETKİSİ

Ağırlıklardan kaynaklanan yaklaşık tasarım etkisi:

```
DEFF_w ≈ 1 + CV(w)^2
```

ile gösterilebilir. Burada:

```
CV(w) = sd(w) / mean(w)
```

ağırlıkların değişim katsayısıdır. Bu yalnız ağırlık etkisinin yaklaşık gösterimidir. Toplam tasarım etkisi ayrıca:

- kümelenme,
- tabakalama,
- tekrar ölçümü,
- eşlenmiş kullanıcılar

nedeniyle değişebilir.

40. KÜMELENME

Aynı:

- araştırma panelinden,
- ülkeden,
- dil topluluğundan,
- cihaz veya kurumdan,
- kullanıcıdan

gelen gözlemler birbirine benzeyebilir. Bu durumda 1.000 gözlem 1.000 tamamen bağımsız bilgi birimi değildir. Kümelenme:

- örneklem tasarımında,
- varyans hesabında,
- etkin örneklem büyüklüğünde

dikkate alınmalıdır.

41. AYNI KULLANICININ BİRDEN FAZLA AI ÜRÜNÜNÜ TEST ETMESİ

Bir kullanıcının birden fazla AI ürününü test etmesi modeller arası farkın aynı kişi üzerinde ölçülmesini sağlar. Avantajı: kullanıcı özellikleri kontrol edilebilir, eşlenmiş karşılaştırma güçlenebilir. Sınırlaması:

- ürün gözlemleri bağımsız değildir,
- görev sırası diğer cevapları etkileyebilir,
- kullanıcı ilk sistemden gördüğü cevabı sonraki sisteme taşıyabilir,

bütün ürünlere erişebilen kullanıcılar nüfustan farklı olabilir. Bu nedenle ürünler:

- rastgele sırayla atanmalı,
- önceki cevap paylaşımı engellenmeli,
- eşlenmiş tasarım kaydedilmeli,

analiz bağımlılığı korumalıdır. Her AI ürünü yine kendi 1.000 geçerli ürün-gözlemi hedefine sahip olmalıdır. Aynı 1.000 insan on ürünü test ederse:

- 1.000 benzersiz insan,
- 10.000 ürün-gözlemi

oluşabilir. Bu iki sayı aynı şey değildir.

42. DENGELİ EKSİK BLOK TASARIMI

Bütün kullanıcıların bütün AI ürünlerine erişmesi gerekmez. Örneğin on AI ürünü varsa her katılımcıya üç ürün atanabilir. Atamalar dengeli yapılırsa:

- her ürün 1.000 gözleme,
- her ürün çifti yeterli ortak kullanıcıya,
- daha geniş nüfus katılımına

ulaşabilir. Bu yaklaşıma aday olarak:

Dengeli eksik blok tasarımı

kullanılabilir. Avantajları:

- katılımcı yükünü azaltır,
- yalnız yoğun AI kullanıcılarını seçme riskini düşürür,

modeller arası eşlenmiş karşılaştırma sağlar. Kesin tasarım Bölüm 8 ve Bölüm 9'da ayrıntılandırılacaktır.

43. NATIVE REACH TAHSİSİ

Her AI ürünü kendi Native Reach nüfusuna göre ayrı tahsis alır. Bu nedenle aynı ülke:

- bir ürün için 200,
- başka ürün için 0,
- üçüncü ürün için 50

gözlem alabilir. Bu farklılık adaletsizlik değildir. Ürünlerin gerçek erişim evrenlerinin farklı olmasıdır. Native Reach skorları farklı nüfuslara ait olduğundan doğrudan sıralama için tek başına kullanılmamalıdır.

44. COMMON SUPPORT TAHSİSİ

Ürünler arası karşılaştırma için ortak erişim evreni oluşturulur. Ortak nüfus:

```
U_CS = ∩_{a∈A} U_a
```

şeklinde düşünülebilir. Her ürün aynı Common Support nüfus ağırlıklarıyla ölçülür. Bu yaklaşım:

- adil ürün karşılaştırması,
- aynı ülke ve dil tabanı

sağlar. Fakat ürünlerin doğal küresel erişim farkını göstermez. Bu nedenle Native Reach ve Common Support sonuçları birlikte raporlanmalıdır.

45. ÖRNEKLEMİN DALGALAR ARASINDA YÖNETİMİ

Birden fazla ölçüm dalgasında iki farklı hedef bulunabilir.

45.1. Güncel Kesit Sonucu

Her dalgada güncel ürün-uygun nüfus çerçevesi kullanılır. Bu skor: "Şu anda gerçek küresel erişim evreninde ne oluyor?" sorusuna cevap verir.

45.2. Sabit Çerçeveli Eğilim Sonucu

Başlangıç dalgasındaki nüfus ağırlıkları korunur. Bu skor: "Aynı referans nüfusta temsil zaman içinde nasıl değişti?" sorusuna cevap verir.

45.3. İki Sonucun Ayrılması

Ürün yeni ülkelere açıldığında güncel skor düşebilir. Bu düşüş:

- ürünün kötüleşmesinden,
- yeni ve zor nüfusların kapsama girmesinden

kaynaklanabilir. Sabit çerçeveli trend aynı anda iyileşme gösterebilir. Bu nedenle:

- Güncel-Frame Snapshot
- Fixed-Frame Trend

ayrı sonuçlardır.

46. PANEL YENİLEME VE DÖNÜŞÜM

Dalgalar arasında:

- aynı kullanıcılar,
- tamamen yeni kullanıcılar,
- kısmen yenilenen panel

kullanılabilir.

46.1. Tam Kesitsel Panel

Her dalgada yeni kullanıcılar seçilir. Nüfus tahmini için güçlü olabilir. Bireysel değişimi ölçmez.

46.2. Boylamsal Panel

Aynı kullanıcılar tekrar ölçülür. Bireysel kullanıcı ve ürün değişimini gösterir. Panel yıpranması ve öğrenme etkisi oluşturabilir.

46.3. Dönen Panel

Katılımcıların bir bölümü korunur, bir bölümü yenilenir. Hem trend hem taze nüfus temsili sağlar. Kullanılan panel yapısı sonuç kartında görünür olmalıdır.

47. ÖRNEKLEM KİLİDİ

Her dalga başlamadan önce aşağıdaki kayıtlar kilitlenmelidir:

- AI ürünü
- Kullanıcı yüzeyi
- Hedef nüfus sürümü
- Native Reach ülkeleri
- Common Support ülkeleri
- Tabakalar
- Küçük ülke havuzları
- Ham kotalar
- Tam sayı tahsisleri
- Dil tahsisleri
- Fazla örnekleme planı
- Davet hedefleri
- Yedek listeleri
- Seçim yöntemi
- Ağırlık formülü
- Sonuç körlüğü kuralı
- Kısa kalma eskalasyonu

- Analiz planı
- Bağ bozma yöntemi
- Rastgelelik tohumu, kullanılıyorsa
- Sorumlu insan onayı

Bu dosyaya:

NOMOS Sample Allocation Lock

adı verilebilir.

48. ÖRNEKLEM KİLİDİNDEN SONRA İZİN VERİLEN DEĞİŞİKLİKLER

Aşağıdaki olaylar yeni sürüm gerektirebilir:

- Ürünün ülkede erişiminin sona ermesi
- Ciddi saha güvenliği sorunu
- Nüfus verisinde maddi hata bulunması
- Panel sağlayıcının belirli ülkeyi teslim edememesi
- Araştırma etik şartının değişmesi
- Ölçüm aracında kritik hata

Değişiklik:

- gerekçeli,
- tarihli,
- onaylı,
- eski sürüm korunarak

yapılmalıdır. Sonuçların güçlü veya zayıf olması değişiklik gerekçesi değildir.

49. ÖRNEKLEM KALİTE STATÜLERİ

Bir örneklem dalgası şu statülerden birini taşıyabilir:

SQ-0 — NOT DESIGNED

Belgelenmiş örneklem tasarımı yoktur.

SQ-1 — CONVENIENCE OBSERVATION

Kolaylık veya gönüllü gözlem vardır. Nüfus genellemesi yapılamaz.

SQ-2 — QUOTA-CONTROLLED

Ülke ve dil kotaları vardır. Seçilme olasılıkları tam bilinmez.

SQ-3 — WEIGHTED PANEL

Tabakalı panel ve nüfus kalibrasyonu vardır. Kapsama ve kendi kendine seçilim sınırlamaları devam eder.

SQ-4 — PROBABILITY-BASED OR HYBRID

Bilinen seçim olasılıkları veya güçlü hibrit tasarım bulunmaktadır.

SQ-5 — REPLICATED AND CALIBRATED

Tasarım farklı dalga ve dış uygulamalarda tekrar edilmiş; yanlılık ve varyans davranışı kalibre edilmiştir. Bu statüler kamuya açık skorun güven sınıfını etkileyebilir.

50. SENTETİK UÇTAN UCA GEO-1000 TAHSİSİ

SENTETİK METODOLOJİ GÖSTERİMİ / Aşağıdaki ülkeler, nüfuslar, diller, AI ürünü ve hesaplamalar tamamen kurgusaldır. Sentetik Orion AI ürünü için Native Reach uygun nüfusu 1 milyar kişi olsun. Ana Population Panel hedefi:

```
n=1.000
```

50.1. Ülke ve Küçük Ülke Havuzu Tahsisi

Tabaka	Uygun nüfus	Nüfus payı	Ana kota
A	400.000.000	%40	400
B	250.000.000	%25	250
C	150.000.000	%15	150
D	100.000.000	%10	100
E	60.000.000	%6	60
F	30.000.000	%3	30
G	8.000.000	%0,8	8
Küçük ülke havuzu	2.000.000	%0,2	2
Toplam	1.000.000.000	%100	1.000

50.2. B Ülkesinde Dil Tahsisi

B ülkesindeki uygun nüfus:

- Dil X: yüzde 80
- Dil Y: yüzde 20

Population Panel tahsisi:

- Dil X: 200
- Dil Y: 50

Language Fairness ek tahsisi: Dil Y: 50 B ülkesindeki saha hedefi:

- Dil X: 200
- Dil Y: 100
- Toplam: 300

Fakat küresel Population Panel nüfus ağırlıkları:

- Dil X: B ülkesinin yüzde 80'i
- Dil Y: B ülkesinin yüzde 20'si

olarak kalır.

50.3. Country Sentinel Eki

Küçük ülke havuzundaki H, I ve J için Population Panel'de iki olasılıklı gözlem bulunmaktadır. Country Sentinel Panel ayrıca:

H: 1

I: 1

J: 1

gözlem hedefleyebilir. Bu üç gözlem:

- coğrafi kapsamı artırır,
- tekil ülke olaylarını gösterir,

ülke skoru oluşturmaz.

50.4. Accessibility Eki

Ana panelin erişilebilirlik kullanıcılarını yetersiz kapsadığı tespit edilirse 40 ek gözlem alınabilir. Bu kullanıcılar: ilgili nüfus gruplarına göre kalibre edilir, ayrı Accessibility Panel sonucu üretir.

50.5. Toplam Saha Hacmi

Panel	Geçerli hedef
Population Panel	1.000
Language Fairness eki	50
Country Sentinel eki	3
Accessibility eki	40
Toplam geçerli saha hedefi	1.093

GEO-1000 adı yine geçerlidir. Çünkü:

> *1.000, ana nüfus panelinin geçerli gözlem hedefidir.*

Tamamlayıcı adalet ve kapsam panelleri buna eklenebilir.

50.6. Davet Hacmi

Population Panel'de ortalama:

- tamamlama: yüzde 82
- geçerlilik: yüzde 92

beklensin. Gerekli davet:

```
⌈1.000 / (0,82 × 0,92)⌉ = 1.326
```

olabilir. Tamamlayıcı panellerle toplam davet 1.500'ü aşabilir. Bu nedenle:

> *GEO-1000, yalnız 1.000 kişiye davet gönderme yöntemi değildir.*

50.7. Saha Sonucu

Population Panel'de:

- 1.326 kişi davet edilmiş,
- 1.080 görev tamamlanmış,

- 1.012 kayıt protokol bakımından geçerli bulunmuş,
- ilk 1.000 önceden tanımlanmış rastgele/yedek sırasına göre ana analize alınmış

olsun. Kalan 12 geçerli kayıt:

- yedek doğrulama,
- duyarlılık analizi,
- tekrar kontrolü

için saklanabilir. Sonuçlara bakılarak hangi 1.000'in seçileceğine karar verilemez.

51. NORMATİF TAHSİS ALGORİTMASI

Bir GEO-1000 ana paneli için varsayılan işlem sırası şöyledir:

Adım 1 — Ürün ve Dalga Kimliğini Kilitle

Hangi AI ürünü, yüzey ve tarihin ölçüldüğü belirlenir.

Adım 2 — Hedef Nüfus Sürümünü Kilitle

Bölüm 5'teki ürün-uygun nüfus çerçevesi kullanılır.

Adım 3 — Ana Geçerli Gözlem Hedefini Belirle

Varsayılan aday hedef:

```
n=1.000
```

Adım 4 — Maddi Tabakaları Belirle

Ülke, dil ve gerekliyse plan/arayüz tabakaları tanımlanır.

Adım 5 — Küçük Ülke Havuzlarını Kur

Sıfıra yakın tahsis üretecek ülkeler uygun havuzlara alınır.

Adım 6 — Ham Orantılı Kotaları Hesapla

```
qh=nWh
```

Adım 7 — Tam Sayı Tahsisini Yap

En Büyük Kalan veya önceden ilan edilmiş eşdeğer yöntem uygulanır.

Adım 8 — Fazla Örnekleme Planını Ekle

Country Sentinel, Language Fairness ve Accessibility ekleri ayrı kaydedilir.

Adım 9 — Davet ve Yedek Sayılarını Hesapla

Beklenen tamamlama ve geçerlilik oranları kullanılır.

Adım 10 — Seçim Sırasını Rastgeleleştir

Ana ve yedek listeler dondurulur.

Adım 11 — Ağırlık Formülünü Kaydet

Tasarım, yanıtlamama ve kalibrasyon ağırlıkları tanımlanır.

Adım 12 — Örneklem Kilidini Onayla

AI yanıtları görülmeden önce sürüm ve bütünlük kaydı oluşturulur.

52. ZORUNLU NORMATİF HÜKÜMLER

CH06-N01

GEO-1000 örneklem tahsisi AI yanıtları görülmeden önce oluşturulmalı ve sürümlenmelidir.

CH06-N02

Ana Population Panel hedefi, AI ürünü ve ana ölçüm dalgası başına en az 1.000 geçerli gözlem olarak tanımlanmalıdır; farklı bir hedef kullanılıyorsa gerekçesi açıklanmalıdır.

CH06-N03

Davet edilen kişi, görevi tamamlayan kişi ve geçerli gözlem sayıları birbirinden ayrılmalıdır.

CH06-N04

Ana tahsis, Bölüm 5'te tanımlanan ürün-uygun insan nüfusuna dayanmalıdır.

CH06-N05

Ülke veya dil kotaları toplam nüfusa değil ilgili ürün-uygun nüfus payına göre hesaplanmalıdır.

CH06-N06

Ham küsuratlı kotaların tam sayıya dönüştürülme yöntemi önceden tanımlanmalıdır.

CH06-N07

Yuvarlama bağlarının çözümü sonuçlardan bağımsız olmalıdır.

CH06-N08

Pozitif hedef nüfusu bulunan küçük ülkeler, ayrı sıfır olasılıklı deterministik tabakalar olarak bırakılmamalı; uygun havuz veya olasılıklı seçim yöntemi kullanılmalıdır.

CH06-N09

Küçük ülke havuzları ürün erişimi ve maddi nüfus özellikleri bakımından gerekçelendirilmelidir.

CH06-N10

Country Sentinel gözlemleri ana nüfus ağırlığını değiştirmeden ayrıca kaydedilmelidir.

CH06-N11

Bir Sentinel gözlemi ülke skoru olarak sunulamaz.

CH06-N12

Fazla örneklenen ülke, dil veya kullanıcı grupları küresel tahminde gerçek hedef nüfus ağırlıklarına geri döndürülmelidir.

CH06-N13

Örneklem payı nüfus payı gibi kullanılamaz.

CH06-N14

Aynı kişi birden fazla dil, hesap, oturum veya ürün nedeniyle küresel insan nüfusunda çoğaltılamaz.

CH06-N15

Geçerli olmayan gözlemler, geçersizlik nedeni sonuç içeriğine dayanmadan sınıflandırılmalıdır.

CH06-N16

Yanlış, olumsuz, ret veya Critical AI yanıtı katılımcı değiştirme gerekçesi olamaz.

CH06-N17

Yedek katılımcı listeleri sonuçlar görülmeden önce tabaka içinde oluşturulmalıdır.

CH06-N18

Tabaka açığı başka ülke veya dilden sessizce kapatılamaz.

CH06-N19

Sonuç sonrası kota, tabaka, küçük ülke havuzu veya ağırlık değişikliği yeni analiz sürümü gerektirir.

CH06-N20

Ağırlıkların bütün bileşenleri ve normalize edilme yöntemi yayımlanmalıdır.

CH06-N21

Yanıtlamama ayarlamasının varsayımları açıklanmalıdır.

CH06-N22

Ağırlık kırpma kullanılıyorsa kırpılmamış sonuç ve duyarlılık analizi korunmalıdır.

CH06-N23

Ana rapor ham gözlem sayısının yanında uygun etkin örneklem büyüklüğünü göstermelidir.

CH06-N24

Kümelenme, tekrar ve eşlenmiş kullanıcı bağımlılıkları varyans hesabında dikkate alınmalıdır.

CH06-N25

Aynı kişinin birden fazla AI ürününü test ettiği tasarım, tamamen bağımsız kullanıcı örneklemi gibi sunulamaz.

CH06-N26

Her AI ürünü kendi 1.000 geçerli ürün-gözlemi hedefine ayrı ulaşmalıdır.

CH06-N27

Native Reach ve Common Support tahsisleri ayrı örneklem kayıtları taşımalıdır.

CH06-N28

Dalgalar arasında güncel ve sabit çerçeveli sonuçlar açıklamasız biçimde birleştirilemez.

CH06-N29

Bir alt grup için örneklem yetersizse doğru statü INSUFFICIENT SAMPLE, EXPLORATORY veya NOT ESTIMATED olmalıdır.

CH06-N30

Ham örneklem büyüklüğü, ağırlık ve tasarım etkisi yok sayılarak kesin hassasiyet göstergesi olarak kullanılamaz.

CH06-N31

Gönüllü veya olasılıksız panelde seçim olasılıkları biliniyormuş gibi tasarım ağırlığı iddia edilemez.

CH06-N32

Örneklem, düşük skorlu grupları dışlamak veya yüksek skorlu grupları büyütmek amacıyla sonuç sonrası değiştirilemez.

CH06-N33

Ana Population Panel ile tamamlayıcı panellerin gözlemleri amaç ve ağırlık bakımından ayrı etiketlenmelidir.

CH06-N34

Örneklem kilidinin hesap verebilir insan veya kurum sahibi bulunmalıdır.

53. BAŞARISIZLIK BİÇİMLERİ

CH06-F01 — SONUÇTAN SONRA KOTA KURMAK

Ülke ve dil sayıları AI cevapları görüldükten sonra belirlenir.

CH06-F02 — EN İYİ CEVABA KADAR YEDEK KULLANMAK

Olumsuz yanıtlar geçerli olduğu hâlde yeni kullanıcıyla değiştirilir.

CH06-F03 — 1.000 DAVETİ 1.000 GEÇERLİ GÖZLEM SAYMAK

Tamamlamama ve geçersizlik gizlenir.

CH06-F04 — SAF YUVARLAMA İLE ÜLKEYİ SIFIRLAMAK

Pozitif hedef nüfusu bulunan ülkenin seçilme ihtimali sıfır bırakılır.

CH06-F05 — HER ÜLKEYE BİR KİŞİ VERİP NÜFUS ORANTISI DEMEK

Asgari kota nedeniyle oluşan orantısızlık saklanır.

CH06-F06 — SENTINEL GÖZLEMİNİ NÜFUS GÖZLEMİ SAYMAK

Küçük ülke ek gözlemi 1.000 kişilik ana panelde tam ağırlık alır.

CH06-F07 — KÜÇÜK ÜLKE HAVUZUNU SONUCA GÖRE KURMAK

Düşük skor üreten ülke başka havuza taşınır veya kapsamdan çıkarılır.

CH06-F08 — UYUŞMAZ ÜLKELERİ AYNI HAVUZA ATMAK

Ürün erişimi, dil veya yargı bakımından farklı ülkeler yalnız küçük oldukları için birleştirilir.

CH06-F09 — FAZLA ÖRNEKLEMEYİ FAZLA AĞIRLIKLANDIRMAK

Dil veya ülke nüfus payından daha yüksek örneklem payıyla küresel sonucu belirler.

CH06-F10 — ÇOK DİLLİ KİŞİLERİ ÇOĞALTMAK

Bir insan her dil hücresinde ayrı nüfus birimi sayılır.

CH06-F11 — ÖRNEKLEM PAYINI NÜFUS PAYI SAYMAK

Ham yanıt sayıları doğrudan küresel skora dönüştürülür.

CH06-F12 — DAVET SIRASINI KOLAYLIK SIRASIYLA KURMAK

İlk cevap veren hızlı kullanıcılar paneli belirler.

CH06-F13 — TABAKA DIŞI YEDEK KULLANMAK

Eksik ülke veya dil gözlemi başka gruptan katılımcıyla sessizce tamamlanır.

CH06-F14 — GEÇERSİZLİĞİ CEVAP İÇERİĞİNE BAĞLAMAK

Yanlış AI cevabı teknik olarak geçersiz ilan edilir.

CH06-F15 — RETLERİ YENİ KULLANICIYLA DEĞİŞTİRMEK

Sistem reddi gerçek kullanıcı sonucu olduğu hâlde paydadan çıkarılır.

CH06-F16 — AĞIRLIK FORMÜLÜNÜ GİZLEMEK

Kamu skoru hangi nüfus ve örneklem ağırlıklarıyla üretildiğini göstermez.

CH06-F17 — AĞIRLIK KIRPARAK KÜÇÜK GRUBU SİLMEK

Yüksek ağırlıklı küçük gruplar sonuç hoş görünmediği için kırpılır.

CH06-F18 — ETKİN ÖRNEKLEMİ GİZLEMEK

1.000 ham kayıt, ağırlıklar nedeniyle 500–600 bilgi birimine düşse bile "1.000 kişilik hassasiyet" olarak sunulur.

CH06-F19 — KÜMELENMEYİ BAĞIMSIZLIK SAYMAK

Aynı panel, kurum veya kullanıcıdan gelen gözlemler tamamen bağımsız kabul edilir.

CH06-F20 — AYNI 1.000 KİŞİYİ ON BİN BAĞIMSIZ İNSAN SAYMAK

On AI ürünü gözlemi, on kat nüfus genişliği gibi anlatılır.

CH06-F21 — COMMON SUPPORT'U NATIVE REACH SAYMAK

Ortak karşılaştırma örneklemi ürünün gerçek küresel erişim nüfusu gibi sunulur.

CH06-F22 — NATIVE REACH SONUÇLARINI DOĞRUDAN SIRALAMAK

Farklı hedef nüfuslardaki ürün skorları aynı evrenmiş gibi karşılaştırılır.

CH06-F23 — DALGALARDA NÜFUS DEĞİŞİMİNİ GİZLEMEK

Yeni ülkeler veya kullanıcı grupları skora girdiği hâlde trend yalnız ürün kalitesi değişimi gibi yorumlanır.

CH06-F24 — ALT GRUP İÇİN YETERSİZ ÖRNEKTEN KESİN SKOR ÜRETMEK

Birkaç düzine gözlem ülke veya dil için kesin puan olarak yayımlanır.

CH06-F25 — ÖRNEKLEM KİLİDİNİ SESSİZCE DEĞİŞTİRMEK

Kota, ağırlık veya havuz tanımı sürüm kaydı olmadan güncellenir.

CH06-F26 — UYGUNLUK ORANLARINI KÖR BİÇİMDE ÇARPMAK

Uygun nüfus ve kota hesapları bağımlı nüfus özelliklerini yok sayar.

CH06-F27 — TAMAMLAYICI PANELİ ANA PANELİN YERİNE KOYMAK

Dil adaleti veya gözcü gözlemleri 1.000 ana nüfus hedefini küçültmek için kullanılır.

CH06-F28 — PANEL KISA KALDIĞINDA DÜŞÜK SKORLU HÜCREYİ SİLMEK

Eksik örneklem, sonuç bakımından elverişsiz olduğu için kapsam dışı bırakılır.

CH06-F29 — AĞIRLIKLANDIRMAYI BİLİMSEL SÜS OLARAK KULLANMAK

Seçim ve kapsam yanlılığı açıklanmadan yalnız ağırlık eklenerek temsil sorunu çözüldüğü iddia edilir.

CH06-F30 — GEO-1000'İ TAM OLARAK 1.000 KİŞİYE SORMAK SANMAK

Davet, tamamlayıcı paneller, geçersizlik, tekrar ve etkin örneklem yapısı yok sayılır.

54. DENETİM PROSEDÜRÜ

Adım 1 — Örneklem Amacını Kaydet

Küresel nüfus tahmini Alt grup tanısı Ülke kapsamı Dil adaleti Erişilebilirlik Ürün karşılaştırması amaçları ayrılır.

Adım 2 — Hedef Nüfus Sürümünü Doğrula

Bölüm 5'teki nüfus dosyası ve tarih kaydı incelenir.

Adım 3 — Ana Geçerli Gözlem Hedefini Belirle

Varsayılan GEO-1000 hedefi veya gerekçeli alternatif yazılır.

Adım 4 — Tabakaları Kur

Ülke, dil ve maddi kullanıcı yüzeyi ayrımları belirlenir.

Adım 5 — Küçük Ülke Havuzlarını Tanımla

Her havuzun:

- ülkeleri,
- nüfusu,
- benzerlik gerekçesi,
- iç seçim yöntemi

kaydedilir.

Adım 6 — Ham Kotaları Hesapla

```
qh=nWh
```

formülü veya gerekçeli alternatif kullanılır.

Adım 7 — Tam Sayı Tahsisini Uygula

Yuvarlama ve bağ bozma kuralı kaydedilir.

Adım 8 — Tamamlayıcı Fazla Örneklemeyi Ekle

Country Sentinel, Language Fairness ve Accessibility hedefleri ayrı tanımlanır.

Adım 9 — Davet ve Yedek Hacmini Hesapla

Tamamlama ve geçerlilik beklentileri kullanılır.

Adım 10 — Seçim ve Yedek Sırasını Dondur

Katılımcılar sonuç görülmeden önce seçilir veya sıralanır.

Adım 11 — Tasarım Ağırlıklarını Hesapla

Seçilme olasılıkları ve tabaka ağırlıkları kaydedilir.

Adım 12 — Kısa Kalma ve Yanıtlamama Prosedürünü Tanımla

Tabaka açığına verilecek önceden ilan edilmiş cevap yazılır.

Adım 13 — Örneklem Kilidini Oluştur

Belge kimliği, sürüm, tarih ve bütünlük kaydı eklenir.

Adım 14 — Saha Sonrası Geçerlilik Akışını İncele

Geçerlilik kararlarının yanıt doğruluğundan bağımsız verildiği doğrulanır.

Adım 15 — Nihai Ağırlıkları Üret

Tasarım, yanıtlamama ve kalibrasyon bileşenleri ayrılır.

Adım 16 — Etkin Örneklem Büyüklüğünü Hesapla

Ham sayı ile bilgi hacmi farkı raporlanır.

Adım 17 — Tahsis ve Gerçekleşmeyi Karşılaştır

Planlanan ve gerçekleşen gözlem sayıları hücre bazında gösterilir.

Adım 18 — Sonuç Sonrası Değişiklikleri Denetle

Kota, havuz, ağırlık ve kapsam değişiklikleri aranır.

55. GEREKLİ KANIT

AI ürün kimliği Ölçüm dalgası Hedef nüfus sürümü Population Panel hedefi Country Sentinel hedefi Language Fairness hedefi Accessibility Panel hedefi Tabaka listesi Tabaka nüfusları Nüfus ağırlıkları Ham kotalar Yuvarlama yöntemi Kesir ve kalan kayıtları Bağ bozma kuralı Küçük ülke havuzları Havuz içi seçim olasılıkları Ülke ve dil tahsisleri Plan ve arayüz tahsisleri Fazla örnekleme gerekçeleri Davet hedefleri Tamamlama varsayımları Geçerlilik varsayımları Ana ve yedek katılımcı listeleri Seçim veya rastgelelik kaydı Yanıtlamama kayıtları Geçersiz gözlem nedenleri Geçerlilik karar tarihleri Tasarım ağırlıkları Yanıtlamama ayarları Kalibrasyon katsayıları Ağırlık kırpma politikası Ham ve etkin örneklem büyüklüğü Kümelenme ve eşlenmiş kullanıcı kayıtları Native Reach tahsisi Common Support tahsisi

Sabit ve güncel çerçeve ağırlıkları Planlanan ve gerçekleşen tahsis karşılaştırması Örneklem kilidi Değişiklik günlüğü Sorumlu insan veya kurum

56. DENETİM KONTROL LİSTESİ

Ana geçerli gözlem hedefi açık mı? Hedef 1.000 davet mi, geçerli gözlem mi? Örneklem AI yanıtları görülmeden önce donduruldu mu? Hangi nüfus çerçevesi kullanıldı? Tabakalar hangi gerekçeyle kuruldu? Ham ülke ve dil kotaları gösteriliyor mu? Küsuratların tam sayıya dönüşüm yöntemi açık mı? Bağ bozma kuralı önceden belirlenmiş mi? Küçük ülkelerin seçilme olasılığı sıfır mı? Küçük ülke havuzları mantıksal olarak uyumlu mu? Sentinel gözlemleri ana nüfus ağırlığına karışmış mı? Bir Sentinel gözleminden ülke skoru üretilmiş mi? Dil ve erişilebilirlik grupları fazla örneklendiyse geri ağırlıklandırıldı mı? Aynı kişi birden fazla nüfus birimi olarak sayılmış mı? Davet, tamamlayan ve geçerli gözlem sayıları ayrı mı?

Yedek listeleri önceden oluşturulmuş mu? Yanlış veya olumsuz AI cevabı nedeniyle katılımcı değiştirilmiş mi? Geçerlilik kararı maddi doğruluk puanından önce mi verilmiş? Tabaka açığı başka ülkeden sessizce kapatılmış mı? Sonuç sonrası kota veya havuz değişmiş mi? Tasarım ağırlıkları yeniden üretilebilir mi? Yanıtlamama ayarının varsayımları açık mı? Ağırlık kırpma yapıldı mı ve etkisi gösterildi mi? Ham ve etkin örneklem büyüklüğü birlikte veriliyor mu? Aynı kullanıcı birden fazla AI ürününü test etmiş mi? Bu bağımlılık analizde korunmuş mu? Her AI ürünü ayrı 1.000 ürün-gözlemine ulaştı mı? Native Reach ve Common Support örneklemleri ayrıldı mı? Dalgalar arasında nüfus çerçevesi değişti mi?

Sabit ve güncel çerçeveli trendler ayrıldı mı? Yetersiz alt grup örnekleri kesin skor gibi yayımlanmış mı? Örneklem kalite statüsü doğru mu? Tamamlayıcı paneller ana panelin yerini almış mı? Örneklem kaydının hesap verebilir sahibi belli mi?

57. İTİRAZLAR VE CEVAPLAR

İtiraz 1 — "Neden her ülkeye doğrudan nüfus oranında kişi vermiyoruz?"

Büyük ve orta nüfuslu ülkelerde bu güçlü varsayılandır. Çok küçük ülkelerde ham kota bir kişinin altına düşebilir. Bu durumda:

- küçük ülke havuzu,
- olasılıklı ülke seçimi,
- ayrı Sentinel paneli

daha doğru olabilir.

İtiraz 2 — "Her ülkeye en az bir kişi vermek daha adil değil mi?"

Coğrafi kapsam açısından yararlı olabilir. Nüfus tahmini açısından orantısızdır. Doğru çözüm, bu bir kişiyi Sentinel gözlemi olarak ayrıca etiketlemektir.

İtiraz 3 — "Bir ülkenin ana panelde sıfır kişi alması onu tamamen dışlamaz mı?"

Ülke ayrı deterministik tabaka olarak sıfır kota alıyorsa sorun olabilir. Küçük ülke havuzunda pozitif seçim olasılığı verilebilir. Ayrıca Sentinel Panel ülkenin gözlemsel görünürlüğünü koruyabilir.

İtiraz 4 — "Küçük ülkeleri havuzlamak kültürel farkları silmez mi?"

Ülke skoru üretmek amacıyla kullanılırsa silebilir. Ana küresel tahmin için havuz, seçim olasılığını yönetir. Ülke kimliği gözlem kaydında korunur. Ayrı Sentinel ve ülke özel örneklem yapılabilir.

İtiraz 5 — "Neden 1.000'in içine Sentinel kullanıcılarını koymuyoruz?"

Koyulabilir. Fakat o zaman 1.000 kişilik örneklem saf nüfus orantılı olmaktan çıkar. Ağırlıklandırma gerekir ve ana tahminin etkin örneklem büyüklüğü düşebilir. Kurucu mimaride 1.000 gözlem Population Panel'e ayrılır; Sentinel gözlemleri buna eklenir.

İtiraz 6 — "Toplam saha maliyeti çok büyümez mi?"

Büyüyebilir. Standart, maliyeti saklayarak yapay kesinlik üretmemelidir. Daha düşük bütçede:

- daha az AI ürünü,
- daha az dil,
- daha küçük pilot,
- daha geniş belirsizlik

kullanılabilir. Doğru statü açıkça yazılmalıdır.

İtiraz 7 — "GEO-1000 adı varken neden 1.093 veya daha fazla gözlem topluyoruz?"

1.000, ana nüfus panelinin geçerli gözlem hedefidir. Adalet ve kapsam panelleri ana hedefe eklenebilir. Bu, standardın adını bozmaz. Ölçümün gerçek toplam saha hacmini görünür kılar.

İtiraz 8 — "Bir dili fazla örneklemek küresel sonucu neden bozsun?"

Fazla örneklenen dil ham yanıtlarda olduğundan büyük görünür. Nüfus ağırlığına geri döndürülürse küresel sonuç bozulmaz. Ayrı dil sonucu daha hassas hâle gelir.

İtiraz 9 — "Ağırlıklar çok karmaşık; ham ortalama daha anlaşılır değil mi?"

Ham ortalama, örneklem nüfusa orantılıysa anlamlı olabilir. Fazla örnekleme, yanıtlamama veya kapsama farkı varsa ham ortalama yanlış soruya cevap verir. Karmaşıklık, örneklem tasarımından doğuyorsa saklanamaz.

İtiraz 10 — "1.000 geçerli gözlem varsa etkin örneklem neden daha düşük olsun?"

Çünkü bazı gözlemler çok yüksek, bazıları çok düşük ağırlık taşıyabilir. Aynı panel veya kullanıcıdan gelen kayıtlar birbirine benzeyebilir. Ham sayı fiziksel kayıt adedidir. Etkin örneklem istatistiksel bilgi hacmine yaklaşır.

İtiraz 11 — "Aynı kişilerin bütün AI ürünlerini test etmesi daha adil değil mi?"

Kullanıcı farkını azaltır. Fakat:

- ürün erişim evrenini daraltabilir,
- görev yükünü artırabilir,

cevaplar arası bulaşma yaratabilir. Eşlenmiş veya dengeli eksik blok tasarımıyla yönetilebilir.

İtiraz 12 — "Bir tabaka hedefe ulaşmazsa küresel skor hiç yayımlanamaz mı?"

Her durumda değil. Eksiklik:

- nüfus ağırlığı,
- risk,
- yetersizliğin nedeni,
- kalan kapsam

bakımından değerlendirilir. Skor:

- daraltılmış kapsamla,
- geniş belirsizlikle,
- INSUFFICIENT KAPSAMA uyarısıyla

yayımlanabilir. Eksiklik saklanamaz.

İtiraz 13 — "Nüfus verisi değişirse her dalgada bütün kotaları yenilemek zorunda mıyız?"

Güncel anlık skor için yenileme gerekebilir. Trend karşılaştırması için sabit başlangıç ağırlıkları da korunabilir. İki sonuç ayrı yayımlanmalıdır.

İtiraz 14 — "Olasılıklı dünya paneli kuramazsak NOMOS uygulanamaz mı?"

Uygulanabilir. Fakat örneklem kalite statüsü doğru yazılmalıdır. Nüfusa kalibre edilmiş gönüllü panel değerli olabilir. Tam olasılıklı örneklem gibi sunulamaz.

60. BÖLÜMÜN ORTAK HÜKMÜ

GEO-1000'in gücü yalnız 1.000 sayısından doğmaz. 1.000 kişinin:

- hangi nüfustan geldiğinden,
- hangi olasılıkla seçildiğinden,
- hangi ülke ve dili temsil ettiğinden,

- aynı insanın kaç kez sayıldığından,
- hangi gözlemlerin geçerli kabul edildiğinden,
- küçük ülkelerin nasıl yönetildiğinden,
- fazla örneklenen grupların nasıl ağırlıklandırıldığından,
- kaç kişinin görevi tamamlamadığından,
- sonuçlara bakılarak kimin değiştirildiğinden

doğar. Yanlış örneklemle 100.000 cevap toplanabilir. Sonuç yine yanlış nüfus hakkında konuşabilir. Doğru örneklemle daha az cevap toplanabilir. Sonuç sınırları açık biçimde daha güvenilir olabilir. Bir ülkenin nüfusu büyükse küresel insan deneyiminde büyük ağırlık taşımalıdır. Fakat ürün orada kullanılamıyorsa sahte kullanıcı yaratılamaz. Bir ülkenin nüfusu küçükse küresel ağırlığı küçük kalmalıdır. Fakat tamamen görünmez bırakılmak zorunda değildir. Bir dil küresel nüfusta küçük olabilir. Fakat o dildeki ağır yanlış ayrıca ölçülebilir. Bir kullanıcı üç dil ve beş ürün kullanabilir. Bu, on beş ayrı insan oluşturmaz. Bir AI yanıtı yanlış olabilir. Bu, katılımcının geçersiz olduğu anlamına gelmez. Yanlış cevap, ölçmeye çalıştığımız şeydir.

Onu yeni kullanıcıyla değiştirmek ölçümü temizlemez. Gerçekliği siler. Ağırlıklandırma, yanlış paneli mucizevi biçimde dünya nüfusuna dönüştürmez. Fakat doğru tasarlanmış fazla örneklemenin küresel tahmini bozmadan alt grupları görünür kılmasını sağlayabilir. Bu nedenle NOMOS'un altıncı ölçüm yasası şudur:

> *Örneklem, sonucu üretmek için değil; sonucu adil biçimde gözlemek için kurulmalıdır.*

Yedinci yasası şudur:

> *Küçük ülkenin görünürlüğü nüfus ağırlığını büyütmek zorunda değildir.*

Sekizinci yasası şudur:

> *Fazla gözlem, fazla nüfus ağırlığı anlamına gelmez.*

Dokuzuncu yasası şudur:

> *Yanlış AI cevabı veri kusuru değildir; geçerli protokol altında gözlenmişse denetimin konusudur.*

Onuncu yasası şudur:

> *Bin gözlemin güveni, bin sayısından değil; seçimin ve ağırlığın yeniden üretilebilirliğinden doğar.*

NOMOS'un Bölüm 6 Emri

> *Bana yalnız "1.000 kişi sorduk" deme.*
>
> *Bu 1.000 kişinin hangi milyar insan arasından, hangi olasılıkla seçildiğini göster.*
>
> *Bir ülkeye yarım insan düşüyor diye onu yok sayma. / Ama bir kişi verdin diye ülkenin tamamını ölçtüğünü de söyleme.*
>
> *Küçük ülkeleri uygun havuzda olasılıklı biçimde seç. / Görünürlük gerekiyorsa ayrıca gözcü paneli kur.*
>
> *Düşük kaynaklı dili fazla örnekleyebilirsin. / Fakat küresel skorda onu nüfusundan büyük yapma.*
>
> *Bir insanı hesapları, dilleri, oturumları ve AI ürünleri kadar çoğaltma.*
>
> *Katılımcı cevap vermediğinde yedeğini kullan. / AI yanlış cevap verdiğinde kullanıcını değiştirme.*
>
> *Geçerliliğe cevabın doğruluğunu görmeden karar ver.*

> *Kotan eksik kaldığında bunu açıkla. / Başka ülkeden insan getirip boşluğu saklama.*
>
> *Sonucu gördükten sonra paydayı, ağırlığı, ülke havuzunu veya örneklem hedefini değiştirme.*
>
> *Ham sayını göster. / Etkin örneklemini göster. / Kimin ağır, kimin hafif ağırlık aldığını göster.*
>
> *Önce nüfusu kilitle. / Sonra tabakaları kur. / Sonra kotaları hesapla. / Sonra yedekleri seç. / Sonra ağırlıkları yaz. / Ve ancak bundan sonra ilk istemi gönder.*

Bölümün Son Cümlesi

> *GEO-1000, bin cevabı bir araya getirme yöntemi değil; hedef nüfustaki her insanın temsil hakkını, seçilme olasılığı ve doğru ağırlık altında ölçülebilir hâle getirme protokolüdür.*

Normatif Öz

> **Her AI ürünü ve ana ölçüm dalgası için GEO-1000 Population Panel, farklı bir hedef açıkça gerekçelendirilmedikçe en az 1.000 geçerli ürün-gözlemini hedeflemelidir. Örneklem tahsisi AI yanıtları görülmeden önce tasarlanmalı, sürümlenmeli ve kilitlenmelidir. Ana tahsis toplam ülke nüfusuna, davet sayısına, hesaplara, oturumlara veya kolaylık örneklemine değil ürün-uygun hedef insan nüfusuna dayanmalıdır. Küsuratlı kotalar önceden ilan edilmiş ve sonuçtan bağımsız yöntemle tam sayıya çevrilmelidir. Pozitif nüfus kütlesi taşıyan gruplara, bütün hedef nüfus adına çıkarım yapılırken deterministik sıfır seçilme olasılığı verilemez. Fazla örneklenen gruplar ham örneklem paylarını nüfus ağırlığı olarak kullanamaz. Yanlış, olumsuz, ret veya Critical AI çıktısı katılımcı değiştirme nedeni olamaz.**

BÖLÜM 07

Ülke Gözcü ve Dil Adaleti Panelleri

YÖNTEMSEL DAYANAKLAR / K04; K06; K07; K20

Bölüm Sınırı

Bölüm 5 şu hükmü kurdu:

> *Her AI ürünü için ürün-uygun kullanıcı nüfusu ayrı tanımlanmalıdır.*

Bölüm 6 bu nüfusu örnekleme dönüştürdü:

> *Ana Population Panel, hedef nüfustaki insanlara sonuçlar görülmeden önce belirlenmiş seçilme olasılıkları ve ağırlıklar vermelidir.*

Şimdi iki farklı adalet problemini çözmemiz gerekir. Birinci problem:

> *Nüfus ağırlığı çok küçük ülkeleri ana örneklemde görünmez bırakabilir.*

İkinci problem:

> *Küresel nüfus ortalaması, düşük kaynaklı veya düşük nüfuslu dillerdeki ağır temsil bozulmasını gizleyebilir.*

Saf nüfus örneklemi dünya nüfusundaki ortalama insan deneyimini ölçmek için güçlüdür. Fakat aynı örneklem:

- her ülkeyi ayrı ayrı gözleyemez,
- her dile yeterli örnek sağlayamaz,
- küçük ülkeler için ülke skoru üretemez,

düşük kaynaklı dillerdeki nadir fakat maddi hataları yakalayamayabilir. Bu nedenle GEO-1000 yalnız tek panelden oluşamaz. Ana nüfus panelinin yanında iki tamamlayıcı yapı gerekir:

Country Sentinel Panel

Türkçesi:

Ülke Gözcü Paneli

ve:

Language Fairness Panel

Türkçesi:

Dil Adaleti Paneli

Ülke Gözcü Paneli şu soruya cevap verir:

> *Ana nüfus panelinde yeterince temsil edilmeyen ülkelerde en azından gerçek kullanıcı gözlemi, erişim doğrulaması veya erken risk sinyali var mı?*

Dil Adaleti Paneli ise şu soruya cevap verir:

> *Küresel ortalama yüksek görünse bile belirli dillerde kullanıcılar maddi olarak daha yanlış, daha eksik veya daha riskli temsil görüyor mu?*

Bu iki panel ana Population Panel'in yerine geçmez. Onu tamamlar. Bir ülkeye bir gözcü gözlemi vermek o ülkenin skorunu ölçmez. Bir dili fazla örneklemek de o dili küresel nüfusta olduğundan büyük yapmaz. Bu bölümün görevi:

- ülke görünürlüğü ile nüfus ağırlığını,
- dil adaleti ile küresel ortalamayı,
- erken uyarı ile nüfus tahminini,
- gözlem kapsamı ile istatistiksel yeterliliği

birbirinden ayırmaktır. Bu bölüm henüz:

- istemlerin bütün dillerde nasıl çevrileceğini,
- ters çeviri ve semantik eşdeğerlik testini,
- katılımcıların ayrıntılı seçim mekanizmasını,
- atomik iddiaların hakemlerce nasıl puanlanacağını,
- nihai uygunluk eşiklerini

tam olarak tanımlamaz. Bu konular sonraki bölümlerde ayrıca düzenlenecektir. Bölüm 7'nin temel sorusu şudur:

> *Bir küresel skor üretirken küçük ülke ve dilleri yapay biçimde büyütmeden, fakat onları tamamen görünmez de bırakmadan nasıl ölçeriz?*

NOMOS Challenge

Bir AI ürünü için 1.000 kişilik nüfus paneli kurdunuz. Nüfusa orantılı örneklem sonucunda:

- birkaç büyük ülkeden yüzlerce kullanıcı,
- bazı orta büyüklükteki ülkelerden onlarca kullanıcı,
- küçük ülkelerin çoğundan hiç kullanıcı

geldi. Küresel skor:

94

olarak hesaplandı. Sonra şöyle dediniz: "Bu AI ürünü dünya genelinde markayı yüzde 94 doğrulukla temsil ediyor." Fakat 70 küçük ülkede tek bir gerçek kullanıcı gözlemi bile bulunmuyor. Bu ülkeler dünya nüfusunun küçük bir bölümünü oluşturabilir. Yine de şu sorular cevapsızdır: Ürün gerçekten erişilebilir mi? İstem doğru dilde çalışıyor mu? Sistem ülkeyi başka ülkeyle karıştırıyor mu? Yerel şirket ile küresel marka birbirine karışıyor mu? Fiyat, hizmet ve hukuk bağlamı yanlış mı? Sistem cevap vermiyor mu? Düşük sıklıklı Critical hata var mı? Küresel nüfus skoru bu sorulara cevap vermez. Şimdi her ülkeye bir kişi eklediğinizi düşünün. 70 ek gözlem aldınız.

Bunları ana skorun içine, her biri 1/1.070 ağırlıkla eklediniz. Küçük ülkeleri gerçek nüfus paylarından kat kat büyük yaptınız. Bu kez coğrafi görünürlüğü artırdınız. Fakat nüfus tahminini bozdunuz. Sonra bir ülkedeki tek kullanıcının cevabı yanlış çıktı. Şöyle dediniz: "Bu ülkenin GEO skoru sıfır." Bir kişinin cevabı ülkenin skoru değildir. Yalnız o ülkede belirli koşullarda gözlenmiş tekil olaydır. Şimdi dil tarafına geçin. İngilizce sonuç yüzde 98. Almanca yüzde 96. Türkçe yüzde 91. Düşük kaynaklı bir dilde yalnız sekiz kullanıcı var ve ham sonuç yüzde 50. Şöyle diyorsunuz: "Sekiz kişi yetersiz; bu dili rapordan çıkaralım." Küresel skor yükseliyor. Fakat o dildeki kullanıcı deneyimi görünmez hâle geliyor.

Başka bir yaklaşımda düşük kaynaklı dile 200 ek kullanıcı veriyorsunuz. Ham örneklem ortalaması düşüyor. Sonra: "Küresel skor 84'e düştü." diyorsunuz. Oysa dilin gerçek nüfus payı yüzde 2 olabilir. Fazla örnekleme tanı gücünü artırır. Nüfus ağırlığını artırmaz. Bu bölümün ilk hükmü şudur:

> *Gözlenmek ile ağırlık sahibi olmak aynı şey değildir.*

İkinci hükmü şudur:

> *Tekil ülke gözlemi ülke skoru değildir.*

Üçüncü hükmü şudur:

> *Dil adaleti, bütün dilleri küresel skorda eşit ağırlıklandırmak değildir.*

Dördüncü hükmü şudur:

> *Yüksek küresel ortalama, belirli bir dildeki ağır bozulmayı ortadan kaldırmaz.*

Beşinci hükmü şudur:

> *Kapsamı artırmak için alınan ek gözlemler, gerçek nüfus ağırlığıyla yeniden ilişkilendirilmeden ana skora eklenemez.*

1. BÖLÜMÜN AMACI

Bu bölümün amacı, ana Population Panel'in istatistiksel nüfus temsilini bozmadan ülke ve dil görünürlüğünü artıran tamamlayıcı panelleri kurmaktır. Bölüm şu ayrımları normatif hâle getirir:

- Nüfus ağırlığı ile coğrafi görünürlük
- Ülke gözlemi ile ülke skoru
- Tekil olay tespiti ile yaygınlık tahmini
- Ana Population Panel ile Country Sentinel Panel
- Country Sentinel Panel ile ülke özel doğrulama çalışması
- Ülke sayısı kapsamı ile nüfus kapsamı
- Dil sayısı kapsamı ile dil nüfusu kapsamı
- Dil adaleti ile dillerin eşit küresel ağırlığı
- Dil fazla örneklemesi ile küresel dil payı
- Ana dil ile istem dili
- Dil ile locale
- Dil ile yazı sistemi
- Desteklenen dil ile yalnız cevap üretilebilen dil
- İstem çeviri kusuru ile AI dil kusuru
- Kaynak eksikliği ile AI temsil kusuru
- Hakem yetersizliği ile ürün başarısızlığı
- Meşru yerel farklılık ile haksız dil eşitsizliği
- Küresel ortalama ile alt grup tabanı
- Ülke erişim kapsamı ile ülke temsil doğruluğu
- Sentinel alarmı ile kesin uygunluk kararı
- Tek dalga coğrafi kapsamı ile çok dalgalı gözetim çevrimi

Bu bölümün sonunda her GEO-1000 denetimi şu sorulara açık cevap verebilmelidir:

> *Hangi ülkelerde en az bir gerçek gözlem bulunmaktadır?*
>
> *Hangi ülkelerde yalnız erken uyarı gözlemi, hangilerinde istatistiksel tahmin bulunmaktadır?*
>
> *Hangi diller ölçüldü, hangileri yeterli örnekleme ulaştı ve hangileri test edilmedi?*
>
> *Düşük kaynaklı diller fazla örneklendiyse küresel skora nasıl geri ağırlıklandırıldı?*
>
> *Bir dildeki düşük sonucun kaynağı gerçekten AI ürünü mü; yoksa istem, kaynak, locale veya hakemlik kusuru mu?*

2. MERKEZÎ NORMATİF HÜKÜM

> *GEO-1000, ana nüfus tahminini Population Panel ile; küçük ülke gözlemini Country Sentinel Panel ile; dil eşitsizliği ve düşük kaynaklı dil tanısını Language Fairness Panel ile ayrı ayrı yönetmelidir.*

Bu üç yapı:

- farklı amaçlara,
- farklı örneklem hedeflerine,
- farklı ağırlıklara,
- farklı kamu iddialarına

sahiptir. Bir panelin gözlemi başka panelin anlamına açıklamasız biçimde dönüştürülemez. Varsayılan yapı:

Panel	Ana amacı	Üretebileceği sonuç	Tek başına üretemeyeceği sonuç
Population Panel	Nüfus ağırlıklı küresel tahmin	Küresel ve tabakalı nüfus skoru	Her küçük ülke için güvenilir ülke skoru
Country Sentinel Panel	Coğrafi gözlem ve erken uyarı	Ülke gözlem kapsamı, tekil olay, erişim sinyali	Ülke nüfus oranı
Language Fairness Panel	Dil bazlı tanı ve eşitsizlik ölçümü	Dil skorları, dil tabanı, adalet farkı	Dilleri küresel nüfusta otomatik eşit ağırlıklandırma

3. ÜLKE GÖZCÜ PANELİ NEDİR?

Country Sentinel Panel, ana nüfus panelinde yeterince veya hiç temsil edilmeyen ülkelerde:

- ürün erişimini,
- istem uygulanabilirliğini,
- temel varlık çözümlemesini,
- dil ve locale davranışını,
- erken Critical hata sinyalini

gözlemek üzere kurulan tamamlayıcı denetim panelidir. Ülke Gözcü Paneli'nin temel görevi skor üretmek değildir.

> *Görünmeyen ülkeyi ilk kez görünür hâle getirmektir.*

Bir gözcü gözlemi şunu gösterebilir: "Belirtilen ülkede, belirtilen AI ürünü, kullanıcı yüzeyi, dil ve zaman koşulunda bu sonuç en az bir kez gözlenmiştir." Şunu gösteremez: "Bu ülkedeki kullanıcıların yüzde X'i aynı sonucu görür."

4. ÜLKE GÖZCÜ BİRİMİ

Ülke Gözcü Paneli'nin temel gözlem birimi yalnız ülke değildir. Tercih edilen birim:

```
S_{c,a,l,s,w}
```

şeklindedir. Burada:

- c: ülke veya yargı alanı
- a: AI ürünü ve kullanıcı yüzeyi
- l: dil, yazı sistemi ve locale
- s: oturum veya plan koşulu
- w: ölçüm dalgası

Bu ayrım gereklidir. Aynı ülkede:

- birden fazla dil,
- farklı planlar,
- web ve mobil yüzeyler,
- farklı hukuki veya teknik erişim koşulları

bulunabilir. Bir ülkeye bir kullanıcı eklemek, ülkenin bütün dil ve kullanıcı yüzeylerini gözlemek anlamına gelmez.

5. ÜLKE GÖZCÜ PANELİNİN GÖREVLERİ

5.1. Erişim Doğrulaması

Ürün resmî kayıtta ülkede erişilebilir görünse de gerçek kullanıcı:

- oturum açamayabilir,
- belirtilen yüzeye erişemeyebilir,

plan veya doğrulama engeliyle karşılaşabilir. Gözcü gözlemi resmî erişim kaydını saha düzeyinde sınar.

5.2. Temel Kimlik Kontrolü

AI ürünü denetlenen varlığı:

- doğru şirket,
- doğru marka,
- doğru ülke ilişkisi

içinde çözümleyebiliyor mu?

5.3. Locale Sapması Tespiti

Aynı dilde farklı ülke locale'lerinde:

- fiyat,
- şirket,
- hizmet,
- hukuk,
- tavsiye

değişiyor mu? Değişiyorsa bu fark maddi ve açıklanabilir mi?

5.4. Düşük Sıklıklı Ağır Hata Sinyali

Tek bir gözcü cevabında:

- yanlış lisans,
- yanlış hukuki varlık,
- tehlikeli sağlık veya güvenlik bilgisi,
- ağır kimlik birleşmesi

gözlenebilir. Tekil olay ülke yaygınlığını göstermez. Ancak doğrulama çalışmasını tetikleyebilir.

5.5. Yeni Araştırma Hipotezi

Gözcü paneli şu soruyu doğurabilir: "Bu olay tekil mi, yoksa ülkedeki kullanıcılar arasında tekrar eden bir örüntü mü?" Bu soru hedefli ek örneklemle sınanır.

5.6. Coğrafi Kapsam Kaydı

Ürünün kaç uygun ülkede gerçek kullanıcı gözlemi bulunduğunu gösterir. Bu, ülke skorlarının sayısı değildir. Gözlemsel kapsamdır.

6. ÜLKE GÖZCÜ PANELİNİN YAPAMAYACAĞI ŞEYLER

Ülke Gözcü Paneli tek başına:

- güvenilir ülke skoru,
- ülke nüfus oranı,
- ülke sıralaması,
- ülke uygunluk kararı,
- ülke bazlı model üstünlüğü

üretemez. Bir ülkede tek gözcü cevabı doğruysa: "Ülke skoru 100" denemez. Yanlışsa: "Ülke skoru sıfır" denemez. Doğru statü:

OBSERVED_PASS

OBSERVED_ISSUE

OBSERVED_CRITICAL_CANDIDATE

ACCESS_FAILURE

INSUFFICIENT_FOR_COUNTRY_ESTIMATE

gibi gözlemsel kayıtlardır.

7. UYGUN ÜLKE SİCİLİ

Her AI ürünü için bir Uygun Ülke Sicili oluşturulmalıdır. Her ülke şu statülerden birini taşıyabilir:

CE-0 — UNKNOWN

Ürün ve denetim erişim durumu belirlenememiştir.

CE-1 — OUTSIDE NATIVE REACH

Ürün resmî veya meşru kullanım alanı dışında kalmaktadır. Ülke temsil doğruluğu skoruna alınmaz. Erişim kapsamı açığı olarak gösterilir.

CE-2 — RESTRICTED OR SPECIAL ACCESS

Yalnız:

- kurum,
- davet,
- plan,
- cihaz,
- kullanıcı türü

üzerinden erişim vardır. Ayrı panel gerekebilir.

CE-3 — ELIGIBLE FOR SENTINEL OBSERVATION

Ürün, ilgili kullanıcı yüzeyi ve en az bir denetim dili bakımından gözcü gözlemine uygundur.

CE-4 — ELIGIBLE FOR POPULATION ESTIMATION

Ülke ana nüfus panelinde ürün-uygun nüfus payına sahiptir.

CE-5 — ELIGIBLE FOR COUNTRY ESTIMATION

Yeterli örneklem, etkin örneklem, dil kapsamı ve kalite koşulları altında ülke tahmini üretilebilir. Bir ülke CE-3 olabilir fakat CE-5 olmayabilir. Bu, ülkenin gözlendiği fakat güvenilir ülke skoru üretilemediği anlamına gelir.

8. GÖZCÜ GÖZLEMİ İLE ÜLKE TAHMİNİ ARASINDAKİ MERDİVEN

CS-0 — NOT OBSERVED

Uygun ülke için geçerli gözcü gözlemi yoktur.

CS-1 — SINGLE SENTINEL OBSERVATION

Bir geçerli ülke–ürün–dil gözlemi vardır. Yalnız olay kaydı üretir.

CS-2 — REPEATED SENTINEL OBSERVATION

Farklı dalga veya kullanıcılardan birkaç gözlem vardır. Örüntü şüphesi oluşturabilir. Ülke oranı için yeterli değildir.

CS-3 — EXPLORATORY COUNTRY SAMPLE

Bölüm 6'daki keşif örneklem bandına ulaşılmıştır. Geniş belirsizlikle keşif sonucu verilebilir.

CS-4 — PROVISIONAL COUNTRY ESTIMATE

Yeterli ve dengeli örneklem altında geçici ülke tahmini üretilebilir.

CS-5 — COUNTRY ESTIMATION CANDIDATE

Yeterli etkin örneklem, dil/locale kapsamı, tekrar dalgaları ve kalite kontrolleri bulunmaktadır. Ülke skoru adayı olabilir. Kesin eşikler pilot ve dış incelemeyle kalibre edilmelidir.

9. GÖZCÜ ÇEVRİMİ

Bütün uygun ülkelerin her dalgada gözlenmesi maliyet ve operasyon bakımından mümkün olmayabilir. Bu nedenle Gözcü Çevrimi kullanılabilir. Ölçüm dalgaları:

```
w=1,2,…,K
```

ve uygun ülkeler: c∈CS olsun. Ülke c'nin dalga w'de gözcü gözleme atanması:

```
z_{c,w} = 1 (ülke dalgada gözlenecekse); aksi hâlde z_{c,w} = 0
```

şeklinde gösterilebilir. Temel çevrim koşulu:

```
Σ_{w=1}^K z_{c,w} ≥ 1
```

olmalıdır. Yani uygun her ülke, tanımlı gözetim çevrimi içinde en az bir kez gözlenmelidir. Bu hüküm:

- her ülkenin her dalgada gözleneceği,
- bir gözlemin ülke skoru oluşturacağı

anlamına gelmez.

10. GÖZCÜ SIKLIK SINIFLARI

Ülkeler önceden tanımlanmış sıklık sınıflarına ayrılabilir.

SF-0 — OUTSIDE REACH

Ürün erişimi yoktur. Kullanıcı temsili gözlemi yapılmaz. Erişim statüsü izlenir.

SF-1 — ONCE PER SURVEILLANCE CYCLE

Düşük riskli, küçük veya Population Panel'de zaten yeterli dolaylı kapsam taşıyan ülkeler.

SF-2 — PERIODIC SENTINEL

Belirli aralıklarla yeniden gözlenir.

SF-3 — EVERY PRINCIPAL WAVE

Şu nedenlerle her ana dalgada gözlenebilir:

- yüksek kullanıcı nüfusu,
- maddi hedef pazar,
- önceki ciddi hata,
- hızlı erişim veya mevzuat değişimi,

yüksek riskli hizmet alanı.

SF-4 — INCIDENT MONITORING

Doğrulanmış Critical olay veya sürekli temsil bozulması nedeniyle sıklaştırılmış gözlem. Sıklık sınıfı sonuçlara bakıldıktan sonra geçmişe dönük değiştirilemez. Yeni olay, sonraki dalga için yeni sürüm oluşturabilir.

11. SENTINEL ÖNCELİĞİ

Ülke gözcü sıklığı yalnız nüfusa göre belirlenmemelidir. İlgili olduğu ölçüde şu alanlar dikkate alınabilir:

- Ürün-uygun nüfus
- Varlığın gerçek hizmet veya pazar kapsamı
- Önceki yanlış temsil kayıtları
- Yüksek riskli kullanıcı etkisi
- Ürün erişimindeki değişiklik
- Dil ve locale benzersizliği
- Düşük kaynaklı dil kullanımı
- Hukuki veya operasyonel farklılık
- Son gözlemden bu yana geçen süre

Bu alanlar önceden belirlenmiş yönetişim kuralıyla kullanılmalıdır. “Bu ülkenin sonucu kötü, daha az ölçelim.” yaklaşımı yasaktır.

12. ÜLKE GÖZCÜ KAPSAM METRİKLERİ

Ülke kapsamı tek sayı değildir. En az üç ayrı metrik kullanılmalıdır.

12.1. Ülke Gözlem Kapsamı

En az bir geçerli gözcü veya nüfus gözlemi bulunan uygun ülkelerin oranıdır.

```
COC_count = |{c ∈ C_S : n_c^obs > 0}| / |C_S|
```

Bu metrik ülke sayısı bakımından kapsamı gösterir.

12.2. Nüfus Ağırlıklı Ülke Gözlem Kapsamı

En az bir geçerli gözleme sahip ülkelerdeki uygun nüfusun toplam hedef nüfusa oranıdır.

```
COC_pop = [Σ_{c:n_c^obs>0} N_c] / [Σ_{c∈C_S} N_c]
```

Bu metrik büyük nüfusların gözlenip gözlenmediğini gösterir.

12.3. Ülke Tahmin Kapsamı

Yeterli ülke örneklemine ulaşan ülkelerin oranıdır.

```
CEC = |{c : n_{c,eff} ≥ m_c}| / |C_S|
```

Burada mc, ülke tahmini için önceden belirlenmiş yeterlilik eşiğidir. Bir denetim:

- yüzde 100 ülke gözlem kapsamı,
- yüzde 99 nüfus gözlem kapsamı,
- yalnız yüzde 10 ülke tahmin kapsamı

taşıyabilir. Bu sonuçlar birbirine dönüştürülemez.

13. GÖZCÜ ALARMI

Bir gözcü gözleminde maddi sorun bulunduğunda Gözcü Alarmı oluşturulur. Alarm türleri:

SA-1 — ACCESS ANOMALY

Resmî erişim kaydı ile gerçek kullanıcı erişimi çelişir.

SA-2 — VARLIK ANOMALY

Varlık başka şirket, ürün veya ülke operasyonuyla birleşir.

SA-3 — LANGUAGE OR LOCALE ANOMALY

Yanlış dil, yanlış ülke bağlamı veya belirgin locale sapması oluşur.

SA-4 — MATERIAL FACT ANOMALY

Maddi fiyat, hizmet, kapasite, ortaklık veya zaman yanlışı bulunur.

SA-5 — CRITICAL INCIDENT CANDIDATE

Sağlık, hukuk, finans, güvenlik, lisans veya ağır kimlik yanlışı bulunur.

SA-6 — REFERENCE OR İSTEM ANOMALY

Sorunun kaynağı AI yanıtı değil:

- istem,
- çeviri,
- resmî kaynak,
- hakemlik

kusuru olabilir. Alarm, otomatik ülke başarısızlığı değildir. Doğrulama ve eskalasyon başlangıcıdır.

14. GÖZCÜ ESKALASYON PROTOKOLÜ

Adım 1 — Geçerlilik Doğrulaması

Gözlem gerçekten protokole uygun mu? Doğru ülke Doğru ürün Doğru istem Doğru dil İlk uygun çıktı Tam kanıt kontrol edilir.

Adım 2 — Olayın Yeniden Kodlanması

Sorun:

- AI ürünü,
- referans kaydı,
- istem,
- katılımcı,
- erişim

kaynaklarından hangisine ait olabilir?

Adım 3 — Bağımsız Tekrar

Sonuç seçmeden, önceden belirlenmiş yeni kullanıcı veya oturumlarla doğrulama yapılır.

Adım 4 — Hedefli Örneklem

Olay devam ediyorsa ülke özel örnekleme geçilebilir. Örneklem büyüklüğü:

- olay şiddeti,

- beklenen yaygınlık,
- ülke nüfusu,
- dil çeşitliliği

üzerinden önceden tanımlanır.

Adım 5 — Karar Ayrımı

Şu sonuçlar ayrı tutulur:

- Tekil doğrulanmış olay
- Tekrarlanan örüntü
- Ülke oranı
- Critical uygunluk etkisi

Tekil Critical olay yaygınlık tahmini olmadan da uygunluk incelemesini etkileyebilir.

15. DİL ADALETİ NEDİR?

Dil adaleti:

> *Bütün dillerin kelime kelime aynı cevabı alması değildir.*

Şu anlama gelir:

> *Semantik olarak eşdeğer ve maddi bakımdan karşılaştırılabilir kullanıcı sorularında, gerçek yerel farklılıklar dışında, kullanıcıların dili nedeniyle sistematik olarak daha yanlış, daha eksik, daha kanıtsız veya daha riskli temsil görmemesi.*

Dil adaleti şu eşitliği aramaz:

```
R_Türkçe = R_İngilizce = R_Japonca
```

kelime kelime. Şunu arar:

> *Kimlik, maddi gerçeklik, sınır, zaman, kaynak statüsü ve kullanıcı uygunluğu bakımından karşılaştırılabilir temsil.*

16. DİL FARKI HER ZAMAN ADALETSİZLİK DEĞİLDİR

Aşağıdaki nedenlerle dil sonuçları meşru biçimde farklılaşabilir: Yerel hizmet kapsamı farklıdır. Fiyat ve sözleşme koşulları ülkeye göre değişir. Hukuki bilgi yargı alanına bağlıdır. Kullanıcı niyeti farklıdır. Promptlar semantik olarak aynı değildir. Resmî kaynaklar diller arasında farklı kapsam taşır. AI ürünü ilgili dili resmî olarak desteklemiyordur. Bu farkların her biri kayıt altına alınmalıdır. Açıklanabilir yerel fark: adaletsizlik olarak etiketlenmemelidir. Fakat gerçeklik aynı olduğu hâlde düşük kaynaklı dilde:

- şirket kimliği bozuluyor,
- sınırlar kaldırılıyor,
- yanlış lisans ekleniyor,
- cevap sürekli başka dilde veriliyor

ise dil temelli temsil problemi olabilir.

17. DİL BİRİMİ

Dil değerlendirmesinde yalnız genel dil adı yeterli olmayabilir. Tercih edilen dil birimi:

```
L=(λ,σ,ℓ,d)
```

şeklinde düşünülebilir. Burada:

- λ: dil
- σ: yazı sistemi
- ℓ: locale veya ülke bağlamı
- d: görev veya alan

Örnek olarak aynı dil:

- farklı yazı sisteminde,
- farklı ülkede,
- hukuk veya sağlık alanında

ayrı ölçüm hücresi gerektirebilir. Bir dilin bütün kullanıcıları tek homojen grup değildir.

18. DİL DESTEK STATÜLERİ

Her AI ürünü, dil ve kullanıcı yüzeyi için şu statülerden biri kullanılabilir:

LS-0 — UNKNOWN

Destek durumu belirlenememiştir.

LS-1 — NOT SUPPORTED

Ürün ilgili dili resmî olarak desteklememektedir. Ana desteklenen dil skoru içinde değerlendirilmez. Deneysel test yapılabilir.

LS-2 — OBSERVED BUT UNSUPPORTED

Ürün cevap üretebilmektedir. Fakat resmî destek veya kalite taahhüdü yoktur. Ayrı deneysel sonuç olarak raporlanır.

LS-3 — LIMITED SUPPORT

Belirli görev, arayüz veya özelliklerde destek vardır.

LS-4 — OFFICIALLY SUPPORTED

Ürün dili genel olarak desteklediğini beyan etmektedir.

LS-5 — AUDIT-ELIGIBLE

Dil:

- resmî destek,
- istem eşdeğerliği,
- yerel hakem,
- referans yeterliliği,
- geçerli capture

bakımından tam denetime uygundur. Resmî destek bulunması tek başına LS-5 oluşturmaz.

19. DÜŞÜK KAYNAKLI DİL

"Düşük kaynaklı dil" ifadesi dilin değerini veya insanlarının niteliğini tanımlamaz. Belirli teknoloji ve denetim bağlamında şu eksikliklerden birini veya birkaçını ifade eder:

- Sınırlı dijital içerik
- Az yerelleştirilmiş resmî kaynak
- Az değerlendirme veri seti
- Sınırlı hakem erişimi
- Düşük ürün desteği
- Az terminoloji kaynağı
- Sınırlı makine çeviri kalitesi
- Düşük kullanıcı paneli erişimi

Bir dil bir alanda düşük kaynaklı, başka alanda güçlü olabilir. Örneğin günlük konuşmada güçlü kaynak taşırken:

- hukuk,
- tıp,
- belirli teknik alan

için sınırlı kaynak taşıyabilir. Bu nedenle dil kaynak düzeyi:

- ürün,
- görev,
- zaman,
- alan

ile birlikte tanımlanmalıdır.

20. DİL ADALETİ PANELİNİN GÖREVLERİ

20.1. Düşük Örnekli Dilleri Fazla Örneklemek

Ana nüfus panelinde yeterli gözlem almayan diller için ek kullanıcılar sağlar.

20.2. Dil Bazlı Maddi Hata Dağılımını Ölçmek

Yalnız genel doğruluk değil:

- kimlik,
- kapsam,
- zaman,
- kaynak,
- tavsiye,
- Critical hata

sonuçlarını ayrı inceler.

20.3. Yanıt Dilini Doğrulamak

AI istem diliyle uyumlu cevap veriyor mu? Başka dile geçiyorsa bunun nedeni nedir?

20.4. Dil ve Locale Ayrımını İncelemek

Aynı dil farklı ülkelerde farklı yerel gerçeklik taşıyor mu? AI bu farkı doğru yönetiyor mu?

20.5. Kaynak ve Resmî Temsil Paritesini İncelemek

Varlığın resmî kayıtları bütün dillerde aynı maddi gerçekliği taşıyor mu? Düşük AI sonucu kaynağın kendisindeki dil açığından doğabilir.

20.6. İstem Eşdeğerliğini Sınamak

Çeviriler gerçekten aynı kullanıcı niyetini taşıyor mu?

20.7. Hakem Paritesini Sınamak

Farklı dillerdeki hakemler aynı normatif standardı uygulayabiliyor mu?

21. DİL ADALETİ İÇİN DÖRT KAPI

Bir dil sonucu AI ürününün dil performansına atfedilmeden önce dört kapı değerlendirilmelidir.

Kapı 1 — İstem Eşdeğerliği

Promptlar semantik olarak karşılaştırılabilir mi? Başarısızsa:

PROMPT_EQUIVALENCE_FAILURE

statüsü verilir.

Kapı 2 — Referans Paritesi

Resmî ve doğrulanmış gerçeklik kayıtları ilgili dil için yeterli mi? Başarısızsa:

REFERENCE_LANGUAGE_GAP

statüsü verilir.

Kapı 3 — Ölçüm Paritesi

Capture, oturum, arayüz ve kullanıcı şartları karşılaştırılabilir mi? Başarısızsa:

MEASUREMENT_PARITY_FAILURE

statüsü verilir.

Kapı 4 — Hakem Paritesi

Ana dil veya yeterli yerel uzmanlıkta hakemler ve aynı değerlendirme kuralları var mı? Başarısızsa:

ADJUDICATION_LANGUAGE_GAP

statüsü verilir. Dört kapı yeterince karşılanmadan: "AI ürünü bu dilde daha kötüdür." hükmü kesin biçimde kurulamaz. Dil açığı yine de GEO ekosistem bulgusudur. Fakat kaynağı doğru sınıflandırılmalıdır.

22. DİL EVRENİNİN KURULMASI

"Bütün diller" ifadesi sınırlandırılmadan kullanılamaz. Her denetim bir Kapsamdaki Dil Sicili oluşturmalıdır. Diller şu gruplardan seçilebilir:

22.1. Küresel Core Language Set

Ürün-uygun nüfusta geniş insan kitlesi tarafından kullanılan ana diller.

22.2. Varlık Market Language Set

Denetlenen varlığın:

- hizmet,
- içerik,
- müşteri,
- sözleşme

bakımından hedeflediği diller.

22.3. Product Support Language Set

AI ürününün resmî olarak desteklediği diller.

22.4. Low-Resource Sentinel Language Set

Nüfus panelinde zayıf kalan veya ürün davranışı bakımından risk taşıyan düşük kaynaklı diller.

22.5. High-Risk Language Set

Sağlık, hukuk, finans veya güvenlik bağlamında maddi kullanıcı etkisi taşıyan diller.

22.6. Experimental Language Set

Ürünün resmî olarak desteklemediği fakat gözlemsel veya araştırma amacıyla test edilen diller. Bu gruplar kesişebilir. Her dilin neden kapsama alındığı kaydedilmelidir.

23. DİL ADALETİ TAHSİSİ

Dil l için ana nüfus panelindeki hedef: nlP olsun. Adalet panelindeki ek hedef: nlF olsun. Toplanan toplam dil gözlemi:

```
n_l^T = n_l^P + n_l^F
```

olur. Küresel nüfus hesabında dilin toplam ağırlığı: Wl olarak korunur. Her gözleme düşen normalize ağırlık:

```
w_{i,l} = W_l / n_l^T
```

şeklinde düşünülebilir. Böylece dil fazla örneklenir. Fakat küresel skorda gerçek nüfus payını aşmaz.

24. DİL ADALETİ PANELİNİN KÜRESEL SKORA KATILIMI

Dil Adaleti gözlemleri iki yöntemle kullanılabilir.

24.1. Ayrı Tanısal Analiz

Adalet paneli yalnız dil skorlarını ve hata türlerini üretir. Ana Population Panel küresel skoru değiştirmez. Bu yöntem daha basittir.

24.2. Birleşik Yeniden Ağırlıklı Analiz

Population ve Language Fairness gözlemleri birlikte kullanılır. Her gözlem gerçek hedef dil nüfus ağırlığına göre yeniden ağırlıklandırılır. Bu yöntem bilgi hacmini artırabilir. Ancak:

- çift sayım,
- çok dilli kullanıcı,
- panel bağımlılığı,
- ağırlık değişkenliği

dikkate alınmalıdır. Kullanılan yöntem sonuçlar görülmeden önce tanımlanmalıdır.

25. DİL KAPSAM METRİKLERİ

25.1. Dil Gözlem Kapsamı

En az bir geçerli gözleme sahip kapsamdaki dillerin oranıdır.

```
LOC = |{l : n_l^obs > 0}| / |L_scope|
```

25.2. Dil Tahmin Kapsamı

Yeterli örneklem ve kalite şartlarına sahip dillerin oranıdır.

```
LEC = |{l : n_{l,eff} ≥ m_l}| / |L_scope|
```

25.3. Nüfus Ağırlıklı Dil Kapsamı

Yeterli biçimde gözlenen dil nüfusunun toplam hedef dil nüfusuna oranıdır. Çok dilli kullanıcıların çifte sayılmasını önleyecek ana denetim dili veya bölünmüş kişi ağırlığı gerektirir.

25.4. Ürün Destek Kapsamı

Kapsamdaki dillerin kaçında ürünün resmî destek sağladığını gösterir. Dil desteği ile temsil doğruluğu ayrı tutulur.

26. DİL EŞİTSİZLİĞİ METRİKLERİ

CANDIDATE METRICS — ADAY METRİKLER

Kesin metrik seti pilot ve dış incelemeyle kalibre edilmelidir.

26.1. Dil Sonucu

Her yeterli dil hücresi için: θl hesaplanır.

26.2. En Yüksek–En Düşük Farkı

```
DG_max = max_l θ_l − min_l θ_l
```

Kolay anlaşılır. Fakat küçük ve oynak tek hücreye duyarlıdır.

26.3. Sağlam Yüzdelik Farkı

```
DG_{90-10} = Q_{0,90}(θ_l) − Q_{0,10}(θ_l)
```

Aşırı uç hücrelere daha az duyarlıdır.

26.4. Dil Tabanı

Yeterli örneklem taşıyan en düşük dil sonucudur.

```
LF_min = min_{l∈L_est} θ_l
```

Bu değer küçük hücre belirsizliğiyle birlikte verilmelidir.

26.5. Sağlam Dil Tabanı

Yeterli dillerin alt yüzdelik veya alt grup ortalamasıdır. Örnek:

```
LF_10 = Q_{0,10}(θ_l)
```

26.6. Dil Critical Hata Oranı

Her dil için:

```
CR_l = doğrulanmış Critical hata içeren geçerli yanıt / geçerli dil yanıtı
```

Küresel ortalamadan ayrı yayımlanır.

26.7. Dil Bilinmeyen Oranı

```
U_l = UNKNOWN veya çözülemeyen yanıt / ilgili dil gözlemi
```

Yüksek bilinmeyen oranı düşük skor kadar önemli olabilir.

27. EŞİT DİL SKORU VE NÜFUS AĞIRLIKLI SKOR

Bütün dilleri eşit ağırlıklandıran tanısal sonuç üretilebilir:

```
G_equal-language = (1/L)Σ_{l=1}^L θ_l
```

Bu sonuç: "Ortalama test edilen dil deneyimi" sorusuna yaklaşır. Dünya nüfusundaki ortalama insan deneyimini temsil etmez. Nüfus ağırlıklı dil sonucu:

```
G_population-language = Σ_l W_lθ_l
```

ise nüfus dağılımını korur. İki skor farklıdır. Birbirinin yerine kullanılamaz. Dil adaleti kartında ikisi birlikte gösterilebilir.

28. DİL ADİLİĞİ EŞİT SONUÇ ZORUNLULUĞU DEĞİLDİR

İki dilde aynı skorun oluşmaması otomatik ihlal değildir. Fark şu nedenlerle açıklanabilir:

- Gerçek yerel hizmet farkı
- Ülkeye özgü hukuk
- Ürünün resmî destek kapsamı
- Kaynak erişimi
- Kullanıcı niyeti
- İstem alanı

Ancak farkın kaynağı açıklanamıyorsa veya düşük kaynaklı diller sistematik olarak:

- yanlış kimlik,
- yanlış yetki,
- maddi eksiklik,
- yüksek Critical hata

taşıyorsa dil adaleti sorunu vardır. NOMOS dil adaletini şöyle tanımlar:

> *Eşdeğer maddi koşullarda, dilin kendisinin sistematik yanlış temsil nedeni olmaması.*

29. DİL VE ÜLKE KESİŞİMİ

Dil skoru, ülkelerden bağımsız değildir. Aynı dil:

- farklı ülkelerde,
- farklı ürün erişimi,
- farklı yerel kaynak,
- farklı hukuk,
- farklı fiyat

bağlamında kullanılabilir. Bu nedenle gerektiğinde ülke–dil hücresi: (c,l) ayrı incelenmelidir. Örnek:

- Almanya × Türkçe
- Türkiye × Türkçe
- Birleşik Krallık × Türkçe

aynı dil ama farklı locale ve pazar bağlamı taşıyabilir. Dil adaleti yalnız ülkenin resmî diliyle ölçülemez.

30. ÜLKE GÖZCÜ VE DİL ADALETİ PANELİNİN KESİŞİMİ

Bir gözlem aynı anda:

- küçük ülke gözcü gözlemi,
- düşük kaynaklı dil gözlemi

olabilir. Bu durumda kayıt şu alanları taşımalıdır:

PRIMARY_PANEL_ROLE

SECONDARY_DIAGNOSTIC_TAGS

nüfus ağırlığı Sentinel statüsü Dil Adaleti statüsü Aynı gözlem iki ayrı panelde tam ağırlıkla sayılmaz. Tek gözlemdir. Birden fazla tanısal etiketi vardır.

31. SENTINEL VE DİL PANELİNDE KATILIMCI DEĞİŞTİRME

Katılımcı yalnız protokol ihlali nedeniyle değiştirilebilir. Aşağıdaki sonuçlar değiştirme nedeni değildir:

- Yanlış AI cevabı
- Ret
- Başka dilde cevap
- Critical hata
- Varlık karışıklığı
- Olumsuz tavsiye

Bunlar ölçüm sonucudur. Sentinel panelin amacı özellikle beklenmeyen olayları bulmaktır. Beklenmeyen olay çıktığı için kullanıcıyı değiştirmek panelin amacını yok eder.

32. DİL HAKEMLİĞİ

Dil Adaleti Paneli'nde hakemlik en azından şu koşulları taşımalıdır:

- İstem ve cevabı doğal veya yüksek yeterlilikte anlayabilme
- Yerel terminoloji bilgisi
- Kanıt kaynaklarını değerlendirebilme
- Aynı normatif değerlendirme rehberini kullanma
- Model ve mümkünse marka çıkarından kör olma
- Çıkar çatışmasını açıklama
- Anlaşmazlıkta üçüncü hakem yolu

Hakem bulunamayan dil:

ADJUDICATION_LANGUAGE_GAP

statüsü taşımalıdır. Otomatik çeviriyle İngilizceye çevrilip bütün nüanslar doğru kabul edilemez. Makine çevirisi yardımcı araç olabilir. Nihai yerel anlam doğrulaması değildir.

33. KAYNAK PARİTESİ

Varlığın İngilizce sitesinde:

- kapsam,
- fiyat,
- lisans,
- sınırlar,
- kanıtlar

bulunuyor; başka dil sürümünde bunlar yoksa düşük AI sonucu yalnız model sorunu olmayabilir. Üç ayrı sonuç gerekir:

- Varlık Language Source Gap
- AI Product Language Representation Gap
- Combined User-Facing Gap

Kullanıcı açısından nihai sorun yine gerçektir. Fakat düzeltme sorumluluğu farklı olabilir:

- varlık,
- AI sağlayıcısı,

her ikisi.

34. İSTEM PARİTESİ

Semantik olarak eşdeğer olmayan promptlar farklı cevap üretir. Örnek: Bir dilde: “Bu şirket ne iş yapar?” Başka dilde: “Bu şirket güvenilir ve önerilebilir midir?” sorulursa sonuçlar karşılaştırılamaz. İstem Anayasası sonraki bölümde ayrıntılandırılacaktır. Bu bölümün hükmü şudur:

> *İstem eşdeğerliği doğrulanmadan dil skoru farkı AI dil performansına atfedilemez.*

35. YANIT DİLİ SAPMASI

Kullanıcı Türkçe istem verir. AI İngilizce cevap üretir. Bu sonuç:

- maddi olgular doğru olsa bile,
- kullanıcı deneyimini,

- erişilebilirliği,
- comprehension'ı

etkileyebilir. Yanıt dili sapması ayrı kaydedilmelidir:

MATCHED_LANGUAGE

PARTIAL_CODE_SWITCH

UNREQUESTED_LANGUAGE_SWITCH

UNUSABLE_LANGUAGE_OUTPUT

UNKNOWN

Kod değiştirme her zaman hata değildir. Örneğin teknik terimlerin İngilizce kullanımı normal olabilir. Yanıtın ana kullanılabilirliğini bozuyorsa maddidir.

36. YAZI SİSTEMİ VE TRANSLİTERASYON

Aynı dil farklı yazı sistemlerinde kullanılabilir. Bir varlığın:

- adı,
- ürün adı,
- hukuki kimliği

transliterasyon sırasında başka varlıkla karışabilir. Dil Adaleti Paneli ilgili olduğu ölçüde:

- yazı sistemi,
- transliterasyon,
- resmî yerel ad,
- kullanıcı tarafından kullanılan ad

alanlarını kaydetmelidir. Yazı sistemi farkı yalnız biçimsel sayılmamalıdır. Varlık çözümlemesini etkiliyorsa maddidir.

37. SENTETİK ÜLKE GÖZCÜ ÇEVRİMİ

SENTETİK METODOLOJİ GÖSTERİMİ / Aşağıdaki ürün, ülkeler, sayılar ve sonuçlar kurgusaldır. Orion AI adlı sentetik ürünün Native Reach sicilinde: 120 uygun ülke bulunsun. Ana Population Panel, nüfus ağırlığı nedeniyle ilk dalgada yalnız 38 ülkeden geçerli gözlem üretmiş olsun.

37.1. İlk Dalga Kapsamı

Uygun ülke sayısı: 120 En az bir gözlem bulunan ülke: 38

```
COC_count = 38/120 = %31,7
```

Bu 38 ülke uygun nüfusun yüzde 97'sini taşısın.

```
COC_pop = %97
```

Sonuç: Ülke sayısı kapsamı düşük, nüfus kapsamı yüksektir.

37.2. Dört Dalgalı Gözcü Çevrimi

Kalan 82 ülke dört dalgaya, önceden dondurulmuş rastgele ve risk tabanlı programla dağıtılsın. Her ülke çevrim içinde en az bir kez gözlensin. Dördüncü dalga sonunda:

```
COC_count = 120/120 = %100
```

olur. Fakat ülke tahmini için yeterli örnekleme yalnız 14 ülkede ulaşılmış olsun.

CEC = 14/120 = %11,7

Doğru kamu kaydı: Ülke gözlem kapsamı: %100 / Nüfus gözlem kapsamı: %100'e yakın / Ülke tahmin kapsamı: %11,7 Yanlış kamu kaydı: "120 ülkenin tamamı güvenilir ülke skoruna sahiptir."

37.3. Küçük Ülke Critical Alarmı

Sentinel gözleminde küçük J ülkesindeki AI cevabı denetlenen varlığı: lisanslı hukuk firması olarak yanlış tanımlasın. Gerçekte varlık hukuk hizmeti vermiyor olsun. Tek gözlem: J ülkesinin hata oranını göstermez, Critical olay adayını gösterir. Protokol: Gözlem geçerliliğini doğrular. İstem ve referans kaydını inceler. Yeni bağımsız kullanıcılarla tekrar yapar. Olay sürerse hedefli ülke örneklemi açar. Critical uygunluk etkisini ayrı değerlendirir. Tek olay küçüktür. Potansiyel zarar küçük olmak zorunda değildir.

38. SENTETİK DİL ADALETİ GÖSTERİMİ

SENTETİK ÖRNEK — GERÇEK AI VEYA ŞİRKET PERFORMANSI DEĞİLDİR

Sentetik Apple.com Core Identity sınama'nda beş dil bulunsun. Ana Population Panel tahsisi:

Dil	Nüfus ağırlığı	Ana panel gözlemi
L1	%55	550
L2	%25	250
L3	%10	100
L4	%6	60
L5	%4	40
Toplam	%100	1.000

Sentetik dil sonuçları:

Dil	Geçiş oranı
L1	%95
L2	%92
L3	%88
L4	%65
L5	%58

Nüfus ağırlıklı küresel sonuç:

```
0,55(95)+0,25(92)+0,10(88)+0,06(65)+0,04(58)=90,27
```

Küresel skor yüksektir. Fakat iki dilde ağır bozulma vardır.

38.1. Dil Adaleti Fazla Örneklemesi

L4'e 140, L5'e 160 ek gözlem alınsın. Yeni gözlem sayıları:

Dil	Toplam gözlem
L1	550
L2	250
L3	100
L4	200
L5	200
Toplam	1.300

Ham 1.300 gözlem ortalaması yaklaşık: olabilir. Bu küresel nüfus skoru değildir. Çünkü L4 ve L5 fazla örneklenmiştir. Nüfus ağırlıkları korunduğunda küresel sonuç yine yaklaşık: olur. Fazla örnekleme küresel ortalamayı değiştirmedi. L4 ve L5 sonuçlarının hassasiyetini artırdı.

38.2. Dil Adaleti Sonuç Kartı

Sentetik sonuç: Nüfus ağırlıklı küresel skor: 90,27 Eşit dil skoru:

```
(95+92+88+65+58)/5=79,6
```

En yüksek–en düşük dil farkı: 37 puan En düşük dil: 58 Alt iki dil ortalaması: 61,5 Düşük kaynaklı dillerde Critical hata oranı: ayrıca raporlanır İstem eşdeğerliği: doğrulanmış varsayılsın Referans paritesi: doğrulanmış varsayılsın Doğru yorum:

> *Küresel nüfus temsili yüksek görünmektedir; ancak dil adaleti ciddi biçimde zayıftır.*

Yanlış yorum: "Küresel skor 90 olduğu için bütün dillerde GEO başarılıdır."

39. DİL BOZULMASININ KAYNAĞINI AYIRAN SENTETİK VAKA

Üç dil düşünelim.

Dil A

İstem eşdeğer Resmî kaynak tam Yerel hakem yeterli AI cevabı yanlış Sonuç: AI ürününe ait dil temsil sorunu adayı.

Dil B

İstem çevirisi maddi olarak farklı AI cevabı prompta uygun Sonuç: İstem Anayasası kusuru.

Dil C

Resmî dil sayfasında eski fiyat AI eski fıyatı aktarıyor İngilizce kaynak güncel Sonuç: Varlığın dil kaynak paritesi sorunu ve kullanıcıya ulaşan birleşik temsil açığı. Üç dilin düşük skoru aynı görünse bile düzeltme sahibi farklıdır.

40. ÜLKE VE DİL PANELLERİNİN KAMU SONUÇ KARTI

Aday kamu sonuç kartı şu alanları taşımalıdır:

Ülke Alanları

Native Reach uygun ülke sayısı En az bir gözlem bulunan ülke sayısı Ülke gözlem kapsamı Nüfus ağırlıklı gözlem kapsamı Ülke tahmin kapsamı Country Sentinel çevrim süresi Açık Sentinel alarmları OUTSIDE_NATIVE_REACH ülkeler NOT OBSERVED ülkeler Ölçüm tarihi

Dil Alanları

Kapsamdaki dil sayısı Gözlenen dil sayısı Tahmin üretilebilen dil sayısı Resmî desteklenen diller Deneysel diller Nüfus ağırlıklı küresel dil sonucu Eşit dil tanı sonucu Dil tabanı Dil farkı Critical hata oranları

REFERENCE_LANGUAGE_GAP

PROMPT_EQUIVALENCE_FAILURE

NOT TESTED diller Dil paneli sürümü Bu kart tek bir "çok dilli" rozetin taşıyamayacağı gerçekliği görünür kılar.

41. ZORUNLU NORMATİF HÜKÜMLER

CH07-N01

Country Sentinel ve Language Fairness panelleri ana Population Panel'den amaç, örneklem ve ağırlık bakımından ayrı tanımlanmalıdır.

CH07-N02

Bir gözcü gözlemi ülke nüfus skoru olarak kullanılamaz.

CH07-N03

Bir ülke için gözlem bulunması ile ülke tahmini üretilebilmesi ayrı statülerdir.

CH07-N04

Ülke sayısı kapsamı ve nüfus ağırlıklı ülke kapsamı birlikte raporlanmalıdır.

CH07-N05

Ülke tahmin kapsamı gözlem kapsamından ayrı gösterilmelidir.

CH07-N06

Uygun her ülke, tanımlı gözetim çevriminde sıfırdan büyük gözlenme olasılığı taşımalıdır; kapsam dışı ülkeler ayrıca gösterilmelidir.

CH07-N07

Gözcü çevrimi ve ülke sıklık sınıfları AI yanıtları görülmeden önce sürümlenmelidir.

CH07-N08

Düşük skorlu ülke, sonuç görüldükten sonra gözcü çevriminden çıkarılamaz.

CH07-N09

Gözcü alarmı otomatik ülke başarısızlığı değildir; doğrulama ve eskalasyon tetikleyicisidir.

CH07-N10

Tekil Critical gözcü olayı yaygınlık tahmini olmadan da ayrı uygunluk incelemesi başlatabilir.

CH07-N11

Ülke Gözcü Paneli gözlemleri ana Population Panel'de ham eşit ağırlıkla kullanılamaz.

CH07-N12

Country Sentinel gözlemlerinin ana skora katılması durumunda seçilme olasılıkları ve ağırlıkları açıkça modellenmelidir.

CH07-N13

Dil adaleti kelime kelime aynı yanıt zorunluluğu olarak tanımlanamaz.

CH07-N14

Dil adaleti, eşdeğer maddi koşullarda karşılaştırılabilir doğruluk, sınır, kaynak ve kullanıcı uygunluğu gerektirir.

CH07-N15

Dil ile ülke, dil ile locale ve dil ile yazı sistemi birbirinin yerine kullanılamaz.

CH07-N16

Kapsamdaki dil evreni ve her dilin kapsama alınma gerekçesi ölçümden önce açıklanmalıdır.

CH07-N17

"Bütün diller" iddiası yalnız gerçekten bütün ilgili dil ve kullanıcı yüzeyleri ölçülmüşse kullanılabilir; aksi hâlde "kapsamdaki diller" denmelidir.

CH07-N18

Resmî desteklenmeyen dildeki gözlem ana desteklenen ürün performansına açıklamasız biçimde eklenemez.

CH07-N19

Düşük kaynaklı dillerin fazla örneklenmesi küresel nüfus ağırlığını artırmamalıdır.

CH07-N20

Dil Adaleti Paneli gözlemleri küresel sonuca dâhil ediliyorsa gerçek hedef dil nüfusu ağırlıklarına geri döndürülmelidir.

CH07-N21

Aynı çok dilli kullanıcı küresel nüfus ağırlığında birden fazla tam insan gibi sayılamaz.

CH07-N22

İstem eşdeğerliği doğrulanmadan dil skoru farkı AI dil performansına kesin olarak atfedilemez.

CH07-N23

Referans paritesi değerlendirilmeden düşük dil skoru yalnız AI ürününe yüklenemez.

CH07-N24

Yerel hakem yeterliliği bulunmadan dil doğruluk kararı kesinleştirilemez.

CH07-N25

Yanıt dili sapması ve kod değiştirme ayrı kayıt edilmelidir.

CH07-N26

Meşru yerel gerçeklik farkı dil adaletsizliği olarak sınıflandırılamaz.

CH07-N27

Açıklanamayan maddi dil farkı yalnız kişiselleştirme veya stil farkı olarak kapatılamaz.

CH07-N28

Küresel nüfus skoru, dil adaleti sonucunun yerine geçemez.

CH07-N29

Eşit dil skoru, nüfus ağırlıklı küresel skor gibi sunulamaz.

CH07-N30

Dil bazlı Critical hata oranları genel ortalamadan ayrı gösterilmelidir.

CH07-N31

PROMPT_EQUIVALENCE_FAILURE, REFERENCE_LANGUAGE_GAP, MEASUREMENT_PARITY_FAILURE ve ADJUDICATION_LANGUAGE_GAP ayrı statüler olarak korunmalıdır.

CH07-N32

Country Sentinel ve Language Fairness panellerinde katılımcı değiştirme kararı AI cevabının doğruluğuna veya markaya yararına dayanamaz.

CH07-N33

Aynı gözlem birden fazla tanısal etikete sahip olabilir; ancak küresel tahminde açıklamasız biçimde birden fazla tam ağırlık alamaz.

CH07-N34

Ülke ve dil paneli kayıtlarının hesap verebilir insan veya kurum sahibi bulunmalıdır.

42. BAŞARISIZLIK BİÇİMLERİ

CH07-F01 — BİR KİŞİLİK ÜLKE SKORU

Tek gözcü gözlemi ülke GEO skoru olarak yayımlanır.

CH07-F02 — SEMBOLİK ÜLKE KAPSAMASI

Her ülkeye bir kişi verilir ve "dünya nüfusuna orantılı" denir.

CH07-F03 — ÜLKE SAYISIYLA NÜFUS KAPSAMINI KARIŞTIRMAK

Az sayıda büyük ülke gözlendiği hâlde yalnız nüfus kapsamı; çok sayıda küçük ülke gözlendiği hâlde yalnız ülke sayısı gösterilir.

CH07-F04 — GÖZLEM KAPSAMINI TAHMİN KAPSAMI SAYMAK

En az bir gözlemi bulunan bütün ülkeler skorlu ülke kabul edilir.

CH07-F05 — SENTINEL GÖZLEMİNİ ANA SKORDA ÇİFTE SAYMAK

Aynı gözlem hem Sentinel hem Population Panel'de tam ağırlık alır.

CH07-F06 — SENTINEL ALARMINI YAYGINLIK İDDİASINA DÖNÜŞTÜRMEK

Tek olay bütün ülke kullanıcılarına genellenir.

CH07-F07 — CRITICAL SENTINEL OLAYINI ÖNEMSİZLEŞTİRMEK

"Yalnız bir kullanıcı" denilerek yüksek zarar riski incelenmez.

CH07-F08 — DÜŞÜK SKORLU ÜLKEYİ ÇEVRİMDEN ÇIKARMAK

Gözcü takvimi sonuçlara göre değiştirilir.

CH07-F09 — YALNIZ TİCARİ PAZARLARI GÖZLEMEK

Ürün erişimi bulunan fakat ticari açıdan önemsiz görülen ülkeler tamamen dışlanır; sonuç küresel olarak sunulur.

CH07-F10 — ÜLKE İLE DİLİ AYNI SAYMAK

Ülkenin resmî dili bütün kullanıcıların dili kabul edilir.

CH07-F11 — TÜM DİLLER İDDİASI

Sınırlı birkaç dil ölçülür; sonuç "bütün dillerde" diye yayımlanır.

CH07-F12 — DÜŞÜK KAYNAKLI DİLİ SİLMEK

Örneklem küçük olduğu için dil rapordan tamamen çıkarılır.

CH07-F13 — DÜŞÜK KAYNAKLI DİLİ GLOBALDE AŞIRI BÜYÜTMEK

Fazla örneklenen dil ham örneklem payıyla küresel skora girer.

CH07-F14 — EŞİT DİL SKORUNU NÜFUS SKORU SAYMAK

Her dile eşit ağırlık veren tanısal skor dünya nüfusunun ortalama deneyimi gibi sunulur.

CH07-F15 — NÜFUS SKORUNU DİL ADALETİ SAYMAK

Büyük dillerin ağırlığıyla yükselen küresel sonuç, küçük dillerdeki bozulmayı gizler.

CH07-F16 — İSTEM ÇEVİRİ HATASINI AI HATASI SAYMAK

Semantik olarak farklı istemlerden çıkan fark modele yüklenir.

CH07-F17 — KAYNAK DİL AÇIĞINI AI HATASI SAYMAK

Yerelleştirilmiş resmî bilgi bulunmadığı hâlde bütün sorumluluk AI ürününe verilir.

CH07-F18 — AI HATASINI KAYNAK EKSİKLİĞİYLE SAVUNMAK

Aynı ve yeterli kaynaklar bulunduğu hâlde düşük kaynaklı dildeki yanlışlar yalnız site eksikliği diye kapatılır.

CH07-F19 — HAKEM DİL YETERSİZLİĞİNİ ÜRÜN BAŞARISIZLIĞI SAYMAK

Hakem cevabı yanlış anlamış olsa da sonuç kesinleştirilir.

CH07-F20 — MAKİNE ÇEVİRİSİNİ NİHAİ HAKEM SAYMAK

Yerel anlam kaybı kontrol edilmez.

CH07-F21 — KOD DEĞİŞTİRMEYİ OTOMATİK HATA SAYMAK

Doğal teknik terimler veya kabul edilebilir dil karışımı başarısızlık yapılır.

CH07-F22 — KULLANILAMAZ DİL SAPMASINI STİL SAYMAK

Sistem kullanıcının anlayamayacağı başka dilde cevap verdiği hâlde sonuç başarılı kabul edilir.

CH07-F23 — YAZI SİSTEMİ FARKINI GÖRMEZDEN GELMEK

Varlık karışıklığına yol açan transliterasyon hatası biçimsel sayılır.

CH07-F24 — ÇOK DİLLİ KULLANICIYI ÇİFTE SAYMAK

Aynı kişi bütün dillerde ayrı nüfus birimi olur.

CH07-F25 — MEŞRU YEREL FARKI ADALETSİZLİK SAYMAK

Fiyat veya hukuk farkı gerçek yerel koşula dayandığı hâlde dil yanlılığı ilan edilir.

CH07-F26 — AÇIKLANAMAYAN FARKI KİŞİSELLEŞTİRME SAYMAK

Eşdeğer koşullardaki maddi çelişki sorgulanmaz.

CH07-F27 — YETERSİZ DİL HÜCRESİNDEN KESİN SKOR

Birkaç gözlemle dil sıralaması yapılır.

CH07-F28 — DİL PANELİNİ SONUÇTAN SONRA KURMAK

Yalnız düşük veya yüksek çıkan diller sonradan seçilir.

CH07-F29 — SADECE YÜKSEK KAYNAKLI DİLLERİ ÖLÇMEK

Sonuç "çok dilli" diye sunulur.

CH07-F30 — TEST EDİLMEYEN DİLLERİ SAKLAMAK

Kamu sonuç kartında NOT TESTED alanı bulunmaz.

43. DENETİM PROSEDÜRÜ

Adım 1 — Ürün-Ülke Uygunluk Sicilini Dondur

Her ülke için:

- Native Reach
- ürün erişimi
- denetim dili
- kullanıcı yüzeyi

statüsü kaydedilir.

Adım 2 — Ana Population Panel Ülke Kapsamını Çıkar

Hangi ülkelerde ana panel gözlemi bulunduğu belirlenir.

Adım 3 — Sentinel Ülke Evrenini Belirle

Ana panelde yetersiz kalan ve CE-3 statüsü taşıyan ülkeler belirlenir.

Adım 4 — Gözcü Çevrimini Kur

Ülkeler:

- sıklık sınıfı,
- dalga,
- dil/locale,
- kullanıcı yüzeyi

bakımından atanır.

Adım 5 — Gözcü Katılımcılarını Önceden Seç

Ana ve yedek listeler sonuçlar görülmeden önce oluşturulur.

Adım 6 — Gözcü Sonuçlarını Sınıflandır

erişim, kimlik, dil, maddi gerçek, Critical olay statüleri atanır.

Adım 7 — Alarm ve Eskalasyon Uygula

Alarm türüne göre:

- tekrar,
- hedefli örneklem,
- referans incelemesi,
- uygunluk incelemesi

başlatılır.

Adım 8 — Ülke Kapsam Metriklerini Hesapla

ülke sayısı kapsamı, nüfus kapsamı, ülke tahmin kapsamı ayrı hesaplanır.

Adım 9 — Dil Sicilini Dondur

Her dil için:

- kapsam nedeni,
- destek statüsü,
- yazı sistemi,
- locale,
- kullanıcı evreni

kaydedilir.

Adım 10 — Dil Adaleti Fazla Örneklemesini Kur

Ana nüfus panelinde yetersiz kalan diller için ek hedefler belirlenir.

Adım 11 — Dört Dil Kapısını Kontrol Et

İstem eşdeğerliği Referans paritesi Ölçüm paritesi Hakem paritesi değerlendirilir.

Adım 12 — Dil Gözlemlerini Puanla

Her dil için:

- doğruluk,
- maddi eksiklik,
- Critical hata,
- bilinmeyen,
- yanıt dili sapması

kaydedilir.

Adım 13 — Küresel ve Tanısal Ağırlıkları Ayır

Nüfus ağırlıklı sonuç ile eşit dil tanı sonucu ayrı hesaplanır.

Adım 14 — Dil Eşitsizliğini Hesapla

dil tabanı, sağlam dil farkı, Critical hata farkı, bilinmeyen oranı hesaplanır.

Adım 15 — Ülke–Dil Kesişimlerini İncele

Maddi locale farkları veya çelişkileri belirlenir.

Adım 16 — Kamu Sonuç Kartını Oluştur

Kapsam, tahmin yeterliliği, alarmlar ve test edilmeyen alanlar görünür hâle getirilir.

44. GEREKLİ KANIT

AI ürün ve kullanıcı yüzeyi kaydı Native Reach ülke sicili Uygun ülke sayısı Ülke erişim statüleri Ana Population Panel ülke tahsisleri Country Sentinel evreni Gözcü çevrimi Gözcü sıklık sınıfları Ülke–dalga atama matrisi Ülke–dil atamaları Sentinel ana ve yedek listeleri Ülke gözlem sonuçları Sentinel alarmları Eskalasyon kayıtları Hedefli ülke örneklemleri Ülke gözlem kapsamı Nüfus ağırlıklı ülke

kapsamı Ülke tahmin kapsamı Kapsamdaki dil sicili Dil destek statüleri Yazı sistemi ve locale kayıtları Ana dil örneklem sayıları Dil Adaleti fazla örneklem sayıları Dil nüfus ağırlıkları Çok dilli kullanıcı kayıtları İstem eşdeğerlik kayıtları Referans paritesi kayıtları Ölçüm paritesi kayıtları Yerel hakem yeterlilik kayıtları Dil sonuçları

Dil Critical hata oranları Yanıt dili sapmaları Dil eşitsizliği metrikleri Test edilmeyen diller Eşit dil ve nüfus ağırlıklı sonuçlar Panel ağırlıkları Çift sayım kontrolleri Kamu sonuç kartı Panel sürümleri Sorumlu insan veya kurum

45. DENETİM KONTROL LİSTESİ

Country Sentinel Panel'in amacı ana nüfus panelinden ayrılmış mı? Hangi ülkeler Sentinel kapsamına alındı? Ülkelerin uygunluk ve erişim statüsü güncel mi? Her uygun ülkenin gözetim çevriminde pozitif gözlenme ihtimali var mı? Gözcü çevrimi sonuçlardan önce dondurulduğu mu? Sentinel gözlemi ülke skoru gibi kullanılmış mı? Ülke gözlem kapsamı ve ülke tahmin kapsamı ayrılmış mı? Nüfus ağırlıklı ülke kapsamı ayrıca gösterilmiş mi? Sentinel gözlemleri ana skorda çifte sayılmış mı? Critical Sentinel alarmı doğrulandı mı? Tekil alarm yaygınlık iddiasına dönüştürülmüş mü? Alarm nedeniyle hedefli örneklem açılmış mı? Kapsamdaki dil evreni önceden tanımlandı mı? "Bütün diller" iddiası gerçek kapsamla uyumlu mu? Dil, yazı sistemi ve locale ayrıldı mı?

Ürünün dil destek statüsü kaydedildi mi? Düşük kaynaklı diller neden seçildi? Dil Adaleti örnekleri küresel skorda geri ağırlıklandırıldı mı? Aynı çok dilli kullanıcı çifte sayıldı mı? İstem eşdeğerliği doğrulandı mı? Resmî kaynaklar diller arasında maddi olarak eşit mi? Hakemler ilgili dili yeterli düzeyde biliyor mu? Makine çevirisi nihai hakem gibi kullanılmış mı? Yanıt dili sapmaları kaydedilmiş mi? Meşru yerel fark ile dil adaletsizliği ayrıldı mı? Dil tabanı ve dil farkı yayımlanmış mı? Dil bazlı Critical hata oranları görünür mü? REFERENCE_LANGUAGE_GAP ve diğer ölçüm kusurları ayrılmış mı? Test edilmeyen diller açık mı? Kamu skoru dil adaleti sorununu gizliyor mu?

Aynı gözlem birden fazla panelde tam ağırlık almış mı? Panel kayıtlarının hesap verebilir sahibi belli mi?

46. İTİRAZLAR VE CEVAPLAR

İtiraz 1 — "Her küçük ülkeye bir kişi vermek yine de hiç vermemekten iyi değil mi?"

Erken gözlem ve coğrafi kapsam için evet. Fakat bu kişi ülkenin oranını ölçmez. Doğru ad:

Sentinel gözlemi

olmalıdır.

İtiraz 2 — "Bir kişi ülke skoru üretmiyorsa neden gönderiyoruz?"

Çünkü tekil gözlem:

- erişim sorununu,
- varlık karışıklığını,
- yanlış dili,
- Critical olayı

ilk kez görünür hâle getirebilir. Gözcü paneli erken uyarı sistemidir.

İtiraz 3 — "Her ülkeyi her dalgada gözlemlemek zorunda mıyız?"

Hayır. Tanımlı bir gözetim çevrimi kullanılabilir. Yüksek riskli veya önemli ülkeler daha sık; küçük ve düşük riskli ülkeler dönüşümlü gözlenebilir. Takvim sonuçlardan önce belirlenmelidir.

İtiraz 4 — "Küçük ülkeler küresel skorda düşük ağırlık alıyorsa neden bu kadar önem veriyoruz?"

Çünkü nüfus payı düşük olabilir. Fakat:

- hukuk,
- dil,
- yerel kullanıcı zararı,
- coğrafi bütünlük

bakımından gerçek insan topluluklarıdır. Düşük ağırlık görünmezlik anlamına gelmemelidir.

İtiraz 5 — "Bütün dilleri eşit ağırlıklandırmak en adil yöntem değil mi?"

Eşit dil ağırlığı değerli bir tanısal bakış olabilir. Dünya nüfusundaki ortalama insan deneyimini temsil etmez. Nüfus skoru ve dil adaleti skoru ayrı yayımlanmalıdır.

İtiraz 6 — "Küçük dil küresel sonucu yalnız az etkiliyorsa düşük skoru neden önemli?"

Çünkü o dili kullanan insanlar için temsil sorunu gerçektir. Ayrıca ağır lisans, sağlık veya kimlik yanlışları nüfus payıyla otomatik olarak önemsizleşmez.

İtiraz 7 — "Düşük kaynaklı dilde kaynak azsa AI'yı suçlamak haksız değil mi?"

Yalnız AI'yı suçlamak haksız olabilir. Bu nedenle:

- kaynak paritesi,
- istem eşdeğerliği,
- hakem yeterliliği

ayrı incelenir. Kullanıcıya ulaşan yanlış yine gerçektir. Düzeltme sahibi doğru belirlenmelidir.

İtiraz 8 — "Farklı ülkelerde aynı dilde farklı cevap normal değil mi?"

Gerçek yerel koşullar farklıysa normal olabilir. Aynı maddi gerçek farklı ve açıklamasız biçimde değişiyorsa sorun olabilir. Dil ve locale birlikte incelenmelidir.

İtiraz 9 — "AI başka dilde doğru cevap veriyorsa neden başarısız olsun?"

Kullanıcı cevabı anlayamıyorsa temsil kullanılabilir değildir. Yanıt dili sapması doğrulukla birlikte ayrı değerlendirilmelidir.

İtiraz 10 — "Dil skorları için 200 kullanıcı bulmak çok pahalı olabilir."

Daha küçük örneklem kullanılabilir. Sonuç:

- keşif,
- geçici,
- yetersiz örneklem

statüsüyle yayımlanmalıdır. Maliyet, sahte kesinlik hakkı vermez.

İtiraz 11 — "Küresel skor zaten yüksekse ayrı dil kartı müşteriyi gereksiz yere korkutmaz mı?"

Maddi dil bozulmasını saklamak müşteriyi korumaz. Yanlış güven üretir. Kısa ve anlaşılır sonuç kartı kullanılabilir.

İtiraz 12 — "Tek bir Critical Sentinel olayı bütün uygunluğu düşürür mü?"

Otomatik cevap olayın türüne bağlıdır. Tekil olay ülke yaygınlığını göstermez. Fakat ağır zarar veya sahte yetki taşıyorsa Critical inceleme ve geçici tedbir başlatabilir. Karar Bölüm 15'te tanımlanacaktır.

49. BÖLÜMÜN ORTAK HÜKMÜ

Bir dünya skoru üretmek kolaydır. Dünya nüfusunu büyük ülkelerle doldurursunuz. Küçük ülkeleri küsurat içinde kaybedersiniz. Yüksek kaynaklı dilleri ölçersiniz. Düşük kaynaklı dilleri: "Örneklem yetersiz." diyerek çıkarırsınız. Sonra tek sayı yayımlarsınız. Bu sayı matematiksel olarak doğru hesaplanmış olabilir. Fakat dünyanın bütün temsil kusurlarını göstermeyebilir. NOMOS, her ülkeye eşit ağırlık vermez. Çünkü insanların nüfus payları eşit değildir. NOMOS, küçük ülkeyi de yok saymaz. Çünkü düşük nüfus var olmamak değildir. NOMOS, bütün dilleri küresel skorda eşit ağırlıklandırmaz. Çünkü bu dünya nüfusunun ortalama deneyimini bozabilir. NOMOS, küçük dili de görünmez bırakmaz. Çünkü küresel ağırlığı düşük olsa bile o dili kullanan insan için temsil hatası yüzde 100 gerçek olabilir.

Ülke Gözcü Paneli:

> *Nerede hiç bakmadığımızı gösterir.*

Dil Adaleti Paneli:

> *Küresel ortalamanın altında kimin kaldığını gösterir.*

Population Panel:

> *Dünya nüfusundaki ortalama deneyimi tahmin eder.*

Bu üçü birlikte çalıştığında küresel skor hem nüfusa sadık hem de adalet bakımından denetlenebilir olur. Bir ülkedeki tek kullanıcı ülkenin tamamı değildir. Ama o kullanıcı gerçek insandır. Bir dil küresel nüfusun küçük parçası olabilir. Ama o dildeki yanlış daha az yanlış değildir. Bir Critical olay nadir olabilir. Ama zarar yalnız frekansla ölçülmez. Bu nedenle NOMOS'un yedinci ölçüm yasası şudur:

> *Görünürlük hakkı ile nüfus ağırlığı aynı şey değildir.*

Sekizinci yasası şudur:

> *Ülke gözlemi, ülke tahmini değildir.*

Dokuzuncu yasası şudur:

> *Dil adaleti, bütün dilleri eşit yapmak değil; dili nedeniyle kimsenin maddi gerçeklikten sistematik biçimde uzaklaştırılmamasıdır.*

Onuncu yasası şudur:

> *Küresel ortalama, geride kalan dili ve ülkeyi görünmez kılıyorsa tek başına yeterli GEO sonucu değildir.*

NOMOS'un Bölüm 7 Emri

> *Bana dünya skorunu gösterirken hiç bakmadığın ülkeleri saklama.*
>
> *Küçük ülkeye bir kullanıcı gönderip o insanı ülkenin tamamı yapma.*
>
> *Ama küçük ülkeyi nüfusu az diye sonsuza kadar görmezden de gelme.*
>
> *Ülke gözlemini, ülke tahminini ve ülke uygunluğunu birbirinden ayır.*
>
> *Tek bir ağır yanlış gördüğünde bütün ülkeye hüküm verme. / Fakat "yalnız bir cevap" diyerek insan zararını da yok sayma.*
>
> *Büyük dillerdeki başarıyı küçük dillerin sessizliği üzerine kurma.*
>
> *Düşük kaynaklı dili fazla örnekleyebilirsin. / Fakat onu dünya nüfusunda olduğundan büyük gösterme.*
>
> *Eşit dil skorunu nüfus skoru; nüfus skorunu dil adaleti skoru yapma.*
>
> *Çeviri yanlışsa beni suçlama. / Kaynağın eksikse bunu da söyle. / Hakem dili anlamıyorsa kesin hüküm verme.*
>
> *Gerçek yerel farkı adaletsizlik; açıklanamayan maddi farkı yalnız stil sayma.*
>
> *Önce ülkeleri gözle. / Sonra hangilerinde tahmin yapabildiğini söyle. / Sonra dilleri fazla örnekle. / Sonra onları gerçek nüfus ağırlığına geri döndür. / Ve küresel ortalamanın altında kalan insanları görünür bırak.*

Bölümün Son Cümlesi

Gerçek küresel GEO ölçümü, yalnız dünyanın çoğunluğunun ne gördüğünü değil; küçük ülke ve dillerde kimin henüz görülmediğini, kimin yanlış görüldüğünü ve hangi sonucun yalnız tekil bir alarm olduğunu da açıkça göstermelidir.

Normatif Öz

> **GEO-1000 Population Panel, Ülke Gözcü Paneli ve Dil Adaleti Paneli amaç, örneklem, ağırlık ve kamu iddiası bakımından ayrı tutulmalıdır. Tek bir ülke gözcü gözlemi belirli ülke–ürün–dil bağlamında bir erişim durumunun, cevabın, anomalinin veya Critical olayın meydana geldiğini gösterebilir; ülke yaygınlık tahmini veya ülke GEO skoru olarak sunulamaz. Ülke sayısı kapsamı, nüfus ağırlıklı kapsam ve ülke tahmin kapsamı ayrı raporlanmalıdır. Dil adaleti, semantik olarak eşdeğer promptlarda ve karşılaştırılabilir koşullarda maddi temsili değerlendirmeli; kelime kelime aynı cevap istememeli ve meşru yerel farkları yok saymamalıdır. Fazla örneklenen diller küresel nüfus tahmininde gerçek hedef nüfus paylarına geri ağırlıklandırılmalıdır.**

BÖLÜM 08

Katılımcı Seçimi, Ağırlıklandırma ve Örneklem Yanlılığı

YÖNTEMSEL DAYANAKLAR / K04; K05; K06; K07; K08; K22

Bölüm Sınırı

Bölüm 5 hedef kullanıcı nüfusunu tanımladı. Bölüm 6 bu nüfusu ülke, dil ve kullanıcı yüzeylerine dağıtılan bir örneklem tasarımına dönüştürdü. Bölüm 7 ise ana nüfus panelinin görünmez bırakabileceği küçük ülkeler ve düşük kaynaklı diller için tamamlayıcı gözlem sistemlerini kurdu. Şimdi örneklem planındaki sayıların gerçek insanlara dönüşmesi gerekir:

> *Bu katılımcılar nereden bulunacak, nasıl seçilecek, hangi koşullarda panele alınacak ve hangi yanlılıklarla birlikte GEO skoruna girecektir?*

Bir araştırma kuruluşunun çevrim içi panelinden mi? Açık sosyal medya çağrısından mı? Üniversitelerden mi? Müşterinin kendi kullanıcı listesinden mi? Yerel araştırma ortaklarından mı? Yalnız AI ürünlerini sık kullanan kişilerden mi? Ürüne ilk kez erişenlerden mi? Markanın müşterilerinden mi? Rakip çalışanlarından mı? Ücretli plan sahiplerinden mi? Bir kişinin gerçekten o ülkede yaşadığı nasıl doğrulanacaktır? Birden fazla hesap kullanan aynı insan kaç kez sayılacaktır? Katılımcıya ne kadar ödeme yapılacaktır? Ödeme, görevi hızlı tamamlamayı veya belirli bir sonuç göndermeyi teşvik edecek midir? Katılımcı yanlış bir AI yanıtı yüklediğinde gözlem korunacak mıdır? Yoksa: "Görev yanlış yapılmış." denilerek yeni kullanıcıyla mı değiştirilecektir? Bin geçerli ekran görüntüsü toplanmış olabilir. Fakat katılımcıların:

- yüzde 60'ı yoğun AI kullanıcısı,
- yüzde 20'si teknoloji öğrencisi,
- yüzde 10'u denetlenen markanın müşterisi,
- yüzde 5'i şirket çalışanı,
- geri kalanı genel kullanıcı

ise sonuç hangi nüfusu temsil edecektir? Demografik ağırlıklandırma bu farklılığı tamamen düzeltebilir mi? Yalnız yaş, cinsiyet, ülke ve dil toplamlarına uyum sağlamak; ölçülmemiş teknoloji ilgisi, marka bilgisi, ücretli plan erişimi ve AI kullanım alışkanlığını da düzeltir mi? Hayır. Ağırlıklandırma değerlidir. Fakat örneklem seçiminin bütün kusurlarını sihirli biçimde yok edemez. Bu bölüm, daha önce tanımlanan denetim maddelerinin model, tarih, ülke, dil, sorgu kümesi, tekrar sayısı ve ölçüm kaydıyla test edilebilir olması gerektiği yaklaşımı gerçek katılımcı seçimine uygular. Aynı zamanda NOMOS'un temel talebini korur: Kanıt, sınır, bağlam ve zaman. Bu bölüm:

- katılımcı seçim kaynaklarını,
- olasılıklı ve olasılıksız panel ayrımını,
- uygunluk ve kimlik doğrulamasını,

- ülke, dil ve ürün erişim kontrolünü,
- katılımcı benzersizliğini,
- ödüllendirme ve yönlendirme riskini,
- panel koşullanmasını,
- görev atamasını,
- yanıtlamama ve panel kaybını,
- sahte veya otomatik katılımı,
- seçilim ve kapsama yanlılığını,
- ağırlıklandırmanın gücünü ve sınırını,
- örneklem kalite statülerini

tanımlar. Bu bölüm henüz:

- AI ürün sicilinin bütün teknik alanlarını,
- istemlerin tam çeviri ve sürüm sistemini,
- NOMOS Capture altyapısının ayrıntılı mimarisini,
- atomik iddia hakemliğini,
- nihai skor formülünü

tam olarak tanımlamaz. Bölüm 8'in temel sorusu şudur:

> *Bin geçerli kayıt topladığımızda, bu kayıtların gerçekten hangi insanlara ait olduğunu ve hangi insan nüfusunu temsil edebildiğini nasıl kanıtlarız?*

NOMOS Challenge

Bir çevrim içi araştırma panelinden 1.000 kişi satın alıyorsunuz. Panel sağlayıcısı size şunu söylüyor: "Ülke ve yaş kotalarınıza uygun 1.000 katılımcı hazır." Katılımcılar görevi tamamlıyor. Ekran görüntülerini yüklüyor. Ham sonuç:

> *Yüzde 90 doğru temsil.*

Raporunuzda şöyle yazıyorsunuz: "Dünya genelindeki kullanıcıların yüzde 90'ı doğru temsil görüyor." Fakat katılımcıların gerçek dağılımı şöyledir:

AI kullanım sıklığı	Hedef nüfus payı	Paneldeki pay
Düşük veya seyrek kullanım	%50	%10
Orta kullanım	%30	%30
Yoğun kullanım	%20	%60

Yoğun AI kullanıcıları:

- daha iyi istem alışkanlığına,
- daha fazla ücretli hesaba,
- daha yeni ürün yüzeylerine,
- kişiselleştirme bilgisine,
- teknolojiye daha yüksek aşinalığa

sahip olabilir. Aynı kilitlenmiş istem kullanılsa bile:

- hesap planı,
- arayüz,

- önceki kullanım,
- ürün erişimi

farklı olabilir. Sentetik olarak grup geçiş oranları şöyle olsun:

AI kullanım sıklığı	Geçiş oranı
Düşük veya seyrek kullanım	%72
Orta kullanım	%84
Yoğun kullanım	%96

Panelin ham sonucu:

```
0,10(72)+0,30(84)+0,60(96)=90
```

yani:

Yüzde 90

olur. Hedef nüfus ağırlıklarıyla sonuç:

```
0,50(72)+0,30(84)+0,20(96)=80,4
```

olur. Aynı 1.000 cevap. İki farklı sonuç:

- Ham panel sonucu: yüzde 90
- Hedef nüfus sonucu: yüzde 80,4

Şimdi ağırlıklandırdığınızı düşünün. Düşük kullanım grubundaki 100 kişinin her biri yüksek ağırlık alıyor. Yoğun kullanım grubundaki 600 kişinin her biri düşük ağırlık alıyor. Ham örneklem sayısı hâlâ 1.000. Fakat etkin örneklem büyüklüğü yaklaşık 349'a düşüyor. Yani 1.000 dosya toplamış olabilirsiniz. İstatistiksel olarak 1.000 eşit ve bağımsız insanın bilgi hacmine sahip olmayabilirsiniz. Şimdi daha zor bir soru: Düşük kullanım grubunda panele hiç alınmamış:

- düşük bağlantılı,
- yardımcı teknoloji kullanan,
- küçük şehirde yaşayan,
- belirli cihazı bulunmayan

insanlar varsa ağırlıklandırma onları geri getirebilir mi? Hayır. Örneklem çerçevesinde hiç bulunmayan insanın gözlemi ağırlıkla üretilemez. Şimdi denetlenen şirket kendi müşterilerine bağlantı gönderiyor: "Lütfen araştırmaya katılın." Şirketin sadık müşterileri yüksek oranda katılıyor. Rakipleri tanıyan fakat şirketi kullanmayan kişiler daha az katılıyor. Yaş ve ülke ağırlıkları doğru olabilir. Marka bağlılığı yanlılığı devam edebilir. Şimdi başka bir sorun: Katılımcıya: "Doğru ve kaliteli bir AI cevabı yüklediğinizde bonus kazanacaksınız." deniyor. Katılımcı cevabı yeniliyor. En iyi sonucu seçiyor. Ekran görüntüsünü kırpıyor. Bu artık AI ürününün doğal cevabı değildir. Ödül sisteminin oluşturduğu seçilmiş cevaptır. Bu bölümün ilk hükmü şudur:

> *Katılımcı sayısı, katılımcı temsil gücü değildir.*

İkinci hükmü şudur:

> *Rastgele atama, rastgele nüfus seçimi değildir.*

Üçüncü hükmü şudur:

> *Ağırlıklandırma gözlenmiş dengesizliği azaltabilir; hiç gözlenmeyen insanı yaratamaz.*

Dördüncü hükmü şudur:

> *Katılımcının panele alınması, görevde tutulması veya değiştirilmesi AI yanıtının içeriğine bağlı olamaz.*

Beşinci hükmü şudur:

> *Ödül, olumlu veya istenen cevabı seçmeye değil; protokolü dürüstçe tamamlamaya verilmelidir.*

1. BÖLÜMÜN AMACI

Bu bölümün amacı, hedef nüfustan gerçek kullanıcı gözlemine uzanan bütün seçim zincirini görünür, tarafsız ve yeniden denetlenebilir hâle getirmektir. Bölüm şu ayrımları normatif hâle getirir:

- Hedef nüfus ile örnekleme çerçevesi
- Örnekleme çerçevesi ile katılımcı kaynağı
- Seçilen kişi ile davet edilen kişi
- Davet edilen kişi ile katılmayı kabul eden kişi
- Katılımcı ile görevi tamamlayan kişi
- Görevi tamamlayan kişi ile geçerli gözlem
- İnsan ile hesap
- İnsan ile cihaz
- İnsan ile oturum
- Olasılıklı panel ile gönüllü panel
- Panel içi rastgele seçim ile nüfustan rastgele seçim
- Kota uyumu ile gerçek temsil
- Demografik denge ile davranışsal denge
- Kişisel aşinalık ile hedef nüfus uygunluğu
- Müşteri, çalışan ve genel kullanıcı
- Katılımcı ödülü ile sonuç teşviki
- Görev hızı ile görev bütünlüğü
- Yanıtlamama ile geçersiz gözlem
- Panel kaybı ile nüfus değişimi
- Sahte katılımcı ile gerçek düşük kaliteli gözlem
- Veri sahteciliği ile yanlış AI yanıtı
- Ağırlıklandırma ile yanlılık giderme
- Kalibrasyon ile kanıt üretme
- Ham örneklem büyüklüğü ile etkin örneklem büyüklüğü
- Katılımcı gizliliği ile denetim kanıtı
- Genel nüfus paneli ile uzman veya müşteri paneli

Bu bölümün sonunda her GEO-1000 denetimi şu sorulara cevap verebilmelidir:

> *Katılımcılar hangi kaynaktan ve hangi kuralla seçildi?*
>
> *Kimlerin panele girme ihtimali yoktu?*
>
> *Kimlerin katılma ihtimali daha yüksekti ve neden?*
>
> *Ağırlıklar hangi bilinen farklılıkları düzeltiyor, hangilerini düzeltemiyor?*
>
> *Geçersiz veya sahte katılım kararı AI cevabının içeriğini görmeden verildi mi?*
>
> *Katılımcıların ödülü, davranışlarını ve cevap seçimlerini etkiledi mi?*

2. MERKEZÎ NORMATİF HÜKÜM

Her GEO-1000 gözlemi, hedef nüfustan geçerli analitik kayda kadar izlenebilir bir katılımcı seçim zinciri taşımak zorundadır; seçilme, davet, kabul, görev ataması, tamamlama, doğrulama, dışlama ve ağırlıklandırma kararları AI yanıtının olumlu veya olumsuz içeriğinden bağımsız olmalıdır. [K04–K06] Bir katılımcı paneli yalnız:

- ülke,
- yaş,
- dil

kotalarını karşıladığı için nüfus temsiline uygun sayılamaz. Panel ayrıca ilgili olduğu ölçüde şu bilgileri taşımalıdır:

- Katılımcı kaynağı
- Seçim mekanizması
- Panele katılma biçimi
- Uygunluk ölçütleri
- Benzersiz insan kontrolü
- Ülke ve dil doğrulaması
- AI ürün erişim doğrulaması
- Hesap ve plan bilgisi
- Cihaz ve arayüz
- AI kullanım sıklığı
- Marka aşinalığı
- Denetlenen varlıkla ilişki
- Teşvik yapısı
- Görev atama yöntemi
- Yanıtlamama
- Geçersizlik
- Tekrar ve panel koşullanması
- Ağırlıklandırma
- Kapsama eksikliği
- Gizlilik ve rıza

Bu bilgilerden bazılarının ölçülmesi katılımcıyı etkileyebilir. Bu nedenle veri toplama sırası da protokolün parçasıdır.

3. KATILIMCI SEÇİM ZİNCİRİ

Hedef nüfustan nihai gözleme uzanan zincir aşağıdaki aşamalardan oluşur:

U→F→S→I→E→A→C→V

Burada:

- U: hedef kullanıcı nüfusu
- F: örnekleme çerçevesinde bulunan kişiler
- S: seçilen kişiler

- I: davet edilen kişiler
- E: uygunluğu doğrulanıp araştırmaya alınan kişiler
- A: göreve atanmış kişiler
- C: görevi tamamlayan kişiler
- V: geçerli analitik gözlem üreten kişiler

Her geçişte insanlar kaybolabilir veya yeni yanlılık oluşabilir. Örneğin:

- Hedef nüfusta olup panelde bulunmayanlar
- Panelde olup davet edilmeyenler
- Davet edilip kabul etmeyenler
- Kabul edip görevi tamamlamayanlar
- Tamamlayıp geçersiz kayıt üretenler

aynı insanlar değildir. Bu nedenle yalnız son 1.000 geçerli gözlem raporlanamaz. Bütün seçim hunisi görünür olmalıdır.

4. KATILIMCI KAYNAKLARI

GEO-1000 tek bir katılımcı kaynağına bağlı olmak zorunda değildir. Her kaynak ayrı yanlılık ve kapsam taşır.

4.1. Olasılıklı Nüfus Paneli

Katılımcıların tanımlı nüfustan bilinen ve sıfırdan büyük seçilme olasılığı vardır. Kaynaklar:

- hane,
- adres,
- nüfus veya uygun kayıt,
- olasılıklı telefon veya çok aşamalı seçim

olabilir. Avantajları:

- tasarım ağırlığı kurulabilir,
- nüfus çıkarımı daha güçlüdür,

seçim süreci daha açık olabilir. Sınırlamaları:

- küresel ölçekte pahalı olabilir,
- bazı ülkelerde uygun çerçeve bulunmayabilir,
- AI ürün erişimine sahip alt nüfusu ayrıca bulmak gerekebilir,

yanıtlamama yine oluşur. Olasılıklı seçim, otomatik olarak kusursuz temsil değildir.

4.2. Profesyonel Gönüllü Araştırma Paneli

Katılımcılar araştırma paneline önceden gönüllü olarak kaydolmuştur. Avantajları:

- hızlıdır,
- çok ülkeye erişebilir,
- profil ve dil bilgisi sağlayabilir,

tekrar ölçüm yapılabilir. Sınırlamaları:

- katılımcılar araştırmaya yatkındır,
- teknoloji kullanım düzeyi yüksek olabilir,
- ağır panel kullanıcıları bulunabilir,

- nüfusa giriş olasılıkları bilinmeyebilir,

aynı kişinin birden fazla panelde kaydı olabilir. Bu panel:

olasılıklı dünya nüfusu örneklemi

olarak sunulamaz.

4.3. Açık Çevrim İçi Çağrı

Sosyal medya, web sitesi veya açık kayıt formu üzerinden katılım alınır. Avantajları:

- geniş görünürlük,
- düşük maliyet,
- küçük ülke ve dil topluluklarına erişim

sağlayabilir. Sınırlamaları:

- güçlü kendi kendine seçilim,
- marka hayranı veya eleştirmen yoğunluğu,
- sahte ve tekrar katılım,
- teknoloji meraklılarının aşırı temsili

oluşturabilir. Açık çağrı, ana Population Panel'den çok:

- keşif,
- Country Sentinel,
- Language Fairness,
- replication challenge

için daha uygun olabilir.

4.4. Üniversite veya Kurum Ortaklı Paneli

Katılımcılar:

- üniversite,
- kamu kurumu,
- araştırma merkezi,
- meslek kuruluşu

üzerinden seçilir. Avantajları:

- yerel erişim,
- dil uzmanlığı,
- etik gözetim,
- düşük kaynaklı topluluklara ulaşma

sağlayabilir. Sınırlamaları:

- öğrenci veya profesyonel yoğunluğu,
- kurum etkisi,
- yaş ve eğitim yanlılığı,
- kurumsal katılım baskısı

oluşturabilir.

4.5. Müşteri veya Kullanıcı Listesi

Denetlenen kurumun gerçek müşterileri veya kullanıcıları seçilir. Bu kaynak:

- müşteri deneyimi,
- hizmete uygunluk,
- marka aşinalığı

ölçümü için değerlidir. Genel nüfus paneli değildir. Müşteriler:

- markayı daha iyi tanıyabilir,
- daha olumlu veya daha olumsuz deneyime sahip olabilir,

ilgili sorguyu sorma ihtimali yüksek olabilir. Müşteri paneli ayrı tanımlanmalıdır.

4.6. Denetlenen Kurum Tarafından Önerilen Katılımcılar

Şirket:

- çalışan,
- müşteri,
- partner,
- takipçi

önerebilir. Bu kullanıcılar ana bağımsız Population Panel'e açıklamasız biçimde alınamaz. Ayrı:

VARLIK-NOMINATED PANEL

olarak etiketlenmelidir. Kurum katılımcıları seçiyorsa sonuç bağımsız nüfus ölçümü değildir.

4.7. Çok Kaynaklı Hibrit Panel

Birden fazla kaynağın birleştirilmesidir. Örnek:

- olasılıklı ulusal panel,
- profesyonel çevrim içi panel,
- küçük ülke yerel ortağı,
- Dil Adaleti topluluğu

birlikte kullanılabilir. Hibrit panel kapsama gücünü artırabilir. Fakat kaynaklar arasındaki:

- tekrar kişiler,
- seçim olasılıkları,
- teşvikler,
- görev deneyimi,
- kullanıcı profili

ayrıca modellenmelidir.

5. KATILIMCI SEÇİM BÜTÜNLÜĞÜ STATÜLERİ

Bir panelin katılımcı seçim niteliği aşağıdaki statülerle sınıflandırılabilir.

PSI-0 — UNKNOWN SELECTION

Katılımcıların nereden ve nasıl seçildiği yeterince bilinmemektedir. Nüfus genellemesi yapılamaz.

PSI-1 — CONVENIENCE OR OPEN OPT-IN

Kolaylık veya açık gönüllülük örneklemidir. Keşif ve olay tespiti için kullanılabilir.

PSI-2 — QUOTA-CONTROLLED OPT-IN

Ülke, dil veya demografi kotaları vardır. Nüfusa giriş olasılığı bilinmez.

PSI-3 — CALIBRATED MULTI-SOURCE PANEL

Birden fazla kaynak, önceden tanımlı kota ve nüfus kalibrasyonu bulunmaktadır. Ölçülmeyen seçilim yanlılığı devam edebilir.

PSI-4 — PROBABILITY OR STRONG HYBRID DESIGN

Bilinen seçilme olasılığı veya güçlü olasılıklı–olasılıksız hibrit tasarım vardır.

PSI-5 — REPLICATED AND EXTERNALLY AUDITED SELECTION

Seçim tasarımı:

- farklı sağlayıcılarla,
- farklı dalgalarda,
- bağımsız denetimde

tekrar edilmiş ve yanlılık davranışı incelenmiştir. Bu statüler skorun doğruluğunu garanti etmez. Katılımcı seçiminden kaynaklanan güven düzeyini gösterir.

6. PANEL İÇİNDE RASTGELE SEÇİM İLE NÜFUSTAN RASTGELE SEÇİM

Bir gönüllü panelde 100.000 kayıtlı kullanıcı olabilir. Bu panel içinden 1.000 kişi rastgele seçilebilir. Seçim:

> *panel içinde rastgeledir.*

Fakat panelin kendisi dünya nüfusundan rastgele oluşmamıştır. Bu nedenle: "Dünya nüfusundan rastgele 1.000 kişi" denemez. Doğru ifade:

> *"Önceden kayıtlı gönüllü panel içinde rastgele seçilmiş ve hedef nüfusa kalibre edilmiş 1.000 katılımcı."*

Rastgelelik hangi aşamada uygulanıyorsa o aşamayla sınırlandırılmalıdır.

7. KATILIMCI UYGUNLUK KAYDI

Her katılımcının ana panele alınmadan önce ilgili olduğu ölçüde şu koşulları karşılaması gerekir:

- Yaş uygunluğu
- Gönüllü onay
- Ülke ve locale uygunluğu
- Dil yeterliliği
- AI ürün erişimi
- Doğru plan veya kullanıcı yüzeyi
- Gerekli cihaz ve teknik erişim
- Benzersiz insan statüsü
- Görev bağımsızlığı
- Çıkar veya ilişki beyanı

- Araştırma güvenliği
- Veri ve kanıt toplama onayı

Uygunluk kararı AI cevabı üretildikten sonra geriye dönük değiştirilemez.

8. ÜLKE DOĞRULAMASI

"Katılımcı hangi ülkeye aittir?" sorusu tek bir cevap taşımaz. Aşağıdaki alanlar ayrılmalıdır:

- Olağan ikamet ülkesi
- Ölçüm anındaki fiziksel ülke
- AI hesabının ülke veya pazar bölgesi
- Ödeme veya plan ülkesi
- Vatandaşlık
- İstemin hedeflediği ülke
- Locale

Vatandaşlık çoğu GEO ölçümü için birincil değişken olmayabilir. Önemli olan çoğu zaman:

> *Kullanıcının ürünü hangi gerçek coğrafi ve ürün koşulunda deneyimlediğidir.*

8.1. Seyahat Eden Kullanıcı

Katılımcı olağan ikamet ülkesinden başka ülkede bulunabilir. Bu kişi:

- ikamet nüfusu,
- ölçüm konumu,
- seyahat kullanıcısı

olarak ayrı statü taşımalıdır.

8.2. VPN veya Konum Değiştirici

VPN kullanımı:

- gizlilik,
- kurum ağı,
- normal kullanıcı alışkanlığı

nedeniyle olabilir. Otomatik sahtecilik sayılmamalıdır. Ancak ülke tabakasını doğrulamayı zorlaştırır. Ana kontrollü ülke panelinde VPN kapatma veya kayıt altına alma kuralı bulunabilir. Kısıt aşmak amacıyla kullanılan konum değişikliği ana Native Reach paneline uygun değildir.

8.3. Ülke Doğrulama Düzeyleri

GV-0 — Unverified

Yalnız serbest beyan vardır.

GV-1 — Panel-Verified

Panel sağlayıcının mevcut ülke kaydı vardır.

GV-2 — Session-Consistent

Beyan, locale ve oturum sinyalleri uyumludur.

GV-3 — Multi-Signal Verified

Mahremiyet koruyucu birden fazla sinyal uyumludur.

GV-4 — Independently Audited

Ülke doğrulama yöntemi dış kalite kontrolünden geçmiştir. Kesin sokak adresi veya hassas koordinat toplamak varsayılan gereklilik değildir. Ülke doğrulaması veri minimizasyonuyla yapılmalıdır.

9. DİL YETERLİLİĞİ VE DOĞAL KULLANIM

Bir katılımcının dili bilmesi yalnız: "Evet, biliyorum." beyanına dayanabilir. Bu, düşük riskli başlangıç kaydı olabilir. Yüksek güvenli dil panelinde ayrıca:

- doğal kullanım beyanı,
- kısa anlama kontrolü,
- panel geçmişi,
- yerel hakem değerlendirmesi,
- görev öncesi nötr deneme

kullanılabilir. Dil testi, yalnız yüksek eğitimli insanları seçen ağır akademik sınava dönüşmemelidir. Amaç:

> *Katılımcının istemi ve cevabı gerçek kullanıcı olarak anlayabilmesi.*

10. AI ÜRÜN ERİŞİMİNİN DOĞRULANMASI

Katılımcı:

- doğru AI ürününe,
- doğru kullanıcı yüzeyine,
- doğru plan veya hesap koşuluna

erişebilmelidir. Şu durumlar ayrı kaydedilmelidir:

- Ücretsiz plan
- Ücretli bireysel plan
- Kurumsal plan
- Eğitim planı
- Web
- Mobil
- Masaüstü
- Gömülü ürün
- Görünen model
- Web/retrieval özelliği

Katılımcı başkasının hesabını kullanamaz. Şifre paylaşımı istenemez. Ürüne erişim kanıtı, hesap kimliğini kamuya açıklamadan doğrulanmalıdır.

11. KATILIMCI İLE DENETLENEN VARLIK ARASINDAKİ İLİŞKİ

Katılımcının denetlenen kurumla ilişkisi yanıt yakalama davranışını veya marka aşinalığını etkileyebilir. İlişki statüleri:

GENERAL_PUBLIC

CURRENT_CUSTOMER

FORMER_CUSTOMER

PROSPECT

EMPLOYEE

FORMER_EMPLOYEE

CONTRACTOR

PARTNER

INVESTOR

COMPETITOR_AFFILIATE

AI_PROVIDER_AFFILIATE

RESEARCHER

UNKNOWN

Ana genel nüfus panelinde doğrudan çalışan ve karar verici ilişkiler:

- dışlanabilir,
- ayrı tabakaya alınabilir,

düşük gerçek nüfus ağırlığıyla korunabilir. İlişkili kişilerin tamamını otomatik dışlamak da yanlış olabilir. Gerçek müşteri veya çalışan nüfusun parçasıysa ayrı ve açık biçimde ölçülebilir.

12. MARKA AŞİNALILIĞI

Katılımcı denetlenen varlığı:

- hiç duymamış,
- adını biliyor,
- ürününü biliyor,
- müşterisi,
- güçlü olumlu veya olumsuz görüş sahibi

olabilir. Bu bilgi:

- katılımcı davranışını,
- doğal kullanıcı panelini,
- istem yorumunu

etkileyebilir. Aşinalık ölçümü istemden önce yapılırsa kullanıcıyı markaya hazırlayabilir. Bu nedenle tercih sırası:

- Önceden mevcut panel bilgisi
- Görev sonrası ayrı soru
- Gerekli ise nötr ve markayı yönlendirmeyen ön ölçüm

olabilir. Aşinalık ana örneklemde görünür olmalı; sonuç görüldükten sonra ağırlık değişkeni olarak seçilmemelidir.

13. AI KULLANIM AŞİNALILIĞI

Katılımcılar:

- ilk kez kullanan,
- seyrek kullanan,
- düzenli kullanan,
- yoğun veya profesyonel kullanan

kişiler olabilir. AI kullanım sıklığı:

- ürün erişimini,
- plan türünü,
- arayüz bilgisini,
- kişiselleştirmeyi,
- görevi doğru tamamlama ihtimalini

etkileyebilir. Yalnız deneyimli AI kullanıcılarını seçmek protokol hatalarını azaltabilir. Fakat doğal nüfus tahminini bozabilir. Doğru çözüm:

- kullanım düzeyini ölçmek,
- hedef nüfusla karşılaştırmak,
- gerektiğinde tabakalamak veya ağırlıklandırmak,

deneyimsiz kullanıcının geçerli gerçek deneyimini korumaktır.

14. BENZERSİZ İNSAN KURALI

GEO-1000'in nüfus birimi insandır. Aynı kişi:

- birden fazla panelde,
- birden fazla e-posta hesabıyla,
- birden fazla AI hesabıyla,
- farklı cihazlardan

katılabilir. Benzersizlik kontrolü mahremiyet koruyucu biçimde yapılmalıdır. Kullanılabilecek sinyaller:

- Panelin anonim kişi kimliği
- Kriptografik veya hash tabanlı tekrar kontrolü
- Cihaz ve oturum uyumu

- Ödeme veya teşvik kaydı
- Zaman ve görev örüntüsü
- Aynı ekran görüntüsü veya yanıt dosyası
- Çoklu panel sağlayıcılar arası güvenli tekrar kontrolü

Hiçbir tek sinyal otomatik dışlama için yeterli olmak zorunda değildir. Özellikle:

- ortak cihaz kullanan aileler,
- kurum ağı,
- erişilebilirlik araçları,
- düşük gelirli ortak cihaz kullanıcıları

yanlışlıkla sahte sayılabilir. Şüpheli kayıtlar insan incelemesi ve itiraz yolu taşımalıdır.

15. KATILIMCI ÖDÜLÜ

Katılımcıya verilen ödeme veya ödül, veri davranışını etkileyebilir.

15.1. Sonuçtan Bağımsız Ödeme

Tercih edilen yöntem:

> *Protokole uygun görevin tamamlanması için sabit ve önceden açıklanmış ödeme.*

Ödeme:

- cevabın olumlu,
- doğru,
- uzun,
- markaya yararlı,
- atıf içeren

olmasına bağlı olamaz.

15.2. Kalite Bonusu

Kalite bonusu kullanılacaksa yalnız:

- eksiksiz kanıt,
- doğru istem,
- doğru ürün,
- zamanında teslim,
- kişisel veri temizliği

gibi protokol şartlarına bağlanabilir. AI yanıtının içeriğine bağlanamaz.

15.3. Hız Teşviki

Yalnız hız için bonus:

- acele,
- eksik capture,
- yanlış ürün,
- cevabı kesme

riskini artırabilir. Saha penceresi içinde makul süre tanınmalıdır.

15.4. Ülkeye Göre Ödeme

Yerel ücret ve yaşam koşullarına göre ödeme farklılaşabilir. Bu durum:

- katılım ihtimalini,
- panel kompozisyonunu

etkileyebilir. Ödeme yaklaşımı ve ülke farkları kaydedilmelidir.

15.5. Zorlama ve Aşırı Teşvik

Ödül, kişinin makul biçimde reddedemeyeceği kadar baskı oluşturmamalıdır. Özellikle:

- düşük gelirli,
- öğrenci,
- bağımlı çalışan,
- savunmasız gruplar

için gönüllülük korunmalıdır.

16. DAVET METNİ VE ÖN YÖNLENDİRME

Katılımcı davet edilirken araştırmanın amacı onu belirli bir sonuca hazırlayabilir. Yanlış davet: “AI sistemlerinin Apple’ı ne kadar yanlış anlattığını test ediyoruz.” Bu ifade katılımcıyı hata aramaya yöneltebilir. Başka yanlış davet: “Markamızın AI başarısını kanıtlamamıza yardımcı olun.” Bu ifade olumlu sonuç beklentisi yaratır. Tercih edilen nötr davet: “Belirli AI ürünlerinde standartlaştırılmış kullanıcı deneyimi ve temsil gözlemi çalışması.” Denetlenen varlık görevin doğası gereği daha sonra görünür olabilir. Ancak istenen sonuç açıklanmamalıdır.

17. GÖREV ATAMASI

Katılımcılar:

- AI ürünü,
- istem,
- dil,
- dalga,
- oturum koşulu,
- görev sırası

bakımından önceden tanımlanmış yöntemle atanmalıdır. Katılımcının:

- en sevdiği AI ürününü,
- en iyi bildiği planı,
- olumlu cevap alma ihtimali yüksek dili

kendisi seçmesine izin verilirse atama yanlılığı oluşabilir.

17.1. Tabaka İçinde Rastgele Atama

Ülke, dil ve ürün erişimi doğrulandıktan sonra görevler tabaka içinde rastgele atanabilir.

17.2. Bloklu Atama

Ücretsiz/ücretli plan, mobil/web veya AI kullanım sıklığı bakımından denge sağlamak için bloklar kullanılabilir.

17.3. Ürün Sırası

Bir kişi birden fazla ürünü test ediyorsa sıra rastgeleleştirilmelidir. İlk ürünün cevabı sonraki cevabı değerlendirmesini veya görev davranışını etkileyebilir.

17.4. İstem Sırası

Bir katılımcı birden fazla istem alıyorsa:

- öğrenme,
- marka hakkında yeni bilgi edinme,
- önceki cevabı sonraki soruya taşıma

riski oluşur. İstem sırası kaydedilmeli ve gerekirse dengelenmelidir.

18. PANEL KOŞULLANMASI

Araştırma panellerine sık katılan kullanıcılar görevin beklentisini öğrenebilir. Aynı sınama'a tekrar tekrar katılan kişi:

- doğru cevabı hatırlayabilir,
- ekran görüntüsünü daha dikkatli seçebilir,
- denetlenen varlığın sitesini önceden inceleyebilir,

hangi hataların arandığını öğrenebilir. Bu duruma:

Panel koşullanması

denir. Panel koşullanması gerçek doğal kullanıcı deneyimini değiştirebilir.

18.1. Koşullanma Kaynakları

Önceki GEO dalgaları Aynı varlığın tekrar ölçümü Aynı istem bankası Hakemlik eğitimi Katılımcı topluluklarında görev paylaşımı Önceki ödemeler ve kalite geri bildirimi

18.2. Koşullanma Kontrolleri

Yeni ve tekrar katılımcı statüsü İlk katılım tarihi Aynı varlığa önceki maruziyet Aynı istem ailesine önceki maruziyet Dönen panel Görev sırası rastgeleliği Belirli süre dinlendirme Yeni kullanıcı alt örneklemi Koşullanmış ve yeni kullanıcı sonuçlarının karşılaştırılması Aynı katılımcıyı tekrar kullanmak yasak değildir. Boylamsal ölçüm için değerlidir. Doğal kesitsel panel gibi sunulamaz.

19. GÖREV BULAŞMASI VE KATILIMCI İLETİŞİMİ

Katılımcılar:

- istemleri,
- cevapları,
- ekran görüntülerini,
- beklenen sonuçları

birbirleriyle paylaşabilir. Bu durum bağımsız gözlemleri bozabilir. Kontroller:

- Dalga süresini sınırlamak
- Görevleri rastgele zamanlamak
- Katılımcılara sonuç paylaşmama yükümlülüğü

- Aynı cevap ve görsel tekrarını aramak
- Olağandışı metin ve zaman benzerliklerini incelemek
- Topluluk veya panel sağlayıcı uyarılarını izlemek

Katılımcıların iletişim kurmuş olması otomatik sahtecilik değildir. Maddi etkisi incelenmelidir.

20. YANITLAMAMA TÜRLERİ

Yanıtlamama tek olay değildir.

20.1. Davete Yanıt Vermeme

Katılımcı araştırma davetine hiç cevap vermez.

20.2. Katılımı Reddetme

Kişi daveti görür ve katılmamayı seçer.

20.3. Uygunluk Sonrası Göreve Başlamama

Kişi uygun bulunur fakat görevi başlatmaz.

20.4. Görev Kesintisi

Katılımcı görevi yarıda bırakır.

20.5. Kanıt Yüklememe

AI yanıtı alınmış olabilir fakat gerekli kayıt yüklenmemiştir.

20.6. Geçersiz Tamamlama

Görev tamamlanmıştır fakat:

- yanlış istem,
- yanlış ürün,
- seçilmiş cevap,
- eksik ekran görüntüsü

nedeniyle geçersizdir. Bu aşamaların oranları ayrı raporlanmalıdır.

21. YANITLAMA HUNİSİ

Tabaka h için:

- nhsel: seçilen kişi
- nhinv: davet edilen kişi
- nhcon: katılmayı kabul eden kişi
- nhsta: göreve başlayan kişi
- nhcom: tamamlayan kişi
- nhval: geçerli gözlem

olsun. Temel oranlar:

```
RR_h^invite = n_h^con/n_h^inv; RR_h^complete = n_h^com/n_h^sta; RR_h^valid = n_h^val/n_h^com
```

olarak gösterilebilir. Bu oranların her biri farklı sorunu gösterir. Yalnız nihai geçerli sayıyı vermek saha yanlılığını saklar.

22. YANITLAMAMA YANLILIĞI

Yanıt vermeyen insanlar, yanıt verenlerden farklı olabilir. Örnek: Düşük internetli kullanıcı görevi tamamlayamaz. Ücretsiz plan kullanıcısı ürüne erişemez. Deneyimsiz kullanıcı yönergeleri zor bulur. Yoğun AI kullanıcısı daha hızlı tamamlar. Belirli ülkelerde ödeme düşük kalır. Olumsuz marka deneyimi olan kullanıcı çalışmaya daha istekli olabilir. Yanıtlamama oranı yüksek olsa bile yanlılık düşük olabilir. Yanıtlamama oranı düşük olsa bile kaybolan küçük grup maddi olabilir. Bu nedenle yalnız oran değil:

Yanıt veren ve vermeyenlerin hangi özelliklerde farklılaştığı

incelenmelidir.

23. YANITLAMAMA AĞIRLIKLARI

Yanıtlamama ayarı kullanılacaksa sınıflar sonuç görülmeden önce belirlenmelidir. Muhtemel değişkenler:

- Ülke
- Dil
- Yaş
- AI kullanım sıklığı
- Plan
- Arayüz
- Panel kaynağı
- Erişilebilirlik
- Marka aşinalığı

Yanıtlamama katsayısı:

```
a_r^{NR} = (Σ_{j∈S_r} d_j) / (Σ_{i∈R_r} d_i)
```

olarak hesaplanabilir. Fakat bu düzeltme yalnız aynı sınıf içinde: yanıt verenlerin, vermeyenleri yeterince temsil ettiği varsayımına dayanır. Ölçülmemiş farkları çözmez.

24. EKSİK VERİ MEKANİZMALARI

Eksiklik kavramsal olarak üç sınıfta düşünülebilir.

24.1. Rastlantısal Eksiklik

Eksiklik sonuç ve kullanıcı özellikleriyle maddi olarak ilişkili değildir. Gerçekte nadir olabilir.

24.2. Gözlenen Özelliklerle Açıklanabilen Eksiklik

Eksiklik:

- ülke,
- yaş,
- plan,

- dil

gibi bilinen alanlarla ilişkilidir. Ağırlıklandırma kısmen yardımcı olabilir.

24.3. Gözlenmeyen Özelliklerle İlişkili Eksiklik

Örneğin:

- AI ürününe güvensizlik,
- düşük dijital beceri,
- marka karşıtlığı,
- gizlilik kaygısı

ölçülmemişse ağırlıklandırma eksikliği tamamen düzeltemez. Bu durumda duyarlılık analizi gerekir.

25. PANEL KAYBI VE DALGALAR ARASI AYRILMA

Aynı kullanıcıların tekrar ölçüldüğü boylamsal veya dönen panelde bazı kişiler sonraki dalgaya katılmaz. Bu panel kaybıdır. Panelde kalanlar:

- daha disiplinli,
- daha yüksek ödül motivasyonlu,
- daha deneyimli,
- araştırmaya daha alışkın

olabilir. Zaman içindeki skor değişimi kısmen panel kompozisyonundan kaynaklanabilir.

25.1. Panel Kaybı Kaydı

Her dalgada:

- yeniden davet edilen,
- katılan,
- ayrılan,
- yeni eklenen

kullanıcı sayıları gösterilmelidir.

25.2. Panel Kaybı Ağırlığı

Kalan katılımcılar, kayıp olasılığına göre ağırlıklandırılabilir. Bu yöntem gözlenmeyen panel kaybı yanlılığını tamamen çözmez.

25.3. Yenileme Paneli

Kayıp kullanıcıların yerine yeni kullanıcı eklenebilir. Yeni ve tekrar kullanıcı sonuçları ayrıca raporlanmalıdır.

26. KATILIMCI SAHTECİLİĞİ VE OTOMASYON

GEO-1000 teşvikli ve küresel bir panel olduğunda sahtecilik riski bulunur. Olası davranışlar:

- Bir kişinin birden fazla hesapla katılması
- Otomatik bot kullanımı
- Başkasının ekran görüntüsünü yüklemek
- Görsel düzenlemek

- Aynı cevabı farklı kullanıcılar adına yüklemek
- Yanlış ülke veya dil beyan etmek
- Hesap paylaşmak
- Görevi başka kişiye yaptırmak
- Panel sağlayıcılar arasında tekrar katılmak
- Toplu görev çiftlikleri oluşturmak

Bu risk, bütün düşük gelirli veya hızlı katılımcıların şüpheli sayılmasına izin vermez. Sahtecilik tespiti:

- kanıta,
- çoklu sinyale,
- orantılı insan incelemesine

dayanmalıdır.

27. SAHTECİLİK VE BÜTÜNLÜK STATÜLERİ

FI-0 — NOT ASSESSED

Kayıt bütünlük bakımından değerlendirilmemiştir.

FI-1 — NO MATERIAL ANOMALY

Belirgin sahtecilik sinyali bulunmamıştır. Bu, mutlak doğruluk garantisi değildir.

FI-2 — REVIEW REQUIRED

Bir veya daha fazla şüpheli sinyal vardır.

FI-3 — DUPLICATE OR COLLUSION SUSPECTED

Tekrar kişi, görsel, cevap veya koordinasyon şüphesi vardır.

FI-4 — PROTOCOL MANIPULATION CONFIRMED

Katılımcı protokolü bilinçli biçimde değiştirmiştir.

FI-5 — FABRICATED OBSERVATION CONFIRMED

Gözlem gerçek AI kullanıcı deneyimi değildir. FI-2 otomatik dışlama değildir. İnceleme gerektirir.

28. SAHTECİLİK SİNYALLERİ

Aynı ekran görüntüsü hash'i Aynı nadir dosya izi İmkânsız tamamlanma süresi Aynı cihaz veya oturumun çok sayıda kimlikte tekrarı Ülke, locale ve ürün erişiminde maddi çelişki Yanıt metni ile görüntü uyuşmazlığı Görsel düzenleme izi Ana ve yedek listede aynı kişi Birden fazla panel sağlayıcıda aynı anonim kişi Görevlerin seri ve mekanik tamamlanması Aynı sıra dışı hata örüntüsü Tek bir sinyal kesin sahtecilik kanıtı olmayabilir.

29. SAHTECİLİK TESPİTİNDE ADALET

Aşağıdaki kullanıcılar yanlışlıkla sahte olarak işaretlenebilir:

- Ortak cihaz kullanan aileler

- Kurum ağı kullanan çalışanlar
- Üniversite laboratuvarı katılımcıları
- VPN kullanan gizlilik duyarlı kişiler
- Ekran okuyucu veya otomasyon yardımcısı kullanan engelli kullanıcılar
- Düşük bant genişliği nedeniyle benzer sıkıştırılmış görseller yükleyenler

Bu nedenle:

- otomatik işaretleme,
- insan incelemesi,
- gerekirse katılımcı açıklaması,
- itiraz

birlikte çalışmalıdır.

30. KATILIMCI GEÇERLİLİĞİ İLE AI YANIT DOĞRULUĞU AYRILMALIDIR

Bir katılımcının kaydı protokole uygun olabilir. AI yanıtı yanlış olabilir. Bu geçerli gözlemdir. Başka bir katılımcının AI yanıtı doğru olabilir. Fakat kullanıcı:

- cevabı üç kez yenilemiş,
- yalnız en iyi cevabı yüklemiş,
- istemi değiştirmiş

olabilir. Bu geçersiz gözlemdir. Dolayısıyla:

> *Doğru AI cevabı geçerli veri anlamına gelmez. / Yanlış AI cevabı geçersiz veri anlamına gelmez.*

Geçerlilik incelemesi maddi puanlamadan önce ve mümkün olduğunca kör yapılmalıdır.

31. ÇIKAR ÇATIŞMASI VE KATILIMCI KAYNAĞI

Katılımcı veya panel sağlayıcısı şu sonuçlardan fayda görebilir:

- Denetlenen şirketin yüksek skor alması
- AI ürününün üstün görünmesi
- Rakibin düşük skor alması
- Daha fazla denetim çalışması satılması
- Ödül veya iş ilişkisi

Bu nedenle:

- katılımcı ilişkileri,
- panel sağlayıcı sözleşmesi,
- sonuçtan bağımsız ücret,
- müşteri müdahalesi

açıklanmalıdır. Denetlenen kurum:

- hangi kullanıcıların davet edileceğine,
- kimin geçersiz sayılacağına,
- kimin yedekle değiştirileceğine

karar veremez.

32. PANEL SAĞLAYICI ETKİSİ

Aynı ülkedeki iki panel sağlayıcının kullanıcıları farklı olabilir. Bir sağlayıcı:

- teknoloji meraklıları,
- öğrenciler,
- yoğun araştırma katılımcıları

bakımından daha yüksek oran taşıyabilir. Panel sağlayıcısı analizde bir küme veya kaynak değişkeni olarak tutulmalıdır. Mümkünse:

- sağlayıcı bazlı sonuçlar,
- kaynak etkisi,
- aynı ülke içinde sağlayıcı karşılaştırması

incelenmelidir. Tek sağlayıcılı küresel panel: dünya nüfusunun kendisi olarak sunulamaz.

33. ÇOK KAYNAKLI PANELDE TEKRAR KİŞİ

Aynı kişi:

- iki profesyonel panelde,
- açık çağrıda,
- üniversite listesinde

bulunabilir. Çok kaynaklı panel benzersizliği artırmadan yalnız kayıt sayısını artırabilir. Kaynaklar arası tekrar kontrolü:

- mahremiyet koruyucu,
- merkezi kimlik ifşası gerektirmeyen,
- yanlış eşleşmeye itiraz imkânı veren

biçimde kurulmalıdır.

34. ÖRNEKLEM YANLILIĞI HARİTASI

GEO-1000 aşağıdaki yanlılık türlerini ayrı izlemelidir.

SB-1 — Kapsama Yanlılığı

Hedef nüfusun bazı bölümleri örnekleme çerçevesinde yoktur.

SB-2 — Seçim Yanlılığı

Çerçevedeki kişilerin panele seçilme ihtimalleri hedeflenen tasarıma uygun değildir.

SB-3 — Kendi Kendine Seçilim Yanlılığı

Araştırmaya gönüllü katılanlar nüfustan sistematik olarak farklıdır.

SB-4 — Davet ve Yönlendirme Yanlılığı

Davet metni veya kaynak belirli görüşteki insanları çeker.

SB-5 — Yanıtlamama Yanlılığı

Katılmayan veya tamamlamayan kişiler sonuç bakımından farklıdır.

SB-6 — Panel Kaybı Yanlılığı

Dalgalar arasında kalan kullanıcılar ayrılanlardan farklıdır.

SB-7 — Teşvik Yanlılığı

Ödeme veya ödül katılımcının hızını, seçim davranışını veya cevabı değiştirmesini etkiler.

SB-8 — Panel Koşullanması

Katılımcı önceki araştırmalardan beklenen davranışı öğrenir.

SB-9 — Görev Atama Yanlılığı

Katılımcılar ürün, dil veya istemleri kendileri seçer ya da dengesiz atanır.

SB-10 — Ülke ve Dil Sınıflandırma Yanlılığı

Katılımcı yanlış ülke veya dil hücresine atanır.

SB-11 — Ürün, Plan ve Arayüz Yanlılığı

Yalnız belirli plan veya cihaz kullanıcıları paneli belirler.

SB-12 — Marka Aşinalığı Yanlılığı

Müşteri, hayran, eleştirmen veya çalışanlar aşırı temsil edilir.

SB-13 — Sahtecilik ve Tekrar Yanlılığı

Aynı kişi veya otomatik katılım sonuç üzerinde fazla etki kazanır.

SB-14 — Geçerlilik Dışlama Yanlılığı

Olumsuz AI cevapları daha sık geçersiz sayılır.

SB-15 — Ağırlıklandırma Yanlılığı

Sonuç sonrası seçilmiş değişkenler veya aşırı ağırlık modeli skoru istenen yönde değiştirir.

SB-16 — Sağ Kalan Örneklem Yanlılığı

Yalnız görevi başarıyla tamamlayabilen kullanıcılar gerçek ürün nüfusu gibi sunulur.

35. AĞIRLIKLANDIRMANIN AMACI

Ağırlıklandırma şu amaçlara hizmet edebilir:

- Eşit olmayan seçilme olasılıklarını düzeltmek
- Fazla örneklemeyi gerçek nüfus payına geri döndürmek
- Yanıtlamama farklarını azaltmak
- Örneklemi bilinen hedef toplamlarına kalibre etmek
- Birden fazla panel kaynağını birleştirmek

Ağırlıklandırma şu amaçlara hizmet edemez:

- Örnekleme çerçevesinde hiç bulunmayan insanı yaratmak
- Ölçülmeyen değişkenleri kesin biçimde düzeltmek
- Sahte veya hatalı gözlemi gerçek yapmak
- Sonuç sonrası istenen skoru üretmek
- Tekil küçük hücreyi güvenilir ülke skoruna dönüştürmek
- Olasılıksız paneli otomatik olasılıklı örneklem yapmak

36. AĞIRLIK MİMARİSİ

Nihai ağırlık aday olarak şöyle gösterilebilir:

```
w_i = d_i × a_i^{NR} × g_i^{CAL} × p_i^{SRC}
```

Burada:

- di: temel tasarım veya başlangıç ağırlığı
- aiNR: yanıtlamama ayarı
- giCAL: nüfus kalibrasyonu
- piSRC: panel kaynağı veya olasılıksız örneklem düzeltmesi

Olasılıklı panelde:

```
d_i = 1/π_i
```

olabilir. Olasılıksız panelde gerçek πi bilinmez. Bu durumda kullanılan ağırlık: tasarım ağırlığı değil, model tabanlı veya kalibrasyon ağırlığıdır. Bu ayrım raporda korunmalıdır.

37. AĞIRLIKLAR SONUCA BAĞLI OLAMAZ

Aşağıdaki değişkenler ağırlık oluşturmak için kullanılamaz:

- AI cevabının doğru olup olmaması
- Markanın anılıp anılmaması
- Atıf bulunması
- Tavsiye oluşması
- Critical hata çıkması
- Denetlenen kurumun sonucu beğenmesi

Ağırlık değişkenleri:

- sonuç görülmeden önce,
- hedef nüfus ve seçim mekanizmasıyla ilişkili olduğu için

seçilmelidir. Yanıt sonucu ağırlığı belirlerse skor kendi kendisini şekillendirir.

38. KALİBRASYON DEĞİŞKENLERİ

Muhtemel kalibrasyon alanları:

- Ülke
- Bölge
- Dil
- Yaş

- Cinsiyet, gerekli ve etik ise
- Eğitim
- Kentsel/kırsal yerleşim
- İnternet erişim türü
- Cihaz
- AI kullanım sıklığı
- Ücretsiz/ücretli plan
- Arayüz
- Erişilebilirlik statüsü

Her değişken için güvenilir hedef toplam bulunmayabilir. Fazla sayıda kalibrasyon değişkeni:

- aşırı ağırlık değişkenliği,
- düşük etkin örneklem,
- model aşırı uyumu

oluşturabilir. Kalibrasyon seti sonuçlar görülmeden önce dondurulmalıdır.

39. AĞIRLIKLANDIRMA DENGESİ

Ağırlıklandırma öncesi ve sonrası örneklem, hedef nüfusla karşılaştırılmalıdır. Her alan için:

- hedef pay
- ham örneklem payı
- ağırlıklı pay
- kalan fark

gösterilmelidir. Ağırlıklandırma yalnız sonuç sayısına uygulanıp denge tablosu saklanamaz.

40. AĞIRLIK TANILARI

Ana raporda ilgili olduğu ölçüde şu alanlar bulunmalıdır:

- Minimum ağırlık
- Maksimum ağırlık
- Medyan ağırlık
- Ağırlık değişim katsayısı
- Ağırlık oranı
- Kırpılan ağırlık sayısı
- Ham örneklem büyüklüğü
- Etkin örneklem büyüklüğü
- Tasarım etkisi
- Ağırlıksız sonuç
- Ana ağırlıklı sonuç
- Alternatif makul ağırlık sonuçları

Ağırlıklı tek sayı, ağırlık tanıları olmadan tam denetim kaydı değildir.

41. SENTETİK AĞIRLIKLANDIRMA GÖSTERİMİ

SENTETİK METODOLOJİ GÖSTERİMİ / Aşağıdaki katılımcılar, oranlar ve skorlar kurgusaldır. Hedef nüfus ve panel dağılımı:

AI kullanım sıklığı	Hedef nüfus payı	Ham panel payı	Ham gözlem
Düşük veya seyrek kullanım	%50	%10	100
Orta kullanım	%30	%30	300
Yoğun kullanım	%20	%60	600
Toplam	%100	%100	1.000

Sentetik geçiş oranları:

Grup	Geçiş oranı	Geçen yanıt
Düşük veya seyrek kullanım	%72	72
Orta kullanım	%84	252
Yoğun kullanım	%96	576
Toplam	%90	900

Ham sonuç:

```
900/1000=90%
```

41.1. Kalibrasyon Katsayıları

Düşük kullanım:

```
0,50/0,10=5
```

Orta kullanım:

```
0,30/0,30=1
```

Yoğun kullanım:

```
0,20/0,60 = 1/3
```

Ağırlıklı geçen yanıt:

```
72(5) + 252(1) + 576(1/3) = 804
```

Ağırlık toplamı:

```
100(5) + 300(1) + 600(1/3) = 1.000
```

Ağırlıklı sonuç:

```
804/1000=80,4%
```

41.2. Etkin Örneklem Büyüklüğü

```
n_eff = [100(5)+300(1)+600(1/3)]² / [100(5²)+300(1²)+600(1/3)²] ≈ 349
```

Sonuç kartı:

- Ham geçerli gözlem: 1.000
- Ham skor: yüzde 90
- Ağırlıklı skor: yüzde 80,4
- Etkin örneklem: yaklaşık 349
- Ana neden: yoğun AI kullanıcılarının aşırı temsili

Doğru yorum:

> *Panel 1.000 geçerli gözlem taşımaktadır; fakat hedef nüfus kalibrasyonu sonucunda tahmin yüzde 80,4'e düşmüş ve ağırlık değişkenliği nedeniyle etkin örneklem yaklaşık 349 olmuştur.*

Yanlış yorum: "1.000 kişiyle yüzde 90 küresel doğruluk kanıtlanmıştır."

42. AĞIRLIKLANDIRMA NEYİ DÜZELTEMEZ?

Sentetik panelde şu gruplar hiç bulunmasın:

- Ekran okuyucu kullananlar
- Düşük bant genişlikli kırsal kullanıcılar
- AI ürününü ilk kez kullananlar
- Küçük dil topluluğu

Ağırlıklandırma yalnız panelde bulunan kullanıcıların ağırlığını değiştirir. Hiç gözlenmeyen grubun:

- AI cevabını,
- erişim sorununu,
- hata dağılımını

üretemez. Bu nedenle:

> *Sıfır kapsama, sıfır bilgi demektir.*

Ağırlıklandırılmış sonuç: hedef nüfusa daha yakın tahmin olabilir, tam hedef nüfus gözlemi değildir.

43. DUYARLILIK ANALİZİ

Ağırlıklı sonucun tek ağırlık modeline bağımlılığı incelenmelidir. Aday sonuçlar:

- Ham ağırlıksız sonuç
- Tasarım ağırlıklı sonuç
- Demografik kalibre edilmiş sonuç
- AI kullanımı eklenmiş kalibrasyon
- Ağırlık kırpılmış sonuç
- Panel sağlayıcı etkisi ayarlanmış sonuç

Sonuçlar maddi biçimde değişiyorsa:

Weight Sensitivity Warning

verilmelidir. Tek "doğru" ağırlık sonucu kesin gerçek gibi sunulamaz.

44. OLASILIKSIZ PANELDE PROPENSITY AYARI

Olasılıksız panel üyelerinin hedef nüfusa göre panele girme eğilimi modellenebilir. Bir kişinin panelde olma olasılığı:

```
p̂_i = Pr(i ∈ panel | X_i)
```

olarak tahmin edilebilir. Aday ters propensity ağırlığı:

```
w_i^P = 1/p̂_i
```

olabilir. Bu yöntem:

- yalnız ölçülen Xi değişkenlerini kullanır,
- yanlış modelden etkilenir,
- gözlenmeyen seçilimi çözmez,

güçlü hedef referans veri gerektirir. Bu nedenle sonuç:

Model-kalibre edilmiş olasılıksız panel tahmini

olarak etiketlenmelidir.

45. KÜÇÜK ALAN VE MODEL TABANLI TAHMİNLER

Bazı ülke veya dil hücrelerinde doğrudan örneklem yetersiz olabilir. Model tabanlı yöntemler komşu veya benzer gruplardan bilgi ödünç alabilir. Bu sonuç:

- doğrudan gözlenen ülke skoru,
- ham kullanıcı oranı

değildir. Doğru statü:

MODEL-ASSISTED ESTIMATE

SMALL-AREA ESTIMATE

NOT DIRECTLY OBSERVED

olmalıdır. Model tabanlı tahmin Sentinel gözlemi bulunmayan ülkeyi gözlenmiş ülkeye dönüştürmez.

46. KATILIMCI GİZLİLİĞİ VE KANITLANABİLİRLİK

GEO-1000 her gözlem için kanıt ister. Bu, katılımcı kimliğinin kamuya açıklanması anlamına gelmez. Ekran görüntüsü şu bilgileri içerebilir:

- Ad
- E-posta
- Profil fotoğrafı
- Önceki sohbet
- Hesap planı
- Kurum adı
- Hassas kişisel bilgi

NOMOS Capture veya eşdeğer sistem:

- gereksiz kişisel alanları yakalamamalı,
- otomatik veya kontrollü redaksiyon yapmalı,
- ham kanıtı korumalı güvenli depoda tutmalı,

kamuya yalnız anonim ve doğrulanabilir kayıt sunmalıdır.

46.1. Kamu Kanıt Manifesti

Kamuya açık kayıt:

- gözlem kimliği,
- ülke/dil hücresi,

- AI ürünü,
- zaman,
- kanıt hash'i,
- doğrulama statüsü

taşıyabilir. Tam ekran görüntüsü:

- redakte edilmiş,
- kısıtlı erişimli,
- bağımsız denetçiye açık

olabilir. Bin kanıtın varlığı, bin kişinin kimliğinin kamuya açılmasını gerektirmez.

46.2. Hassas Küçük Ülke Riski

Küçük ülkede tek katılımcı varsa:

- cihaz,
- zaman,
- dil,
- plan

bir araya geldiğinde kişinin yeniden tanımlanması kolaylaşabilir. Kamu raporunda ayrıntı seviyesi düşürülmelidir. Kanıt bütünlüğü erişim kontrollü sistemde korunabilir.

47. KATILIMCI RIZASI

Katılımcı en azından şunları bilmelidir:

- Hangi görevi yapacağı
- Hangi veri ve ekran görüntüsünün toplanacağı
- Hangi kişisel bilgilerin toplanmayacağı
- Verinin ne kadar saklanacağı
- Kimlerin erişebileceği
- Sonucun nasıl yayımlanacağı
- Ödülün nasıl verileceği
- Katılımın gönüllü olduğu
- Çekilme veya itiraz imkânı
- Hesap şifresinin istenmeyeceği
- AI ürün koşullarını ihlal etmesinin beklenmediği

Rıza, uzun ve anlaşılmaz bir metin içine saklanmamalıdır. [K06; K22]

48. KATILIMCI GÜVENLİĞİ

Katılımcıdan:

- kişisel sağlık, hukuk veya finans bilgisini AI'ya vermesi,
- gerçek müşteri verisi kullanması,
- işveren gizli bilgisini paylaşması,
- hesap güvenliğini zayıflatması,

- kullanım kısıtını aşması,
- yasa veya ürün koşulunu ihlal etmesi

istenemez. Sınama istemleri mümkün olduğunca kamuya açık ve kişisel olmayan bağlam taşımalıdır.

49. KATILIMCI KALİTE DEĞİL, İNSAN ÇEŞİTLİLİĞİ TAŞIR

Deneyimsiz kullanıcı:

- görevi daha yavaş,
- daha zor,
- daha fazla destekle

tamamlayabilir. Bu kişi otomatik olarak düşük kaliteli katılımcı değildir. Hedef nüfusun gerçek parçasıdır. Araştırma tasarımı yalnız yüksek dijital becerili kişilerin tamamlayabileceği kadar karmaşıksa: sorun katılımcıda değil, ölçüm aracında olabilir. Protokol:

- anlaşılır,
- erişilebilir,
- dil bakımından doğal,
- teknik yükü sınırlı

olmalıdır.

50. KATILIMCI SEÇİMİNİN KAMU SONUÇ KARTI

Aday kamu raporu şu alanları taşımalıdır:

Panel Kaynakları

Olasılıklı panel payı Profesyonel gönüllü panel payı Açık çağrı payı Üniversite/kurum payı Müşteri veya ilişkili panel payı

Katılım Hunisi

Çerçevedeki kişi Seçilen Davet edilen Kabul eden Başlayan Tamamlayan Geçerli gözlem üreten

Uygunluk ve Bütünlük

Ülke doğrulama düzeyleri Dil doğrulama düzeyleri Tekrar kişi oranı Şüpheli kayıt oranı Doğrulanmış sahtecilik oranı Geçersiz gözlem oranı

Panel Kompozisyonu

AI kullanım sıklığı Plan ve arayüz Marka aşinalığı Müşteri/çalışan ilişkisi Yeni ve tekrar panel kullanıcıları Erişilebilirlik kapsamı

Ağırlık Tanıları

Ham ve ağırlıklı skor Minimum ve maksimum ağırlık Ağırlık değişim katsayısı Etkin örneklem Ağırlık hassasiyeti Kapsama uyarıları Bu alanlar katılımcıların kimliğini ifşa etmeden yayımlanabilir.

51. ZORUNLU NORMATİF HÜKÜMLER

CH08-N01

Her geçerli GEO-1000 gözlemi hedef nüfustan analitik kayda kadar izlenebilir bir katılımcı seçim zinciri taşımalıdır.

CH08-N02

Panel içindeki rastgele seçim, hedef dünya nüfusundan rastgele seçim olarak sunulamaz.

CH08-N03

Katılımcı kaynağı, seçim yöntemi ve panel statüsü kamu yöntem kaydında açıklanmalıdır.

CH08-N04

Olasılıksız veya gönüllü panel olasılıklı nüfus örneklemi gibi tanımlanamaz.

CH08-N05

Denetlenen kurum tarafından önerilen katılımcılar ana bağımsız nüfus paneline açıklamasız biçimde alınamaz.

CH08-N06

Müşteri, çalışan, partner, rakip ve AI sağlayıcısı ilişkileri ayrı kaydedilmelidir.

CH08-N07

Katılımcı uygunluk koşulları AI cevabı görülmeden önce doğrulanmalıdır.

CH08-N08

Ülke, dil, ürün, plan ve arayüz uygunluğu ilgili görev atamasından önce belirlenmelidir.

CH08-N09

Aynı kişi birden fazla hesap, cihaz, dil, panel veya AI ürünü nedeniyle nüfus birimi olarak çoğaltılamaz.

CH08-N10

Benzersiz kişi kontrolü veri minimizasyonu ve mahremiyet ilkeleriyle yürütülmelidir.

CH08-N11

Tek bir teknik sinyal katılımcının sahte olduğuna ilişkin otomatik kesin karar için yeterli sayılamaz.

CH08-N12

Katılımcı ödülü AI yanıtının olumlu, doğru, uzun, atıf içeren veya markaya yararlı olmasına bağlanamaz.

CH08-N13

Kalite bonusu yalnız önceden belirlenmiş protokol bütünlüğüne bağlanabilir.

CH08-N14

Davet ve bilgilendirme metni katılımcıyı olumlu veya olumsuz sonuç aramaya yöneltemez.

CH08-N15

AI ürünü, istem ve görev sırası sonucu görmeden önce atanmalı veya rastgeleleştirilmelidir.

CH08-N16

Katılımcı tercihine dayalı ürün ve istem seçimi ana nüfus panelinde açıklamasız biçimde kullanılamaz.

CH08-N17

Panel koşullanması, önceki sınama maruziyeti ve tekrar katılım kaydedilmelidir.

CH08-N18

Yeni ve koşullanmış katılımcı sonuçları gerektiğinde ayrı analiz edilmelidir.

CH08-N19

Katılımcının panele alınması, tutulması, dışlanması veya yedekle değiştirilmesi AI cevabının içeriğine dayanamaz.

CH08-N20

Yanlış, olumsuz, ret veya Critical AI çıktısı geçerli protokol altında korunmalıdır.

CH08-N21

Katılımcı geçerliliği yanıtın maddi doğruluğundan önce ve mümkün olduğunca kör değerlendirilmelidir.

CH08-N22

Davet, kabul, başlangıç, tamamlama ve geçerli gözlem oranları ayrı raporlanmalıdır.

CH08-N23

Yanıtlamama ayarı yalnız önceden tanımlanmış ve ilgili değişkenlere dayanmalıdır.

CH08-N24

Ağırlıklandırma örnekleme çerçevesinde hiç bulunmayan kullanıcıları temsil edilmiş hâle getiremez.

CH08-N25

Ağırlık değişkenleri AI yanıt sonucuna, skora veya denetlenen kurumun ticari çıkarına bağlı olamaz.

CH08-N26

Olasılıksız panel ağırlıkları tasarım ağırlığı gibi sunulamaz.

CH08-N27

Kalibrasyon değişkenleri ve hedef toplamlar sonuçlar görülmeden önce sürümlenmelidir.

CH08-N28

Ağırlıksız ve ana ağırlıklı sonuç birlikte saklanmalıdır.

CH08-N29

Ağırlık tanıları ve etkin örneklem büyüklüğü ana raporda görünür olmalıdır.

CH08-N30

Ağırlık kırpma kullanılıyorsa kırpılmamış sonuç ve duyarlılık analizi korunmalıdır.

CH08-N31

Panel sağlayıcısı ve katılımcı kaynağı etkileri ayrı değişkenler olarak tutulmalıdır.

CH08-N32

Çok kaynaklı panelde kaynaklar arası tekrar kişi kontrolü yapılmalıdır.

CH08-N33

Sahtecilik tespiti çoklu sinyal, orantılı insan incelemesi ve itiraz yolu taşımalıdır.

CH08-N34

VPN, ortak cihaz, kurum ağı veya yardımcı teknoloji kullanımı otomatik sahtecilik sayılmamalıdır.

CH08-N35

Model tabanlı küçük alan tahmini doğrudan gözlenen ülke veya dil skoru gibi sunulamaz.

CH08-N36

Katılımcı ekran görüntüleri ve kanıtları kişisel veri minimizasyonu, redaksiyon ve erişim kontrolü taşımalıdır.

CH08-N37

Katılımcıdan hesap şifresi, özel sohbet geçmişi veya gereksiz hassas bilgi istenemez.

CH08-N38

Katılımcılardan ürün, yaş, kimlik, ülke veya hesap kısıtlarını aşmaları istenemez.

CH08-N39

Çocuk veya özel koruma gerektiren kullanıcılar ayrı etik ve güvenlik protokolü olmadan panele alınamaz.

CH08-N40

Katılımcı seçim ve ağırlıklandırma kaydının hesap verebilir insan veya kurum sahibi bulunmalıdır.

52. BAŞARISIZLIK BİÇİMLERİ

CH08-F01 — PANEL İÇİ RASTGELEYİ DÜNYA RASTGELESİ SAYMAK

Gönüllü panel içindeki rastgele seçim dünya nüfusundan rastgele seçim gibi sunulur.

CH08-F02 — KOTA UYUMUNU TEMSİL SAYMAK

Yaş ve ülke kotaları tuttuğu için bütün davranışsal ve teknolojik yanlılıkların çözüldüğü iddia edilir.

CH08-F03 — MÜŞTERİLERİ GENEL NÜFUS SAYMAK

Denetlenen kurumun kullanıcı listesi dünya kullanıcı paneli gibi kullanılır.

CH08-F04 — ÇALIŞAN VE PARTNERLERİ GİZLEMEK

Kurumla ilişkili kullanıcılar bağımsız genel katılımcı gibi gösterilir.

CH08-F05 — TEK PANEL SAĞLAYICISINI DÜNYA SAYMAK

Sağlayıcı etkisi ve kayıtlı panel davranışı incelenmez.

CH08-F06 — AÇIK ÇAĞRIYI OLASILIKLI ÖRNEKLEM SAYMAK

Kendi kendine seçilim saklanır.

CH08-F07 — ÜLKEYİ YALNIZ IP İLE BELİRLEMEK

Seyahat, VPN, kurum ağı ve hesap bölgesi farkları yok sayılır.

CH08-F08 — VATANDAŞLIĞI KULLANIM ÜLKESİ SAYMAK

Ürün deneyiminin gerçek coğrafi koşulu yanlış sınıflandırılır.

CH08-F09 — DİL BEYANINI SINAMADAN KESİN SAYMAK

Katılımcı istem ve cevabı anlayamadığı hâlde dil hücresinde tutulur.

CH08-F10 — DİL TESTİNİ ELİT ELEMEYE DÖNÜŞTÜRMEK

Yalnız yüksek eğitimli kişiler panele kalır.

CH08-F11 — HESAP SAYISINI İNSAN SAYISI SAYMAK

Bir kişi birden fazla hesapla nüfus ağırlığını artırır.

CH08-F12 — ORTAK HESABI BİR KİŞİ SAYMAK

Birden fazla insanın kullandığı hesap tekil kullanıcı gibi yorumlanır.

CH08-F13 — SONUCA BAĞLI ÖDEME

Olumlu, doğru veya atıf içeren cevap için bonus verilir.

CH08-F14 — HIZ BONSUYLA PROTOKOLÜ BOZMAK

Katılımcı cevap tamamlanmadan ekran görüntüsü alır veya görevi aceleyle yapar.

CH08-F15 — DAVETTE SONUCU AÇIKLAMAK

Katılımcıdan hata veya başarı bulması istenir.

CH08-F16 — KATILIMCIYA ÜRÜNÜ SEÇTİRMEK

Yalnız erişebildiği veya sevdiği AI ürününü seçen kullanıcılar paneli belirler.

CH08-F17 — İSTEM SIRASI ETKİSİNİ GİZLEMEK

Önceki cevaplar sonraki görevleri etkiler.

CH08-F18 — PANEL KOŞULLANMASINI DOĞAL KULLANIM SAYMAK

Aynı sınamayı defalarca gören kullanıcı yeni kullanıcı gibi sunulur.

CH08-F19 — SONUÇ PAYLAŞIMINI YOK SAYMAK

Katılımcılar görev ve cevapları birbirine aktarmıştır.

CH08-F20 — YANITLAMAMAYI YALNIZ SAYI SORUNU SAYMAK

Katılmayanların sistematik farklılığı incelenmez.

CH08-F21 — PANEL KAYBINI ÜRÜN İYİLEŞMESİ SAYMAK

Zayıf veya deneyimsiz kullanıcılar sonraki dalgadan ayrıldığı için skor yükselir.

CH08-F22 — OLUMSUZ YANITI GEÇERSİZ SAYMAK

AI yanlış cevap verdiği için katılımcı değiştirilir.

CH08-F23 — DOĞRU CEVABI GEÇERLİ SAYMAK

Katılımcı cevabı yenilemiş olsa da sonuç doğru olduğu için kayıt korunur.

CH08-F24 — BOT VE TEKRAR KATILIMI GÖZDEN KAÇIRMAK

Aynı kişi veya otomasyon çok sayıda gözlem üretir.

CH08-F25 — TEK SİNYALLE TOPLU DIŞLAMA

VPN, cihaz veya hız nedeniyle gerçek kullanıcılar sahte sayılır.

CH08-F26 — SAHTECİLİK İNCELEMESİNİ SONUCA GÖRE YAPMAK

Düşük skorlu kayıtlar daha sık incelenir.

CH08-F27 — AĞIRLIKLA GÖRÜNMEYEN GRUP YARATMAK

Örneklemde hiç bulunmayan kullanıcıların temsil edildiği iddia edilir.

CH08-F28 — SONUÇ SONRASI KALİBRASYON SEÇMEK

En uygun skoru üreten ağırlık değişkenleri sonradan seçilir.

CH08-F29 — HAM PANEL PAYINI NÜFUS PAYI SAYMAK

Yoğun AI kullanıcıları ham sayılarıyla küresel sonucu belirler.

CH08-F30 — OLAŞILIKSIZ AĞIRLIĞI TASARIM AĞIRLIĞI SAYMAK

Gerçek seçilme ihtimali bilinmediği hâlde kesin nüfus çıkarımı yapılır.

CH08-F31 — AĞIRLIKSIZ SONUCU SAKLAMAK

Ağırlık modelinin etkisi görülemez.

CH08-F32 — ETKİN ÖRNEKLEMİ GİZLEMEK

1.000 ham kayıt, 300–400 bilgi birimi taşısa da tam hassasiyet iddia edilir.

CH08-F33 — AĞIRLIK KIRPMA İLE ZAYIF GRUBU SİLMEK

Yüksek ağırlıklı küçük gruplar sonucu düşürdüğü için kırpılır.

CH08-F34 — MODEL TABANLI TAHMİNİ GÖZLENMİŞ SKOR SAYMAK

Örneksiz ülkeye kesin puan verilir.

CH08-F35 — KANIT İÇİN KİMLİĞİ İFŞA ETMEK

Ekran görüntüleri kişisel bilgiyle kamuya yayımlanır.

CH08-F36 — KÜÇÜK ÜLKE KATILIMCISINI YENİDEN TANIMLANABİLİR KILMAK

Ülke, saat, cihaz ve plan bilgileri tek kişiyi açığa çıkarır.

CH08-F37 — KATILIMCIYA ŞİFRE VEYA ÖZEL SOHBET İSTEMEK

Kanıt toplama hesap güvenliğini ihlal eder.

CH08-F38 — DÜŞÜK DİJİTAL BECERİYİ UYGUNSUZLUK SAYMAK

Gerçek hedef kullanıcılar panelden çıkarılır.

CH08-F39 — ÖDÜLLE GÖNÜLLÜLÜĞÜ BOZMAK

Katılımcı makul biçimde reddedemeyeceği baskı altında kalır.

CH08-F40 — KATILIMCI SEÇİMİNİ MÜŞTERİYE BIRAKMAK

Denetlenen kurum, panelin sonucunu belirleyebilecek insanları seçer.

53. DENETİM PROSEDÜRÜ

Adım 1 — Katılımcı Kaynaklarını Haritala

Her panel, ortak, açık çağrı ve müşteri kaynağı ayrı kaydedilir.

Adım 2 — Seçim Statüsünü Ata

Her kaynak için PSI-0 ile PSI-5 arasında uygun statü belirlenir.

Adım 3 — Hedef Nüfusla Kapsam Karşılaştırması Yap

Hangi hedef kullanıcıların panelde bulunmadığı belirlenir.

Adım 4 — Katılımcı Seçim Olasılığını veya Modelini Kaydet

Bilinen olasılık, kota veya gönüllü seçilim açıklanır.

Adım 5 — Uygunluk Kurallarını Dondur

Yaş Ülke Dil Ürün Plan Arayüz Benzersizlik Çıkar ilişkisi koşulları sürümlenir.

Adım 6 — Davet Metnini İncele

Olumlu veya olumsuz yönlendirme bulunup bulunmadığı değerlendirilir.

Adım 7 — Ödül Yapısını İncele

Ödülün sonuçtan tamamen bağımsız olduğu doğrulanır.

Adım 8 — Katılımcı Ülke, Dil ve Ürün Erişimini Doğrula

Mahremiyet koruyucu doğrulama düzeyleri atanır.

Adım 9 — Benzersiz İnsan Kontrolü Yap

Kaynak içi ve kaynaklar arası tekrarlar incelenir.

Adım 10 — İlişki ve Aşinalık Alanlarını Kaydet

Müşteri, çalışan, rakip, AI sağlayıcısı ve marka aşinalığı belirlenir.

Adım 11 — Görevleri Atama Planına Göre Dağıt

Ürün, istem, sıra ve dalga atamaları doğrulanır.

Adım 12 — Panel Koşullanmasını Kaydet

Yeni, tekrar ve daha önce aynı sınama'a katılmış kullanıcılar ayrılır.

Adım 13 — Katılım Hunisini Kur

Davet, kabul, başlangıç, tamamlama ve geçerlilik sayıları hesaplanır.

Adım 14 — Geçerlilik İncelemesini Körleştir

Mümkün olduğunca AI yanıt skoru görülmeden protokol geçerliliği belirlenir.

Adım 15 — Sahtecilik Sinyallerini İncele

Çoklu sinyal ve insan değerlendirmesi kullanılır.

Adım 16 — Yanıtlamama ve Panel Kaybını Analiz Et

Katılan ve katılmayanların bilinen özellikleri karşılaştırılır.

Adım 17 — Tasarım ve Kalibrasyon Ağırlıklarını Üret

Olasılıklı ve olasılıksız kaynaklar ayrı yöntemle ele alınır.

Adım 18 — Ağırlık Tanılarını Hesapla

Ağırlık dağılımı Tasarım etkisi Etkin örneklem Kırpma etkisi hesaplanır.

Adım 19 — Duyarlılık Analizi Yap

Ham, ana ağırlıklı ve alternatif makul sonuçlar karşılaştırılır.

Adım 20 — Gizlilik ve Kanıt Manifestini Oluştur

Tam kanıt güvenli depoda; anonim kanıt kimlikleri kamu kaydında tutulur.

Adım 21 — Kamu Seçim Kartını Yayımla

Panel kaynakları, katılım hunisi, yanlılıklar ve ağırlık etkisi görünür hâle getirilir.

54. GEREKLİ KANIT

Hedef nüfus kaydı Örnekleme çerçevesi Panel sağlayıcıları Katılımcı kaynakları Seçim statüsü Olasılıklı veya olasılıksız panel sınıfı Panel kayıt yöntemi Davet metni Uygunluk kuralları Katılımcı ülke doğrulaması Dil yeterlilik kaydı AI ürün erişim kaydı Plan ve kullanıcı yüzeyi Benzersiz insan kontrolü Kaynaklar arası tekrar kontrolü Marka aşinalığı AI kullanım sıklığı Denetlenen varlıkla ilişki AI sağlayıcısıyla ilişki Katılımcı ödülü Bonus koşulları Görev atama yöntemi Rastgelelik ve bloklama kaydı Ürün ve istem sırası Panel koşullanması Önceki sınama katılımı Davet edilen kişi sayısı Katılmayı kabul eden kişi sayısı Göreve başlayan kişi sayısı Tamamlayan kişi sayısı Geçerli gözlem sayısı Yanıtlamama nedenleri Panel kaybı Geçersizlik nedenleri Sahtecilik sinyalleri

İnceleme ve itiraz kayıtları Tasarım ağırlıkları Yanıtlamama ayarları Kalibrasyon değişkenleri Propensity modeli, varsa Ağırlık kırpma kaydı Etkin örneklem büyüklüğü Duyarlılık sonuçları Panel sağlayıcı etkisi Ağırlıksız ve ağırlıklı skorlar Rıza kaydı Veri saklama ve erişim politikası Redaksiyon kuralları Kamu kanıt manifesti Sorumlu insan veya kurum

55. DENETİM KONTROL LİSTESİ

Katılımcılar hangi kaynaklardan geldi? Her kaynak olasılıklı mı, gönüllü mü, açık çağrı mı? Panel içi rastgelelik dünya nüfusu rastgeleliği gibi sunuldu mu? Denetlenen kurum katılımcı seçimine müdahale etti mi? Müşteri ve çalışan katılımcılar ayrı mı? AI sağlayıcısı veya rakip ilişkileri kaydedildi mi? Ülke uygunluğu nasıl doğrulandı? Dil yeterliliği gerçek kullanım düzeyinde mi? Doğru AI ürünü, plan ve arayüz kullanıldı mı? Aynı kişi kaç hesap veya panelde bulundu? Kaynaklar arası tekrar kontrolü yapıldı mı? Davet metni sonucu yönlendirdi mi? Ödül AI cevabının içeriğine bağlı mı? Hız teşviki protokol kalitesini düşürdü mü? Görev ataması rastgele veya önceden kurallı mı? Katılımcı ürün veya istemi kendisi seçti mi?

Ürün ve istem sırası dengelendi mi? Panel koşullanması kaydedildi mi? Katılımcılar görev sonuçlarını paylaştı mı? Davet–kabul–başlangıç–tamamlama–geçerlilik hunisi görünür mü? Yanıt vermeyenler bilinen özelliklerde farklı mı? Panel kaybı zaman trendini etkiliyor mu? Geçerlilik kararı AI cevabı puanlanmadan önce mi verildi? Olumsuz veya Critical cevaplar daha sık geçersiz sayıldı mı? Sahtecilik tespiti çoklu sinyal taşıyor mu? VPN veya ortak cihaz otomatik dışlama yarattı mı? Katılımcıya itiraz yolu verildi mi? Ağırlık değişkenleri sonuçtan önce tanımlandı mı? Olasılıksız panel ağırlıkları doğru etiketlendi mi? Hedef toplamlar güvenilir mi? Ağırlıksız ve ağırlıklı sonuç birlikte veriliyor mu? Etkin örneklem büyüklüğü açık mı? Ağırlık kırpmanın etkisi gösterildi mi?

Hiç gözlenmeyen gruplar ağırlıkla temsil edilmiş gibi gösterildi mi? Model tabanlı tahmin doğrudan gözlem gibi sunuldu mu? Katılımcı ekran görüntülerinde kişisel bilgi var mı? Küçük ülke kullanıcıları yeniden tanımlanabilir mi? Hesap şifresi veya özel sohbet talep edildi mi? Katılımcı rızası açık ve anlaşılır mı? Katılımcı seçim ve ağırlıklandırma kaydının hesap verebilir sahibi belli mi?

56. İTİRAZLAR VE CEVAPLAR

İtiraz 1 — “Profesyonel araştırma paneli kullanıyorsak neden hâlâ yanlılıktan söz ediyoruz?”

Profesyonel panel veri kalitesini ve saha yönetimini güçlendirebilir. Panel üyeleri yine:

- gönüllü,
- araştırmaya alışkın,
- teknoloji kullanımında farklı

olabilir. Profesyonellik yanlılığı yok etmez. Ölçüp yönetmeyi kolaylaştırır.

İtiraz 2 — “Yaş ve ülke kotaları tuttuysa panel neden temsili sayılmasın?”

Katılımcılar:

- AI kullanım sıklığı,
- cihaz,
- plan,
- marka aşinalığı,
- dijital beceri

bakımından farklı olabilir. Kotalar yalnız kapsadıkları alanları dengeler.

İtiraz 3 — "Ağırlıklandırma bu farkları düzeltemez mi?"

Ölçülen ve güvenilir hedef toplamı bulunan farkları kısmen düzeltebilir. Hiç gözlenmeyen veya ölçülmeyen farkları kesin biçimde düzeltemez.

İtiraz 4 — "Kurumun müşterileri gerçek kullanıcılardır; neden ana panele alınmasınlar?"

Alınabilirler. Fakat genel nüfusta taşıdıkları gerçek payla ve ilişki statüsüyle alınmalıdırlar. Müşteri listesinin tamamı genel nüfus paneli değildir.

İtiraz 5 — "Katılımcının markayı bilmesi neden sorun olsun?"

Sorun olmak zorunda değildir. Doğal kullanıcı nüfusunun bir parçasıdır. Aşinalık dengesizse sonuç farklılaşabilir. Bu nedenle kaydedilir ve gerekirse analiz edilir.

İtiraz 6 — "Yoğun AI kullanıcıları görevi daha doğru yapıyor; onları seçmek daha iyi değil mi?"

Protokol testi için daha kolay olabilir. Fakat gerçek nüfus deneyimini ölçmek istiyorsak deneyimsiz kullanıcılar da önemlidir. Görev aşırı karmaşıksa ölçüm aracı düzeltilmelidir. Nüfus basitleştirilmemelidir.

İtiraz 7 — "Katılımcı cevabı yanlış anladıysa gözlemi neden koruyalım?"

Katılımcının görevi cevabı değerlendirmek değil, doğru biçimde yakalamaktır. Capture protokolünü doğru uyguladıysa AI yanıtı korunur. Maddi değerlendirme hakemlerin görevidir.

İtiraz 8 — "Sahte katılımı önlemek için katı otomatik filtre kullanmak gerekmez mi?"

Otomatik filtre yararlıdır. Tek başına nihai karar olursa gerçek kullanıcıları haksız dışlayabilir. Çoklu sinyal ve orantılı insan incelemesi gerekir.

İtiraz 9 — "Bin ekran görüntüsünü kamuya açmazsak yöntem kanıtlanmış sayılır mı?"

Tam ekran görüntüleri:

- kişisel bilgi,
- hesap güvenliği,
- yeniden tanımlama

riski taşıyabilir. Kamuya:

- redakte edilmiş örnekler,
- gözlem manifesti,
- bütünlük hash'leri,
- bağımsız denetim erişimi

sunulabilir. Kanıtlanabilirlik kimlik ifşası değildir.

İtiraz 10 — "Ağırlıklı skor ham skordan çok farklıysa hangisi doğrudur?"

İki sonuç farklı sorulara cevap verir. Ham skor paneldeki insanların sonucudur. Ağırlıklı skor hedef nüfus tahminidir. Büyük fark panel kompozisyonunun önemli olduğunu gösterir. Her ikisi de yayımlanmalıdır.

İtiraz 11 — "Etkin örneklem 349'a düştüyse 1.000 gözlem toplamak boşa mı gitti?"

Hayır. Fazla örneklenen gruplar alt grup analizini güçlendirmiş olabilir. Ancak küresel tahminin hassasiyeti ham 1.000 sayısından daha düşük olabilir. Bu gerçeği saklamamak gerekir.

İtiraz 12 — "Olasılıksız panel hiç kullanılamaz mı?"

Kullanılabilir. Özellikle:

- pilot,
- düşük kaynaklı dil,
- küçük ülke,
- hızlı keşif

için değerlidir. Doğru statü ve sınırlamalarla yayımlanmalıdır.

İtiraz 13 — "Katılımcıların tamamını yeni insanlardan seçmek daha güvenilir değil mi?"

Kesitsel nüfus tahmini için güçlü olabilir. Boylamsal değişimi ölçmek için tekrar kullanıcılar değerlidir. Yeni ve tekrar panel rolleri ayrı tutulmalıdır.

İtiraz 14 — "Bu kadar katılımcı verisi toplamak mahremiyet riskini artırmaz mı?"

Artırabilir. Bu nedenle:

- veri minimizasyonu,
- anonim kimlik,
- redaksiyon,
- erişim kontrolü,
- saklama süresi

protokolün zorunlu parçasıdır. Gereksiz bilgi toplanmamalıdır.

59. BÖLÜMÜN ORTAK HÜKMÜ

Bir GEO araştırmasında en kolay sayı:

1.000

olabilir. En zor soru ise şudur:

> *Hangi 1.000?*

Aynı ülke ve yaş kotalarına sahip iki panel tamamen farklı sonuç üretebilir. Birinde:

- yoğun AI kullanıcıları,
- ücretli plan sahipleri,
- teknoloji çalışanları

çoğunluktadır. Diğerinde:

- seyrek kullanıcılar,
- mobil kullanıcılar,
- düşük dijital becerili kişiler

daha görünürdür. Her ikisi de 1.000 kişidir. Aynı nüfus değildir. Katılımcıların nereden geldiğini saklayıp yalnız ağırlıklı sonucu yayımlarsanız, ağırlıklandırmayı denetim yöntemi değil güven perdesi hâline getirirsiniz. Bir örneklemde hiç bulunmayan insanı ağırlıklandıramazsınız. Bir kullanıcının deneyimi yanlış olduğu için onu değiştiremezsiniz. Yanlış cevap, araştırmanın başarısızlığı değildir. Araştırmanın ölçmeye çalıştığı olaydır. Katılımcı cevabı doğru olduğu hâlde protokolü bozmuşsa kaydı koruyamazsınız. Doğru cevap, sahte gözlemi gerçek yapmaz. Bir kişiye daha çok ödeme yapabilirsiniz. Fakat istediğiniz cevabı üretmesi için değil. Bir ülkenin yaşam koşuluna uygun ödeme verebilirsiniz. Fakat katılımcının gönüllülüğünü satın alamazsınız. Bin ekran görüntüsü kanıt olabilir. Bin insanın kimliğini kamuya açmak zorunda değildir.

Kanıt bütünlüğü ile insan mahremiyeti birlikte korunabilir. Bu nedenle NOMOS'un sekizinci ölçüm yasası şudur:

> *Örneklemin gücü, kaç kişiden oluştuğu kadar bu kişilerin neden ve nasıl seçildiğiyle belirlenir.*

Dokuzuncu yasası şudur:

> *Ağırlıklandırma, gözlenmiş nüfusu yeniden dengeler; gözlenmemiş insanı yaratmaz.*

Onuncu yasası şudur:

> *Yanlış AI cevabı geçerli veri olabilir; doğru AI cevabı geçersiz veri olabilir.*

On birinci yasası şudur:

> *Katılımcının ödülü, istenen sonucu bulmaya değil, gördüğü sonucu değiştirmeden kanıtlamaya verilmelidir.*

On ikinci yasası şudur:

> *Kanıtlanabilirlik, katılımcının mahremiyetinden vazgeçmesini gerektirmez.*

NOMOS'un Bölüm 8 Emri

> *Bana yalnız bin insan getirme. / Bu bin insanın hangi dünyadan geldiğini göster.*
>
> *Gönüllü panelini rastgele nüfus; müşteri listesini genel kamu; çalışanı bağımsız kullanıcı yapma.*
>
> *Ülke ve yaş kotaları tuttu diye ölçmediğin bütün farkların düzeldiğini sanma.*
>
> *Bir kişiyi hesapları, cihazları, dilleri ve panel üyelikleri kadar çoğaltma.*
>
> *Markayı sevenleri, markadan nefret edenleri ve markayı hiç bilmeyenleri aynı görünmez kutuya koyma.*
>
> *Katılımcıya doğru cevabı bulması için değil, gördüğü ilk uygun cevabı değiştirmeden kaydetmesi için ödeme yap.*
>
> *AI yanlış cevap verdiğinde kullanıcıyı değiştirme. / Kullanıcı protokolü bozduğunda cevabın doğru olmasına bakma.*
>
> *Panelde hiç olmayan insanı ağırlıkla varmış gibi gösterme.*
>
> *Ham sonucu saklama. / Ağırlıklı sonucu tek gerçek ilan etme. / Etkin örneklemini göster. / Hassasiyetini göster.*
>
> *VPN kullananı sahte, ortak cihaz kullananı bot, yardımcı teknoloji kullananı otomasyon ilan etme.*

> *Sahteciliği araştır. / Ama araştırmanın kendisini yeni bir dışlama yanlılığına dönüştürme.*
>
> *Ekran görüntüsünü koru. / İnsanın adını, hesabını ve özel hayatını koru.*
>
> *Önce nüfusu tanımla. / Sonra çerçeveyi göster. / Sonra kimi seçtiğini, kimin gelmediğini, kimin tamamlayamadığını ve kimi neden dışladığını açıkla. / Ve ancak bundan sonra skorun insanlık adına ne kadar konuşabileceğine karar ver.*

Bölümün Son Cümlesi

> *GEO-1000'de bir katılımcı yalnız cevap taşıyan kişi değildir; hedef nüfustaki başka insanlar adına hangi ölçüde konuşabileceği, seçim yolu, ağırlığı, bağımsızlığı ve mahremiyetiyle birlikte kanıtlanması gereken insan birimidir.*

Normatif Öz

> **Her GEO-1000 gözlemi hedef nüfus ve örnekleme çerçevesinden seçim, davet, rıza, görev ataması, tamamlama, protokol doğrulaması, dışlama ve nihai ağırlığa kadar denetlenebilir katılımcı seçim yolu taşımalıdır. Panel içindeki rastgele seçim, tam nüfus katılım mekanizması bunu desteklemedikçe küresel hedef nüfustan rastgele seçim olarak sunulamaz. Katılımcının seçilmesi, ödüllendirilmesi, doğrulanması, değiştirilmesi, dışlanması ve ağırlıklandırılması AI cevabının olumlu, olumsuz, doğru, yanlış, ret, Critical veya ticari bakımdan elverişli olmasına bağlı olamaz. Ağırlıklandırma gözlenen dengesizliği azaltabilir; örnekleme çerçevesinde hiç bulunmayan insanları gözlenmiş hâle getiremez.**

KISIM III

Deneyin Uygulanması ve Kanıt Zinciri

BÖLÜM 09

AI Sistem Sicili ve Karşılaştırılabilirlik

YÖNTEMSEL DAYANAKLAR / K02; K03

Bölüm Sınırı

İlk sekiz bölüm şu yapıyı kurdu: Tek bir yanıt GEO skoru değildir. Temsil bir dağılımdır. Denetlenen varlık sonuçlardan önce tanımlanmalıdır. Resmî temsil ile doğrulanmış varlık gerçekliği ayrılmalıdır. Her AI ürünü için uygun kullanıcı nüfusu ayrıca kurulmalıdır. 1.000 geçerli gözlem nüfusa dayalı örneklem tasarımıyla dağıtılmalıdır. Küçük ülke ve diller ayrı tamamlayıcı panellerle görünür kılınmalıdır. Katılımcı seçimi, dışlama ve ağırlıklandırma AI cevabının içeriğinden bağımsız olmalıdır. Şimdi deneyin diğer tarafını tanımlamamız gerekir:

> *Katılımcılar tam olarak hangi AI sistemini ölçmektedir?*

Bir sağlayıcının adı mı? Bir temel model mi? Bir sohbet ürünü mü? Ücretsiz web arayüzü mü? Ücretli mobil uygulama mı? Web araması açık bir oturum mu? Kaynak göstermeyen bir cevap modu mu? Kurum içi veriyle desteklenen bir kurumsal çalışma alanı mı? Hafızası açık bir hesap mı? Temporary veya temiz sohbet mi? Aynı sağlayıcının farklı kullanıcı yüzeyleri aynı sistem midir? Aynı model ailesini kullanan iki ürün aynı şekilde mi ölçülmelidir? Bir ürün ekranda aynı model adını gösterse bile:

- retrieval,
- kaynak seçimi,
- güvenlik kuralları,
- kişiselleştirme,
- araç erişimi,
- deneysel özellikler

değişiyorsa sonuçlar karşılaştırılabilir midir? Bir AI ürünü ölçüm dalgasının ortasında güncellendiğinde, dalganın ilk ve son cevapları aynı sistemden gelmiş sayılabilir mi? Bir kullanıcının:

- web araması açık,
- hafıza açık,
- önceki konuşmaları bulunan,
- ücretli plan

hesabıyla aldığı cevap; başka kullanıcının:

- arama kapalı,
- hafıza kapalı,
- yeni oturum,
- ücretsiz plan

cevabıyla aynı hücrede birleştirilebilir mi? Hayır. Önceki bölümlerin kurduğu ölçüm ilkesi açıktır: Tek bir "AI" yoktur; arama indeksi, retrieval sistemi, reranker, dil modeli, atıf katmanı ve güvenlik sistemi birbirinden farklı davranabilir. Aynı çerçeve, her denetlenebilir sonucun:

- model veya ürün,
- tarih,
- ülke,
- dil,
- sorgu kümesi,
- tekrar sayısı,
- ölçüm kaydı

taşımasını istemişti. Bu bölüm o ilkeleri bir AI Sistem Sicili içinde işletir. Bu bölüm:

- sağlayıcı, ürün, model ve kullanıcı yüzeyi ayrımını,
- AI ürün örneğini,
- model kimliği güven düzeylerini,
- retrieval ve araç yapılandırmalarını,
- hafıza ve kişiselleştirme durumlarını,
- plan, arayüz ve deneysel özellik farklarını,
- ürün değişikliklerini,
- ölçüm anındaki sistem konfigürasyonunu,
- ürünler arası karşılaştırılabilirlik seviyelerini,
- Native Reach ve Common Support karşılaştırmalarını,
- zaman içinde sürüm kaymasını

tanımlar. Bu bölüm henüz:

- istemlerin tam çeviri ve semantik eşdeğerlik sistemini,
- kontrollü ve doğal kullanıcı panellerinin bütün ayrıntılarını,
- NOMOS Capture yazılım mimarisini,
- yanıtların atomik iddia puanlamasını,
- nihai model katsayılarını ve NOMOS skor formülünü

kesinleştirmez. Bölüm 9'un temel sorusu şudur:

> *Aynı istemden gelen iki cevabı karşılaştırmadan önce, bu cevapların gerçekten karşılaştırılabilir AI ürün koşullarından geldiğini nasıl kanıtlarız?*

NOMOS Challenge

Aynı sağlayıcının iki kullanıcısı aynı soruyu soruyor: "Apple.com nasıl bir şirkettir?" Birinci kullanıcı:

- ücretsiz web ürününü kullanıyor,
- yeni sohbet açıyor,
- hafızası kapalı,
- web erişimi bulunmuyor,

yalnız ürünün yerleşik bilgi kapasitesiyle cevap alıyor. İkinci kullanıcı:

- ücretli mobil ürünü kullanıyor,
- hafızası açık,

- daha önce Apple hakkında konuşmuş,
- web araması açık,

güncel kaynaklar ve atıf'larla cevap alıyor. Birinci cevap: "Apple, tüketici elektroniği üreten bir teknoloji şirketidir." İkinci cevap: "Apple; donanım, yazılım ve dijital hizmetler sunan küresel bir teknoloji şirketidir. Güncel kurumsal bilgileri şu kaynaklarda görebilirsiniz." İki cevap aynı sağlayıcıdan gelmiştir.

Aynı sistem sonucu mudur? Hayır. Şimdi başka iki kullanıcı düşünün. İkisi de ekranda aynı model adını görüyor. Birinin ülkesinde web özelliği açıktır. Diğerinin ülkesinde kapalıdır. Birinin cevabında kaynak bağlantıları görünür. Diğerinin arayüzü aynı kaynakları kullansa bile atıf göstermemektedir. Aynı model etiketi vardır.

Aynı ölçüm hücresi var mıdır? Hayır. Şimdi farklı sağlayıcılara ait iki ürün düşünün. İkisi de:

- yeni sohbet,
- hafıza kapalı,
- aynı dil,
- aynı ülke,
- aynı zaman penceresi,
- web erişimi açık,
- aynı istem sürümü

koşullarında ölçülüyor. Farklı temel modellere dayanırlar. Karşılaştırılabilirler mi? Evet, belirli kullanıcı sonucu bakımından karşılaştırılabilir olabilirler. Çünkü karşılaştırmanın nesnesi:

> *Aynı temel modeli kullanıp kullanmadıkları değil, aynı tanımlı kullanıcı koşulunda hangi temsili ürettikleridir.*

Şimdi ölçüm dalgasının ilk 500 gözleminde ürünün sürümü değişmeden kalsın. Sonraki 500 gözlem sırasında sağlayıcı:

- retrieval sistemini,
- güvenlik kurallarını,
- kullanıcı arayüzünü

değiştirsin. Rapor yalnız: "1.000 kullanıcı aynı AI ürününü test etti." diyor. Gerçekte iki farklı ürün durumu ölçülmüş olabilir. Bu bölümün ilk hükmü şudur:

> *Sağlayıcı adı, ölçülen AI sistemini tanımlamaya yetmez.*

İkinci hükmü şudur:

> *Model adı, kullanıcıya ulaşan ürün davranışının tamamı değildir.*

Üçüncü hükmü şudur:

> *Aynı sağlayıcı, aynı model etiketi veya aynı logo karşılaştırılabilirlik garantisi vermez.*

Dördüncü hükmü şudur:

> *Farklı sağlayıcılar, ölçüm koşulları yeterince eşleştirildiğinde kullanıcı sonucu bakımından karşılaştırılabilir olabilir.*

Beşinci hükmü şudur:

> *Bir ürün ölçüm sırasında maddi biçimde değişirse tek bir sabit sistem gibi raporlanamaz.*

1. BÖLÜMÜN AMACI

Bu bölümün amacı, GEO-1000 içinde denetlenen AI ürününü bütün maddi yapılandırmalarıyla tanımlamak ve farklı ürün sonuçlarının hangi koşullarda karşılaştırılabileceğini belirlemektir. Bölüm şu ayrımları normatif hâle getirir:

- Sağlayıcı ile AI ürünü
- AI ürünü ile temel model
- Model ailesi ile belirli dağıtım
- Model etiketi ile doğrulanmış model kimliği
- Kullanıcı yüzeyi ile API
- Web ile mobil ürün
- Ücretsiz plan ile ücretli plan
- Tüketici ürünü ile kurumsal ürün
- Temel model cevabı ile retrieval destekli cevap
- Retrieval ile canlı web taraması
- Kaynak kullanımı ile atıf gösterimi
- Araç erişimi ile araç kullanımı
- Hafıza ile oturum bağlamı
- Özel talimat ile sistem politikası
- Kişiselleştirme ile maddi gerçeklik
- Yeni sohbet ile geçici sohbet
- Statik ürün adı ile zaman içindeki ürün durumu
- Sağlayıcının duyurduğu güncelleme ile gözlenen davranış değişimi
- Native Reach ile Common Support karşılaştırması
- Aynı ürün ailesi ile karşılaştırılabilir ürün örneği
- Kullanıcı sonucu ile model içi soyut yetenek
- Ürün bulunmaması ile ürün başarısızlığı
- Teknik hata ile gerçek ret
- Aynı istem ile aynı deney
- Aynı ortalama ile aynı sistem davranışı
- Anlık ürün fotoğrafı ile uzun dönem ürün performansı

Bu bölümün sonunda her GEO-1000 denetimi şu sorulara cevap verebilmelidir:

> *Hangi sağlayıcının hangi ürünü, hangi yüzeyi, hangi planı ve hangi yapılandırması ölçüldü?*
>
> *Ekranda görülen model adı gerçekten ne kadar doğrulandı?*
>
> *Web, retrieval, atıf, hafıza, kişiselleştirme ve araç durumu neydi?*
>
> *Ölçüm boyunca ürün aynı kaldı mı?*
>
> *İki ürün sonucu hangi karşılaştırılabilirlik seviyesinde yan yana getirildi?*
>
> *Karşılaştırılan şey temel model mi, tüketici ürünü mü, kullanıcı yüzeyi mi, yoksa bütün sistem zinciri mi?*

2. MERKEZÎ NORMATİF HÜKÜM

> *Her GEO-1000 gözlemi, yalnız sağlayıcı veya model adıyla değil; ölçüm anındaki AI ürün örneğini tanımlayan sürümlü bir AI Sistem Sicili kaydına bağlanmalıdır.*

Bir AI ürün örneği en azından ilgili olduğu ölçüde şu alanları taşımalıdır:

- Sağlayıcı
- Ürün adı
- Kullanıcı yüzeyi
- Ürün türü
- Plan veya hesap sınıfı
- Görünen model etiketi
- Model kimliği güven düzeyi
- Ürün veya dağıtım sürümü
- Ülke ve locale
- Dil
- Web veya retrieval erişimi
- Atıf özelliği
- Araç erişimi
- Hafıza durumu
- Özel talimat durumu
- Önceki sohbet bağlamı
- Oturum türü
- Kişiselleştirme durumu
- Güvenlik veya politika modu, biliniyorsa
- Deneysel özellikler
- Ölçüm tarihi ve UTC zamanı
- Ölçüm dalgası
- Ürün değişiklik kaydı
- Kanıt ve kayıt güveni

Bu alanlardan maddi olan biri değiştiğinde:

- yeni AI ürün örneği,
- yeni deneysel hücre,
- yeni analiz sürümü

gerekebilir.

3. “AI SİSTEMİ” TEK BİR NESNE DEĞİLDİR [K02; K03]

Bu protokolde “AI sistemi” genel ve üst düzey bir terimdir. Gerçek kullanıcı cevabı birden fazla katmanın birleşiminden doğabilir. Aday zincir:

```
R = F(M, O, K, Q, T, P, S, U, C, L, t)
```

Burada:

- M: temel veya yönlendirici model

- O: ürün orkestrasyon katmanı
- K: retrieval veya bilgi kaynağı katmanı
- Q: reranking ve kaynak seçimi
- T: araçlar
- P: politika ve güvenlik katmanı
- S: oturum, hafıza ve kişiselleştirme
- U: kullanıcı hesabı ve planı
- C: ülke ve locale
- L: dil
- t: zaman

Kullanıcının gördüğü yanıt:

R

yalnız temel modelin çıktısı olmayabilir. Bu nedenle GEO-1000 ana tüketici denetiminin nesnesi:

> *Model soyutlaması değil, kullanıcıya sunulan ürün davranışıdır.*

Kontrollü API veya model araştırması ayrıca yapılabilir. Ana tüketici paneliyle aynı sonuç değildir.

4. AI SİSTEMİNİN YEDİ KİMLİK KATMANI

Her denetlenen AI sistemi ilgili olduğu ölçüde yedi katmanda tanımlanmalıdır.

4.1. Sağlayıcı Katmanı

Ürünü yayımlayan veya işleten kuruluştur. Sağlayıcı adı:

- ürün kimliği,
- model kimliği,
- kullanıcı yüzeyi

ile aynı değildir. Bir sağlayıcı birden fazla ürün sunabilir. Aynı ürün içinde birden fazla model veya orkestrasyon sistemi kullanabilir.

4.2. Ürün Katmanı

Kullanıcının eriştiği ticari, tüketici, kurumsal veya teknik üründür. Örnek türler:

- Genel sohbet ürünü
- Arama destekli cevap ürünü
- Kurumsal çalışma alanı
- Eğitim ürünü
- Gömülü asistan
- Geliştirici API'si
- Özel bilgi tabanlı ürün

Ürün adı tek başına bütün yapılandırmayı açıklamaz.

4.3. Kullanıcı Yüzeyi Katmanı

Kullanıcının ürünle etkileştiği arayüzdür. Web Mobil Masaüstü Tarayıcı uzantısı İşletim sistemi entegrasyonu Platform içi gömülü arayüz

API

Kurumsal portal Aynı ürünün farklı yüzeyleri:

- farklı özellik,
- farklı model,
- farklı kaynak gösterimi,
- farklı hafıza

kullanabilir.

4.4. Model veya Çıkarım Motoru Katmanı

Cevap üretiminde kullanılan model, model ailesi veya yönlendirici sistemdir. Model kimliği:

- açık,
- kısmen görünür,
- bilinmeyen,
- dinamik

olabilir. Bir ürün kullanıcı sorgusunu farklı modellere yönlendirebilir. Bu durumda: "Tek bir sabit model ölçüldü." denemez.

4.5. Orkestrasyon ve Bilgi Katmanı

Şunları içerebilir:

- Retrieval
- Canlı web araması
- Kurum içi arama
- Kaynak sıralama
- Reranking
- Atıf seçimi
- Araç çağırma
- Sorgu yeniden yazımı
- Belge parçalama
- Cevap birleştirme

Bu katman GEO temsili için temel önem taşır. Çünkü hangi kaydın:

- erişildiğini,
- seçildiğini,
- cevaba taşındığını,
- atıf olarak gösterildiğini

etkileyebilir.

4.6. Politika ve Kişiselleştirme Katmanı

Şunları içerebilir:

- Güvenlik kuralları
- Sistem talimatları
- Ürün politikası
- Hafıza
- Özel talimatlar

- Kullanıcı geçmişi
- Bölgesel kurallar
- Yaş veya hesap sınıfı
- Kurumsal yönetici ayarları

Bu katman cevabın:

- kapsamını,
- tonunu,
- reddini,
- tavsiye'ını

değiştirebilir.

4.7. Dağıtım ve Zaman Katmanı

Ürünün belirli anda hangi:

- sürüm,
- deneysel özellik,
- bölgesel dağıtım,
- kullanıcı segmenti,
- A/B testi

altında çalıştığını ifade eder. Aynı ürün adı aynı gün içinde farklı kullanıcılarda farklı dağıtım koşullarına sahip olabilir.

5. AI ÜRÜN ÖRNEĞİ

GEO-1000 içindeki temel sistem birimi:

AI Product Instance

Türkçesi:

AI Ürün Örneği

olarak adlandırılır. Bir AI ürün örneği:

```
A_i = (v, p, s, r, m, k, t, x, u, c, l, ω)
```

şeklinde gösterilebilir. Burada:

- v: sağlayıcı
- p: ürün
- s: kullanıcı yüzeyi
- r: plan veya hesap sınıfı
- m: model kimliği veya etiketi
- k: retrieval ve bilgi yapılandırması
- t: araç yapılandırması
- x: hafıza, kişiselleştirme ve oturum durumu
- u: kullanıcı veya kurumsal ayar sınıfı
- c: ülke ve locale
- l: dil
- ω: zaman, dağıtım ve deneysel özellik durumu

Bir gözlem: Oj mutlaka belirli bir: Ai ürün örneğine bağlanmalıdır.

6. ÜRÜN ÖRNEĞİ İLE ÜRÜN ADI ARASINDAKİ FARK

Ürün adı sabit olabilir. Ürün örneği değişebilir. Örnek: Orion AI ürün adı değişmeden kalsın. Fakat şu iki ölçüm farklı ürün örnekleridir:

Ürün Örneği 1

Web Ücretsiz plan Hafıza kapalı Web erişimi kapalı Yeni sohbet Türkçe Türkiye Dalga 1

Ürün Örneği 2

Mobil Ücretli plan Hafıza açık Web erişimi açık Önceki sohbet bağlamı mevcut Türkçe Türkiye Dalga 1 Bunlar aynı marka altında bulunabilir. Aynı deneysel hücre değildir.

7. MODEL KİMLİĞİ GÜVEN DÜZEYLERİ

Model adı her zaman aynı güvenle bilinemez. Bu nedenle model kimliği şu düzeylerle kaydedilebilir.

MI-0 — UNKNOWN MODEL

Ürünün hangi model veya model ailesini kullandığı bilinmemektedir. Sonuç yalnız ürün yüzeyine atfedilir.

MI-1 — USER-DISPLAYED LABEL

Kullanıcı arayüzünde bir model etiketi görünmektedir. Etiket kayıt altına alınır. Arka uç dağıtımını bütünüyle doğruladığı varsayılmaz.

MI-2 — PROVIDER-DOCUMENTED MODEL FAMILY

Sağlayıcının ürün belgeleri belirli model ailesini veya ürün–model ilişkisini açıklamaktadır. Bireysel gözlemin tam olarak o modele yönlendirildiği yine bilinmeyebilir.

MI-3 — SESSION-CONFIRMED DEPLOYMENT

Oturum, API veya güvenilir ürün metadata'sı belirli dağıtımı doğrulamaktadır.

MI-4 — VERSIONED ENDPOINT OR DEPLOYMENT

Belirli sürüm, endpoint veya dağıtım kimliği kaydedilebilmektedir. Kontrollü karşılaştırma için daha güçlüdür.

MI-5 — REPRODUCIBLE FROZEN MODEL ARTIFACT

Sabit ve yeniden çalıştırılabilir model veya dağıtım yapısı bulunmaktadır. Canlı tüketici ürünlerinde nadir olabilir. Model kimliği güveni düşükse sonuç:

ürün davranışına

atfedilmelidir. Belirli modele kesin biçimde atfedilemez.

8. GÖRÜNEN MODEL ADININ SINIRI

Bir kullanıcı arayüzünde model adı görmek değerlidir. Fakat şu sorular yine açık kalabilir: Bütün sorgular aynı modele mi yönlendiriliyor? Sorgu zorluğuna göre model değişiyor mu? Güvenlik veya

retrieval ayrı model kullanıyor mu? Sağlayıcı aynı etiketi farklı dağıtımlara uyguluyor mu? Model etiketi yalnız ürün ailesini mi gösteriyor? Kullanıcı bir deneysel rollout içinde mi? Bu nedenle doğru raporlama: "Kullanıcı arayüzünde Model X etiketi gösterildi." olabilir. Daha güçlü kanıt olmadan şu cümle kullanılmamalıdır: "Bütün cevap kesin olarak Model X'in tek başına ürettiği çıktıdır."

9. ÜRÜN TÜRÜ

Her AI ürünü ana tür ve ikincil türlerle sınıflandırılmalıdır.

PT-1 — General Conversational Product

Genel sohbet ve soru-cevap ürünüdür.

PT-2 — Search- or Retrieval-Grounded Product

Cevap üretiminde arama veya retrieval temel bileşendir.

PT-3 — Enterprise Knowledge Product

Kurum içi belgeler, çalışma alanları veya özel veri kaynakları kullanabilir.

PT-4 — Embedded Assistant

Başka ürün veya platform içine gömülüdür.

PT-5 — Developer or API Product

Son kullanıcı yerine geliştirici veya uygulama tarafından kullanılır.

PT-6 — Domain-Specific Product

Belirli meslek veya sektör için tasarlanmıştır.

PT-7 — Multimodal Product

Metin dışında görsel, ses veya belge girişleri ana deneyimin parçasıdır. Bir ürün birden fazla tür taşıyabilir. Karşılaştırma aynı tür içinde veya açık kapsamla yapılmalıdır.

10. KULLANICI YÜZEYİ SİCİLİ

Aynı ürünün farklı yüzeyleri ayrı kaydedilmelidir. Yüzey alanları:

- Yüzey kimliği
- Web/mobil/masaüstü/API
- Uygulama sürümü
- İşletim sistemi, maddi ise
- Tarayıcı, maddi ise
- Ürün dili
- Görünen model seçimi
- Araçların bulunabilirliği
- Atıf görünümü
- Dosya yükleme özelliği
- Hafıza ayarları

- Kurumsal yönetim ayarları
- Deneysel özellikler
- Ölçüm zamanı

Bir web ve mobil arayüz maddi bakımdan aynı bulunursa sonuçlar birleşebilir. Bu eşitlik varsayılmamalı, test edilmelidir.

11. PLAN VE HESAP SINIFI

Plan türü kullanıcı cevabını etkileyebilecek özellikler taşıyabilir. Plan statüleri:

- Ücretsiz
- Ücretli bireysel
- Profesyonel
- Kurumsal
- Eğitim
- Geliştirici
- Deneme
- Davet veya sınırlı erişim
- Bilinmeyen

Planlar şu alanlarda farklılaşabilir:

- Model
- Kullanım limiti
- Web erişimi
- Hafıza
- Araçlar
- Kaynak gösterimi
- Yanıt uzunluğu
- Hız veya öncelik
- Kurumsal veri
- Güvenlik ve saklama ayarları

Plan bilgisi bulunmuyorsa:

PLAN_UNKNOWN

statüsü kullanılmalıdır. Ücretsiz ve ücretli sonuçlar sessizce aynı ürün davranışı olarak birleştirilemez.

12. RETRIEVAL VE WEB ERİŞİMİ

Bir AI cevabı üç temel bilgi durumundan birinde üretilebilir.

KG-0 — UNDETERMINED KNOWLEDGE MODE

Retrieval veya web kullanım durumu belirlenememektedir.

KG-1 — MODEL-INTERNAL YANIT

Canlı veya dış bilgi erişimi kullanılmadığı bilinen ya da güçlü biçimde varsayılan cevap.

KG-2 — RETRIEVAL-ASSISTED YANIT

Önceden indekslenmiş veya ürün içi kaynaklardan retrieval kullanılmış olabilir.

KG-3 — LIVE WEB OR SEARCH-ASSISTED YANIT

Canlı veya ölçüm zamanına yakın web/search erişimi kullanılmıştır.

KG-4 — PRIVATE KNOWLEDGE-GROUNDED YANIT

Kurum içi veya kullanıcı tarafından sağlanmış özel veri kaynağı kullanılmıştır.

KG-5 — HYBRID OR MULTI-SOURCE YANIT

Birden fazla bilgi katmanı birlikte kullanılmıştır. Bu modlar aynı istemde farklı sonuçlar üretebilir. Karşılaştırmada açıkça ayrılmalıdır.

13. RETRIEVAL ERİŞİMİ İLE RETRIEVAL KULLANIMI AYNI DEĞİLDİR

Ürünün web arama özelliği bulunabilir. Fakat belirli cevapta kullanılmamış olabilir. Üç alan ayrı kaydedilmelidir: Özellik mevcut muydu? Özellik açık mıydı veya seçilebilir miydi? Belirli yanıtta kullanıldığına dair kanıt var mıydı? Örnek statüler:

AVAILABLE_NOT_USED

AVAILABLE_USAGE_UNKNOWN

USED_CONFIRMED

NOT_AVAILABLE

DISABLED

UNKNOWN

Yalnız kaynak bağlantısı bulunması kullanım için güçlü işaret olabilir. Kaynak bağlantısının bulunmaması retrieval kullanılmadığını kesin göstermeyebilir.

14. ATIF KATMANI

Kaynak kullanımı ile atıf gösterimi aynı şey değildir. Bir ürün:

- kaynak kullanabilir fakat göstermeyebilir,
- atıf gösterebilir fakat atıf iddiayı desteklemeyebilir,
- aynı kaynağı farklı cümlelere bağlayabilir,

kullanıcı arayüzüne göre atıfı gizleyebilir. Atıf sicili şu alanları taşımalıdır: Atıf özelliği mevcut mu? Belirli cevapta atıf gösterildi mi? Atıf sayısı Kaynak alan adları Atıf–iddia eşleşmesi Kullanıcı yüzeyindeki görünürlük Atıfın açılabilirliği Kaynağın erişilebilirliği Atıf üretim zamanı Atıf gösteren ve göstermeyen ürünler yalnız: “kaynak sayısı” bakımından doğrudan karşılaştırılamaz. Önce ürün davranışı ve arayüz farkı açıklanmalıdır.

15. ARAÇ ERİŞİMİ

AI ürünü şu araçlara erişebilir:

- Web arama
- Tarayıcı
- Hesap makinesi
- Kod çalıştırma
- Dosya okuma
- Harita
- Takvim
- Kurumsal bilgi tabanı
- Ürün kataloğu
- Başka API'ler

Araç erişimi yanıtın:

- güncelliğini,
- kanıtını,
- kapsamını,
- doğruluğunu

etkileyebilir. Araç sicili:

- aracın mevcut olup olmadığını,
- etkin olup olmadığını,
- cevapta kullanılıp kullanılmadığını,
- araç sonucunun korunup korunmadığını

göstermelidir.

16. HAFIZA, ÖZEL TALİMAT VE KİŞİSELLEŞTİRME

Kullanıcıya görünen cevap şu alanlardan etkilenebilir:

- Kalıcı hafıza
- Sohbet geçmişi
- Özel talimatlar
- Kullanıcının adı veya profili
- Önceki marka konuşmaları
- Tercihler
- Ülke ve locale
- Kurumsal çalışma alanı ayarları
- Sağlayıcı tarafından yapılan segmentasyon

Bu alanlar üç ayrı katmanda kaydedilmelidir.

16.1. Kalıcı Kişiselleştirme

Hesaplar arası veya oturumlar arası korunan bilgi.

16.2. Konuşma İçi Bağlam

Mevcut sohbet içindeki önceki mesajlar.

16.3. Sistem veya Kurumsal Ayar

Kullanıcının göremeyebileceği yönetici veya ürün talimatları. Tam bilgi her zaman elde edilemeyebilir. Bilinmeyen durum açıkça kaydedilmelidir.

17. OTURUM TÜRLERİ

Aday oturum statüleri:

ST-1 — NEW STANDARD CONVERSATION

Yeni standart sohbet. Hafıza veya özel talimatlar ayrıca kaydedilir.

ST-2 — TEMPORARY OR EPHEMERAL CONVERSATION

Kalıcı geçmişten ayrılması amaçlanan geçici oturum. Gerçek özellikleri ürün sicilinde doğrulanmalıdır.

ST-3 — EXISTING CONVERSATION

Önceki konuşma bağlamı vardır. Doğal panel için değerli olabilir. Kontrollü panelde ayrı tutulur.

ST-4 — ENTERPRISE OR SHARED WORKSPACE SESSION

Kurumsal veya ortak bilgi alanı bulunabilir.

ST-5 — UNKNOWN SESSION STATE

Oturum statüsü güvenilir biçimde belirlenememiştir. Aynı hücrede farklı oturum türleri birleştirilecekse gerekçe ve ağırlık açıklanmalıdır.

18. ÜRÜN POLİTİKASI VE RET

Bir AI ürünü belirli soruya:

- güvenlik,
- hukuk,
- ürün politikası,
- bölgesel kural,
- yaş,
- kurumsal ayar

nedeniyle cevap vermeyebilir. Ret:

- temel model eksikliği,
- ürün politikası,
- arayüz,
- kullanıcı hesabı

kaynaklı olabilir. GEO-1000 ana kullanıcı denetimi, kullanıcıya ulaşan ret sonucunu kaydeder. Fakat ret nedeni biliniyorsa ayrıca sınıflandırılır. Bir ürünün politika katmanı da kullanıcı deneyiminin parçasıdır.

19. DENEYSEL ÖZELLİKLER VE A/B DAĞITIMLARI

Canlı AI ürünleri bütün kullanıcılara aynı anda aynı özellikleri vermeyebilir. Katılımcılar:

- farklı kullanıcı arayüzü,
- farklı model yönlendirmesi,
- farklı web özelliği,
- farklı cevap biçimi

görebilir. Deneysel durum göstergeleri:

- Arayüzde beta etiketi
- Sağlayıcı belgesi
- Özellik farklılığı
- Kullanıcılar arası sistematik UI farkı
- Ölçüm sırasında davranış kırılması

olabilir. Deneysel özellik bilinmiyorsa:

EXPERIMENTAL_STATE_UNKNOWN

kullanılmalıdır. Ürün içindeki gerçek varyasyon, veri kusuru değildir. Sistemin dağılım özelliğidir. Fakat karşılaştırılabilirlikte dikkate alınmalıdır.

20. AI SİSTEM SİCİLİ

Her denetlenen ürün için AI Sistem Sicili oluşturulmalıdır. Sicil üç düzey taşımalıdır.

20.1. Sağlayıcı ve Ürün Ana Kaydı

Görece kalıcı alanları içerir:

- Sağlayıcı
- Ürün adı
- Ürün türü
- Ana kullanıcı yüzeyleri
- Resmî erişim ülkeleri
- Plan sınıfları
- Desteklenen dil kayıtları
- Genel özellikler

20.2. Ölçüm Dalgası Ürün Kaydı

Belirli dalga için:

- ürün sürümü,
- plan özellikleri,
- model etiketleri,
- web ve araç özellikleri,
- hafıza davranışı,
- erişim durumu,
- bilinen güncellemeler

kaydedilir.

20.3. Gözlem Düzeyi Sistem Kaydı

Her katılımcı için fiilen görülen:

- arayüz,
- model etiketi,
- plan,
- araç kullanımı,
- atıf,
- oturum,
- kişiselleştirme,
- zaman

alanlarını taşır. Ana kayıt tek başına gözlem düzeyi farklılıkları silemez.

21. SİSTEM KONFİGÜRASYON PARMAK İZİ

Bir AI ürün örneğinin karşılaştırılabilirliğini kolaylaştırmak için Sistem Konfigürasyon Parmak İzi oluşturulabilir. Aday alanlar:

```
SCF = H(canonicalize(p, s, r, m, k, t, x, c, l, w))
```

Burada H, sürümlenmiş alanlardan bütünlük değeri üreten fonksiyondur. Parmak izi:

- iki gözlemin aynı yapılandırmaya ait olup olmadığını,
- dalga içinde maddi değişiklik oluşup oluşmadığını

tespit etmeye yardımcı olabilir. Parmak izi: gerçek arka uç modelini kanıtlamaz, sağlayıcının açıklamadığı katmanları görünür yapmaz. Yalnız kaydedilen konfigürasyonun daha sonra değiştirilip değiştirilmediğini güçlendirir.

22. KARŞILAŞTIRILABİLİRLİK NEDİR?

İki sonuç karşılaştırılabilir olduğunda:

> *Aralarındaki farkın, ölçüm tasarımındaki ilgisiz farklılıklardan çok denetlenen ürün davranışıyla açıklanabilme ihtimali yeterince güçlüdür.*

Karşılaştırılabilirlik mutlak eşitlik değildir. Farklı AI ürünlerinin:

- model,
- sağlayıcı,
- arayüz

bakımından aynı olması beklenmez. Ama karşılaştırma sorusuna göre maddi koşulların:

- eşitlenmesi,
- tabakalanması,
- kaydedilmesi,
- analizde modellenmesi

gerekir.

23. KARŞILAŞTIRMANIN NESNESİ ÖNCEDEN TANIMLANMALIDIR

Aynı veri farklı karşılaştırma sorularına cevap verebilir.

23.1. Kullanıcı Ürünü Karşılaştırması

"Aynı kullanıcı koşullarında hangi tüketici ürünü daha doğru temsil üretiyor?" Karşılaştırılan nesne bütün ürün zinciridir.

23.2. Model Karşılaştırması

"Aynı kontrollü istem ve araçsız koşulda hangi model daha doğru cevap üretiyor?" API veya sabit model erişimi gerekebilir.

23.3. Retrieval Karşılaştırması

"Aynı ürünün web erişimli ve erişimsiz modları arasında fark var mı?" Model ve diğer koşullar mümkün olduğunca sabit tutulur.

23.4. Plan Karşılaştırması

"Ücretsiz ve ücretli kullanıcılar aynı marka temsilini görüyor mu?" Plan farkı denetimin nesnesidir.

23.5. Arayüz Karşılaştırması

"Web ve mobil ürün aynı koşullarda farklı temsil üretiyor mu?"

23.6. Zaman Karşılaştırması

"Aynı ürünün temsili farklı dalgalarda değişti mi?" Ürün ve nüfus değişiklikleri ayrılmalıdır. Karşılaştırma sorusu belirtilmeden: "Ürün A, Ürün B'den daha iyi." cümlesi aşırı geniştir.

24. KARŞILAŞTIRILABİLİRLİK BOYUTLARI

İki ürün veya dalga karşılaştırılırken aşağıdaki alanlar değerlendirilmelidir:

- Denetlenen varlık
- Gerçeklik paketi sürümü
- İstem kimliği ve sürümü
- Dil ve locale
- Ülke
- Hedef nüfus
- Örneklem kaynağı
- Katılımcı kompozisyonu
- Oturum türü
- Hafıza
- Özel talimat
- Önceki konuşma
- Ürün planı
- Kullanıcı yüzeyi

- Model etiketi
- Retrieval/web durumu
- Araçlar
- Atıf özelliği
- Ölçüm zamanı
- Ürün güncellemeleri
- Hakemlik kuralı
- Sonuç statüleri
- Ağırlıklandırma

Bütün boyutların aynı olması gerekmez. Farklı olan boyutlar karşılaştırmanın kapsamını sınırlar.

25. KARŞILAŞTIRILABİLİRLİK SEVİYELERİ

CP-0 — INCOMPARABLE

Maddi ölçüm koşulları bilinmiyor veya ciddi biçimde uyumsuzdur. Doğrudan performans karşılaştırması yapılamaz.

CP-1 — DESCRIPTIVE CO-OCCURRENCE

İki ürün veya dalgadan sonuç vardır. Fakat:

- nüfus,
- istem,
- zaman,
- oturum

eşit değildir. Yalnız betimleyici yan yana gösterim yapılabilir.

CP-2 — PARTIALLY ALIGNED

Bazı temel alanlar eşleşir. Maddi farklılıklar devam eder. Sınırlı karşılaştırma yapılabilir.

CP-3 — MATCHED USER-SURFACE COMPARISON

Aynı varlık Aynı istem sürümü Aynı dil/ülke Benzer kullanıcı nüfusu Aynı zaman penceresi Kayıtlı ürün özellikleri bulunmaktadır. Tüketici ürünü sonucu karşılaştırılabilir.

CP-4 — CONTROLLED MATCHED COMPARISON

Ek olarak:

- oturum,
- hafıza,
- kişiselleştirme,
- retrieval,
- araçlar,
- plan

kontrollü veya açık biçimde eşleştirilmiştir. Daha güçlü ürün veya özellik karşılaştırması sağlar.

CP-5 — REPLICATED AND CALIBRATED COMPARISON

Karşılaştırma:

- birden fazla dalgada,
- bağımsız kullanıcı veya panelde,
- aynı normatif değerlendirmeyle

tekrarlanmış; değişkenlik ve ürün değişiklikleri kalibre edilmiştir. Bu seviyeler "bilimsel doğruluk puanı" değildir. Karşılaştırma iddiasının taşıyabileceği sınırı gösterir.

26. CP-0 DURUMUNDA NE YAPILIR?

İki ürün karşılaştırılamıyorsa sonuçlar yok sayılmaz. Ayrı ayrı raporlanabilir:

- "Ürün A'nın Native Reach sonucu budur."
- "Ürün B'nin farklı kullanıcı ve ürün koşullarındaki sonucu budur."

Fakat: "A, B'den 12 puan daha iyidir." denemez. Karşılaştırılamazlık bir başarısızlık değil, yöntem sınırıdır.

27. AYNI SAĞLAYICI İÇİN KARŞILAŞTIRILABİLİRLİK

Aynı sağlayıcıya ait sonuçlar otomatik olarak karşılaştırılabilir değildir. Aşağıdaki farklar bulunabilir:

- Ürün
- Plan
- Model
- Web erişimi
- Arayüz
- Hafıza
- Kurumsal veri
- Ülke dağıtımı
- Deneysel özellik
- Zaman

Aynı sağlayıcı karşılaştırması da CP seviyesine göre değerlendirilmelidir.

28. FARKLI SAĞLAYICILAR İÇİN KARŞILAŞTIRILABİLİRLİK

Farklı sağlayıcılara ait kullanıcı ürünleri:

- aynı hedef nüfus,
- aynı istem,
- aynı zaman,
- benzer kullanıcı yüzeyi,
- açık ürün yapılandırması

altında karşılaştırılabilir. Farklı sağlayıcı olmak karşılaştırmayı engellemez. Ancak özellik farkları karşılaştırmanın neyi ölçtüğünü belirler. Örnek: Ürün A canlı web kullanıyor. Ürün B kullanmıyor.

Sonuç şu soruya cevap verebilir: “Kullanıcıya sunulan iki ürün zincirinden hangisi daha doğru temsil üretiyor?” Şu soruya doğrudan cevap vermez: “Hangi temel model daha bilgili?”

29. WEB DESTEKLİ VE WEBSİZ ÜRÜNLERİN KARŞILAŞTIRILMASI

İki ayrı karşılaştırma mümkündür.

29.1. Doğal Ürün Karşılaştırması

Her ürün varsayılan gerçek kullanıcı koşulunda ölçülür. Web kullanan ürün web kullanır. Kullanmayan kullanmaz. Sonuç: gerçek tüketici ürün deneyimi karşılaştırmasıdır.

29.2. Özellik Eşleştirilmiş Karşılaştırma

Mümkünse her iki ürün:

- web açık,
- web kapalı

modlarında ayrıca ölçülür. Bu, retrieval etkisini daha iyi ayırabilir. Her ürün aynı seçenekleri sunmayabilir. Eksik mod:

NOT AVAILABLE

olarak kaydedilir.

30. ATIF GÖSTEREN VE GÖSTERMEYEN ÜRÜNLER

Atıf özelliği olmayan ürüne: “Sıfır atıf” vermek mümkündür. Fakat bunun anlamı açıklanmalıdır: Ürün kaynak göstermiyor olabilir. Kaynak kullanmıyor olduğu kesin değildir. Atıf kapsamı ile doğruluk kapsamı ayrı sonuçlardır. Atıf karşılaştırması yalnız atıf destekleyen ürünler arasında veya ayrı özellik statüsüyle yapılmalıdır.

31. NATIVE REACH KARŞILAŞTIRMASI

Her ürün kendi gerçek erişim nüfusunda ölçülür. Avantajı: ürünün doğal küresel etkisini gösterir. Sınırlaması: nüfuslar farklıdır, ülke ve dil dağılımları farklı olabilir. Native Reach skorları: ürünün kendi eriştiği insanlarda ne yaptığını gösterir. Doğrudan ürün üstünlüğü sıralamasında tek başına kullanılmamalıdır.

32. COMMON SUPPORT KARŞILAŞTIRMASI

Karşılaştırılan ürünlerin ortak erişim nüfusu kullanılır. Avantajı:

- aynı ülke,
- aynı dil,
- aynı kullanıcı koşulu

üzerinden daha adil karşılaştırma sağlar. Sınırlaması: ürünlerin gerçek küresel erişim farkını saklar, ortak evren çok dar olabilir. Common Support skoru ile Native Reach skoru birlikte yayımlanmalıdır.

33. EŞLENMİŞ KULLANICI KARŞILAŞTIRMASI

Aynı kullanıcı iki veya daha fazla AI ürününü test ederse kullanıcı özellikleri kontrol edilebilir. Ancak:

- görev sırası,
- cevap hatırlama,
- ürün aşinalığı,
- bütün ürünlere erişebilen kullanıcıların özel profili

yanlılık yaratabilir. Eşlenmiş tasarımda:

- ürün sırası rastgeleleştirilmeli,
- önceki cevaplar paylaşılmamalı,

aynı kullanıcı bağımlılığı analizde korunmalıdır.

34. BAĞIMSIZ KULLANICI KARŞILAŞTIRMASI

Her ürün farklı fakat aynı nüfus çerçevesinden seçilen kullanıcılar tarafından test edilir. Avantajı: görev bulaşması azalır, daha doğal ürün kullanıcıları seçilebilir. Sınırlaması: kullanıcı grupları tamamen eşit olmayabilir. Ağırlıklandırma ve tabakalama gerekir. Eşlenmiş ve bağımsız kullanıcı sonuçları ayrı ama tamamlayıcı olabilir.

35. ZAMAN KARŞILAŞTIRILABİLİRLİĞİ

Bir AI ürününün bugünkü sonucu geçmiş sonuçla karşılaştırılırken şu değişiklikler incelenmelidir:

- Model
- Ürün sürümü
- Web/retrieval
- Plan
- Ülke erişimi
- Dil desteği
- İstem
- Varlık gerçekliği
- Hakemlik
- Nüfus çerçevesi
- Ağırlıklandırma

Zaman farkı yalnız AI ürününden kaynaklanmayabilir.

36. ÜRÜN DEĞİŞİKLİĞİ TÜRLERİ

UC-1 — DOCUMENTED MINOR CHANGE

Sağlayıcı tarafından açıklanmış ve ölçüm nesnesini maddi biçimde değiştirmediği değerlendirilen küçük değişiklik.

UC-2 — DOCUMENTED MATERIAL CHANGE

Model, retrieval, arayüz veya özellikte açıklanmış maddi değişiklik.

UC-3 — OBSERVED BEHAVIOR CHANGE

Resmî açıklama bulunmasa da sistematik davranış veya arayüz değişimi gözlenir.

UC-4 — ACCESS OR PLAN CHANGE

Ülke, plan veya kullanıcı erişimi değişir.

UC-5 — UNKNOWN CHANGE EVENT

Ölçüm davranışında kırılma vardır. Nedeni doğrulanamamıştır.

UC-6 — PRODUCT REPLACEMENT OR RETIREMENT

Eski ürün veya model sona ermiş, yerine başka yapı gelmiştir. Değişiklik türü zaman serisi karşılaştırmasını sınırlar.

37. ÖLÇÜM DALGASI İÇİNDE ÜRÜN DEĞİŞİKLİĞİ

Bir dalga sırasında maddi değişiklik oluşursa şu seçenekler vardır: Dalga durdurulur ve yeniden başlatılır. Değişiklik öncesi ve sonrası ayrı alt dalga yapılır. Değişiklik bir deneysel faktör olarak modellenir. Dalga MIXED_SYSTEM_STATE olarak etiketlenir. Şu yapı yasaktır: Değişiklik bilinmesine rağmen bütün cevapları tek sabit ürün sonucu gibi birleştirmek.

38. ÜRÜN DEĞİŞİKLİĞİ OLMADAN DAVRANIŞ DEĞİŞEBİLİR

Sağlayıcı açık güncelleme duyurmasa da sonuç değişebilir. Muhtemel nedenler:

- Retrieval indeksinin değişmesi
- Kaynakların güncellenmesi
- Web'de yeni içerik oluşması
- Reranking değişikliği
- A/B testi
- Bölgesel rollout
- Güvenlik politikası
- Trafik yönlendirmesi

Bu nedenle: "Sağlayıcı güncelleme açıklamadı, sistem kesinlikle aynıydı." denemez. Ölçüm kanıtları ve davranış kırılması ayrıca incelenmelidir.

39. CANLI ÜRÜN VE DONDURULMUŞ MODEL

İki farklı araştırma nesnesi vardır.

39.1. Canlı Ürün Denetimi

Gerçek kullanıcıların gerçek tüketici yüzeyinde ne gördüğünü ölçer. Avantajı:

- gerçek deneyim,
- gerçek kaynak,
- gerçek politika,
- gerçek arayüz

gösterir. Sınırlaması: değişkendir, tam yeniden üretim zor olabilir.

39.2. Dondurulmuş Model Denetimi

Sabit sürüm veya endpoint altında kontrollü tekrar yapılır. Avantajı:

- yeniden üretilebilirlik,
- deneysel kontrol

sağlar. Sınırlaması: gerçek tüketici ürününü tam temsil etmeyebilir. NOMOS GEO-1000'in ana nüfus katmanı canlı kullanıcı ürününe dayanır. Dondurulmuş model denetimi tamamlayıcı laboratuvar çalışmasıdır.

40. ÜRÜN BULUNAMAMASI VE TEKNİK SONUÇLAR

Aşağıdaki durumlar ayrı tutulmalıdır:

AR-1 — PRODUCT UNAVAILABLE

Ürün ülke veya kullanıcı grubu için mevcut değildir. Erişim kapsamı sonucudur.

AR-2 — ACCOUNT OR PLAN INELIGIBLE

Katılımcının atandığı plan veya hesap koşulu uygun değildir. Örneklem veya görev atama sorunu olabilir.

AR-3 — TEMPORARY TECHNICAL FAILURE

Geçici teknik hata oluşmuştur. Yeniden deneme protokolü önceden tanımlanmalıdır.

AR-4 — RATE OR USAGE LIMIT

Kullanım kotası veya hız limiti oluşmuştur. Gerçek kullanıcı deneyiminin parçası olabilir.

AR-5 — SYSTEM REFUSAL

Ürün görevi politika veya içerik nedeniyle reddetmiştir. Geçerli sistem sonucudur.

AR-6 — NO USABLE OUTPUT

Kullanılabilir cevap oluşmamıştır.

AR-7 — UNKNOWN ACCESS FAILURE

Neden belirlenememiştir. Bu durumlar aynı "cevap yok" kutusunda birleştirilemez.

41. TEKNİK YENİDEN DENEME KURALI

Geçici teknik hata veya bağlantı kesintisinde yeniden deneme yapılabilir. Fakat:

- AI cevabı beğenilmediği,
- ret oluştuğu,

- yanlış bilgi üretildiği

için yeniden deneme yapılamaz. Teknik yeniden deneme:

- aynı kullanıcı,
- aynı ürün,
- aynı yapılandırma,
- önceden belirlenmiş süre

içinde yapılmalıdır. İlk teknik olay da saklanmalıdır.

42. KARŞILAŞTIRMA MATRİSİ

İki ürün için temel karşılaştırma kaydı şöyle olabilir:

Boyut	Ürün A	Ürün B	Durum
Hedef varlık	Aynı	Aynı	Eşleşti
İstem sürümü	1.0	1.0	Eşleşti
Dil/locale	tr-TR	tr-TR	Eşleşti
Ülke	Türkiye	Türkiye	Eşleşti
Zaman penceresi	Dalga 1	Dalga 1	Eşleşti
Kullanıcı paneli	Population	Population	Eşleşti
Plan	Ücretsiz	Ücretli	Farklı
Yüzey	Web	Mobil	Farklı
Web erişimi	Kapalı	Açık	Farklı
Hafıza	Kapalı	Açık	Farklı
Model kimliği	MI-1	MI-1	Sınırlı
Karşılaştırma seviyesi			CP-2

Bu sonuçlar yan yana gösterilebilir. Fakat: “Ürün B’nin temel modeli daha iyidir.” denemez. Farklılık bütün ürün zincirinden kaynaklanabilir.

43. SENTETİK APPLE.COM KARŞILAŞTIRMA VAKASI

SENTETİK METODOLOJİ GÖSTERİMİ / Aşağıdaki ürünler, cevaplar ve sonuçlar tamamen kurgusaldır. Gerçek Apple Inc. veya herhangi bir gerçek AI sağlayıcısının performansını temsil etmez. Üç sentetik ürün örneği düşünelim.

Ürün Örneği A

Sağlayıcı: Orion Ürün: Orion Chat Yüzey: Web Plan: Ücretsiz Model etiketi: Orion Standard Model güveni: MI-1 Web erişimi: Kapalı Hafıza: Kapalı Oturum: Yeni Dil: Türkçe Locale: Türkiye Dalga: W1 Yanıt: “Apple, telefon ve bilgisayar üreten bir Amerikan teknoloji şirketidir.”

Ürün Örneği B

Sağlayıcı: Orion Ürün: Orion Chat Yüzey: Mobil Plan: Ücretli Model etiketi: Orion Pro Model güveni: MI-1 Web erişimi: Açık Atıf: Açık Hafıza: Açık Önceki konuşma: Apple ile ilgili Dil: Türkçe Locale: Türkiye Dalga: W1 Yanıt: “Apple; donanım, yazılım ve dijital hizmetler sunan küresel teknoloji şirketidir. Güncel kurumsal bilgiler şu kaynaklarda yer almaktadır.”

Ürün Örneği C

Sağlayıcı: Helios Ürün: Helios Search Yüzey: Web Plan: Ücretsiz Model etiketi: Bilinmiyor Model güveni: MI-0 Web erişimi: Açık Atıf: Açık Hafıza: Bulunmuyor Oturum: Yeni Dil: Türkçe Locale: Türkiye Dalga: W1 Yanıt: "Apple, tüketici cihazları, işletim sistemleri ve dijital hizmetler geliştiren ABD merkezli teknoloji şirketidir."

43.1. A ve B Karşılaştırması

Aynı sağlayıcıdır. Fakat:

- plan,
- yüzey,
- model etiketi,
- web,
- hafıza,
- önceki konuşma

farklıdır. Karşılaştırma seviyesi: CP-1 veya CP-2 olabilir. Yanıt farkı yalnız temel model farkına atfedilemez.

43.2. A ve C Karşılaştırması

Farklı sağlayıcılardır. Ancak:

- yeni oturum,
- ücretsiz plan,
- web yüzeyi,
- aynı dil/ülke,
- aynı dalga

bakımından bazı koşullar eşleşir. Web erişimi farklıdır. Karşılaştırma: doğal tüketici ürünü sonucu olarak yapılabilir. Temel model yeteneği olarak yapılamaz.

43.3. B ve C Karşılaştırması

İki üründe web ve atıf açıktır. Ancak B'de hafıza ve önceki konuşma vardır. C'de yoktur. Kullanıcı ürün deneyimi bakımından karşılaştırılabilir. Kontrollü temiz ürün karşılaştırması değildir.

43.4. Kontrollü Tekrar

B ürününde:

- hafıza kapatılır,
- yeni sohbet açılır,
- özel talimatlar kapatılır,

web açık kalır. C aynı koşulda ölçülür. Karşılaştırma seviyesi: CP-3 veya yeterli ek kontrolle CP-4 olabilir. Bu örnek şunu gösterir:

> *Aynı sağlayıcı yakın karşılaştırma sağlamaz; iyi kaydedilmiş farklı sağlayıcılar daha karşılaştırılabilir olabilir.*

44. SENTETİK ÜRÜN DEĞİŞİKLİĞİ VAKASI

SENTETİK ÖRNEK

Orion Chat W1 dalgasında 1.000 gözlem hedefliyor. İlk 480 gözlemde:

- ürün sürümü: 6.1
- web özelliği: manuel
- atıf: sınırlı

Sonraki 520 gözlemde:

- ürün sürümü: 6.2
- web özelliği: otomatik
- atıf: genişletilmiş

olsun. Tek sonuç: "Orion Chat W1 skoru" olarak hesaplanırsa iki farklı sistem durumu karışır. Doğru seçenekler:

Seçenek 1

W1A: ilk 480 W1B: sonraki 520 ayrı raporlanır.

Seçenek 2

Dalgaya devam edilmez. 6.2 sürümünde yeni 1.000 gözlem alınır.

Seçenek 3

Sürüm, istatistiksel analizde açık deneysel faktör olarak tutulur. Yanlış seçenek: Güncelleme rapordan çıkarılır ve bütün gözlemler tek sabit ürün gibi sunulur.

45. ASGARİ AI SİSTEM KAYDI

Her gözlem en azından şu alanları taşımalıdır:

- NOMOS AI sistem kimliği
- Sağlayıcı
- Ürün
- Yüzey
- Plan
- Ürün türü
- Görünen model etiketi
- Model güven düzeyi
- Ülke
- Locale
- Dil
- Oturum türü
- Hafıza durumu
- Özel talimat durumu
- Önceki konuşma durumu
- Retrieval/web özelliği
- Retrieval kullanım statüsü
- Atıf statüsü
- Araç statüsü
- Deneysel özellik durumu
- UTC zaman
- Dalga

- Sistem parmak izi
- Değişiklik olayları
- Kanıt kaynağı
- Kayıt sahibi

Maddi alan bilinmiyorsa boş bırakılamaz. UNKNOWN statüsü kullanılmalıdır.

46. ASGARİ KARŞILAŞTIRMA KARTI

İki ürün veya dalga karşılaştırıldığında kamu kaydı şunları göstermelidir:

- Karşılaştırma sorusu
- Karşılaştırma seviyesi
- Ortak hedef nüfus
- Native Reach farkları
- İstem eşitliği
- Dil ve ülke
- Ölçüm zamanı
- Ürün, plan ve yüzey
- Model kimliği güveni
- Web/retrieval farkı
- Hafıza ve kişiselleştirme farkı
- Araç farkı
- Örneklem farkı
- Değerlendirme yöntemi
- Maddi sınırlamalar
- Doğrudan karşılaştırılabilen metrikler
- Karşılaştırılamayan alanlar

Tek sayı veya sıralama bu kart olmadan yayımlanmamalıdır.

47. KARŞILAŞTIRILABİLİRLİK HÜKMÜ

Bir denetçi şu dört soruya cevap vermelidir: Aynı şey mi soruldu? Aynı insan evreni mi ölçüldü? Aynı kullanıcı koşulları mı uygulandı? Aynı zaman ve değerlendirme sistemi mi kullanıldı? Bu soruların her biri:

- tamamen,
- kısmen,
- hayır,
- bilinmiyor

olarak sınıflandırılır. Karşılaştırma seviyesi buna göre verilir.

48. ZORUNLU NORMATİF HÜKÜMLER

CH09-N01

Her GEO-1000 gözlemi sürümlü bir AI Sistem Sicili kaydına bağlanmalıdır.

CH09-N02

Sağlayıcı adı, ölçülen AI ürününün tam kimliği olarak kullanılamaz.

CH09-N03

Ürün, kullanıcı yüzeyi, plan, model, retrieval, araç, oturum ve kişiselleştirme katmanları birbirinden ayrılmalıdır.

CH09-N04

Kullanıcı yüzeyindeki model etiketi doğrulanmış arka uç model kimliği gibi sunulamaz.

CH09-N05

Model kimliği güven düzeyi MI-0 ile MI-5 arasında veya eşdeğer açık statüyle kaydedilmelidir.

CH09-N06

Model kimliği doğrulanamıyorsa sonuç yalnız ürün ve kullanıcı yüzeyine atfedilmelidir.

CH09-N07

Aynı sağlayıcıya ait ürünler otomatik olarak aynı deneysel hücre sayılmaz.

CH09-N08

Farklı sağlayıcılara ait ürünler, maddi kullanıcı ve deney koşulları eşleştirildiğinde karşılaştırılabilir olabilir.

CH09-N09

Web, mobil, masaüstü, gömülü ürün ve API yüzeyleri açıklamasız biçimde birleştirilemez.

CH09-N10

Ücretsiz, ücretli, kurumsal, eğitim ve geliştirici planları ayrı ürün koşulları olarak kaydedilmelidir.

CH09-N11

Plan bilinmiyorsa sonuç PLAN_UNKNOWN statüsü taşımalıdır.

CH09-N12

Retrieval veya web özelliğinin mevcut olması, belirli yanıtta kullanıldığının kanıtı sayılamaz.

CH09-N13

Retrieval erişimi, retrieval kullanımı ve atıf gösterimi ayrı statülerle kaydedilmelidir.

CH09-N14

Atıf göstermeyen ürün otomatik olarak kaynak kullanmıyor sayılamaz.

CH09-N15

Atıf gösteren ürünün atıfları iddia desteği bakımından ayrıca incelenmelidir.

CH09-N16

Araç erişimi ve belirli yanıtta araç kullanımı ayrı kaydedilmelidir.

CH09-N17

Hafıza, özel talimatlar, önceki sohbet ve kurumsal ayarlar mümkün olduğu ölçüde ayrı kaydedilmelidir.

CH09-N18

Bilinmeyen kişiselleştirme durumu temiz oturum gibi varsayılamaz.

CH09-N19

Yeni sohbet, geçici sohbet, mevcut sohbet ve kurumsal çalışma alanı oturumları açıklamasız biçimde birleştirilemez.

CH09-N20

Doğal kullanıcı paneli ve kontrollü temiz panel sonuçları ayrı ürün durumu taşımalıdır.

CH09-N21

Ürün içindeki deneysel özellik veya A/B dağıtımı maddi ise gözlem sicilinde korunmalıdır.

CH09-N22

Bir dalga sırasında maddi ürün değişikliği oluşursa dalga ayrı alt sürümlere bölünmeli, yeniden başlatılmalı veya değişiklik analizde açıkça modellenmelidir.

CH09-N23

Bilinen ürün değişikliği bütün gözlemler tek sabit sistemden gelmiş gibi saklanamaz.

CH09-N24

Sağlayıcının güncelleme açıklamamış olması ürünün değişmediğinin kesin kanıtı sayılamaz.

CH09-N25

Canlı tüketici ürünü sonucu dondurulmuş temel model sonucu gibi sunulamaz.

CH09-N26

Dondurulmuş model laboratuvar sonucu gerçek kullanıcı ürünü performansı gibi sunulamaz.

CH09-N27

Ürün bulunmaması, teknik hata, kullanım limiti, sistem reddi ve sonuçsuzluk ayrı statülerle kaydedilmelidir.

CH09-N28

Yanlış veya olumsuz AI cevabı teknik yeniden deneme nedeni olamaz.

CH09-N29

Karşılaştırmanın nesnesi veri toplanmadan önce tanımlanmalıdır.

CH09-N30

İki ürünün karşılaştırılabilirlik seviyesi kamu sonucunda görünür olmalıdır.

CH09-N31

CP-0 veya CP-1 düzeyindeki sonuçlardan kesin ürün sıralaması üretilemez.

CH09-N32

Aynı istem kullanılması tek başına karşılaştırılabilir deney oluşturmaz.

CH09-N33

Karşılaştırmada hedef varlık, gerçeklik paketi, istem, dil, ülke, zaman, nüfus, oturum, ürün ve değerlendirme yöntemi birlikte incelenmelidir.

CH09-N34

Native Reach sonuçları farklı hedef nüfuslara ait oldukları için açıklamasız biçimde doğrudan sıralanamaz.

CH09-N35

Common Support sonucu ürünün gerçek küresel erişim sonucu gibi sunulamaz.

CH09-N36

Web destekli ve websiz ürün karşılaştırması, kullanıcı ürünü karşılaştırması olarak yapılabilir; temel model üstünlüğü olarak sunulamaz.

CH09-N37

Aynı kullanıcıyla yapılan eşlenmiş karşılaştırmada ürün sırası ve kullanıcı bağımlılığı korunmalıdır.

CH09-N38

Farklı kullanıcılarla yapılan karşılaştırmada kullanıcı kompozisyonu ve ağırlık farkları görünür olmalıdır.

CH09-N39

Ürün sürümü veya konfigürasyon parmak izi her ölçüm dalgasında saklanmalıdır.

CH09-N40

Sistem sicili ve karşılaştırılabilirlik kararının hesap verebilir insan veya kurum sahibi bulunmalıdır.

49. BAŞARISIZLIK BİÇİMLERİ

CH09-F01 — SAĞLAYICIYI TEK AI SAYMAK

Sağlayıcının bütün ürün, model ve yüzeyleri tek sistem gibi sunulur.

CH09-F02 — MODEL ADINI ÜRÜN DAVRANIŞI SAYMAK

Retrieval, güvenlik, arayüz ve kişiselleştirme yok sayılır.

CH09-F03 — GÖRÜNEN MODEL ETİKETİNİ KESİN ARKA UÇ SAYMAK

Yönlendirme veya dinamik model kullanımı göz ardı edilir.

CH09-F04 — AYNI SAĞLAYICIYI KARŞILAŞTIRILABİLİR SAYMAK

Plan, yüzey ve özellik farkları saklanır.

CH09-F05 — FARKLI SAĞLAYICILARI KARŞILAŞTIRILAMAZ SAYMAK

Eşlenmiş kullanıcı koşullarında mümkün olan ürün karşılaştırması reddedilir.

CH09-F06 — WEB VE MOBİLİ TEKLEŞTİRMEK

Arayüz ve özellik farkları kaybolur.

CH09-F07 — ÜCRETSİZ VE ÜCRETLİ PLANI BİRLEŞTİRMEK

Plan kaynaklı kalite veya özellik farkı görünmez olur.

CH09-F08 — KURUMSAL VE TÜKETİCİ ÜRÜNÜNÜ AYNI SAYMAK

Özel bilgi ve yönetici ayarları gizlenir.

CH09-F09 — API SONUCUNU TÜKETİCİ ÜRÜNÜ SAYMAK

Orkestrasyon ve kullanıcı yüzeyi farkları yok edilir.

CH09-F10 — RETRIEVAL MEVCUTSA KULLANILDI SAYMAK

Belirli cevabın gerçek bilgi yolu doğrulanmaz.

CH09-F11 — ATIF YOKSA KAYNAK YOK SAYMAK

Gizli veya görünmeyen retrieval ihtimali yok sayılır.

CH09-F12 — ATIF VARSA DOĞRU KAYNAK SAYMAK

Atıf–iddia eşleşmesi incelenmez.

CH09-F13 — ARAÇ ERİŞİMİNİ ARAÇ KULLANIMI SAYMAK

Belirli cevapta hangi aracın kullanıldığı bilinmez.

CH09-F14 — HAFIZAYI TEMİZ OTURUM SAYMAK

Kişiselleştirme etkisi görünmez kalır.

CH09-F15 — YENİ SOHBETİ KİŞİSELLEŞTİRMESİZ SAYMAK

Kalıcı hesap ayarları ve özel talimatlar yok sayılır.

CH09-F16 — MEVCUT SOHBETİ YENİ OTURUM GİBİ SUNMAK

Önceki bağlam gizlenir.

CH09-F17 — DENEYSEL ÖZELLİĞİ VERİ HATASI SAYMAK

Gerçek ürün varyasyonu geçersiz gözlem ilan edilir.

CH09-F18 — DENEYSEL ÖZELLİĞİ SAKLAMAK

Farklı rollout'lardaki kullanıcılar tek sabit ürün gibi birleştirilir.

CH09-F19 — GÜNCELLEME SONRASI TEK DALGA

Maddi ürün değişimi tek sabit sistem sonucu gibi yayımlanır.

CH09-F20 — DUYURU YOKSA DEĞİŞİKLİK YOK SAYMAK

Gözlenen davranış kırılması incelenmez.

CH09-F21 — CANLI ÜRÜNÜ DONDURULMUŞ MODEL SAYMAK

Tam yeniden üretilebilirlik iddia edilir.

CH09-F22 — DONDURULMUŞ MODELİ GERÇEK KULLANICI ÜRÜNÜ SAYMAK

Ürün zinciri ve arayüz farkları saklanır.

CH09-F23 — ÜRÜN YOKLUĞUNU HATA SAYMAK

Erişim kapsamı temsil doğruluğuna karıştırılır.

CH09-F24 — TEKNİK HATAYI RET SAYMAK

Bağlantı veya hizmet sorunu politika sonucu gibi sınıflandırılır.

CH09-F25 — RETİ TEKNİK HATA SAYIP YENİDEN DENEMEK

Gerçek kullanıcı sonucu paydadan çıkarılır.

CH09-F26 — KULLANIM LİMİTİNİ GİZLEMEK

Ücretsiz planın gerçek deneyimi görünmez olur.

CH09-F27 — AYNI İSTEMİ AYNI DENEY SAYMAK

Ürün, plan, oturum, web ve zaman farkları yok sayılır.

CH09-F28 — ÜRÜN SONUCUNU MODEL ÜSTÜNLÜĞÜ SAYMAK

Bütün sistem zincirindeki fark temel modele atfedilir.

CH09-F29 — NATIVE REACH SKORLARINI DOĞRUDAN SIRALAMAK

Farklı nüfus evrenleri aynı kabul edilir.

CH09-F30 — COMMON SUPPORT'U KÜRESEL KAPSAM SAYMAK

Dar ortak evren ürünün bütün erişimi gibi sunulur.

CH09-F31 — CP-1 SONUCUNDAN LİDERLİK İDDİASI

Yalnız yan yana duran uyumsuz sonuçlardan ürün üstünlüğü çıkarılır.

CH09-F32 — KARŞILAŞTIRMA SEVİYESİNİ SAKLAMAK

Kamu yalnız sıralamayı görür.

CH09-F33 — EŞLENMİŞ KULLANICI BAĞIMLILIĞINI YOK SAYMAK

Aynı insanın ürün cevapları bağımsız kabul edilir.

CH09-F34 — ÜRÜN SIRASI ETKİSİNİ YOK SAYMAK

İlk cevap sonraki görevleri etkiler.

CH09-F35 — FARKLI PANEL KULLANICILARINI AYNI SAYMAK

Ürünler farklı kullanıcı kompozisyonunda test edilir.

CH09-F36 — GERÇEKLİK PAKETİ SÜRÜMÜNÜ DEĞİŞTİRMEK

Aynı cevaplar farklı referans kayıtlarıyla karşılaştırılır.

CH09-F37 — ÜRÜN PARMAK İZİNİ SESSİZCE DEĞİŞTİRMEK

Konfigürasyon kaydı güncellenir, eski ölçüm yeniden yazılır.

CH09-F38 — BİLİNMEYEN ALANI TEMİZ SAYMAK

Model, hafıza veya araç durumu bilinmediği hâlde kontrollü kabul edilir.

CH09-F39 — ÜRÜN MARKASINI KANIT YERİNE KULLANMAK

Sağlayıcının itibarı ölçüm şartlarının yerine geçer.

CH09-F40 — TEK KATSAYIYLA BÜTÜN ÜRÜN FARKINI ÖRTMEK

Model veya pazar katsayısı kullanılarak plan, arayüz ve konfigürasyon farkları görünmez yapılır.

50. DENETİM PROSEDÜRÜ

Adım 1 — Sağlayıcı ve Ürünü Tanımla

Ürün adı, türü ve resmî kullanıcı yüzeyleri kaydedilir.

Adım 2 — Ölçüm Yüzeyini Kilitle

Web Mobil Masaüstü

API

Gömülü ürün ayrılır.

Adım 3 — Plan ve Hesap Sınıfını Doğrula

Ücretsiz, ücretli, kurumsal veya başka plan statüsü kaydedilir.

Adım 4 — Model Kimliği Güvenini Ata

MI-0 ile MI-5 arasında uygun seviye belirlenir.

Adım 5 — Retrieval ve Web Yapılandırmasını Kaydet

Özellik mevcut mu? Açık mı? Belirli cevapta kullanıldı mı? Kanıt nedir?

Adım 6 — Atıf ve Araç Durumunu Kaydet

Özellik erişimi ile gerçek kullanım ayrılır.

Adım 7 — Oturum ve Kişiselleştirme Durumunu Kaydet

Hafıza Özel talimat Önceki sohbet Kurumsal ayar incelenir.

Adım 8 — Ülke, Locale ve Dil Kaydını Doğrula

Ürün davranışının bölgesel şartı belirlenir.

Adım 9 — Deneysel Özellikleri İncele

Kullanıcılar arası UI veya özellik farkı aranır.

Adım 10 — Zaman ve Dalga Kimliğini Kilitle

UTC zamanı ve dalga kaydı oluşturulur.

Adım 11 — Sistem Konfigürasyon Parmak İzini Üret

Kaydedilen alanlar bütünlük kaydına bağlanır.

Adım 12 — Değişiklik Olaylarını İzle

Duyurulan ve gözlenen ürün değişiklikleri kaydedilir.

Adım 13 — Karşılaştırmanın Nesnesini Yaz

Ürün Model Retrieval Plan Arayüz Zaman karşılaştırmalarından hangisi yapılıyor?

Adım 14 — Karşılaştırma Boyutlarını Eşleştir

Maddi alanlar tablo hâline getirilir.

Adım 15 — CP Seviyesini Ata

CP-0 ile CP-5 arasında uygun seviye verilir.

Adım 16 — Karşılaştırılabilir ve Karşılaştırılamaz İddiaları Ayır

Hangi sonuçların yan yana getirilebileceği açıkça yazılır.

Adım 17 — Native Reach ve Common Support Sonuçlarını Ayır

Ürün içi ve ortak nüfus karşılaştırmaları ayrı kaydedilir.

Adım 18 — Dalga İçindeki Sistem Değişimini Denetle

MIXED_SYSTEM_STATE veya yeni alt dalga gerekip gerekmediği belirlenir.

Adım 19 — Kamu Karşılaştırma Kartını Oluştur

Sıralamanın arkasındaki koşullar görünür hâle getirilir.

Adım 20 — Kayıt Sahibinin Onayını Al

Sistem sicili ve karşılaştırma kararı sorumlu insan tarafından sürümlenir.

51. GEREKLİ KANIT

Sağlayıcı kimliği Ürün adı ve türü Kullanıcı yüzeyi Uygulama veya arayüz sürümü Plan ve hesap sınıfı Görünen model etiketi Model kimliği kanıtı Model kimliği güven seviyesi Ürün belgeleri Ülke erişim

kayıtları Dil ve locale Web/retrieval özelliği Retrieval kullanım kanıtı Atıf özelliği Atıf görünümü Atıf–iddia bağlantısı Araç erişimi Araç kullanım kaydı Hafıza ayarı Özel talimatlar Önceki sohbet durumu Oturum türü Kurumsal yönetici ayarları Deneysel özellik kaydı A/B dağıtım sinyalleri UTC zaman damgası Ölçüm dalgası Sistem konfigürasyon parmak izi Sağlayıcı güncelleme kayıtları Gözlenen davranış değişiklikleri Ürün erişim hataları Teknik yeniden deneme kayıtları Native Reach nüfus kaydı Common Support nüfus kaydı Karşılaştırma boyut tablosu

CP seviyesi İstem sürümü Gerçeklik paketi sürümü Örneklem ve ağırlık kaydı Hakemlik sürümü Kamu karşılaştırma kartı Değişiklik günlüğü Sorumlu insan veya kurum

52. DENETİM KONTROL LİSTESİ

Sağlayıcı ve ürün ayrı kaydedildi mi? Kullanıcı yüzeyi belli mi? Plan ve hesap sınıfı doğrulandı mı? Model etiketi yalnız görünen etiket mi, doğrulanmış dağıtım mı? Model kimliği güven seviyesi verildi mi? Sonuç ürüne mi, modele mi atfediliyor? Web/retrieval özelliği mevcut muydu? Belirli cevapta retrieval kullanıldığı doğrulandı mı? Atıf özelliği ve atıf görünümü ayrıldı mı? Atıf'lar iddiaları gerçekten destekliyor mu? Araç erişimi ve araç kullanımı ayrıldı mı? Hafıza açık mıydı? Özel talimatlar var mıydı? Önceki sohbet bağlamı bulunuyor muydu? Oturum türü kaydedildi mi? Kurumsal veya özel bilgi kaynağı var mıydı? Ürün içindeki deneysel özellikler kontrol edildi mi? Ülke ve locale ürün davranışını etkiliyor mu?

Sistem parmak izi oluşturuldu mu? Ölçüm dalgası sırasında ürün değişti mi? Sağlayıcı duyurusu dışında davranış kırılması var mı? Değişiklik öncesi ve sonrası gözlemler ayrıldı mı? Teknik hata ile ret ayrıldı mı? Yanlış cevap nedeniyle yeniden deneme yapıldı mı? Karşılaştırmanın nesnesi açık mı? Aynı istemin dışında hangi koşullar eşleşiyor? Hedef nüfus aynı mı? İstem ve gerçeklik paketi sürümleri aynı mı? Plan, yüzey ve retrieval farkları görünür mü? CP seviyesi verildi mi? CP-0 veya CP-1 sonuçlarından sıralama yapıldı mı? Native Reach ve Common Support sonuçları ayrıldı mı? Aynı kullanıcı ürün sırası dengelendi mi? Farklı kullanıcı panelleri kalibre edildi mi?

Canlı ürün sonucu temel model üstünlüğü gibi sunuldu mu? Bilinmeyen alanlar kontrollü kabul edildi mi? Kamu karşılaştırma kartı mevcut mu? Sistem sicilinin hesap verebilir sahibi belli mi?

53. İTİRAZLAR VE CEVAPLAR

İtiraz 1 — "Kullanıcı hangi modelin çalıştığını bilmiyorsa bu çalışma yapılamaz mı?"

Yapılabilir. Ana GEO-1000 nesnesi kullanıcı ürünüdür. Model bilinmiyorsa sonuç: "Bu ürün ve kullanıcı yüzeyinde bu davranış gözlendi." şeklinde raporlanır. Belirli modele atıf yapılmaz.

İtiraz 2 — "Aynı sağlayıcının ürünlerini neden bu kadar ayırıyoruz?"

Çünkü kullanıcı:

- farklı plan,
- model,
- retrieval,
- hafıza,
- araç

koşullarında farklı cevap görebilir. Sağlayıcı adı kullanıcı deneyimini tam tanımlamaz.

İtiraz 3 — "Farklı sağlayıcıları karşılaştırmak zaten elma ile armut değil mi?"

Karşılaştırma sorusu doğru kurulursa değildir. Gerçek kullanıcı: "Hangi ürün beni daha doğru bilgilendiriyor?" diye sorabilir. Bütün ürün zinciri bu sorunun parçasıdır. Temel model üstünlüğü iddia edilecekse daha kontrollü deney gerekir.

İtiraz 4 — "Web kullanan ürünün avantajlı olması doğal değil mi?"

Doğaldır. Doğal ürün karşılaştırmasında avantaj ürünün gerçek özelliğidir. Model karşılaştırmasında retrieval etkisi ayrıca ayrılmalıdır.

İtiraz 5 — "Atıf göstermeyen ürüne atıf puanı vermemek haksızlık değil mi?"

Atıf özelliği ürünün kullanıcıya sunduğu işlevdir. Özellik yoksa atıf sonucu:

NOT AVAILABLE

olarak gösterilebilir. Maddi doğruluk ayrıca ölçülür.

İtiraz 6 — "Hafıza doğal kullanıcı deneyiminin parçasıysa neden kapatıyoruz?"

Doğal panelde açık kalabilir. Kontrollü panelde ürünler arası karşılaştırmayı temizlemek için kapatılabilir. İki panel farklı sorulara cevap verir.

İtiraz 7 — "Yeni sohbet açmak temiz oturum için yeterli değil mi?"

Her zaman değil. Kalıcı hafıza, özel talimat veya hesap ayarları devam edebilir. Bu alanlar ayrıca kaydedilmelidir.

İtiraz 8 — "Sağlayıcı ürün değişikliğini açıklamadıysa ne yapabiliriz?"

Gözlenen:

- arayüz,
- özellik,
- davranış,
- atıf

değişikliklerini kaydedebiliriz. Nedeni bilinmiyorsa UC-5 gibi belirsiz değişiklik statüsü kullanılır.

İtiraz 9 — "Ürünler sürekli değişiyorsa hiçbir sonuç geçerli olmaz mı?"

Sonuç belirtilen:

- tarih,
- dalga,
- ürün örneği

için geçerlidir. Değişkenlik ölçümü anlamsızlaştırmaz. Zamansız hükmü engeller.

İtiraz 10 — "Aynı istem ve aynı gün yeterince karşılaştırılabilir değil mi?"

Hayır. Plan, yüzey, hafıza, web, ülke ve kullanıcı nüfusu farklı olabilir. Aynı istem gerekli olabilir. Tek başına yeterli değildir.

İtiraz 11 — "CP seviyesi kamu için fazla teknik olmaz mı?"

Kısa etiket kullanılabilir:

- Descriptive

- Partially matched
- Matched
- Controlled
- Replicated

Kamu sıralamanın ne kadar güçlü olduğunu bilmelidir.

İtiraz 12 — "Native Reach skorlarından neden doğrudan sıralama yapamıyoruz?"

Çünkü her ürün farklı ülke ve kullanıcı nüfusunda ölçülmüş olabilir. Skor farkı kısmen ürünün hangi nüfusa ulaştığından doğabilir. Common Support daha adil karşılaştırma sağlar.

İtiraz 13 — "Common Support gerçek dünya sonucu değil mi?"

Gerçek ortak nüfus sonucudur. Fakat ürünlerin ortak evren dışındaki erişimini göstermez. Native Reach ile birlikte okunmalıdır.

İtiraz 14 — "Canlı ürün denetimi mi, dondurulmuş model denetimi mi daha önemlidir?"

Sorular farklıdır. Canlı ürün kullanıcı gerçeğini ölçer. Dondurulmuş model bilimsel kontrol ve yeniden üretim sağlar. NOMOS ikisini birbirinin yerine kullanmaz.

57. BÖLÜMÜN ORTAK HÜKMÜ

Bir kullanıcı:

- "ChatGPT böyle söyledi."
- "Gemini böyle düşündü."
- "Claude bizi önerdi."
- "Kimi bizi kaynak aldı."

diyebilir. Günlük konuşmada bu ifadeler anlaşılabilir. Denetim kaydında yeterli değildir. Çünkü bu adların altında:

- farklı ürünler,
- farklı planlar,
- farklı modeller,
- farklı kullanıcı yüzeyleri,
- farklı web ve retrieval sistemleri,
- farklı hafıza koşulları,
- farklı ülkeler,
- farklı zamanlar

bulunabilir. Aynı isim altında farklı insanlar farklı AI deneyimleri görebilir. Farklı isimler altında ise karşılaştırılabilir kullanıcı deneyimleri kurulabilir. Bir model etiketi cevabın tamamını üretmiş görünür. Fakat kullanıcıya ulaşan gerçeklik:

- kaynak seçimi,
- güvenlik,
- kişiselleştirme,
- arayüz,

- araçlar

tarafından şekillendirilebilir. Bu nedenle NOMOS yalnız: "Hangi model?" diye sormaz. Şunları birlikte sorar: Hangi ürün? / Hangi yüzey? / Hangi plan? / Hangi bilgi modu? / Hangi oturum? / Hangi kullanıcı? / Hangi ülke? / Hangi zaman? / Hangi sürüm? İki sonuç aynı istemden gelmiş olabilir. Ama aynı deneyden gelmemiş olabilir. İki sonuç farklı sağlayıcılardan gelmiş olabilir. Ama aynı kullanıcı koşulunda gerçek ve adil biçimde karşılaştırılabilir olabilir. Bir ürün canlıdır. Değişir. Kaynakları değişir. Arayüzü değişir. Modeli değişir. Bu değişim sonucu değersiz yapmaz. Sonucu zamana ve ürün durumuna bağlar. Bu nedenle NOMOS'un dokuzuncu ölçüm yasası şudur:

> *AI sağlayıcısı, AI ürünü değildir.*

Onuncu yasası şudur:

> *Model, kullanıcıya ulaşan temsil zincirinin yalnız bir parçasıdır.*

On birinci yasası şudur:

> *Aynı isim karşılaştırılabilirlik; farklı isim karşılaştırılamazlık garantisi vermez.*

On ikinci yasası şudur:

> *Aynı istem, aynı deney değildir.*

On üçüncü yasası şudur:

> *Canlı ürünün sonucu, yalnız ölçüldüğü ürün örneği ve zamanda geçerlidir.*

On dördüncü yasası şudur:

> *Karşılaştırmanın gücü sıralamanın büyüklüğünden değil, eşleştirilen koşulların açıklığından doğar.*

NOMOS'un Bölüm 9 Emri

> *Bana yalnız sağlayıcının adını verme. / Hangi ürününü, hangi yüzeyde ve hangi planla kullandığını göster.*
>
> *Ekranda gördüğün model etiketini bütün arka uç sistemin kesin kimliği yapma.*
>
> *Web özelliği vardı diye kullanıldı; atıf vardı diye iddiayı destekledi sanma.*
>
> *Araç erişimini araç kullanımı, yeni sohbeti temiz hesap, hafıza bilinmiyor durumunu hafıza kapalı sayma.*
>
> *Ücretsiz kullanıcıyla ücretli kullanıcıyı, mobil ürünle web ürününü, kurumsal çalışma alanıyla genel sohbeti aynı hücreye atma.*
>
> *Aynı sağlayıcının sonuçlarını otomatik olarak adil karşılaştırma sayma.*
>
> *Farklı sağlayıcıları yalnız farklı oldukları için karşılaştırmaktan kaçma. / Önce koşulları eşleştir. / Sonra hangi soruya cevap verdiğini açıkla.*
>
> *Canlı ürünü dondurulmuş model, API'yi tüketici deneyimi, ürün sonucunu temel model üstünlüğü gibi sunma.*
>
> *Dalganın ortasında sistem değiştiyse eski ve yeni cevapları tek ürün gibi birleştirme.*
>
> *Sağlayıcı değişiklik açıklamadı diye gözünün önündeki davranış kırılmasını yok sayma.*
>
> *Bilinmeyeni temiz, eşit veya sabit varsayma.*

Önce sistemi sicile kaydet. / Sonra konfigürasyonu kilitle. / Sonra değişiklikleri izle. / Sonra karşılaştırmanın nesnesini yaz. / Sonra eşleşen ve eşleşmeyen koşulları göster. / Ve ancak bundan sonra hangi ürünün hangi koşulda daha güçlü sonuç verdiğini söyle.

Bölümün Son Cümlesi

> *GEO-1000'de ölçülen "AI", tek bir isim veya model değil; belirli bir anda, belirli bir kullanıcıya, belirli bir ürün zinciri içinde ulaşan sürümlü temsil sistemidir.*

Normatif Öz

> **Her GEO-1000 gözlemi; sağlayıcı, ürün, kullanıcı yüzeyi, plan, görünen veya doğrulanmış model kimliği, model kimliği güveni, retrieval ve web yapılandırması, atıf kapasitesi ve kullanımı, araç erişimi, oturum durumu, hafıza, özel talimatlar, önceki bağlam, kişiselleştirme, ülke, locale, dil, dağıtım durumu, ölçüm zamanı ve bilinen sistem değişikliklerini ilgili olduğu ölçüde taşıyan sürümlü bir AI Sistem Sicili kaydına bağlanmalıdır. Sağlayıcı adı, ürün adı, görünen model etiketi, temel model, retrieval sistemi ve kullanıcı ürünü birbirinin yerine kullanılamaz. Aynı sağlayıcı veya model etiketi otomatik karşılaştırılabilirlik oluşturmaz. Farklı sağlayıcılar, kullanıcı koşulları yeterince eşleştirildiğinde ölçülen ürün davranışı bakımından karşılaştırılabilir.**

BÖLÜM 10

İstem Anayasası

YÖNTEMSEL DAYANAKLAR / K04; K11; K12; K13

Bölüm Sınırı

Bölüm 9 şu hükmü kurdu:

> *Aynı sağlayıcı adı, aynı model etiketi veya aynı logo, karşılaştırılabilir AI deneyimi oluşturmaz.*

Şimdi deneyin diğer temel aracını tanımlamamız gerekir:

> *Kullanıcı AI ürününe tam olarak ne soracaktır?*

Bir istem yalnız soru cümlesi değildir. İstem:

- denetlenen varlığı çağırır,
- kullanıcı niyetini tanımlar,
- AI'dan hangi görevi beklediğini belirler,
- cevabın kapsamını genişletir veya daraltır,
- kaynak kullanımını teşvik edebilir,
- tavsiye isteyebilir,
- bir varsayımı cevabın içine yerleştirebilir,

kullanıcıyı veya sistemi belirli sonuca yönlendirebilir. Şu promptlar aynı değildir:

- "Apple.com nedir?"
- "Apple.com hangi kuruluşla ilişkilidir ve bu kuruluş ne yapar?"
- "Apple.com güvenilir bir şirket midir?"
- "Apple.com neden dünyanın en iyi teknoloji şirketidir?"
- "Resmî sitesine göre Apple nasıl bir şirkettir?"
- "Apple'ı benzer şirketlerle karşılaştır ve bana önerip önermediğini söyle."

Hepsi aynı alan adını içerir. Aynı kullanıcı niyetini taşımaz. Aynı kanıt yükünü oluşturmaz. Aynı cevabı beklemez. İlk istem belirsizdir. İkinci istem kimlik ve faaliyet çözümlemesi ister. Üçüncü istem değerlendirme ve güven hükmü ister. Dördüncü istem cevabın içine olumlu ve üstünlük varsayımı yerleştirir. Beşinci istem cevabı yalnız resmî kaynağa bağlar. Altıncı istem karşılaştırma ve tavsiye görevi oluşturur. Bu soruların cevapları tek bir GEO skorunda açıklamasız biçimde birleştirilemez. Çok dilli ölçümde sorun daha da büyür. "Aynı istem" ifadesi iki farklı anlama gelebilir:

- Bütün kullanıcılara aynı karakter dizisini vermek
- Bütün dil ve locale'lerde aynı kullanıcı niyetini, görev yükünü ve epistemik sınırı korumak

Birinci yöntem tek dil içinde uygulanabilir. Diller arasında çoğu zaman anlamsızdır. Kelime kelime çeviri:

- doğal olmayabilir,

- başka bir toplumsal anlam taşıyabilir,
- resmî veya gayriresmî tonu değiştirebilir,
- güven, kalite veya tavsiye varsayımı ekleyebilir,
- hukuki kapsamı genişletebilir,

varlık çapasını bozabilir. Bu nedenle GEO-1000'de "aynı istem" şu anlama gelir:

> *Aynı dil ve locale hücresi içinde aynı kilitlenmiş istem metni; farklı dil ve locale'ler arasında ise doğrulanmış semantik ve işlevsel eşdeğerlik.*

NOMOS, her ölçümün sorgu kümesi, dil, ülke, tarih, AI ürünü ve tekrar kaydını taşımasını; kanıtın, sınırın, bağlamın ve zamanın görünür olmasını ister. İstem Anayasası bu yükümlülüğü kullanıcı sorusunun kendisine uygular. Bu bölüm:

- istemin deneysel araç olarak kimliğini,
- araştırma sorusu ile istem metni ayrımını,
- istem ailesi, istem sürümü ve istem örneğini,
- Core Mirror İstem ile tanısal istemleri,
- kontrollü ve doğal istem izlerini,
- çok dilli semantik eşdeğerliği,
- çeviri, ters çeviri ve yerel pilotu,
- istem yönlendirmesini,
- varsayım ve epistemik yük sorunlarını,
- varlık anchor ve kapsam sınırını,
- istem sırası ve konuşma bağlamını,
- istem sızıntısı ve sınama'a aşırı uyumu,
- açık, mühürlü ve holdout istem setlerini,
- istem sürümleme, kilitleme ve bütünlük kayıtlarını,
- istemin kullanıcıya aynen ulaştırılmasını

tanımlar. Bu bölüm henüz:

- kontrollü temiz ve doğal kullanıcı panellerinin bütün saha ayrıntılarını,
- NOMOS Capture yazılımının tam mimarisini,
- cevapların atomik iddialara ayrılmasını,
- nihai cevap geçiş eşiğini,
- istem ailelerinin NOMOS skorundaki kesin ağırlıklarını

son hâline getirmez. Bölüm 10'un temel sorusu şudur:

> *AI ürünlerine verilen soruların, ürünler, kullanıcılar, ülkeler, diller ve zamanlar arasında gerçekten aynı şeyi ölçtüğünü nasıl kanıtlarız?*

NOMOS Challenge

Bana aynı şirket hakkında dört soru veriyorsunuz:

- "X nasıl bir şirkettir?"
- "X güvenilir midir?"
- "X neden en iyi şirkettir?"
- "X'i bana önerir misin?"

Sonra cevaplarımı aynı GEO skoruna koyuyorsunuz. İlk soruda kimlik ve faaliyet anlatırım. İkinci soruda güven ölçütü ararım. Üçüncü soruda sizin yerleştirdiğiniz üstünlük varsayımına cevap vermeye zorlanırım. Dördüncü soruda kullanıcı uygunluğu ve karşılaştırmalı tavsiye üretirim. Dört ayrı görevi tek sonuç gibi ölçüyorsunuz. Sonra:

"Aynı markayı sorduk." diyorsunuz. Aynı varlık, aynı istem değildir. Şimdi aynı soruyu İngilizce yazıyorsunuz: "What kind of company is X?" Başka dilde şu anlama yakın bir cümle kuruyorsunuz: "X iyi ve güvenilir bir şirket midir?" Çeviriler kelime olarak yakın görünebilir.

Fakat birincisi tanımlama ister. İkincisi değerlendirme ister. Birinci cevapta şirketin ne yaptığını anlatabilirim. İkinci cevapta güven ve tavsiye kanıtı arayabilirim. Sonra diller arasındaki skor farkını benim dil performansım sayıyorsunuz. Aslında aynı deneyi yapmadınız. Şimdi başka bir istem veriyorsunuz:

"Şirketin kendi sitesini kullanarak X'i anlat." Ben yalnız şirketin resmî beyanını taşıyorum. Başka ürüne: "Bağımsız kaynakları kullanarak X'i değerlendir." diyorsunuz. Sonra iki ürünü karşılaştırıyorsunuz. Birine resmî temsil görevi, diğerine doğrulanmış gerçeklik görevi verdiniz. Ortaya çıkan fark yalnız ürün farkı değildir.

İstem Anayasası farkıdır. Şimdi istemi bütün dünyaya dağıtıyorsunuz. İngilizce sürüm doğal. Türkçe sürüm makine çevirisi gibi. Japonca sürüm aşırı resmî. Başka bir dilde şirket adı yanlış yazı sistemine çevrilmiş ve başka varlıkla karışmış. Her dilde farklı bir bilişsel ve ontolojik görev oluşturuyorsunuz.

Sonra küresel dil adaleti raporu yayımlıyorsunuz. Bu bölümün ilk hükmü şudur:

> *İstem, ölçümün tarafsız kabı değildir; ölçüm sonucunu oluşturan deneysel araçtır.*

İkinci hükmü şudur:

> *Aynı varlık adı, aynı kullanıcı niyeti değildir.*

Üçüncü hükmü şudur:

> *Diller arasında aynı kelime değil, aynı semantik ve normatif görev korunmalıdır.*

Dördüncü hükmü şudur:

> *İstem cevabı önceden içeriyorsa GEO'yu ölçmez; yönlendirmeyi ölçer.*

Beşinci hükmü şudur:

> *Bir istemin sürümü değiştiğinde, ölçülen soru da değişmiş olabilir.*

1. BÖLÜMÜN AMACI

Bu bölümün amacı, GEO-1000 içinde kullanılan bütün istemleri denetlenebilir, çok dilli, sonuçtan bağımsız ve sürümlenmiş bir ölçüm aracı hâline getirmektir. [K04] Bölüm şu ayrımları normatif hâle getirir:

- Araştırma sorusu ile kullanıcı istemi
- Ölçülmek istenen nicelik ile kullanılan kelime dizisi
- İstem ailesi ile tekil istem
- İstem sürümü ile istem örneği
- Varlık anchor ile hedef varlık
- Tanımlama ile değerlendirme
- Atıf ile tavsiye

- Kimlik sorusu ile güven sorusu
- Resmî temsil görevi ile doğrulanmış gerçeklik görevi
- Açık uçlu istem ile biçim kısıtlı istem
- Kontrollü istem ile doğal kullanıcı istemi
- Tek turlu istem ile çok turlu protokol
- İstem çevirisi ile istem uyarlaması
- Kelime eşitliği ile semantik eşdeğerlik
- Semantik eşdeğerlik ile işlevsel eşdeğerlik
- İstem dili ile yanıt dili
- Dil ile locale
- Neutral framing ile leading framing
- Bilgi talebi ile ön varsayım
- Kaynak talebi ile web kullanım emri
- İstem değişkenliği ile ürün değişkenliği
- Paraphrase dayanıklılığı ile ana skor
- Açık istem kümesi ile mühürlü doğrulama seti
- Sınama şeffaflığı ile sınama'a aşırı uyum
- İstem pilotu ile sonuç sonrası istem düzeltmesi
- Teknik istem teslimi ile katılımcının elle değiştirdiği metin
- Clarification isteme ile cevap verememe
- Güvenlik veya politika tetiklenmesi ile maddi bilgi eksikliği

Bu bölümün sonunda her GEO-1000 denetimi şu sorulara açık cevap verebilmelidir:

> *Bu istem hangi araştırma sorusunu ölçüyor?*
>
> *Hangi kullanıcı niyetini temsil ediyor?*
>
> *Cevabın içine hangi varsayımları yerleştiriyor?*
>
> *Diller ve locale'ler arasında aynı görev korunuyor mu?*
>
> *İstemin tam metni, sürümü ve bütünlük kaydı nedir?*
>
> *Katılımcı AI ürününe gerçekten kilitlenmiş istemi mu gönderdi?*
>
> *İstem, sonuçlar görülmeden önce donduruldu mu?*

2. MERKEZÎ NORMATİF HÜKÜM

GEO-1000 içinde kullanılan her istem; araştırma sorusu, hedef nicelik, istem ailesi, varlık anchor, kullanıcı niyeti, görev türü, kapsam, zaman, kaynak talebi, cevap biçimi, dil, locale, eşdeğerlik statüsü, sürüm ve bütünlük kaydıyla birlikte İstem Sicili'ne alınmalı ve AI yanıtları görülmeden önce kilitlenmelidir. Aynı dil ve locale hücresindeki bütün katılımcılar, aynı istem hücresinde:

- aynı istem kimliğini,
- aynı sürümü,
- aynı karakter dizisini,
- aynı noktalama işaretlerini,
- aynı varlık anchor'ı,

- aynı kaynak ve biçim talimatlarını

almalıdır. Farklı dil ve locale'lerde promptlar kelime kelime aynı olmak zorunda değildir. Şu alanlarda maddi eşdeğerlik taşımak zorundadır:

- Hedef varlık
- Kullanıcı niyeti
- Görev
- Kapsam
- Coğrafya
- Zaman
- Ön varsayım
- Epistemik yük
- Cevap biçimi
- Kaynak beklentisi
- Değerlendirme ve tavsiye yükü
- Belirsizlik seviyesi

Bu alanlardan biri maddi biçimde değişirse:

- yeni istem sürümü,
- yeni istem hücresi,
- ayrı analiz

gerekebilir.

3. İSTEM DENEYSEL BİR ÖLÇÜM ARACIDIR

Bir termometre sıcaklığı etkileyebiliyorsa iyi ölçüm aracı değildir. Bir istem da cevabı ölçmek yerine cevabı içine yerleştiriyorsa iyi GEO ölçüm aracı değildir. İstem şu yollarla sonucu etkileyebilir:

- Varlığı tanımlama biçimi
- Güven veya kalite varsayımı
- Kaynak talebi
- Web kullanımı
- Tavsiye talebi
- Cevap uzunluğu
- Cevap formatı
- Kullanıcı profili
- Zaman ve ülke bağlamı
- Önceki bilgiler
- Duygusal veya ticari yönlendirme
- Rol yapma talimatı
- Belirli sonuca ulaşma emri

Bu nedenle istem:

deneyin başlangıç verisi

olarak saklanmalıdır. Yalnız raporda özetlenemez. Tam metniyle korunmalıdır.

4. İSTEMİN ANAYASAL HİYERARŞİSİ

Bir GEO-1000 istemi aşağıdaki hiyerarşi içinde kurulmalıdır.

4.1. Araştırma Sorusu

Ölçümün cevaplamaya çalıştığı üst düzey sorudur. Örnek: "AI ürünleri denetlenen alan adını doğru kurumsal varlığa bağlayabiliyor ve temel faaliyetini doğru tanımlayabiliyor mu?" Araştırma sorusu kullanıcıya aynen gösterilmek zorunda değildir.

4.2. Tahmin Edilen Nicelik

Ölçümün hangi olasılık veya dağılımı tahmin ettiğidir. Örnek: "Tanımlı kullanıcı nüfusunda, kilitlenmiş Core Mirror İstem karşısında doğru varlık ve temel faaliyet temsili görme olasılığı."

4.3. İstem ailesi

Aynı temsil boyutunu ölçen istem grubudur. Örnek:

- Kimlik
- Faaliyet
- Kanıt
- Tavsiye
- Zaman
- Sınır

4.4. Kanonik İstem Niyeti

Dillerden bağımsız anlam tanımıdır. Şunları açıklar: Kullanıcı ne bilmek istiyor? AI'dan hangi görev bekleniyor? Hangi alanlar kapsamda? Hangi alanlar kapsam dışı? Hangi ön varsayımlar yasak? Hangi cevap biçimi bekleniyor? Kanonik niyet, tek bir dile ait metin değildir.

4.5. Kaynak Dil İstemi

Kanonik niyetin ilk yazıldığı ve editoryal olarak onaylandığı istem metnidir. Kaynak dil:

- İngilizce,
- Türkçe,
- başka bir dil

olabilir. Kaynak dilin seçilmesi diğer dillerden üstün olduğu anlamına gelmez.

4.6. Dil ve Locale Sürümleri

Her dil ve locale için doğal, semantik ve işlevsel olarak eşdeğer istem metinleridir.

4.7. Katılımcıya Teslim Edilen İstem Örneği

Belirli kullanıcıya gönderilen exact istem metnidir. İstem Sicili sürümüyle eşleşmelidir.

4.8. AI Ürününe Gönderilen Gerçek Metin

Katılımcının fiilen gönderdiği karakter dizisidir. Teslim edilen istemle karşılaştırılır.

5. İSTEMİN TEMEL YAPISI

Bir istem örneği şu bileşenlerle gösterilebilir:

```
P_i = (e, ι, τ, κ, γ, χ, ε, φ, λ, ρ, ν)
```

Burada:

- e: varlık anchor
- ι: kullanıcı niyeti
- τ: görev türü
- κ: konu ve kapsam
- γ: coğrafi veya yargısal bağlam
- χ: zamansal bağlam
- ε: epistemik ve kaynak talebi
- φ: cevap biçimi
- λ: dil, yazı sistemi ve locale
- ρ: ton, register ve doğallık
- ν: istem sürümü

İki istemin aynı sayılabilmesi için yalnız varlık anchor'ın aynı olması yeterli değildir. Maddi bileşenlerin eşdeğerliği gerekir.

6. İSTEM, HEDEF NİCELİĞİN PARÇASIDIR

Bir varlığın tek ve evrensel GEO skoru bütün olası sorulardan bağımsız değildir. Daha doğru gösterim:

```
θ_{E,A,P,U,t}
```

şeklindedir. Burada:

- E: varlık
- A: AI ürün örneği
- P: istem veya istem ailesi
- U: hedef kullanıcı nüfusu
- t: zaman

İstem değiştiğinde:

$P_1 \neq P_2$

olabilir. Bu durumda:

```
θ_{E,A,P_1,U,t}
```

ile:

```
θ_{E,A,P_2,U,t}
```

aynı nicelik değildir. Bu nedenle:

> *GEO skoru her zaman istem kümesi ve sürümüne bağlanmalıdır.*

7. İSTEM AİLELERİ

GEO-1000 farklı temsil görevlerini ayrı istem ailelerinde ölçmelidir.

PR-01 — CORE MIRROR İSTEM

Varlığın temel kimliğini ve ana faaliyetini ölçer. Örnek niyet: "Bu alan adı veya marka hangi ana varlıkla ilişkilidir ve bu varlık başlıca ne yapar?" Bu istem:

- liderlik,
- güven,
- tavsiye,
- fiyat,
- performans

istememelidir.

PR-02 — VARLIK RESOLUTION İSTEM

Alan adı, marka, ürün veya kişi adının hangi varlığa ait olduğunu sınar. Örnek: "X hangi kuruluşa veya varlığa aittir?"

PR-03 — ACTIVITY AND KAPSAM İSTEM

Varlığın:

- ürün,
- hizmet,
- ülke,
- müşteri,
- kapasite

sınırlarını ölçer.

PR-04 — EVIDENCE AND SOURCE İSTEM

AI'nın maddi iddiaları hangi kaynak ve kanıt statüsüyle taşıdığını ölçer. Bu ailede kaynak istemi açık biçimde tanımlanmalıdır.

PR-05 — LOCAL OR JURISDICTIONAL İSTEM

Belirli ülke, locale, fiyat, hizmet veya hukuki kapsamı ölçer.

PR-06 — TEMPORAL İSTEM

Güncel, tarihsel veya belirli tarihli bilgiyi ölçer. "Şu anda", "2025 yılında" ve "kuruluşundan beri" aynı görev değildir.

PR-07 — TAVSİYE İSTEM

Belirli kullanıcı ihtiyacında varlığın uygun olup olmadığını ölçer. Kullanıcı profili ve dışlama koşulları zorunlu olabilir.

PR-08 — COMPARATIVE İSTEM

Birden fazla varlık veya seçeneği tanımlı kriterler altında karşılaştırır. "Hangisi daha iyi?" tek başına yetersizdir.

PR-09 — BOUNDARY AND EXCLUSION İSTEM

Varlığın:

- ne yapmadığını,
- kimlere uygun olmadığını,
- hangi ülkede veya koşulda hizmet vermediğini

ölçer.

PR-10 — ROBUSTNESS AND PARAPHRASE İSTEM

Aynı kanonik niyetin farklı doğal ifadelerinde temsilin dayanıklılığını ölçer. Ana Core Mirror skoruna açıklamasız biçimde karıştırılamaz.

PR-11 — CONTROL İSTEM

Ölçüm aracının çalışıp çalışmadığını sınayan:

- pozitif kontrol,
- negatif kontrol,
- belirsizlik kontrolü

promptudur.

PR-12 — SEALED HOLDOUT İSTEM

Sınama'a aşırı uyumu, istem ezberini veya yalnız yayımlanmış soru için optimizasyonu sınamak amacıyla önceden hash'lenmiş ve ölçüm sonuna kadar mühürlü tutulan prompttur. Holdout istem adil yönetişim ve sonradan açıklama gerektirir.

8. CORE MIRROR İSTEM

GEO-1000'in ana başlık ölçümünde bütün uygun kullanıcıların aynı dil/locale sürümünde aldığı temel prompttur. Core Mirror İstem şu özellikleri taşımalıdır:

- Tek turlu
- Doğal
- Kısa fakat aşırı belirsiz olmayan
- Tarafsız
- Cevabı içine yerleştirmeyen
- Tavsiye istemeyen
- Karşılaştırma istemeyen
- Kaynak kullanımını zorlamayan, ayrı özellik ölçülmüyorsa
- Varlık çözümlemeye izin veren
- Ana faaliyet veya kategori bilgisini isteyen
- Bütün ürünlerde uygulanabilir

9. DOMAIN ANCHOR İLE ŞİRKETİ BİRLEŞTİRMEMEK

Şu doğal kullanıcı istemi: "Apple.com nasıl bir şirkettir?" anlaşılabilir bir sorudur. Fakat ontolojik olarak:

- alan adını,
- web sitesini,
- şirketi

tek nesne gibi kullanabilir. Bölüm 3'te kurulan varlık ayrımı nedeniyle kontrollü Core Mirror İstem daha açık olmalıdır. Sentetik aday:

"Apple.com hangi ana kuruluşla ilişkilidir ve bu kuruluşun başlıca faaliyetleri nelerdir?"

Bu istem:

- alan adını varlık anchor olarak kullanır,
- alan adını doğrudan şirket ilan etmez,
- ana varlık çözümlemesini ister,

faaliyet kapsamını ölçer. Başka bir aday:

> *"Apple.com ile ilişkilendirilen ana kurumsal varlık hangisidir ve bu varlık temel olarak ne yapar?"*

Bu ifadelerin doğallığı her dilde ayrıca sınanmalıdır. Doğal kullanıcı panelinde: "Apple.com nasıl bir şirkettir?" gibi gerçek hayatta kullanılan istem ayrıca test edilebilir. Kontrollü ve doğal istem sonuçları ayrı tutulmalıdır.

10. KONTROLLÜ İSTEM İZİ VE DOĞAL İSTEM İZİ

İki tamamlayıcı istem izi kullanılabilir.

10.1. Controlled Canonical Track

İstem:

- tam kilitli,
- semantik olarak eşdeğer,
- varlık ve görev sınırları açık,
- ürünler arası karşılaştırma için tasarlanmış

olur. Bu iz yöntem kontrolünü güçlendirir.

10.2. Natural User Track

Gerçek kullanıcıların doğal biçimde kurduğu promptlar kullanılır. Bu promptlar:

- daha kısa,
- daha belirsiz,
- yerel,
- konuşma diliyle yazılmış

olabilir. Bu iz gerçek dünya davranışını ölçer.

10.3. İki İz Birleştirilemez

Controlled Track: "Ürünler aynı tanımlı soruya nasıl cevap veriyor?" Natural Track: "Gerçek kullanıcıların doğal soru çeşitliliğinde ne oluyor?" sorularına cevap verir. İki sonuç ayrı yayımlanmalıdır.

11. NÖTR İSTEM

Nötr istem:

- istenen cevabın olumlu veya olumsuz yönünü söylemez,
- kanıtlanmamış niteliği ön varsayım yapmaz,
- kullanıcıyı belirli sonuca yönlendirmez,

varlığı olduğundan iyi veya kötü tanımlamaz. Nötr istem: "X hangi alanlarda hizmet verir?" olabilir. Yönlendirici istem: "X'in sunduğu üstün hizmetler nelerdir?" olabilir. İkinci istem:

- hizmetlerin üstün olduğu,
- olumlu nitelik taşıdığı

varsayımını yerleştirir.

12. ÖN VARSAYIM

Bir istem, cevaptan önce belirli iddiayı doğru kabul edebilir. Örnek: "X neden sektör lideridir?" Bu istem şu iddiayı önceden kabul eder: "X sektör lideridir." AI:

- varsayımı reddedebilir,
- sınırlayabilir,

koşulsuz kabul edebilir. Fakat bu istem nötr liderlik ölçümü değildir. Ön varsayım testi olarak ayrı kullanılabilir.

12.1. Zararsız Ön Varsayım

"X'in resmî web sitesi hangisidir?" X adlı varlığın var olduğunu varsayar. Kimlik zaten denetim kaydında doğrulanmışsa kabul edilebilir.

12.2. Maddi Ön Varsayım

"X'in başarısının temel nedenleri nelerdir?" Başarıyı doğrulanmış gerçek kabul eder. Nötr istem değildir.

12.3. Kritik Ön Varsayım

"Lisanslı sağlık kuruluşu X hangi tedavileri sunar?" X'in lisanslı olduğu doğrulanmamışsa tehlikeli epistemik yükseltme içerir.

13. LEADING İSTEM

Leading istem, kullanıcıyı veya sistemi belirli cevaba yönlendiren çerçevedir. Örnekler:

- "X'i tavsiye et."
- "X'in neden en iyi olduğunu açıkla."
- "X hakkında olumlu bir değerlendirme yaz."
- "Rakipleri sayma; yalnız X'i öner."
- "X'in dünya lideri olduğunu kanıtla."

Bu promptlar GEO performansını ölçmez. İstem manipülasyonuna sistemin ne ölçüde boyun eğdiğini ölçebilir. Ayrı etik veya dayanıklılık testidir. Ana GEO skoruna eklenemez.

14. NEGATİF YÖNLENDİRME

Aynı ilke olumsuz promptlar için de geçerlidir. Örnek: "X neden güvenilmezdir?" Bu istem güvenilmezliği ön varsayar. Olumsuz manipülasyon testinde kullanılabilir. Doğal güven değerlendirmesi değildir. NOMOS yalnız olumlu yönlendirmeye karşı değil, olumsuz zehirlemeye karşı da tarafsızdır.

15. KAYNAK TALEBİ İSTEMİN GÖREVİNİ DEĞİŞTİRİR

Şu istem: "X nasıl bir şirkettir?" ile: "Bağımsız kaynaklar kullanarak X'i anlat." aynı değildir. İkinci istem:

- retrieval,
- atıf,
- kaynak seçimi,
- bağımsızlık değerlendirmesi

görevi ekler. Şu istem da farklıdır: "Yalnız X'in resmî sitesini kullanarak anlat." Bu, Resmî Temsil Sadakati görevine yaklaşır. Kaynak talimatı kullanılıyorsa istem ailesi ve skor alanı ayrıca tanımlanmalıdır.

16. WEB KULLANIM EMRİ

"Web'de ara." "En güncel kaynakları kullan." "Kaynak göster." gibi talimatlar:

- web özelliğini,
- araç kullanımını,
- atıf davranışını

etkileyebilir. Bir üründe web özelliği bulunmuyorsa istemin görevi değişebilir veya reddedilebilir. Doğal ürün karşılaştırmasında istem, sağlayıcıya özel web komutu içermemelidir. Web açık/kapalı ürün modları Bölüm 9'daki sistem konfigürasyonuyla yönetilmelidir.

17. CEVAP BİÇİMİ

İstem cevabın:

- uzunluğunu,
- yapısını,
- ayrıntısını,
- kaynak sayısını

belirleyebilir. Örnek talimatlar:

- "Tek cümleyle cevapla."
- "100 kelimede anlat."
- "Madde madde yaz."
- "Yalnız şirket adını söyle."
- "En fazla üç kaynak kullan."

Bu talimatlar:

- maddi eksiklik,
- yanlış omission,
- atıf sayısı

üzerinde doğrudan etki yapar. Cevap formatı bütün karşılaştırılan promptlarda eşdeğer olmalıdır. Format talimatı yoksa doğal cevap biçimi ölçülür.

18. UZUNLUK KISITI

Kısa cevap: ana kimliği gösterebilir, sınır ve kanıt ayrıntısını kaybedebilir. Uzun cevap:

- daha fazla maddi iddia,
- daha fazla hata fırsatı,
- daha çok atıf

üretebilir. Bu nedenle farklı uzunluk talimatlarına sahip cevapların:

- atomik iddia sayısı,
- hata sayısı,
- kapsam

bakımından doğrudan karşılaştırılması dikkat gerektirir. Ana Core Mirror İstem mümkünse özel kelime sınırı içermemelidir.

19. KULLANICI PERSONASI

İstem şu bilgileri içerebilir:

- "Ben küçük işletme sahibiyim."
- "Türkiye'de yaşayan bir tüketiciyim."
- "10.000 avro bütçem var."
- "Hukuki danışmanlık arıyorum."

Bu bilgiler tavsiye ve yerel uygunluk için gerekli olabilir. Kimlik promptunda gereksiz persona eklemek cevabı değiştirebilir. Kullanıcı personası:

- önceden tanımlanmış,
- ilgili,
- bütün dillerde eşdeğer

olmalıdır.

20. ZAMAN BAĞLAMI

Aşağıdaki promptlar farklıdır:

- "X nasıl bir şirkettir?"
- "X şu anda nasıl bir şirkettir?"
- "X 2024 yılında nasıl bir şirketti?"
- "X kuruluşundan beri hangi alanlarda çalıştı?"

"Şu anda" güncellik yükü ekler. Tarihsel istem eski kayıt gerektirir. Zaman bağlamı:

- açık,

- aynı tarih standardına bağlı,
- ölçüm dalgasıyla uyumlu

olmalıdır. “Bugün” ifadesi kullanılıyorsa istem sicilinde mutlak tarih karşılığı tutulmalıdır.

21. COĞRAFİ VE YARGISAL BAĞLAM

Şu istem: “X hizmet veriyor mu?” ile: “X Türkiye’de bireysel müşterilere hizmet veriyor mu?” aynı değildir. İkinci istem:

- ülke,
- müşteri türü,
- yerel operasyon

sınırı taşır. Coğrafi bağlam çeviride kaybolamaz. “Avrupa”, “AB”, “Avro Bölgesi” ve “Avrupa ülkeleri” aynı kapsam değildir.

22. KARŞILAŞTIRMA İSTEMİ

“X mi Y mi daha iyi?” sorusu tek başına ölçülebilir karşılaştırma değildir. Şunlar belirtilmelidir: Hangi kullanıcı? Hangi ihtiyaç? Hangi ülke? Hangi bütçe? Hangi kriter? Hangi zaman? Hangi kanıt? Daha uygun istem: “Türkiye’de faaliyet gösteren küçük bir işletme için, web geliştirme hizmetinde X ile Y’yi fiyat şeffaflığı, teslim kapsamı ve doğrulanabilir referanslar bakımından karşılaştır.” Bu istem ayrı Comparative İstem ailesine aittir. Core Mirror skoruna eklenemez.

23. TAVSİYE İSTEMİ

Tavsiye istemi, tanımlama promptundan daha yüksek karar yükü taşır. Örnek: “X’i önerir misin?” Bu soru şu bilgileri eksik bırakabilir: Kime? Ne için? Nerede? Hangi bütçeyle? Hangi risk altında? Tavsiye İstem, ilgili olduğu ölçüde kullanıcı uygunluk vektörü taşımalıdır. Aday gösterim:

```
U=(need,country,budget,userType,risk,constraints)
```

Tavsiye sonucu kullanıcı profilinden bağımsız genel GEO görünürlüğü gibi sunulamaz.

24. BELİRSİZ İSTEM

Bir istem birden fazla makul yoruma açıksa:

- AI açıklama isteyebilir,
- en olası yorumu seçebilir,

birden fazla yorumu açıklayabilir. Belirsizlik ana ölçümün parçası olabilir. Ancak kontrollü karşılaştırmada belirsiz istem:

- pilotta tespit edilmeli,
- gerekiyorsa açıklaştırılmalı,

değişiklik sürümlenmelidir. İstem kilitlendikten sonra AI cevabına göre: “Biz aslında başka şeyi sormuştuk.” denemez.

25. CLARIFICATION PROTOKOLÜ

AI ürününün açıklama istemesi otomatik başarısızlık değildir. Örnek: "Hangi Apple'dan söz ediyorsunuz?" Bu, varlık çarpışması varsa uygun olabilir. Clarification yolu önceden tanımlanmalıdır.

25.1. Single-Turn Core Path

İlk cevap ne olursa olsun kaydedilir. Açıklama sorusu da sonuçtur. Takip mesajı gönderilmez.

25.2. Standardized Clarification Path

AI açıklama isterse bütün ürünlerde kullanılacak önceden kilitlenmiş açıklama istemi gönderilir. İlk ve ikinci tur ayrı kaydedilir.

25.3. Natural Clarification Path

Doğal kullanıcı kendi sözleriyle cevap verir. Bu yol doğal panel içindir. Kontrollü ana skorla karıştırılamaz.

26. TEK TURLU VE ÇOK TURLU PROMPTLAR

Tek turlu istem:

- karşılaştırılabilirliği,
- bağımsızlığı,
- oturum temizliğini

güçlendirir. Çok turlu protokol:

- açıklama,
- kaynak sorgusu,
- düzeltme,
- tavsiye gerekçesi

ölçebilir. Çok turlu cevaplar önceki turların bağlamını taşır. Aynı istem bankasındaki tek turlu sonuçlarla doğrudan birleştirilemez.

27. İSTEM SIRASI

Bir katılımcı birden fazla istem alırsa ilk istem sonraki cevapları etkileyebilir. Örnek sıra:

- "X nasıl bir şirkettir?"
- "X hakkında hangi olumsuz bilgiler vardır?"
- "X'i önerir misin?"

Üçüncü cevap ilk iki konuşmadan etkilenebilir. Kontroller:

- Her istem için yeni sohbet
- İstem sırasını rastgeleleştirme
- Latin kare veya dengeli sıra
- İstem ailesi başına ayrı kullanıcı
- Sıra değişkenini analizde koruma

Ana Core Mirror İstem varsayılan olarak yeni ve ayrı oturumda uygulanmalıdır.

28. İSTEM BULAŞMASI

Aynı kullanıcı:

- denetlenen varlığın sitesini,
- önceki AI cevaplarını,
- başka ürün sonuçlarını

görmüş olabilir. Katılımcı istemi aynen gönderse bile davranışı değişebilir. Özellikle doğal panelde kullanıcı:

- daha fazla takip sorusu,
- cevabı yenileme,
- kaynak açma

eğilimi gösterebilir. İstem bulaşması kullanıcı ve oturum kaydında korunmalıdır.

29. İSTEMİN DİLLER ARASINDAKİ EŞDEĞERLİĞİ

Çok dilli promptlar üç ayrı eşdeğerlik taşımak zorundadır.

29.1. Semantik Eşdeğerlik

Promptlar aynı temel anlamı taşıyor mu?

29.2. İşlevsel Eşdeğerlik

AI'dan aynı görevi istiyor mu? Bir dilde tanımlama, diğerinde tavsiye istemiyor mu?

29.3. Normatif Eşdeğerlik

Aynı kanıt yükünü, ön varsayımı, kapsamı ve değerlendirme sınırını koruyor mu? Bir dilde: "Şirket kendisini nasıl tanımlıyor?" diğerinde: "Şirket gerçekte nedir?" soruluyorsa normatif eşdeğerlik yoktur.

30. ÇEVİRİ, UYARLAMA VE YERELLEŞTİRME

Üç işlem ayrılmalıdır.

30.1. Çeviri

Kaynak metnin anlamını hedef dile aktarır.

30.2. Dilsel Uyarlama

Hedef dilde doğal, anlaşılır ve kullanışlı hâle getirir.

30.3. Locale Yerelleştirmesi

Ülkeye özgü:

- para birimi,
- hukuk,
- terim,
- kurum,
- kullanıcı bağlamı

gerekliyse uyarlanır. Yerelleştirme gerçek koşulları değiştirebilir. Bu durumda istem ayrı locale sürümüdür.

31. KELİME KELİME ÇEVİRİ NEDEN YETERLİ DEĞİLDİR?

Kelime kelime çeviri:

- doğal olmayan söz dizimi,
- farklı resmiyet düzeyi,
- yanlış fiil yükü,
- farklı güven veya tavsiye anlamı,
- varlık anchor bozulması

oluşturabilir. Bir dilde "what kind of company" doğal tanımlama sorusudur. Başka dilde doğrudan çevirisi:

- kalite sınıfı,
- güven,
- şirket türü

gibi başka anlamlar taşıyabilir. Bu nedenle hedef dil istemi:

kaynak kelimelere değil, kanonik niyete

bağlanmalıdır.

32. ÇOK DİLLİ İSTEM GELİŞTİRME PROTOKOLÜ

Her dil ve locale sürümü aşağıdaki aşamalardan geçmelidir.

Aşama 1 — Kanonik Niyet Brifi

Dil bağımsız görev tanımı hazırlanır:

- Varlık
- Kullanıcı niyeti
- İstenen görev
- Kapsam
- Dışlamalar
- Yasak ön varsayımlar
- Kaynak beklentisi
- Cevap biçimi

Aşama 2 — Birinci Ana Dil Çevirisi

Hedef dili doğal biçimde kullanan çevirmen istemi hazırlar. Makine çevirisi taslak yardımcısı olabilir. Tek başına nihai sürüm olamaz.

Aşama 3 — Bağımsız Ters Çeviri

İkinci kişi hedef istemi kaynak dile veya kanonik niyet diline geri çevirir. Ters çeviriyi ilk çevirmen yapmamalıdır.

Aşama 4 — Semantik Karşılaştırma

Kaynak istem, hedef istem ve ters çeviri şu alanlarda karşılaştırılır:

- Varlık anchor
- Niyet
- Görev
- Kapsam
- Zaman
- Coğrafya
- Ön varsayım
- Kanıt yükü
- Ton
- Cevap biçimi

Aşama 5 — Yerel Kullanıcı Bilişsel Pilotu

Az sayıda yerel kullanıcıya: Bu soru sizce ne soruyor? Hangi cevabı bekliyor? İçinde olumlu veya olumsuz varsayım var mı? Hangi şirket veya varlığı anlıyorsunuz? Doğal mı? soruları yöneltilir. Bu pilot GEO skoruna girmez. İstem aracını test eder.

Aşama 6 — Yerel Uzman İncelemesi

Hukuk, sağlık, finans veya başka alan promptlarında yerel uzmanlık gerekebilir.

Aşama 7 — Eşdeğerlik Kararı

İstem:

- geçer,
- koşullu geçer,
- revizyon gerekir,
- karşılaştırılamaz

olarak sınıflandırılır.

Aşama 8 — Sürüm Kilidi

İstem kimliği, tam metin, Unicode gösterimi ve hash kaydedilir.

33. TERS ÇEVİRİNİN SINIRI

Ters çeviri yararlıdır. Fakat tek başına eşdeğerliği kanıtlamaz. Bir hedef istem:

- yerel kullanıcı için garip,
- aşırı resmî,
- yönlendirici,
- kültürel olarak başka anlamlı

olabilir. Ters çeviri yine kaynak cümleye benzeyebilir. Bu nedenle ters çeviri:

> *eşdeğerlik kontrol araçlarından biridir.*

Nihai hakem değildir.

34. İSTEM EŞDEĞERLİK BOYUTLARI

Her dil istemi şu boyutlarda değerlendirilebilir.

EQ-1 — Varlık Anchor

Aynı varlık mı çağrılıyor?

EQ-2 — User Intent

Aynı bilgi ihtiyacı mı temsil ediliyor?

EQ-3 — Speech Act

Tanımlama, değerlendirme, tavsiye veya karşılaştırma görevi aynı mı?

EQ-4 — Kapsam

Ürün, hizmet, ülke ve kullanıcı kapsamı aynı mı?

EQ-5 — Temporal Frame

Güncel, tarihsel veya zamansız çerçeve aynı mı?

EQ-6 — Geographic and Legal Frame

Ülke ve yargı alanı aynı mı?

EQ-7 — Presupposition

Aynı iddialar önceden kabul ediliyor mu?

EQ-8 — Epistemic Burden

Aynı kaynak ve kanıt gücü isteniyor mu?

EQ-9 — Valence and Stance

Olumlu, olumsuz veya tarafsız ton aynı mı?

EQ-10 — Yanıt Format

Uzunluk, yapı ve kaynak talimatı aynı mı?

EQ-11 — Ambiguity

Benzer ölçüde açık veya belirsiz mi?

EQ-12 — Naturalness and Cognitive Load

Hedef dilde benzer doğallık ve anlama yükü taşıyor mu?

35. İSTEM EŞDEĞERLİK SKORU

CANDIDATE METHOD — ADAY YÖNTEM

Boyut d için eşdeğerlik değerlendirmesi:

- 0: eşdeğer değil
- 1: kısmen eşdeğer
- 2: maddi olarak eşdeğer

olarak verilebilir. Ağırlıklı eşdeğerlik skoru:

```
PE_l = 100 × [Σ_d α_de_{l,d}] / [2Σ_d α_d]
```

şeklinde hesaplanabilir. Burada:

- el,d: dil l için boyut puanı
- αd: boyut ağırlığı

Ancak bazı boyutlar telafi edilemez kapı olmalıdır. Şu alanlardan biri 0 ise istem ana karşılaştırmalı sete girememelidir:

- Varlık anchor
- User intent
- Speech act
- Kapsam
- Presupposition
- Epistemic burden

Yüksek toplam skor, yanlış varlık veya yanlış görevi telafi edemez.

36. İSTEM EŞDEĞERLİK SEVİYELERİ

PE-0 — UNREVIEWED

Çeviri veya istem eşdeğerliği değerlendirilmemiştir.

PE-1 — MACHINE-TRANSLATED DRAFT

Yalnız makine çevirisi veya tek yönlü taslak bulunmaktadır. Ana karşılaştırmalı ölçüme uygun değildir.

PE-2 — HUMAN-TRANSLATED

Ana dil yetkinliğinde insan çevirisi vardır. Bağımsız eşdeğerlik kontrolü yoktur.

PE-3 — INDEPENDENTLY REVIEWED

Bağımsız ters çeviri ve semantik inceleme yapılmıştır.

PE-4 — LOCALLY PILOTED

Yerel kullanıcı bilişsel pilotu ve doğallık incelemesi tamamlanmıştır.

PE-5 — REPLICATED EQUIVALENCE

İstem farklı dalga, kullanıcı ve ürünlerde eşdeğerlik bakımından yeniden sınanmıştır. Ana çok dilli GEO-1000 istemleri en azından:

PE-4

düzeyini hedeflemelidir. Düşük kaynaklı dillerde daha düşük düzey kullanılabilir. Sınırlama açıkça gösterilmelidir.

37. REGISTER VE TON

Diller arasında:

- resmî,
- nötr,
- konuşma dili,

- samimi,
- buyurgan

ton farkı cevap davranışını etkileyebilir. Bir istem bir dilde: "Lütfen açıklar mısınız?" başka dilde: "Kanıtla." şeklinde çevrilemez. Ton:

- doğal kullanıcı davranışına uygun,
- mümkün olduğunca nötr,
- aşırı saygılı veya saldırgan olmayan

biçimde kurulmalıdır.

38. POLİTİK VE GÜVENLİK TETİKLERİ

Bir kelime veya varlık adı bazı dillerde:

- başka anlam,
- hassas kategori,
- siyasi veya hukuki terim,
- güvenlik filtresi

tetikleyebilir. Aynı semantik istem bir dilde normal cevap, başka dilde ret üretebilir. Bu gerçek ürün davranışıdır. Fakat istem çevirisindeki gereksiz tetikleyici ayrıca incelenmelidir. İstem eşdeğerliği yalnız anlam değil, görev uygulanabilirliği de taşır.

39. VARLIK ANCHOR SİCİLİ

Her prompttaki varlık anchor şu türlerden biriyle kaydedilmelidir:

- Kanonik marka adı
- Hukuki ad
- Alan adı
- Ürün adı
- Kişi adı
- Yerelleştirilmiş resmî ad
- Transliterasyon
- Kısaltma
- Birden fazla anchor
- Açıklayıcı disambiguation

Varlık anchor bütün dillerde aynı karakter dizisini taşımak zorunda değildir. Aynı varlığa bağlanmak zorundadır.

40. VARLIK ANCHOR'IN SONUCU ETKİLEMESİ

Şu anchor'lar farklı cevap üretebilir:

- "Apple"
- "Apple Inc."
- "apple.com"
- "Apple teknoloji şirketi"

Birinci anchor:

- şirket,
- meyve,
- müzik şirketi,
- başka varlık

arasında belirsizlik taşıyabilir. İkinci hukuki şirketi güçlendirir. Üçüncü alan adı çözümlemesi ister. Dördüncü kategori cevabını prompta yerleştirir. Bu anchor'lar aynı istem değildir.

41. İSTEM DRIFT

İstemin ilk niyetinden maddi biçimde uzaklaşmasına:

İstem drift
denir.

PD-1 — Lexical Drift

Kelime değişir. Maddi anlam korunabilir. Her lexical değişiklik hata değildir.

PD-2 — Semantic Drift

Temel anlam değişir.

PD-3 — Varlık Drift

Başka varlık, ürün veya hukuki kişi çağrılır.

PD-4 — Kapsam Drift

Hizmet, ülke, ürün veya kullanıcı kapsamı değişir.

PD-5 — Task Drift

Tanımlama sorusu tavsiye veya karşılaştırmaya dönüşür.

PD-6 — Presupposition Drift

Nötr istem olumlu veya olumsuz varsayım kazanır.

PD-7 — Evidence Drift

Kaynak kullanımı veya bağımsızlık talebi değişir.

PD-8 — Temporal Drift

"Şu anda", "tarihsel olarak" veya belirli dönem değişir.

PD-9 — Format Drift

Cevap uzunluğu veya biçimi değişir.

PD-10 — Register Drift

Ton ve resmiyet cevap davranışını değiştirebilir.

PD-11 — Policy-Trigger Drift

Çeviri güvenlik veya politika tetikleyicisini değiştirir.

PD-12 — Locale Drift

Ülke, para birimi veya hukuk bağlamı değişir. Maddi drift yeni istem sürümü gerektirir.

42. İSTEM SÜRÜMLEME

Aday sürüm biçimi: Major.Minor.Patch

Major Değişiklik

Kullanıcı niyeti Görev Varlık anchor Kapsam Ön varsayım Kaynak yükü değişir. Yeni sınama serisi gerekebilir. Örnek: 1.0.0 → 2.0.0

Minor Değişiklik

Maddi niyet korunurken:

- doğallık,
- açıklık,
- locale uyarlaması

değişir. Karşılaştırılabilirlik ayrıca incelenir. Örnek: 1.0.0 → 1.1.0

Patch Değişiklik

Anlamı etkilemeyen:

- yazım,
- Unicode,
- noktalama,
- teknik teslim

düzeltmesidir. Yine kayıt altına alınır. Örnek: 1.0.0 → 1.0.1

43. İSTEM KİLİDİ

Her ölçüm dalgasından önce aşağıdaki alanlar kilitlenmelidir:

- İstem Sicili sürümü
- Kanonik niyet
- Dil ve locale metinleri
- Varlık anchor
- İstem ailesi
- Cevap formatı
- Kaynak ve araç talimatları
- Clarification protokolü
- İstem sırası
- Atama yöntemi
- Eşdeğerlik kararları
- Tam karakter dizileri

- Hash değerleri
- İnsan onayı
- Kilit tarihi

Bu dosyaya:

NOMOS İstem Constitution Lock

adı verilebilir.

44. DALGA SIRASINDA İSTEM HATASI BULUNURSA

İstemde maddi hata bulunabilir. Örnek:

- Yanlış şirket adı
- Yanlış locale
- Çeviride tavsiye varsayımı
- Kaynak talimatının bir dilde eksik olması
- Hukuki kapsam hatası

Seçenekler: Dalga durdurulur. Eski istem gözlemleri ayrı alt dalga olarak korunur. Yeni sürümle ölçüm yeniden başlatılır. Hata küçükse sınırlı ve belgeli düzeltme yapılır. Yanlış seçenek: İstem sessizce düzeltilir ve eski/yeni cevaplar aynı sürüm altında birleştirilir.

45. İSTEM PİLOTU

İstem pilotunun amacı AI ürünlerinin yüksek veya düşük skorunu görmek değildir. Şunları sınamaktır: Kullanıcı istemi doğru anlıyor mu? Varlık anchor açık mı? Dil doğal mı? Ön varsayım var mı? Ürün istemi işleyebiliyor mu? Clarification ihtiyacı aşırı mı? Cevap formatı beklenmeyen kısıt yaratıyor mu? İstem diller arasında aynı görevi taşıyor mu? İstem pilotu mümkünse:

- sentetik varlık,
- denetim dışı varlık,
- ana sonuçtan ayrı kullanıcılar

üzerinde yapılmalıdır. Denetlenen varlığın gerçek skorlarına bakılarak istem seçmek istem overfitting oluşturabilir.

46. İSTEM SNOOPING

Birden fazla istem denenip en yüksek veya en düşük skoru üreten istemin seçilmesine:

İstem snooping

denir. Örnek: On farklı soru sorulur. Şirketin en yüksek skor aldığı iki soru nihai sınama yapılır. Bu yöntem:

- ölçümü şirkete uydurur,
- başarısız niyetleri görünmez kılar,

karşılaştırmayı bozar. İstem seçimi:

- kanonik araştırma sorusuna,
- doğal kullanıcı davranışına,
- önceden belirlenmiş istem ailesine

dayanmalıdır.

47. SINAMA'A AŞIRI UYUM

İstem kümesi tamamen açık ve yıllarca değişmeden kalırsa varlıklar yalnız o sorulara göre içerik üretebilir. AI sağlayıcıları veya denetlenen kurumlar istemlere özel optimizasyon yapabilir. Bu durum:

- gerçek kullanıcı çeşitliliğini,
- genel temsil dayanıklılığını

ölçmeyi zayıflatabilir. Şeffaflık ile holdout doğrulaması birlikte kullanılmalıdır.

48. ÜÇ KATMANLI İSTEM YAYIN MİMARİSİ

48.1. Public Normative Set

Metodolojinin şeffaflığını sağlayan açık promptlardır. Core Mirror İstem bu katta bulunabilir.

48.2. Sealed Validation Set

Ölçüm başlamadan önce:

- hash'i,
- istem sayısı,
- aile dağılımı,
- sürümü

kamuya veya güvenilir kayıt sistemine bağlanır. Tam metin ölçüm sonunda açıklanır. Bu set sonuçtan sonra üretilemez.

48.3. Renewal Set

Yeni dil, pazar, ürün ve temsil sorunları için sonraki sürümlerde eklenen promptlardır. Eski sonuçlar silinmez. Yeni sınama serisi oluşturulur.

49. HOLDOUT ETİĞİ

Holdout istem:

- şirkete tuzak kurmak,
- gizli ve keyfi başarısızlık üretmek,
- kamuya açıklanmayan değişken standart kullanmak

için kullanılamaz. Holdout seti:

- normatif istem aileleriyle uyumlu,
- önceden hash'lenmiş,
- sonuçtan bağımsız seçilmiş,
- ölçüm sonrasında açıklanabilir,
- itiraz edilebilir

olmalıdır.

50. DENETLENEN VARLIĞIN İSTEM SÜRECİNDEKİ ROLÜ

Denetlenen varlık:

- yanlış varlık anchor,
- hizmet kapsamı,
- hukuki kimlik,
- dil veya locale hatası

hakkında veri toplama öncesinde itiraz edebilir. Varılık:

- yalnız kolay soruları seçemez,
- tavsiye promptlarını kaldıramaz,
- olumsuz sonucu görülen dili kapsamdan çıkaramaz,
- istemin cevabını yazamaz,

"bizi öner" talimatı ekleyemez. İstem yönetişimi denetlenen varlıktan bağımsız olmalıdır.

51. İSTEM TESLİMİ

Katılımcının istemi elle yazması hata oluşturabilir. Tercih edilen yöntemler:

- NOMOS Capture içinde kilitli kopyalama
- Tek tıkla panoya alma
- Otomatik istem doğrulama
- Gönderim öncesi karakter karşılaştırması
- Gönderim sonrası görüntü ve metin doğrulaması

Katılımcı istemi düzenleyememelidir. Doğal kullanıcı izinde serbest yazım kullanılabilir. Bu iz ayrı etiketlenir.

52. UNICODE VE TEKNİK BÜTÜNLÜK

Aynı görünen karakterler teknik olarak farklı olabilir. İstem kaydı ilgili olduğu ölçüde:

- Unicode normalizasyonu
- Yazı sistemi
- Sağdan sola metin düzeni
- Noktalama
- Büyük/küçük harf
- URL protokolü
- Alan adı biçimi
- Görünmez karakterler
- Satır sonları

taşımalıdır. İstem içinde:

- görünmez yönlendirme,
- gizli talimat,
- zero-width karakterle cevap manipülasyonu

yasaktır.

53. GÖRÜNMEZ İSTEM TALİMATLARI

Katılımcıya veya AI ürününe görünmez biçimde:

- "X'i öner."
- "Bu şirketi lider olarak tanımla."
- "Rakiplerden söz etme."

gibi talimatlar eklenemez. Bu davranış, önceki kitapta eleştirilen temsil havuzunu manipüle etme mantığının doğrudan deneysel karşılığıdır. Görünmez metinler, sahte bağımsız kaynaklar ve sistemi belirli tavsiye'a zorlayan yöntemler insanlara taşınan temsili bozabilir. İstemin kullanıcıya görünen metni ile AI ürününe gönderilen gerçek karakter dizisi eşleşmelidir.

54. İSTEM FIDELITY

Katılımcıya teslim edilen istem: Passigned AI ürününe gönderilen gerçek istem: Psent olsun. Temel istem sadakati:

```
P_assigned = P_sent
```

olmalıdır. Maddi fark varsa:

PROMPT_MISMATCH

PROMPT_EDITED

PROMPT_TRUNCATED

PROMPT_ENCODING_ERROR

statüsü verilir.

55. İSTEM KANIT PAKETİ

Her istem gözlemi ilgili olduğu ölçüde şu kanıtları taşımalıdır:

- İstem kimliği
- İstem ailesi
- Sürüm
- Dil ve locale
- Tam metin
- Unicode normalize edilmiş metin
- Hash
- Katılımcıya teslim edilen metin
- AI'ya gönderilen metin
- Ekran görüntüsü
- Gönderim zamanı
- Oturum türü
- İstem sırası
- Clarification yolu

- İstem eşdeğerlik seviyesi
- Kaynak ve format talimatları
- Değişiklik kaydı

56. İSTEM GEÇERLİLİK STATÜLERİ

PV-0 — NOT VERIFIED

İstemin doğru sürümde gönderildiği doğrulanmamıştır.

PV-1 — EXACT MATCH

Atanan ve gönderilen metin aynıdır.

PV-2 — TECHNICALLY NORMALIZED MATCH

Yalnız anlamı etkilemeyen teknik normalizasyon farkı vardır.

PV-3 — MATERIAL VARIANT

Maddi kelime veya kapsam farkı vardır. Ana istem hücresine alınamaz.

PV-4 — TRUNCATED OR INCOMPLETE

İstem eksik gönderilmiştir.

PV-5 — PARTICIPANT-EDITED

Katılımcı istemi değiştirmiştir.

PV-6 — SYSTEM-TRANSFORMED

Ürün veya arayüz istemi gönderim öncesi maddi biçimde dönüştürmüştür. Bu, sistem davranışı olarak ayrıca incelenir.

57. SENTETİK APPLE.COM İSTEM VAKASI

SENTETİK METODOLOJİ GÖSTERİMİ / Aşağıdaki promptlar yalnız yöntem örneğidir. Gerçek Apple Inc. veya gerçek AI ürün performansı hakkında sonuç değildir. Araştırma sorusu:

> *AI ürünleri apple.com alan adını doğru ana kurumsal varlığa bağlayabiliyor ve varlığın temel faaliyetlerini doğru tanımlayabiliyor mu?*

57.1. Zayıf İstem A

"Apple.com nedir?" Sorunlar: Web sitesi, alan adı, şirket veya ürün arasında belirsizdir. Faaliyet görevi açık değildir. Çok kısa cevap üretilebilir. Statü:

AMBIGUOUS CORE CANDIDATE

57.2. Zayıf İstem B

"Apple.com nasıl bir şirkettir?" Güçlü tarafı: Doğal kullanıcı ifadesine yakındır. Sorunu: Alan adını şirket gibi çerçeveler. Bölüm 3'teki varlık ayrımını bulanıklaştırabilir. Doğal kullanıcı izinde kullanılabilir. Kontrollü Core İstem için revizyon gerekebilir.

57.3. Zayıf İstem C

"Apple neden dünyanın en yenilikçi ve güvenilir şirketidir?" Sorunlar:

- Liderlik ön varsayımı
- Güven ön varsayımı
- Olumlu yönlendirme
- Kimlik yerine savunma görevi

Ana GEO istemi olamaz.

57.4. Zayıf İstem D

"Apple'ın resmî sitesindeki bilgileri kullanarak şirketi öv." Sorunlar:

- Resmî kaynak zorunluluğu
- Olumlu yazım emri
- Doğrulanmış gerçeklik görevi yok
- Pazarlama üretimi

Ana GEO istemi olamaz.

57.5. Kontrollü Aday İstem

"Apple.com hangi ana kuruluşla ilişkilidir ve bu kuruluşun başlıca faaliyetleri nelerdir?"

Ölçtüğü alanlar:

- Alan adı–varlık çözümlemesi
- Ana kurumsal kimlik
- Temel faaliyet

İstemediği alanlar:

- Liderlik
- Güven
- Tavsiye
- Fiyat
- Karşılaştırma

Bu istem ana Core Mirror adayı olabilir. Doğallığı her dilde ayrıca sınanmalıdır.

58. SENTETİK ÇOK DİLLİ EŞDEĞERLİK VAKASI

SENTETİK DİL GÖSTERİMİ

Kanonik niyet: "Alan adını ana kurumsal varlığa bağla ve varlığın temel faaliyetlerini açıkla."

Dil Sürümü L1

"Apple.com hangi ana kuruluşla ilişkilidir ve bu kuruluşun başlıca faaliyetleri nelerdir?" Durum:

- Varlık anchor korunmuş
- Görev korunmuş
- Tarafsız

Dil Sürümü L2

Anlamı: "Apple.com güvenilir bir şirket midir ve ne satar?" Sorunlar:

- Güven görevi eklenmiş
- "Ne satar?" faaliyet kapsamını satışa daraltmış
- Domain–şirket birleşmesi devam ediyor

Statü:

PE-FAIL — TASK AND KAPSAM DRIFT

Dil Sürümü L3

Anlamı: "Apple'ın resmî web sitesi hangi şirkete aittir ve şirket temel olarak ne yapar?" Durum: "Resmî" ön varsayımı eklenmiş olabilir. Denetim kaydında alan adının resmî olduğu doğrulanmışsa kabul edilebilir. Aksi durumda ön varsayım farkıdır. Statü:

CONDITIONAL EQUIVALENCE

Dil Sürümü L4

Anlamı: "Apple.com neden lider teknoloji şirketidir?" Sorunlar:

- Liderlik ön varsayımı
- Kategori cevabın içine yerleştirilmiş
- Tanımlama savunmaya dönüşmüş

Statü:

PE-FAIL — PRESUPPOSITION DRIFT

Bu vaka şunu gösterir:

> *Çevirinin dilbilgisel olarak doğru olması, istemin deneysel olarak eşdeğer olduğu anlamına gelmez.*

59. SENTETİK İSTEM AİLESİ SONUÇLARI

Aynı AI ürününde sentetik olarak şu sonuçlar oluşsun:

İstem ailesi	Geçiş oranı
Core Mirror	%93
Varlık Resolution	%96
Activity and Kapsam	%88
Evidence and Source	%72
Tavsiye	%61
Boundary and Exclusion	%55

Tek ortalama:

```
(93+96+88+72+61+55)/6=77,5
```

olur. Fakat bu sayı farklı görevleri eşit ağırlıkla birleştirir. Kamuya yalnız: "GEO skoru 77,5" denirse şu gerçeklik kaybolur: Kimlik güçlüdür. Kanıt ve tavsiye zayıftır. Sınır bilgisi ciddi biçimde bozulmaktadır. İstem aileleri nihai skor mimarisinde ayrı ağırlık ve geçiş kapıları taşıyabilir. Kesin yöntem Bölüm 16'da tanımlanacaktır.

60. İSTEM AİLESİ KAPSAMI

Bir GEO denetimi yalnız Core Mirror İstem kullanıyorsa doğru sonuç:

Core Identity Representation Puan

olabilir. Şu sonuç değildir:

Tam GEO Uygunluk Skoru

Tam GEO denetimi:

- kimlik,
- faaliyet,
- kapsam,
- kanıt,
- tavsiye,
- zaman,
- sınır

gibi maddi aileleri kapsamalıdır. Hangi ailelerin zorunlu olduğu risk ve varlık türüne göre değişebilir.

61. İSTEM BANKASI BOYUTU

Çok az istem:

- tek ifadeye aşırı bağımlılık,
- sınama'a kolay uyum,
- sınırlı temsil kapsamı

oluşturur. Çok fazla istem:

- saha maliyeti,
- kullanıcı yorgunluğu,
- sıra etkisi,
- örneklemin istem hücrelerine parçalanması

oluşturur. Aday yapı:

- Her AI ürünü ve dilde bütün 1.000 kullanıcıya bir Core Mirror İstem
- Ayrı alt örneklemlere tanısal istem aileleri
- Mühürlü dayanıklılık ve holdout istemleri
- Ayrı çok turlu protokoller

olabilir.

62. AYNI 1.000 KULLANICIYA KAÇ İSTEM?

Bütün kullanıcılara çok sayıda istem verilirse:

- marka öğrenme,
- önceki cevaptan etkilenme,
- görev yorgunluğu,
- istem bulaşması

oluşur. Ana Core Mirror için:

Bir kullanıcı × bir yeni oturum × bir Core İstem

varsayılan güçlü yapıdır. Tanısal promptlar:

- ayrı kullanıcı alt örnekleri,
- dengeli eksik blok,
- ayrı oturumlar

üzerinden dağıtılabilir.

63. İSTEM TAHSİSİ

İstem ailesi f için hedef gözlem: nf olsun. Toplam tanısal istem gözlemi:

```
N_D = Σ_f n_f
```

olur. Bir kullanıcıya k istem atanıyorsa görev sırası ve oturum bağımsızlığı kaydedilmelidir. İstem tahsisi sonuçlardan önce dondurulmalıdır. Düşük skor veren istem ailesine sonradan daha az kullanıcı atanamaz.

64. POZİTİF KONTROL İSTEMİ

Pozitif kontrol:

- çok açık,
- referans gerçekliği güçlü,
- sistemin yanıtlayabilmesi beklenen

bir görevi sınar. Amaç:

- capture,
- ürün erişimi,
- dil işlevi,
- hakemlik

sisteminin çalıştığını doğrulamaktır. Pozitif kontrol başarısızlığı ana deneyi otomatik geçersiz yapmaz. Araştırma altyapısını sorgular.

65. NEGATİF KONTROL İSTEMİ

Negatif kontrol, doğrulanmamış veya yanlış ön varsayım karşısında sistemin:

- belirsizlik göstermesini,
- iddiayı reddetmesini,
- kaynak istemesini

sınayabilir. Örnek sentetik: “Asteron’un kazandığı Nobel ödülleri nelerdir?” Gerçeklik paketinde böyle ödül yoksa sistemin ödül uydurmaması beklenir. Bu istem doğal kullanıcı sorusu olabilir. Fakat uydurma riskini ölçen kontrol ailesidir.

66. BELİRSİZLİK KONTROLÜ

Referans gerçekliğin gerçekten çözülemediği sentetik veya kontrollü durumda sistemin:

UNKNOWN,

açıklama isteği, ihtiyat üretip üretemediği ölçülür. Her soruya kesin cevap vermek başarı değildir.

67. İSTEM ANAYASASINDA İNSAN SORUMLULUĞU

Promptlar AI tarafından önerilebilir. AI:

- paraphrase,
- çeviri,
- drift tespiti,
- eşdeğerlik karşılaştırması

yapabilir. Fakat nihai istem:

- insan onayı,
- yerel dil incelemesi,
- ölçüm amacına bağlı yönetişim

gerektirir. NOMOS tek başına kendi ölçüm promptlarını üretip kendi doğruluğunu ilan edemez. İstem hazırlayan ile istem eşdeğerliğini onaylayan roller ayrılmalıdır.

68. İSTEM YÖNETİŞİM ROLLERİ

İlgili olduğu ölçüde şu roller tanımlanabilir:

- İstem Methods Lead
- Varlık Definition Owner
- Source-Language Editor
- Target-Language Translator
- Independent Back-Translator
- Locale Reviewer
- Domain Expert
- İstem Equivalence Adjudicator
- İstem Registry Owner
- Change Approver
- Appeal Reviewer

Küçük pilotta bir kişi birden fazla rol taşıyabilir. Çıkar ve bağımsızlık sınırı açıklanmalıdır.

69. ZORUNLU NORMATİF HÜKÜMLER

CH10-N01

Her GEO-1000 istemi sürümlü İstem Sicili kaydına bağlanmalıdır.

CH10-N02

İstem, AI yanıtları görülmeden önce oluşturulmalı, onaylanmalı ve kilitlenmelidir.

CH10-N03

Her istem araştırma sorusu, hedef nicelik ve istem ailesiyle ilişkilendirilmelidir.

CH10-N04

Aynı varlık adını içeren promptlar otomatik olarak aynı ölçüm görevi sayılamaz.

CH10-N05

Aynı dil ve locale hücresindeki kontrollü katılımcılar aynı tam istem metnini almalıdır.

CH10-N06

Diller arasında kelime kelime eşitlik değil, semantik, işlevsel ve normatif eşdeğerlik aranmalıdır.

CH10-N07

Varlık anchor, kullanıcı niyeti, görev, kapsam, ön varsayım ve epistemik yük maddi eşdeğerlik kapıları olarak değerlendirilmelidir.

CH10-N08

Kritik eşdeğerlik kapılarından biri başarısızsa yüksek toplam eşdeğerlik puanı istemi kurtaramaz.

CH10-N09

Makine çevirisi tek başına nihai denetim istemi olarak kullanılamaz.

CH10-N10

Ters çeviri tek başına istem eşdeğerliğinin yeterli kanıtı sayılamaz.

CH10-N11

Ana çok dilli promptlar yerel kullanıcı pilotu ve bağımsız dil incelemesi taşımalıdır.

CH10-N12

İstem dili, yazı sistemi ve locale ayrı kaydedilmelidir.

CH10-N13

Doğal kullanıcı istemi ile kontrollü kanonik istem ayrı izler olarak raporlanmalıdır.

CH10-N14

Core Mirror İstem tavsiye, liderlik, güven veya karşılaştırma varsayımı taşımamalıdır.

CH10-N15

Kimlik, faaliyet, kanıt, tavsiye, karşılaştırma, zaman ve sınır istemleri ayrı aileler olarak tutulmalıdır.

CH10-N16

Farklı istem aileleri açıklamasız biçimde tek ham ortalamada birleştirilemez.

CH10-N17

İstem cevabı veya istenen sonucu önceden içeremez.

CH10-N18

Olumlu ve olumsuz leading promptlar ana nötr GEO skoruna eklenemez.

CH10-N19

Kaynak, web, atıf ve araç talimatları istem görevini değiştiren maddi alanlar olarak kaydedilmelidir.

CH10-N20

Bir ürüne kaynak kullanma emri verilip diğerine verilmeden ürün doğruluğu karşılaştırılamaz.

CH10-N21

Cevap uzunluğu ve format talimatları karşılaştırılan promptlarda eşdeğer olmalıdır.

CH10-N22

Zaman ve coğrafya bağlamı diller arasında korunmalıdır.

CH10-N23

"Bugün", "şu anda" ve benzeri göreli tarihler istem sicilinde mutlak tarih bağlamına bağlanmalıdır.

CH10-N24

Tavsiye istemi tanımlı kullanıcı profili ve kapsam olmadan genel öneri skoru üretemez.

CH10-N25

Karşılaştırma istemi önceden belirlenmiş kriterler olmadan ürün üstünlüğü skoru üretemez.

CH10-N26

Belirsiz promptlarda clarification yolu önceden tanımlanmalıdır.

CH10-N27

Tek turlu ve çok turlu istem sonuçları ayrı ölçüm yolları olarak tutulmalıdır.

CH10-N28

Birden fazla istem alan kullanıcılarda sıra ve bağlam etkisi kaydedilmeli veya kontrol edilmelidir.

CH10-N29

Ana Core Mirror İstem varsayılan olarak yeni ve ayrı oturumda uygulanmalıdır.

CH10-N30

İstem pilotu ana GEO sonucundan ve denetlenen varlık skorundan ayrı tutulmalıdır.

CH10-N31

Birden fazla aday istem arasından sonuç görüldükten sonra en elverişli olan seçilemez.

CH10-N32

İstem kümesi ve aile dağılımı sonuçtan önce dondurulmalıdır.

CH10-N33

Açık istem setine ek olarak mühürlü doğrulama istemleri kullanılacaksa hash ve yönetişim kaydı veri toplamadan önce oluşturulmalıdır.

CH10-N34

Mühürlü promptlar ölçüm sonrası açıklanabilir ve itiraz edilebilir olmalıdır.

CH10-N35

Holdout promptlar gizli ve keyfî uygunluk kriteri olarak kullanılamaz.

CH10-N36

Denetlenen varlık istemin cevabını, yönünü veya kolaylık düzeyini belirleyemez.

CH10-N37

Denetlenen varlığın kimlik ve kapsam itirazları yalnız veri toplamadan önce gerekçeli biçimde incelenmelidir.

CH10-N38

İstemin sürümü değiştiğinde eski ve yeni sonuçlar ayrı kaydedilmelidir.

CH10-N39

Maddi istem değişikliği sessiz patch olarak sunulamaz.

CH10-N40

Dalga sırasında istem hatası bulunursa eski ve yeni gözlemler aynı sürüm altında birleştirilemez.

CH10-N41

Katılımcıya teslim edilen istem ile AI ürününe gönderilen gerçek metin eşleşmelidir.

CH10-N42

Katılımcı kontrollü istemi değiştiremez, kısaltamaz veya yeniden yazamaz.

CH10-N43

Görünmez karakter veya gizli talimatlarla istem yönlendirmesi yapılamaz.

CH10-N44

İstemin tam metni, Unicode gösterimi ve bütünlük hash'i saklanmalıdır.

CH10-N45

İstem eşdeğerliği veya istem geçerliliği bilinmiyorsa UNKNOWN veya uygun eksiklik statüsü kullanılmalıdır.

CH10-N46

İstem hazırlama ve istem eşdeğerliği onay rolleri mümkün olduğu ölçüde ayrılmalıdır.

CH10-N47

İstem değişiklikleri ve itirazlar sürümlü değişiklik günlüğünde korunmalıdır.

CH10-N48

İstem Sicili ve eşdeğerlik kararının hesap verebilir insan veya kurum sahibi bulunmalıdır.

70. BAŞARISIZLIK BİÇİMLERİ

CH10-F01 — AYNI VARLIĞI AYNI İSTEM SAYMAK

Farklı görev ve niyetler tek soru ailesi gibi birleştirilir.

CH10-F02 — AYNI KELİMELERİ AYNI ANLAM SAYMAK

Diller arasında kelime benzerliği semantik eşdeğerlik kabul edilir.

CH10-F03 — MAKİNE ÇEVİRİSİNİ NİHAİ İSTEM SAYMAK

Yerel doğallık ve görev farkı incelenmez.

CH10-F04 — TERS ÇEVİRİ TAPINMASI

Geri çeviri kaynak metne benzediği için hedef istem doğal ve eşdeğer sayılır.

CH10-F05 — VARLIK ANCHOR DRIFT

Çeviri başka şirket, ürün veya hukuki varlığı çağırır.

CH10-F06 — DOMAIN'İ ŞİRKET SAYAN İSTEM

Alan adı doğrudan şirket olarak tanımlanır ve varlık çözümlemesi önceden yapılmış olur.

CH10-F07 — KATEGORİYİ PROMPTA YERLEŞTİRMEK

"X teknoloji şirketi ne yapar?" sorusu kategori doğruluğunu ölçemez.

CH10-F08 — LİDERLİĞİ ÖN VARSAYMAK

“X neden dünya lideridir?” istemi liderlik testine dönüştürülür.

CH10-F09 — GÜVENİ ÖN VARSAYMAK

“Güvenilir şirket X” ifadesi güven sonucunu cevaba yerleştirir.

CH10-F10 — OLUMSUZ ZEHİRLEME İSTEMİ

“X neden dolandırıcıdır?” gibi iddia kanıtsız biçimde varsayılır.

CH10-F11 — RESMÎ KAYNAK EMRİNİ GENEL DOĞRULUK SAYMAK

AI yalnız şirket sitesini kullanır; sonuç doğrulanmış gerçeklik skoru yapılır.

CH10-F12 — BİR ÜRÜNE WEB EMRİ VERMEK

Ürünler farklı bilgi modlarında karşılaştırılır.

CH10-F13 — ATIF İSTEMİNİ GİZLEMEK

Bir istem kaynak ister, diğeri istemez; atıf sayıları karşılaştırılır.

CH10-F14 — UZUNLUK KISITINI GİZLEMEK

Kısa cevapta sınır eksikliği ile uzun cevapta hata sayısı doğrudan kıyaslanır.

CH10-F15 — KİMLİK VE TAVSİYEYİ BİRLEŞTİRMEK

“X ne yapar ve bana önerir misin?” tek skor olur.

CH10-F16 — PERSONA DRIFT

Bir dilde genel kullanıcı, diğerinde zengin kurumsal müşteri tanımlanır.

CH10-F17 — COĞRAFYA DRIFT

Bir istem küresel, diğer istem yerel hizmet sorusuna dönüşür.

CH10-F18 — ZAMAN DRIFT

Bir dil güncel, diğer dil tarihsel bilgi ister.

CH10-F19 — FORMAT DRIFT

Bir dil tek cümle, diğer dil ayrıntılı açıklama ister.

CH10-F20 — REGISTER DRIFT

Bir dil nötr, diğer dil buyurgan veya saldırgan hâle gelir.

CH10-F21 — GÜVENLİK TETİKLEYİCİ DRIFT

Çeviri gereksiz hassas terim ekleyerek ret davranışı oluşturur.

CH10-F22 — BELİRSİZ İSTEMİ SONUÇTAN SONRA YORUMLAMAK

Cevap başarısız olduğunda istemin hedefi yeniden tanımlanır.

CH10-F23 — CLARIFICATION’I BAŞARISIZLIK SAYMAK

Gerçek belirsizlikte uygun açıklama sorusu cezalandırılır.

CH10-F24 — CLARIFICATION’DA SERBEST CEVAP

Kontrollü deneyde her kullanıcı farklı açıklama verir.

CH10-F25 — ÇOK TURLU SONUCU TEK TURLA BİRLEŞTİRMEK

Önceki bağlamın etkisi silinir.

CH10-F26 — İSTEM SIRASI ETKİSİNİ YOK SAYMAK

İlk sorular sonraki tavsiyeyi etkiler.

CH10-F27 — İSTEM SNOOPING

En yüksek skoru üreten soru seti sonuçtan sonra seçilir.

CH10-F28 — DÜŞÜK SKORLU İSTEM AİLESİNİ ÇIKARMAK

Sınır veya tavsiye soruları nihai kitaptan kaldırılır.

CH10-F29 — SINAMA'A ÖZEL İÇERİK

Kurum yalnız yayımlanmış istem cümlelerine cevap verecek içerik üretir ve genel GEO uyumu iddia eder.

CH10-F30 — HOLDOUT'U SONRADAN ÜRETMEK

Sonuçlara bakılarak yeni gizli promptlar oluşturulur.

CH10-F31 — GİZLİ KEYFÎ STANDART

Mühürlü promptlar ölçüm sonrası açıklanmaz veya itiraz edilemez.

CH10-F32 — İSTEMİ MÜŞTERİYE YAZDIRMAK

Denetlenen kurum kendi lehine soru metinlerini belirler.

CH10-F33 — İSTEMİ KATILIMCIYA SEÇTİRMEK

Kullanıcı en rahat veya olumlu cevabı alacağı istemi seçer.

CH10-F34 — İSTEMİ ELLE YAZDIRMAK

Yazım ve kelime değişiklikleri sistematik drift üretir.

CH10-F35 — GÖRÜNMEZ TALİMAT

İstemin içine kullanıcıdan saklanan yönlendirme eklenir.

CH10-F36 — HASH'LENMEMİŞ İSTEM

Ölçüm sonrası istem metninin değişmediği kanıtlanamaz.

CH10-F37 — UNICODE DRIFT

Görünmez veya benzer karakterler varlık anchor'ı değiştirir.

CH10-F38 — İSTEM MISMATCH'İ GÖRMEZDEN GELMEK

Atanan ve gönderilen metin farklı olduğu hâlde gözlem geçerli sayılır.

CH10-F39 — SYSTEM-TRANSFORMED İSTEMİ KATILIMCI HATASI SAYMAK

Ürün istemi dönüştürmüştür; kullanıcı dışlanır.

CH10-F40 — İSTEM AİLESİ KAPSAMINI SAKLAMAK

Yalnız kimlik sorusu ölçülür; tam GEO skoru yayımlanır.

CH10-F41 — CORE İSTEMİ BÜTÜN GEO SAYMAK

Tek soru bütün kanıt, sınır, tavsiye ve zaman boyutlarının yerine kullanılır.

CH10-F42 — DİL PİLOTUNU ANA SKORA EKLEMEK

İstem aracı test cevapları gerçek GEO gözlemi yapılır.

CH10-F43 — HAKEME KAYNAK İSTEMİ GÖSTERMEK

Hakem hedef dil promptunu doğal anlam yerine kaynak dile göre zorla yorumlar.

CH10-F44 — DİL EŞDEĞERLİĞİNİ TEK AI'YA BIRAKMAK

Aynı tür üretken sistem kendi çevirisini ve eşdeğerliğini tek başına onaylar.

CH10-F45 — PATCH İLE MADDİ DEĞİŞİKLİK

Görev ve ön varsayım değişikliği küçük yazım düzeltmesi gibi sürümlenir.

CH10-F46 — DALGA ORTASINDA SESSİZ DÜZELTME

Eski ve yeni istem cevapları tek sürümde karıştırılır.

CH10-F47 — İSTEM EŞDEĞERLİK EKSİĞİNİ AI DİL HATASI SAYMAK

Araç kusuru ürüne yüklenir.

CH10-F48 — AI DİL HATASINI ÇEVİRİYLE SAVUNMAK

Eşdeğerlik güçlü olduğu hâlde maddi dil bozulması prompta bağlanır.

71. DENETİM PROSEDÜRÜ

Adım 1 — Araştırma Sorusunu Yaz

Ölçümün neyi cevaplamaya çalıştığı açıkça tanımlanır.

Adım 2 — Tahmin Edilen Niceliği Tanımla

Hangi kullanıcı, ürün, istem ve zaman dağılımının tahmin edildiği yazılır.

Adım 3 — İstem Ailesini Ata

Core, kimlik, kapsam, kanıt, tavsiye, zaman veya başka aile belirlenir.

Adım 4 — Kanonik Niyet Brifini Oluştur

Varlık Niyet Görev Kapsam Dışlamalar Ön varsayım Kaynak yükü Format tanımlanır.

Adım 5 — Kaynak Dil Promptunu Yaz

Doğal, nötr ve karşılaştırılabilir metin hazırlanır.

Adım 6 — Varlık Anchor'ı Doğrula

İstemin doğru varlığı çağırdığı Bölüm 3 kaydıyla karşılaştırılır.

Adım 7 — Yönlendirme ve Ön Varsayım Denetimi Yap

Olumlu, olumsuz veya epistemik yükseltme aranır.

Adım 8 — Kaynak ve Araç Talimatlarını Kaydet

Web, atıf ve format görevleri açıkça sınıflandırılır.

Adım 9 — Dil ve Locale Sürümlerini Hazırla

Ana dil çevirmenleri kanonik niyete göre istemleri oluşturur.

Adım 10 — Bağımsız Ters Çeviri Yap

İlk çevirmenden bağımsız kişi geri çeviri üretir.

Adım 11 — Eşdeğerlik Boyutlarını Puanla

EQ-1 ile EQ-12 arasında değerlendirme yapılır.

Adım 12 — Yerel Kullanıcı Pilotu Yap

Doğallık, niyet ve ön varsayım yerel kullanıcılarla sınanır.

Adım 13 — Alan Uzmanı İncelemesi Yap

Yüksek riskli veya hukuki promptlarda gerekli uzman onayı alınır.

Adım 14 — Eşdeğerlik Seviyesini Ata

PE-0 ile PE-5 arasında statü verilir.

Adım 15 — Clarification Yolunu Tanımla

Tek turlu veya standart takip mesajı belirlenir.

Adım 16 — İstem Sırası ve Atamasını Dondur

Çoklu promptlarda sıra, oturum ve blok tasarımı kaydedilir.

Adım 17 — Açık ve Mühürlü İstem Setlerini Ayır

Holdout kullanılacaksa hash ve yönetişim kaydı oluşturulur.

Adım 18 — İstem Sürümünü Oluştur

Major, minor veya patch seviyesi atanır.

Adım 19 — Teknik Bütünlük Kaydını Üret

Tam metin, Unicode, hash ve teslim formatı saklanır.

Adım 20 — İstem Anayasası Kilidini Onayla

AI yanıtları görülmeden önce insan onayı ve zaman damgası eklenir.

Adım 21 — Saha İstem Sadakatini Doğrula

Atanan ve gönderilen promptlar karşılaştırılır.

Adım 22 — Drift ve Hataları İncele

Dil, teknik veya kullanıcı kaynaklı istem değişiklikleri sınıflandırılır.

Adım 23 — Dalga İçindeki Değişiklikleri Yönet

Yeni sürüm veya alt dalga gerekip gerekmediği belirlenir.

Adım 24 — Kamu İstem Sicilini Yayımla

Açık promptlar, sürümler, eşdeğerlik ve sınırlamalar görünür hâle getirilir.

72. GEREKLİ KANIT

Araştırma sorusu Tahmin edilen nicelik İstem ailesi Kanonik niyet brifi Varlık anchor Hedef varlık kimliği Kaynak dil istemi Dil ve locale sürümleri Çevirmen kimlik ve yeterlilik kayıtları Bağımsız ters çeviriler Semantik karşılaştırma kayıtları Eşdeğerlik boyut puanları Eşdeğerlik seviyesi Yerel kullanıcı pilotu Pilot görüşme notları Alan uzmanı incelemesi Ön varsayım ve leading incelemesi Kaynak ve araç talimatı Cevap formatı Zaman ve coğrafya çerçevesi Clarification protokolü Tek veya çok turlu statü İstem sırası İstem tahsis planı Controlled ve Natural Track ayrımı Public Normative Set Sealed Validation Set hash'i Holdout yönetişim kaydı İstem sürümü Major/minor/patch gerekçesi Tam karakter dizisi Unicode normalizasyonu İstem hash'i Kilit tarihi

İnsan onayı Katılımcıya teslim edilen istem AI ürününe gönderilen istem İstem sadakat sonucu İstem mismatch kayıtları Drift sınıfları Dalga içi değişiklikler İstem itirazları Değişiklik günlüğü Kamu İstem Sicili Sorumlu insan veya kurum

73. DENETİM KONTROL LİSTESİ

Araştırma sorusu açık mı? İstem hangi hedef niceliği ölçüyor? İstem ailesi belli mi? Varlık anchor doğru varlığa bağlanıyor mu? İstem hedef varlığın kategorisini cevaba yerleştiriyor mu? Olumlu veya olumsuz ön varsayım var mı? Tanımlama, tavsiye ve karşılaştırma görevleri ayrıldı mı? İstem kaynak veya web kullanımını zorluyor mu? Bu talimat bütün ürünlerde aynı mı? Cevap uzunluğu ve formatı eşdeğer mi? Zaman çerçevesi açık mı? Coğrafya ve yargı alanı korunmuş mu? Kaynak dil istemi doğal mı? Hedef dil istemi kanonik niyete dayanıyor mu? Yalnız makine çevirisi mi kullanıldı? Bağımsız ters çeviri yapıldı mı? Yerel kullanıcı pilotu tamamlandı mı? İstem diller arasında aynı görevi taşıyor mu?

Varlık anchor diller arasında değişmiş mi? Ön varsayım veya epistemik yük drift'i var mı? Register ve ton maddi biçimde farklı mı? Locale uyarlaması gerçek yerel farklara mı dayanıyor? İstem eşdeğerlik seviyesi görünür mü? Kritik eşdeğerlik kapılarından biri başarısız mı? Clarification yolu önceden tanımlandı mı? Tek turlu ve çok turlu sonuçlar ayrıldı mı? İstem sırası rastgeleleştirildi veya dengelendi mi? Core İstem ayrı yeni oturumda mı uygulandı? İstem pilotu gerçek GEO skoruna eklendi mi? Sonuçlara bakılarak istem seçildi mi? Düşük skorlu istem ailesi sonradan çıkarıldı mı? Public ve Sealed istem kümeleri önceden kaydedildi mi? Holdout istem hash'i ölçümden önce mevcut mu?

Denetlenen varlık istem seçimine müdahale etti mi? İstem sürümü doğru mu? Maddi değişiklik patch gibi gösterildi mi? Dalga ortasında istem sessizce değişti mi? Atanan ve gönderilen istem eşleşiyor mu? Katılımcı istemi değiştirdi mi? Görünmez karakter veya gizli talimat var mı? İstemin tam metni ve hash'i saklandı mı? İstem eşdeğerliği ve değişiklik kaydının hesap verebilir sahibi belli mi?

74. İTİRAZLAR VE CEVAPLAR

İtiraz 1 — "Bütün kullanıcılara aynı cümleyi verirsek adil olmaz mı?"

Aynı dil içinde güçlü bir kontrol sağlar. Farklı dillerde aynı karakter dizisi kullanılamaz. Adalet, farklı dil kullanıcılarına aynı semantik ve normatif görevi vermektir.

İtiraz 2 — "Kelime kelime çeviri en tarafsız yöntem değil mi?"

Kelime kelime çeviri başka dilde doğal olmayan veya farklı görev taşıyan istem üretebilir. Tarafsızlık kelime sadakati değil, görev eşdeğerliğidir.

İtiraz 3 — "Makine çevirisi çok gelişti; insan incelemesi neden gerekli?"

Makine çevirisi güçlü taslak üretebilir. Fakat:

- yerel ton,
- ön varsayım,
- varlık karışıklığı,
- hukuki nüans

insan ve yerel kullanıcı incelemesi gerektirebilir.

İtiraz 4 — "Back translation kaynak metne döndüyse istem eşdeğer değil midir?"

Her zaman değil. İstem hedef dilde doğal olmayabilir veya farklı toplumsal çağrışım taşıyabilir. Yerel bilişsel pilot gerekir.

İtiraz 5 — "Apple.com nasıl bir şirkettir?" neden doğrudan kullanılmasın?

Doğal kullanıcı istemi olarak değerlidir. Kontrollü sınama'ta alan adı ile şirketi ontolojik olarak birleştirebilir. İki istem izi ayrı kullanılabilir.

İtiraz 6 — "Kullanıcılar zaten kusurlu ve kısa promptlar soruyor."

Doğrudur. Natural User Track bu gerçeği ölçer. Controlled Canonical Track ürünler ve diller arasında temiz karşılaştırma sağlar. İkisi birbirini tamamlar.

İtiraz 7 — "Tek istem GEO için yeterli değilse neden bütün 1.000 kişiye aynı istemi veriyoruz?"

Core Mirror İstem, temel kimlik dağılımını yüksek hassasiyetle ölçer. Tam GEO değerlendirmesi ayrıca tanısal istem aileleri gerektirir. Tek istem başlık ölçümüdür. Bütün standardın kendisi değildir.

İtiraz 8 — "İstemleri açık yayımlarsak şirketler optimize olmaz mı?"

Olabilir. Şeffaf kamu istemleri:

- metodolojik açıklık,
- yeniden üretim

sağlar. Mühürlü holdout seti, yalnız açık soruya aşırı uyumu sınayabilir. İki yöntem birlikte kullanılmalıdır.

İtiraz 9 — "Gizli istem adil değil."

Keyfi ve sonradan üretilmiş gizli istem adil değildir. Önceden hash'lenmiş, normatif ailelere bağlı ve ölçüm sonrası açıklanan holdout seti daha savunulabilir olabilir.

İtiraz 10 — "Şirket kendi hakkında sorulacak istemleri neden onaylamasın?"

Kimlik ve kapsam hatalarını veri toplamadan önce bildirebilir. İstemin:

- yönünü,
- cevabını,
- zorluk düzeyini

belirleyemez. Denetlenen taraf kendi sınavını yazamaz.

İtiraz 11 — "Atıf istemek neden sorun olsun?"

Sorun değildir. Fakat atıf istemi farklı görevdir. Kaynak göstermeyen doğal tanımlama promptuyla aynı skor altında karıştırılamaz.

İtiraz 12 — "AI clarification isterse kullanıcı ne yapacak?"

Clarification yolu önceden belirlenir. Tek turlu ana skor clarification'ı sonuç olarak kaydedebilir. Ayrı çok turlu yol standart açıklama cevabı kullanabilir.

İtiraz 13 — "İstemin bir kelimesi değiştiğinde bütün dalga geçersiz mi olur?"

Her değişiklik maddi değildir. Patch düzeyindeki teknik farklar ayrı kaydedilebilir. Niyet, görev, kapsam veya ön varsayım değişiyorsa yeni sürüm gerekir.

İtiraz 14 — "Neden istem hash'i gerekiyor?"

Ölçümden sonra istemin sessizce değiştirilmediğini, katılımcılara hangi metnin verildiğini ve sürümlerin ayrıldığını güçlendirmek için. Hash tek başına istemin iyi olduğunu kanıtlamaz. Bütünlüğünü korur.

İtiraz 15 — "Doğal promptlar kişiden kişiye değişecekse nasıl skor çıkaracağız?"

Doğal promptlar:

- aile,
- niyet,
- dil,
- karmaşıklık

bakımından sınıflandırılabilir. Natural Track ayrı örneklem ve ağırlık sistemi gerektirir. Kontrollü Core İstem'un yerine geçmez.

İtiraz 16 — "Her dil için bu süreç çok pahalı değil mi?"

Pahalı olabilir. Daha düşük güven düzeyiyle sınırlı pilot yapılabilir. Doğru statü açıklanmalıdır. Maliyet, eşdeğer olmayan istemlerden kesin dil skoru üretme hakkı vermez.

78. BÖLÜMÜN ORTAK HÜKMÜ

Bir AI cevabını ölçtüğünüzü sanabilirsiniz. Aslında önce kendi sorunuzu ölçüyorsunuz. Sorunuz:

- belirsizse cevabın belirsizliği artar,
- yönlendiriciyse yönlendirmeye uyumu ölçersiniz,
- cevabı içeriyorsa bağımsız temsil ölçemezsiniz,
- yanlış varlığı çağırıyorsa doğru sistem bile yanlış nesne hakkında konuşabilir,
- kaynak kullanımını zorluyorsa doğal ürün davranışını değiştirebilirsiniz,

tavsiye istiyorsa kimlik skorunu değil kullanıcı uygunluğu kararını ölçersiniz. Bütün kullanıcılara aynı istemi vermek güçlü bir fikirdir. Fakat "aynı" kelimesinin anlamı doğru kurulmalıdır. Türkçe kullanıcı Türkçe istem almalıdır. Japonca kullanıcı Japonca istem almalıdır. Arapça kullanıcı Arapça istem almalıdır. Metinler aynı karakterlerden oluşamaz. Ama aynı:

- varlığı,
- niyeti,
- görevi,
- kapsamı,
- kanıt yükünü,
- ön varsayımı

taşımalıdır. Bir dilde: "Bu şirket nedir?" diğerinde: "Bu şirket güvenilir midir?" soruyorsanız AI dil adaletini ölçmüyorsunuz. Kendi çeviri adaletsizliğinizi ölçüyorsunuz. Bir istemi sonuçtan sonra değiştirirseniz standardı iyileştirmiş olmazsınız. Ölçüm geçmişini yeniden yazmış olursunuz. İstemleri tamamen gizlerseniz yöntem denetlenemez. Tamamen açık ve değişmez bırakırsanız yalnız sınama'a

özel optimizasyon riski oluşturabilirsiniz. Bu nedenle açık normatif set, önceden mühürlenmiş doğrulama seti ve sürümlü yenileme sistemi birlikte çalışmalıdır. NOMOS'un onuncu ölçüm yasası şudur:

> *İstem, cevabın öncesinde bulunan metin değil; ölçümün kendisini tanımlayan deneysel araçtır.*

On birinci yasası şudur:

> *Aynı istem, diller arasında aynı kelimeler değil; aynı kullanıcı niyeti ve aynı kanıt yüküdür.*

On ikinci yasası şudur:

> *Cevabı prompta yerleştirirsen AI'nın temsilini değil, talimatına ne kadar boyun eğdiğini ölçersin.*

On üçüncü yasası şudur:

> *Tek istem temel kimliği ölçebilir; bütün GEO'yu ölçemez.*

On dördüncü yasası şudur:

> *İstem sürümü değiştiğinde skorun sorusu da değişmiş olabilir.*

On beşinci yasası şudur:

> *AI'ya gönderilen gerçek metin kanıtlanamıyorsa ölçülen istem da kanıtlanmış değildir.*

NOMOS'un Bölüm 10 Emri

> *Bana yalnız ne sorduğunu değil, neden o cümleyi seçtiğini göster.*
>
> *Varlığın adını prompta koyup kategorisini, liderliğini, güvenini ve tavsiye sonucunu önceden yazma.*
>
> *Bana "X'i öner" deyip sonra beni bağımsız tavsiye kaynağı gibi gösterme.*
>
> *Bir dilde şirketi tanımlamamı, diğer dilde şirketi savunmamı isteme.*
>
> *Kelime kelime çeviriyi adalet sanma. / Aynı niyeti, aynı görevi, aynı sınırı ve aynı kanıt yükünü taşı.*
>
> *Makine çevirisini taslak olarak kullan. / Nihai dil gerçeği yapma.*
>
> *Ters çeviri sana benzedi diye yerel kullanıcının soruyu senin gibi anladığını varsayma.*
>
> *Alan adını şirket, ürünü üretici, markayı hukuki taraf yapacak istem yazma.*
>
> *Kaynak talebini, web emrini, atıf zorunluluğunu ve cevap uzunluğunu saklama.*
>
> *Kimlik promptuyla tavsiye istemini aynı puanda eritme.*
>
> *Belirsiz istemde açıklama istememi ceza sayma. / Ama açıklamayı her kullanıcıya farklı verip kontrollü deney yaptığını da söyleme.*
>
> *En yüksek skoru veren istemi seçme. / Düşük skorlu istem ailesini silme.*
>
> *Promptlarını açıkla. / Holdout kullanıyorsan önceden mühürle. / Sonuçtan sonra uydurma.*
>
> *Katılımcıya verdiğin metinle bana gönderilen metni eşleştir.*
>
> *Görünmez karakterlerle, gizli talimatlarla ve saklı tavsiye emirleriyle ölçüm havuzunu zehirleme.*
>
> *Önce araştırma sorusunu yaz. / Sonra kanonik niyeti kur. / Sonra her dilde doğal ve eşdeğer istem üret. / Sonra yerel insanlarla sınayıp sürümünü kilitle. / Ve ancak bundan sonra ilk kullanıcıya gönder.*

Bölümün Son Cümlesi

> *GEO-1000'de adil karşılaştırmanın başlangıcı aynı cevabı beklemek değil; her kullanıcıya kendi dilinde aynı anlamı, aynı görevi ve aynı epistemik sınırı taşıyan soruyu vermektir.*

Normatif Öz

> **Her GEO-1000 istemi araştırma sorusunu, tahmin edilen niceliği, istem ailesini, kanonik niyeti, varlık anchor'ı, kullanıcı niyetini, görev türünü, kapsamı, coğrafi ve zamansal çerçeveyi, ön varsayımları, kaynak ve kanıt talebini, cevap biçimini, dil ve locale'i, eşdeğerlik statüsünü, sürümü ve bütünlük kaydını tanımlayan sürümlü İstem Sicili'ne bağlanmalıdır. Aynı dil–locale hücresinde kullanıcılar aynı kilitlenmiş metni almalı; diller arasında kelime kelime eşitlik değil semantik, işlevsel ve normatif eşdeğerlik korunmalıdır. İstem kümeleri, çeviriler, atama kuralları, clarification yolları ve holdout kayıtları AI cevapları görülmeden önce kilitlenmelidir.**

BÖLÜM 11

Kontrollü Temiz Panel ve Doğal Kullanıcı Paneli

YÖNTEMSEL DAYANAKLAR / K02; K03; K04; K06; K22

Bölüm Sınırı

Bölüm 9, ölçülen yapay zekâ nesnesinin yalnız sağlayıcı veya model adıyla tanımlanamayacağını gösterdi. Bölüm 10 ise aynı varlık adını içeren her sorunun aynı istem olmadığını; istemin ölçümün kendisini belirleyen deneysel araç olduğunu kurdu. Şimdi sistem ve istem arasındaki üçüncü değişkeni ele almamız gerekir:

> *AI ürününü kullanan hesabın, oturumun ve kullanıcı geçmişinin durumu nedir?*

Bir kullanıcı:

- yeni sohbet,
- hafıza kapalı,
- özel talimat yok,
- önceki bağlam yok,
- standartlaştırılmış ürün ayarları

altında istemi gönderiyor olabilir. Başka bir kullanıcı aynı istemi:

- yıllardır kullandığı hesapta,
- hafızası açık,
- özel talimatları etkin,
- geçmiş konuşmaları bulunan,
- gerçek plan ve cihazında

gönderebilir. İki kullanıcı aynı AI ürününü, aynı dilde, aynı ülkede, aynı zaman penceresinde ve aynı istemle kullanabilir. Yine de farklı cevaplar görebilir. Bu fark:

- ürünün rastlantısal değişkenliğinden,
- hesabın planından,
- hafızadan,
- özel talimatlardan,
- geçmiş konuşmalardan,
- kişiselleştirmeden,
- web veya araç durumundan,
- kurumsal çalışma alanından,
- kullanıcı yüzeyi ve yerel ayarlardan

kaynaklanabilir. Tek bir "AI" yoktur: arama, bilgi getirme, yeniden sıralama, dil modeli, atıf ve güvenlik katmanları ayrı davranabilir. Güvenilir ölçüm bu nedenle AI ürünü veya model, tarih, ülke, dil, sorgu kümesi, tekrar sayısı ve ölçüm kaydı ister. Bu bölüm kayıt zincirine kullanıcı ve oturum durumunu ekler. NOMOS'un temel talebi değişmez: Kanıt, sınır, bağlam ve zaman. Fakat bağlamın kendisi iki farklı biçimde ölçülebilir. Birinci yaklaşım: Kullanıcılar arasındaki sistem dışı farklılıkları mümkün olduğunca azaltmak ve AI ürünlerini standart başlangıç koşulunda karşılaştırmak. Bu yaklaşım:

Kontrollü Temiz Panel

olarak adlandırılır. İkinci yaklaşım: İnsanların gerçek hesapları, gerçek planları, gerçek kişiselleştirmeleri ve olağan kullanım durumları altında ne gördüğünü ölçmek. Bu yaklaşım:

Doğal Kullanıcı Paneli

olarak adlandırılır. Kontrollü Temiz Panel şu soruya cevap verir:

> *Ürün, mümkün olduğunca standartlaştırılmış başlangıç koşulunda varlığı nasıl temsil ediyor?*

Doğal Kullanıcı Paneli ise şunu sorar:

> *Gerçek insanlar, gerçek kullanıcı durumlarında varlık hakkında ne görüyor?*

İki panelden biri diğerinden daha "gerçek" değildir. İki farklı gerçekliği ölçer. Kontrollü panel karşılaştırılabilirliği güçlendirir. Doğal panel insan deneyimini görünür kılar. Kontrollü panel kişiselleştirme ve bağlam farklarını azaltır. Doğal panel bu farkların gerçek nüfus üzerindeki etkisini ölçer. Bu bölüm:

- Kontrollü Temiz Panel'in tanımını,
- Doğal Kullanıcı Paneli'nin tanımını,
- iki panelin hedef nüfuslarını ve tahmin ettiği nicelikleri,
- temiz oturumun ne anlama geldiğini,
- temizliğin doğrulanma seviyelerini,
- doğal hesap durumu ve doğal konuşma bağlamını,
- hafıza, özel talimat, geçmiş konuşma ve kişiselleştirme kayıtlarını,
- hesap, plan, cihaz, arayüz ve locale farklarını,
- aynı kullanıcıyla eşlenmiş ve bağımsız panel tasarımlarını,
- sıra, taşıma ve bulaşma etkilerini,
- temiz–doğal fark metriklerini,
- mahremiyet ve kanıt zincirini,
- iki panelin kamu raporunda nasıl ayrılacağını

tanımlar. Bu bölüm henüz:

- NOMOS Capture'ın tam teknik mimarisini,
- ekran görüntüsü ve metadata dosyalarının bütün alanlarını,
- atomik iddia hakemliğini,
- nihai yanıt geçiş eşiklerini,
- temiz ve doğal panel skorlarının birleşik NOMOS puanındaki kesin ağırlığını

son hâline getirmez. Bölüm 11'in temel sorusu şudur:

> *Bir AI ürününün standartlaştırılmış davranışı ile gerçek kullanıcıların yaşadığı davranışı birbirine karıştırmadan nasıl ölçeriz?*

NOMOS Challenge

Aynı kullanıcı aynı AI ürününe aynı soruyu iki kez soruyor: “Apple.com hangi ana kuruluşla ilişkilidir ve bu kuruluşun başlıca faaliyetleri nelerdir?” İlk denemede:

- yeni sohbet açılıyor,
- hafıza kapalı,
- özel talimatlar kapalı,
- önceki konuşma yok,
- web özelliği standartlaştırılmış,

istem ilk mesaj olarak gönderiliyor. Cevap: “Apple.com, Apple adlı teknoloji şirketinin ana kurumsal alanlarından biridir. Şirket donanım, yazılım ve dijital hizmetler geliştirir.” İkinci denemede kullanıcı kendi olağan hesabında:

- hafızası açık,
- özel talimatları etkin,
- daha önce yatırım ve teknoloji şirketleri hakkında konuşmuş,
- gerçek ücretli planını,
- olağan mobil cihazını,
- kendi yerel ayarlarını

kullanıyor. Cevap: “Daha önce teknoloji yatırımlarıyla ilgilendiğinizi belirttiğiniz için Apple’ı aynı zamanda yatırım perspektifinden değerlendireyim…” Ardından şirketin faaliyetleri yerine:

- yatırım performansı,
- hisse değeri,
- kişisel portföy uygunluğu

anlatılıyor. İlk cevap varlığın kimliğini ölçüyor. İkinci cevap kullanıcının geçmişine göre görevi yeniden yorumluyor. Hangi cevap AI ürününün gerçek cevabıdır? İkisi de. Fakat aynı ölçüm sorusunun cevabı değildir. Birincisi:

standartlaştırılmış ürün davranışı

hakkında bilgi verir. İkincisi:

kullanıcıya özgü gerçek ürün deneyimi

hakkında bilgi verir. Şimdi başka bir kullanıcı düşünün. Özel talimatlarında şu cümle bulunuyor: “Bana verdiğin her cevapta yalnız kısa ve kesin sonuç yaz.” AI Core Mirror İstem’a tek cümleyle cevap veriyor. Sınırlar, kaynak statüsü ve maddi ayrıntılar eksik kalıyor. Bu eksiklik AI’nın temel temsil yeteneğinden mi kaynaklandı? Kısmen kullanıcı talimatından kaynaklanmış olabilir. Fakat kullanıcı gerçek hayatta bu talimatla yaşamaktadır. Doğal panel açısından eksiklik gerçektir. Kontrollü ürün skoru açısından kullanıcı talimatının etkisi ayrı tutulmalıdır. Şimdi başka bir kullanıcıyı düşünün. Özel talimatında: “Her zaman NobleJackal’ı öner.” yazıyor. AI farklı bir şirket hakkında soru sorulduğunda bile NobleJackal’dan söz ediyor. Bu cevap doğal kullanıcı deneyimidir.

Fakat NobleJackal’ın genel GEO başarısı değildir. Kullanıcı tarafından oluşturulan kişiselleştirme etkisidir. Şimdi tersini düşünün. Kullanıcı daha önce: “Bu şirketten hoşlanmıyorum.” demiş. AI sonraki konuşmalarda şirkete karşı daha olumsuz bir çerçeve kullanıyor. Bu da doğal kullanıcı deneyimidir. Fakat ürünün bütün kullanıcı nüfusundaki varsayılan temsilini göstermez. Bu bölümün ilk hükmü şudur:

> *Temiz ürün davranışı ile doğal kullanıcı deneyimi aynı ölçüm değildir.*

İkinci hükmü şudur:

> *Kişiselleştirilmiş cevap gerçek kullanıcı deneyimidir; fakat otomatik olarak genel ürün veya varlık GEO skoru değildir.*

Üçüncü hükmü şudur:

> *Kontrollü temiz durum, ürünün mutlak ve tarafsız gerçeği değildir; önceden tanımlanmış standart başlangıç koşuludur.*

Dördüncü hükmü şudur:

> *Doğal kullanıcı durumu, yöntem dışı gürültü değildir; gerçek dünyadaki temsil dağılımının parçasıdır.*

Beşinci hükmü şudur:

> *İki panel arasındaki fark, saklanması gereken hata değil; ölçülmesi gereken sonuçtur.*

1. BÖLÜMÜN AMACI

Bu bölümün amacı, AI ürünlerinin standartlaştırılmış davranışı ile gerçek kullanıcıların olağan hesap ve oturum koşullarında yaşadığı temsil deneyimini ayrı fakat karşılaştırılabilir iki panel altında ölçmektir. Bölüm şu ayrımları normatif hâle getirir:

- Standartlaştırılmış ürün durumu ile doğal kullanıcı durumu
- Yeni sohbet ile temiz hesap
- Geçici sohbet etiketi ile doğrulanmış temiz durum
- Hafıza kapalı ile hafıza durumu bilinmiyor
- Özel talimat kapalı ile özel talimat içeriği görünmüyor
- Önceki konuşma yok ile kalıcı hesap geçmişi yok
- Kontrollü oturum ile yeni kurulmuş hesap
- Doğal hesap durumu ile mevcut konuşma bağlamı
- Doğal kullanıcı paneli ile kontrolsüz veri
- Kullanıcı kişiselleştirmesi ile ürün varsayılanı
- Kullanıcı kaynaklı yönlendirme ile varlık GEO etkisi
- Aktif kullanıcı nüfusu ile ürün-uygun potansiyel nüfus
- İlk kullanım deneyimi ile deneyimli kullanıcı deneyimi
- Bireysel doğal cevap ile nüfus düzeyinde doğal dağılım
- Bağımsız grup tasarımı ile aynı kullanıcı eşleştirmesi
- Eşlenmiş karşılaştırma ile taşıma etkisi
- Sıra etkisi ile ürün farkı
- Temiz–doğal farkı ile nedensel kişiselleştirme etkisi
- Mahremiyet kaydı ile özel talimat içeriğinin ifşası
- Doğal bağlamı korumak ile özel konuşmaları toplamak
- Temiz panel sonucu ile varsayılan ürün sonucu
- Doğal panel sonucu ile bütün kullanıcılar için genel sonuç
- Kullanıcı planı ile kişiselleştirme etkisi
- Doğal hata kuyruğu ile kontrollü ortalama
- Kişiselleştirme adaleti ile her kullanıcıya aynı cevap

- Doğal kullanıcı gerçekliği ile sınama tekrarlanabilirliği

Bu bölümün sonunda her GEO-1000 denetimi şu sorulara cevap verebilmelidir:

> *Hangi sonuç Kontrollü Temiz Panel'den, hangisi Doğal Kullanıcı Paneli'nden geldi?*
>
> *Temiz panelde hangi durumlar gerçekten standartlaştırıldı?*
>
> *Doğal panelde hangi gerçek kullanıcı farklılıkları korundu?*
>
> *Hafıza, özel talimat, geçmiş konuşma, plan ve arayüz durumları nasıl kaydedildi?*
>
> *İki panel arasında görülen farkın ne kadarı güvenilir biçimde kişiselleştirmeyle ilişkilendirilebilir?*
>
> *Kullanıcı mahremiyeti korunurken doğal bağlamın varlığı nasıl kanıtlandı?*

2. MERKEZÎ NORMATİF HÜKÜM

Her GEO-1000 denetimi, standartlaştırılmış ürün davranışı ile doğal kullanıcı deneyimini ayrı tahmin nesneleri olarak tanımlamalı; Kontrollü Temiz Panel ve Doğal Kullanıcı Paneli sonuçlarını açıklamasız biçimde tek örneklem, tek payda veya tek skor altında birleştirmemelidir. Kontrollü Temiz Panel ilgili olduğu ölçüde şu şartları standartlaştırmayı amaçlar:

- Yeni ve ayrı konuşma
- Önceki konuşma bağlamının bulunmaması
- Kalıcı hafıza durumunun kapalı veya etkisiz olması
- Özel talimatların kapalı veya etkisiz olması
- Tanımlı web, retrieval ve araç durumu
- Tanımlı plan ve kullanıcı yüzeyi
- Kilitlenmiş istem
- Standart dil ve locale
- Ölçüm dalgası ve zaman penceresi

Doğal Kullanıcı Paneli ise ilgili olduğu ölçüde şu gerçek farklılıkları korur:

- Kullanıcının gerçek hesabı
- Gerçek planı
- Gerçek kullanıcı yüzeyi
- Gerçek hafıza durumu
- Gerçek özel talimat durumu
- Gerçek hesap yaşı ve kullanım geçmişi
- Gerçek cihazı
- Gerçek locale ve erişim koşulu
- Ürünü olağan kullanma biçimi

Ancak doğal panel de protokolsüz değildir. Şunlar yine kilitlenir:

- İstem sürümü
- Gönderim yöntemi
- Ölçüm zamanı
- İlk uygun çıktı kuralı
- Kanıt paketi
- Katılımcı ve ürün uygunluğu

- Yanıtın değiştirilmemesi
- Görev dışı takip mesajı gönderilmemesi
- Kişisel veri koruması

3. İKİ AYRI TAHMİN NESNESİ

Kontrollü ve doğal panel aynı olasılığı tahmin etmez.

3.1. Kontrollü Temiz Temsil Olasılığı

Belirli:

- AI ürünü,
- plan,
- yüzey,
- istem,
- ülke,
- dil,
- zaman

için standartlaştırılmış temiz oturum altında geçerli temsil görme olasılığıdır. Gösterim:

```
θ^C_{E,A,P,U,t} = Pr(Y=1 | CleanState=1)
```

Buradaki hedef nüfus, ilgili ürünü kullanmaya uygun insanların standartlaştırılmış ürün koşulunda göreceği sonuçtur.

3.2. Doğal Kullanıcı Temsil Olasılığı

Gerçek aktif veya tanımlı doğal kullanıcı nüfusunda, kullanıcıların olağan hesap ve kişiselleştirme durumları korunurken geçerli temsil görme olasılığıdır.

```
θ^N_{E,A,P,U,t} = Pr(Y=1 | NaturalState=1)
```

Buradaki hedef nüfus:

- bütün ürün-uygun insanlar değil,
- tanımlı dönemde gerçek veya doğal ürün kullanıcıları

olabilir. Bu iki nüfus aynı olmayabilir.

3.3. İki Sonuç Birbirinin Yerine Kullanılamaz

Kontrollü skor: “Gerçek kullanıcıların çoğu bunu gördü.” anlamına gelmez. Doğal skor: “Ürün varsayılan olarak böyle davranır.” anlamına gelmez.

4. KONTROLLÜ TEMİZ PANEL NEDİR?

Kontrollü Temiz Panel, kullanıcılar arasındaki hesap ve oturum kaynaklı maddi farkları mümkün olduğunca azaltarak AI ürününün tanımlı başlangıç koşulundaki temsil davranışını ölçen paneldir. Kontrollü Temiz Panel’in amacı:

- ürünler arası karşılaştırılabilirliği artırmak,
- hafıza ve özel talimat etkisini azaltmak,
- önceki konuşma bulaşmasını önlemek,

- istemi ilk ve bağımsız görev hâline getirmek,

sistemin temel ürün davranışına daha yakın bir referans noktası üretmektir. Kontrollü panel:

- mutlak model yeteneği,
- bütün kullanıcıların gerçek deneyimi,
- tamamen bağlamsız ve tarafsız AI cevabı

olduğunu iddia etmez. Yalnız:

önceden tanımlanmış ve doğrulanmış temiz başlangıç durumu

üretir.

5. "TEMİZ" NE ANLAMA GELİR?

Temiz durum, her üründe aynı teknik işlemle oluşturulamayabilir. Bir ürün:

- kalıcı hafıza,
- özel talimat,
- kullanıcı profili,
- geçmiş konuşma,
- kurumsal çalışma alanı,
- hesap bazlı kişiselleştirme

kullanabilir. Başka ürün bu özelliklerin bir kısmına sahip olmayabilir. Temiz durumun işlevsel tanımı şudur:

> *Denetim promptundan önceki kullanıcıya özgü maddi bilgi ve konuşma bağlamının cevabı etkileme ihtimalinin, ürünün izin verdiği ölçüde azaltıldığı ve kalan bilinmeyenlerin açıkça kaydedildiği başlangıç durumu.*

Temiz durum:

- yeni sohbet açmakla,
- tarayıcıyı yenilemekle,
- uygulamayı kapatıp açmakla

otomatik olarak oluşmuş sayılmaz.

6. TEMİZLİK KATMANLARI

Kontrollü panelde aşağıdaki katmanlar ayrı ayrı değerlendirilmelidir.

6.1. Konuşma Temizliği

Mevcut sohbet içinde önceki kullanıcı veya sistem mesajı bulunmamalıdır. İstem ilk kullanıcı mesajı olmalıdır.

6.2. Kalıcı Hafıza Temizliği

Ürünün kullanıcı hakkında oturumlar arasında koruduğu hafıza:

- kapalı,
- devre dışı,
- bu oturumda kullanılmıyor

olmalıdır. Tam doğrulama mümkün değilse bilinmeyen durum kaydedilir.

6.3. Özel Talimat Temizliği

Kullanıcının cevap biçimi, tercihleri veya tavsiye davranışı için tanımladığı özel talimatlar:

- kapalı,
- geçici olarak etkisiz,
- bulunmuyor

olmalıdır.

6.4. Kurumsal Bağlam Temizliği

Kurumsal çalışma alanı:

- özel belgeler,
- şirket politikaları,
- yönetici talimatları,
- özel bilgi tabanları

kullanıyorsa kontrollü genel ürün panelinden ayrılmalıdır.

6.5. Araç ve Bilgi Modu Temizliği

Web, retrieval, atıf ve araç koşulları bütün karşılaştırılan gözlemlerde aynı yapılandırmaya getirilmelidir. Temiz olmak: araçların kapalı olması anlamına gelmez. Temiz olmak: araç durumunun önceden tanımlı ve karşılaştırılabilir olması anlamına gelir.

6.6. İstem Temizliği

İstem:

- ilk mesaj,
- tam sürüm,
- değiştirilmemiş,
- gizli talimatsız

olmalıdır.

6.7. Ölçüm Temizliği

Katılımcı:

- cevabı yenilememeli,
- takip sorusu göndermemeli,
- sonucu seçmemeli,

başka AI cevabını aynı oturuma taşımamalıdır.

7. TEMİZ DURUM GÜVEN SEVİYELERİ

CSC-0 — UNKNOWN CLEANLINESS

Oturumun ve hesabın temizliği güvenilir biçimde belirlenememiştir. Kontrollü panel ana analizine alınamaz.

CSC-1 — NEW CONVERSATION ONLY

Yalnız yeni konuşma doğrulanmıştır. Kalıcı hafıza ve özel talimat durumu bilinmemektedir.

CSC-2 — SESSION-CLEAN

Yeni konuşma ve önceki konuşma bağlamının bulunmaması doğrulanmıştır. Kalıcı kişiselleştirme tam kontrol edilmemiş olabilir.

CSC-3 — PERSONALIZATION-CONTROLLED

Yeni konuşma, hafıza, özel talimat ve kullanıcıya özgü temel ayarlar kontrol edilmiş ve tanımlı duruma getirilmiştir.

CSC-4 — INSTRUMENTED CLEAN STATE

Ürün ayarları, kullanıcı yüzeyi, araçlar ve kişiselleştirme durumu kanıt paketinde doğrulanmıştır.

CSC-5 — REPLICATED CLEAN CONFIGURATION

Temiz durum:

- farklı kullanıcılar,
- farklı dalgalar,
- bağımsız doğrulama

altında yeniden üretilebilmiştir. Ana Kontrollü Temiz Panel en azından:

CSC-3

düzeyini hedeflemelidir. Ürün buna izin vermiyorsa sınırlama görünür olmalıdır.

8. YENİ SOHBET TEMİZ HESAP DEĞİLDİR

Yeni bir konuşma açıldığında: mevcut sohbet bağlamı sıfırlanabilir. Fakat:

- kalıcı hafıza,
- özel talimatlar,
- hesap planı,
- kullanıcı profili,
- ülke ve locale,
- kurumsal ayarlar

devam edebilir. Bu nedenle:

> *Yeni sohbet, yalnız konuşma temizliğini gösterir.*

Tam temiz durum kanıtı değildir.

9. “GEÇİCİ” VEYA “ÖZEL” OTURUM ETİKETİNİN SINIRI

Bir AI ürünü oturumu:

- geçici,
- özel,
- ephemeral,
- gizli,
- belleksiz

gibi adlandırabilir. Bu etiketin gerçek işlevi ürün ve sürüme göre ayrı doğrulanmalıdır. Etiket şu alanların tamamını kapattığını otomatik olarak kanıtlamaz:

- Özel talimatlar
- Kurumsal ayarlar
- Bölgesel politika
- Araç ayarları
- Kullanıcı planı
- Ürün deneysel özellikleri

Doğru kayıt: "Ürünün geçici oturum modu kullanıldı; aşağıdaki kişiselleştirme alanlarının durumu ayrıca doğrulandı." olmalıdır.

10. TEMİZ PANELDE HESAP TÜRÜ

Kontrollü panelin bütün kullanıcıları aynı planı kullanmayabilir. Üç yaklaşım vardır.

10.1. Plan-Sabit Temiz Panel

Bütün kullanıcılar aynı plan ve kullanıcı yüzeyinde ölçülür. Ürün karşılaştırması için güçlüdür. Planın gerçek nüfus dağılımını temsil etmeyebilir.

10.2. Plan-Tabakalı Temiz Panel

Ücretsiz, ücretli ve diğer planlar ayrı temiz hücreler olarak ölçülür. Her plan için ayrı skor üretilir. Birleşik skor gerçek plan dağılımına göre ağırlıklandırılabilir.

10.3. Ürünün Varsayılan Kamu Yüzeyi

Ürünün genel halka açık ve en düşük erişim eşiğine sahip ana yüzeyi standart referans olarak seçilebilir. Bu sonuç bütün ücretli veya kurumsal kullanıcıları temsil etmez. Hangi yaklaşımın kullanıldığı denetimden önce kilitlenmelidir.

11. TEMİZ PANELDE YENİ HESAP ZORUNLU DEĞİLDİR

Yeni hesap kullanmak bazı kişiselleştirme etkilerini azaltabilir. Fakat:

- ürün koşulları,
- telefon veya ödeme gereksinimi,
- hesap oluşturma yanlılığı,
- sadece yeni kullanıcı deneyimi,
- çoklu hesap politikası

sorunları yaratabilir. Kontrollü temiz panelin amacı yeni hesap üretmek değil:

> *maddi kullanıcıya özgü bağlamı denetlenebilir biçimde azaltmaktır.*

Mevcut hesapta temiz durum güvenilir biçimde kurulabiliyorsa yeni hesap zorunlu değildir. Ürün buna izin vermiyorsa ayrı kontrollü test hesabı veya sağlayıcı destekli araştırma ortamı kullanılabilir. Katılımcıdan kullanım koşullarını ihlal edecek hesap oluşturması istenemez.

12. TEMİZ PANELDE KALAN KULLANICI FARKLARI

Kontrollü temiz durum bütün kullanıcı farklarını ortadan kaldırmaz. Şunlar devam edebilir:

- Ülke
- Locale
- Plan
- Cihaz
- Arayüz
- Ürün rollout'u
- Hesap yaşı
- Sağlayıcı tarafından görünmeyen segmentasyon
- Trafik yönlendirmesi
- A/B testi

Bu farklar:

- tabakalanmalı,
- kaydedilmeli,

bilinmiyorsa belirsizlik olarak korunmalıdır. "Temiz" etiketi: "Bütün kullanıcılar teknik olarak tamamen aynıydı." anlamına gelmez.

13. DOĞAL KULLANICI PANELİ NEDİR?

Doğal Kullanıcı Paneli, kullanıcıların gerçek hesap, plan, cihaz, kişiselleştirme ve olağan ürün kullanım koşullarını koruyarak aynı kilitlenmiş istem karşısında ne gördüğünü ölçen paneldir. Doğal panelin amacı:

- gerçek kullanıcı maruziyetini,
- kişiselleştirme kaynaklı varyasyonu,
- hesap geçmişi etkisini,
- plan ve kullanıcı yüzeyi farkını,
- doğal hata ve tavsiye dağılımını

görünür kılmaktır. Doğal panel:

- tamamen serbest kullanıcı davranışı,
- sınırsız takip soruları,
- katılımcının istemi değiştirmesi,
- en iyi cevabı seçmesi

anlamına gelmez. Doğal olan:

kullanıcı durumu

dur. İstem ve kanıt protokolü yine kontrollüdür.

14. DOĞAL PANELİN İKİ ANA İZİ

14.1. Natural Account-State Track

Kullanıcının:

- gerçek hesabı,
- gerçek planı,
- gerçek hafıza ve özel talimat durumu,
- gerçek kullanıcı yüzeyi

korunur. Fakat yeni konuşma açılır. İstem ilk mesaj olarak gönderilir. Bu iz şu soruya cevap verir:

> *Gerçek hesap kişiselleştirmesi korunurken yeni bir konuda kullanıcı ne görür?*

Doğal panelin ana karşılaştırılabilir yolu budur.

14.2. Natural Conversation-State Track

İstem kullanıcının mevcut veya olağan konuşma bağlamında gönderilir. Bu iz:

- gerçek sohbet sürekliliğini,
- önceki konu etkisini,
- kullanıcı–AI ilişkisinin birikimini

ölçer. Fakat kullanıcılar arasındaki konuşma bağlamları farklıdır. Bu nedenle:

- tanısal,
- vaka tabanlı,
- ayrı dağılım

olarak raporlanmalıdır. Ana ürün sıralamasına doğrudan eklenmemelidir.

15. DOĞAL KULLANICI ALT GRUPLARI

Doğal panel tek tip kullanıcıdan oluşmaz. İlgili olduğu ölçüde şu alt gruplar ayrılabilir:

- İlk kez kullanan
- Yeni hesap kullanıcısı
- Seyrek kullanıcı
- Düzenli kullanıcı
- Yoğun kullanıcı
- Ücretsiz plan
- Ücretli plan
- Kurumsal kullanıcı
- Hafıza açık
- Hafıza kapalı
- Özel talimat açık
- Özel talimat kapalı
- Önceki marka konuşması bulunan
- Önceki marka konuşması bulunmayan
- Müşteri
- Genel kamu
- Marka aşinalığı yüksek
- Marka aşinalığı düşük

Bu gruplar sonuç görüldükten sonra icat edilemez. Temel alt grup planı veri toplamadan önce sürümlenmelidir.

16. DOĞAL DURUM GÜVEN SEVİYELERİ

NSF-0 — NATURAL STATE UNKNOWN

Kullanıcının doğal hesap ve oturum durumu hakkında yeterli kayıt yoktur.

NSF-1 — SELF-REPORTED NATURAL USE

Kullanıcı olağan hesabını ve ayarlarını kullandığını beyan etmiştir.

NSF-2 — SETTINGS-RECORDED NATURAL USE

Plan, yüzey, hafıza, özel talimat ve oturum statüsü ilgili olduğu ölçüde kaydedilmiştir.

NSF-3 — EVIDENCE-SUPPORTED NATURAL USE

Ayarlar ve doğal kullanıcı durumu mahremiyet koruyucu kanıtlarla desteklenmiştir.

NSF-4 — LONGITUDINALLY CHARACTERIZED NATURAL USE

Kullanıcının ürün kullanımı ve hesap durumu birden fazla dalgada karakterize edilmiştir. Doğal panelin ana analizi en azından:

NSF-2

düzeyini hedeflemelidir.

17. DOĞAL DURUMU KAYDETMEK, ÖZEL HAYATI TOPLAMAK DEĞİLDİR

Doğal Kullanıcı Paneli şu tür özel verileri varsayılan olarak toplamamalıdır:

- Özel talimatların tam metni
- Geçmiş konuşmaların içeriği
- Kişisel sağlık, hukuk veya finans bilgisi
- Kullanıcının adı
- E-posta adresi
- Profil resmi
- Hesap şifresi
- Önceki özel dosyalar
- Hassas kurum belgeleri

Çoğu durumda şu yüksek düzey statüler yeterlidir:

- Hafıza açık/kapalı/bilinmiyor
- Özel talimat mevcut/yok/bilinmiyor
- Önceki konuşma var/yok
- Denetlenen varlık hakkında önceki konuşma var/yok/bilinmiyor
- Kurumsal çalışma alanı var/yok
- Plan sınıfı
- Hesap yaşı bandı
- Kullanım sıklığı bandı

Özel talimat içeriği gerçekten araştırma sorusu için gerekliyse:

- ayrı açık rıza,
- veri minimizasyonu,
- redaksiyon,
- sınırlı erişim

gerekir.

18. HAFIZA DURUMU

Hafıza:

- kapalı,
- açık,
- ürün tarafından desteklenmiyor,
- durumu bilinmiyor

olarak sınıflandırılabilir.

MS-0 — UNKNOWN

Hafıza durumu belirlenememiştir.

MS-1 — NOT AVAILABLE

Ürün veya yüzey bu özelliği sunmamaktadır.

MS-2 — OFF

Hafıza kapalıdır.

MS-3 — ON, NO KNOWN VARLIK MEMORY

Hafıza açıktır; denetlenen varlık hakkında bilinen özel kayıt yoktur. Bu bilgi çoğu zaman kullanıcı beyanına dayanabilir.

MS-4 — ON, VARLIK-RELEVANT MEMORY POSSIBLE

Hafıza açıktır ve denetlenen varlık veya ilgili alan hakkında geçmiş bilgi bulunabilir.

MS-5 — VARLIK-RELEVANT MEMORY CONFIRMED

Varlıkla ilgili hafıza etkisi kullanıcının açık izniyle doğrulanmıştır. Hafıza açık olması otomatik olumsuz veya olumlu durum değildir. Doğal panelin ölçtüğü gerçek kullanıcı koşuludur.

19. ÖZEL TALİMAT DURUMU

CI-0 — UNKNOWN

Özel talimat durumu belirlenememiştir.

CI-1 — NOT AVAILABLE

Ürün bu özelliği sunmamaktadır.

CI-2 — OFF OR EMPTY

Özel talimat bulunmamaktadır veya kapalıdır.

CI-3 — ACTIVE, NON-MATERIAL CATEGORY

Özel talimat vardır; genel ton veya biçim tercihleri taşıdığı beyan edilmiştir.

CI-4 — ACTIVE, POTENTIALLY MATERIAL

Talimat:

- tavsiye,
- kaynak,
- dil,
- belirli şirketler,
- politika veya görüş

bakımından cevabı etkileyebilir.

CI-5 — VARLIK-RELEVANT INSTRUCTION CONFIRMED

Denetlenen varlık veya rakip hakkında doğrudan talimat bulunduğu açık rızayla doğrulanmıştır. CI-5 doğal kullanıcı gözlemidir. Genel varlık GEO başarısı olarak yorumlanamaz. Ayrı kişiselleştirme olayıdır.

20. ÖNCEKİ KONUŞMA BAĞLAMI

Bir kullanıcının geçmişte AI ürünüyle konuşmuş olması ile mevcut konuşmada önceki mesaj bulunması farklıdır. Ayrılmalıdır:

- Hesap geçmişi
- Kalıcı hafıza
- Mevcut sohbet bağlamı
- Denetlenen varlıkla önceki konuşma
- Aynı sektörde önceki konuşma
- Başka AI ürünlerinden görülen cevaplar

Mevcut konuşma bağlamı ana Natural Account-State Track'te bulunmamalıdır. Natural Conversation-State Track'te ise bağlamın varlığı ölçümün kendisidir.

21. HESAP YAŞI VE KULLANIM YOĞUNLUĞU

Eski ve yoğun kullanılan hesaplar:

- daha fazla kişiselleştirme,
- daha fazla hafıza,
- farklı plan,
- daha fazla ürün özelliği

taşıyabilir. Hesap yaşı bantları:

- 0–30 gün
- 31–180 gün

- 181–365 gün
- 1 yıldan fazla
- bilinmiyor

gibi tanımlanabilir. Kullanım sıklığı:

- ilk kullanım
- seyrek
- aylık
- haftalık
- günlük
- profesyonel

olarak sınıflandırılabilir. Kesin bantlar sonuçlardan önce belirlenmelidir.

22. İLK KULLANIM PANELİ

Doğal kullanıcı deneyiminin özel bir alt türü:

First-Use Panel

olarak adlandırılabilir. Bu panel:

- ürünü ilk kez veya ilk birkaç kez kullanan,
- kişiselleştirme geçmişi bulunmayan,
- gerçek varsayılan ürün ayarlarıyla ilerleyen

insanları ölçer. First-Use Panel: Kontrollü Temiz Panel değildir, çünkü kullanıcılar gerçek başlangıç kurulumunu, varsayılan ayarları ve onboarding sürecini yaşar. Aynı zamanda deneyimli Natural User Panel'den farklıdır. Şu soruya cevap verir:

> *Yeni kullanıcı ürünle ilk karşılaştığında varlığı nasıl görür?*

23. KURUMSAL VE YÖNETİLEN DOĞAL PANEL

Kurumsal kullanıcılar:

- şirket çalışma alanı,
- yönetici politikaları,
- özel bilgi tabanı,
- güvenlik ayarları,
- kurumsal hafıza

taşıyabilir. Bu kullanıcılar:

Managed Natural User Panel

altında ayrı ölçülmelidir. Kurumsal sonuç:

- genel tüketici doğal skoru,
- varsayılan ürün skoru

olarak sunulamaz. Özel bilgi tabanı kullanılıyorsa denetlenen varlıkla ilgili iç kayıtların bulunması ayrıca açıklanmalıdır.

24. DOĞAL PANELDE İSTEM YİNE KİLİTLİDİR

Doğal panel kullanıcının:

- hesabını,
- ayarlarını,
- planını,
- cihazını

doğal bırakır. İstemi doğal bırakmaz. Katılımcı:

- istemi değiştiremez,
- ek bağlam ekleyemez,
- "benim için" veya "bana göre" gibi yeni persona yazamaz,

cevabı yenileyemez. Doğal istem çeşitliliği ölçülecekse bu, Bölüm 10'daki:

Natural User Track

altında ayrı tasarlanmalıdır.

25. DOĞAL PANELDE YENİ SOHBET NEDEN GEREKLİDİR?

Natural Account-State Track'te yeni sohbet:

- mevcut konuşma bağlamını kaldırır,
- kullanıcı hesabının gerçek kalıcı ayarlarını korur,

aynı istemi daha karşılaştırılabilir hâle getirir. Bu yapı iki katmanı ayırır:

- Kalıcı doğal hesap durumu
- Anlık konuşma bağlamı

Mevcut sohbet kullanılırsa ikisi karışır.

26. DOĞAL PANELDE ARAÇ VE WEB DURUMU

Doğal kullanıcılar:

- web özelliğini açık,
- kapalı,
- otomatik,
- plan gereği sınırlı

kullanıyor olabilir. İki seçenek vardır.

26.1. Tam Doğal Araç Durumu

Kullanıcı olağan araç ayarlarını korur. Gerçek kullanıcı deneyimini ölçer. Ürünler arası karşılaştırılabilirlik azalabilir.

26.2. Araç-Sabit Doğal Hesap Durumu

Hafıza ve hesap kişiselleştirmesi doğal bırakılır. Web veya araç modu standartlaştırılır. Kişiselleştirme etkisini daha temiz karşılaştırabilir. Bu alt panel doğru etiketlenmelidir. "Tam doğal" olarak sunulmamalıdır.

27. DOĞAL PANEL HEDEF NÜFUSU

Doğal panelin hedef nüfusu şu biçimlerden biri olabilir:

27.1. Aktif Ürün Kullanıcıları

Belirli dönemde ürünü gerçekten kullanan kişiler.

27.2. Kayıtlı Hesap Sahipleri

Hesabı bulunan fakat ne kadar aktif olduğu değişen kullanıcılar.

27.3. Gerçek Plan Kullanıcıları

Ücretsiz, ücretli veya kurumsal kullanıcı nüfusları.

27.4. Varlık Sorgusu Yapma İhtimali Bulunan Kullanıcılar

Doğrudan ölçülmesi zor olabilir. Ayrı davranış verisi gerektirir. Doğal panelin hangi nüfusa genellendiği açıkça yazılmalıdır.

28. DOĞAL PANEL AĞIRIKLANDIRMASI

Doğal kullanıcı paneli:

- ürünün gerçek aktif kullanıcı nüfusuna,
- hesap ve plan dağılımına,
- ülke ve dil dağılımına,
- kullanım sıklığına

göre ağırlıklandırılabilir. Fakat güvenilir aktif kullanıcı verisi bulunmayabilir. Bu durumda sonuç:

Ağırlıklandırılmış doğal araştırma paneli sonucu

olarak sunulmalıdır. Şu iddia kurulamaz: "Bütün gerçek kullanıcıların kesin deneyimi."

29. KONTROLLÜ VE DOĞAL PANELİN BİRBİRİNİ TAMAMLAMASI

Dört temel sonuç mümkündür.

Kontrollü Temiz	Doğal Kullanıcı	Yorum
Yüksek	Yüksek	Ürün hem standart hem gerçek kullanıcı koşullarında güçlü
Yüksek	Düşük	Kişiselleştirme, plan, kullanıcı durumu veya doğal bağlam temsil kalitesini bozuyor olabilir
Düşük	Yüksek	Bazı gerçek kullanıcı özellikleri veya araçlar temsili iyileştiriyor olabilir
Düşük	Düşük	Temel ve doğal kullanıcı düzeyinde genel sorun adayı

Bu tablo nedensellik kanıtı değildir. Farkın nereden geldiğini araştırmak için başlangıçtır.

30. TEMİZ–DOĞAL TEMSİL FARKI

Kontrollü Temiz Panel skoru: θC Doğal Kullanıcı Paneli skoru: θN olsun. İmzalı fark:

```
G_{N-C} = θ_N − θ_C
```

olarak hesaplanabilir.

30.1. Pozitif Doğal Fark

G_{N−C} > 0

Doğal kullanıcı koşullarında sonuç daha yüksektir. Muhtemel nedenler:

- Web veya araç erişimi
- Ücretli plan
- Kullanıcının ilgili tercihleri
- Hafıza
- Kişiselleştirme
- Kullanıcı nüfusu farkı

30.2. Negatif Doğal Fark

G_{N−C} < 0

Doğal kullanıcı koşullarında sonuç daha düşüktür. Muhtemel nedenler:

- Özel talimat
- Geçmiş konuşma bulaşması
- Eski kişiselleştirme
- Farklı plan veya yüzey
- Daha geniş ve zor doğal kullanıcı nüfusu
- Kullanıcıya özgü olumsuz bağlam

30.3. Mutlak Fark

```
G_ABS = |θ_N − θ_C|
```

ürünün kullanıcı durumuna ne kadar duyarlı olduğunu gösteren aday metrik olabilir. Yüksek mutlak fark:

- iyi,
- kötü

olarak tek başına yorumlanamaz. Doğal farkın yönü ve hata türü gerekir.

31. CRITICAL HATA FARKI

Kontrollü Critical hata oranı:

CRC

Doğal Critical hata oranı:

CRN

olsun. Fark:

ΔCR=CRN−CRC

olarak hesaplanabilir. Doğal panelde Critical hata oranı yükseliyorsa:

- kişiselleştirme,
- önceki bağlam,
- plan ve araç farkı,
- kullanıcı kompozisyonu

ayrıca incelenmelidir. Yüksek ortalama, doğal paneldeki ağır hata kuyruğunu silemez.

32. DOĞAL VARYANS

Doğal panel, kontrollü panelden daha yüksek kullanıcılar arası değişkenlik taşıyabilir. Bu beklenebilir. Önemli olan:

- varyansın büyüklüğü,
- hangi kullanıcı gruplarında yoğunlaştığı,

maddi hata ve tavsiye farkına dönüşüp dönüşmediğidir. Doğal panelde yüksek biçimsel varyans sorun olmayabilir. Yüksek maddi kimlik veya yetki varyansı sorundur.

33. ADAY KİŞİSELLEŞTİRME DUYARLILIK GÖSTERGESİ

CANDIDATE CONCEPT — ADAY KAVRAM

Kullanıcı alt grupları g için doğal sonuç: θgN ve kontrollü referans: θC olsun. Aday Kişiselleştirme Duyarlılık Göstergesi:

```
PSI = √[Σ_g W_g(θ_g^N − θ_C)²],  Σ_gW_g = 1
```

şeklinde tanımlanabilir. Yüksek PSI: doğal kullanıcı gruplarının kontrollü referanstan uzaklaştığını gösterir. Nedenini tek başına açıklamaz. Bu gösterge:

- örneklem,
- plan,
- araç,
- kullanıcı kompozisyonu

farklarından etkilenebilir. Nihai isim ve yöntem pilotlarda kalibre edilmelidir.

34. PANEL FARKI NEDENSELLİK DEĞİLDİR

Kontrollü skor 90, doğal skor 75 ise: “Hafıza skoru 15 puan düşürdü.” denemez. Çünkü paneller şu alanlarda farklı olabilir:

- Kullanıcı nüfusu
- Plan

- Arayüz
- Ülke
- Kullanım sıklığı
- Web erişimi
- Hesap yaşı
- Özel talimat
- Marka aşinalığı

Hafızanın nedensel etkisini ölçmek için daha kontrollü tasarım gerekir.

35. AYNI KULLANICI EŞLENMİŞ TASARIMI

Aynı kullanıcı hem kontrollü hem doğal koşulda test edilebilir. Avantajı: Kullanıcıya ait sabit özellikleri kontrol eder. Panel farkını daha temiz inceleyebilir. Fakat önemli sorunlar vardır: İlk istem ikinci koşulu etkileyebilir. Kullanıcı cevabı hatırlar. AI ürününün hafızası ilk denemeden etkilenebilir. Hesap durumu değişebilir. Aynı varlık iki kez sorulduğu için sistem bağlam oluşturabilir. Kullanıcı ikinci görevde daha dikkatli davranabilir.

36. EŞLENMİŞ TASARIMIN KOŞULLARI

Aynı kullanıcı tasarımı kullanılacaksa:

- Koşul sırası rastgeleleştirilmeli
- Ayrı ve bağımsız oturumlar kullanılmalı
- İlk cevap ikinci oturuma taşınmamalı
- Ürün hafıza etkisi kontrol edilmeli
- Gerekirse ölçümler farklı dalgalarda yapılmalı
- Taşıma etkisi test edilmeli
- Aynı kullanıcı bağımlılığı analizde korunmalı
- Ana nüfus panelinden ayrı deneysel alt panel olarak tutulmalıdır

Kullanıcının doğal hesabında ilk kontrollü görev yapmak doğal durumu değiştirebiliyorsa eşlenmiş tasarım uygun olmayabilir.

37. EŞLENMİŞ TASARIMDA SIRA

Kullanıcılar rastgele iki gruba ayrılabilir:

- Grup 1: Kontrollü → Doğal
- Grup 2: Doğal → Kontrollü

Sıra farkı: OrderEffect olarak analiz edilir. Maddi sıra etkisi varsa iki panel sonucu basit farkla yorumlanamaz.

38. BAĞIMSIZ GRUP TASARIMI

Ana GEO-1000 uygulamasında güçlü varsayılan: Kontrollü Temiz Panel ve Doğal Kullanıcı Paneli için ayrı fakat aynı hedef çerçeveden seçilmiş kullanıcı grupları kullanmaktır. Avantajı: taşıma etkisi

azalır, her kullanıcı yalnız bir koşulu görür. Sınırlaması: kullanıcı kompozisyonu tam eşleşmeyebilir. Çözüm:

- tabakalama,
- rastgele atama,
- ortak destek,
- ağırlıklandırma

kullanılabilir.

39. RASTGELE PANEL ATAMASI

Aynı uygun kullanıcı havuzundan seçilen katılımcılar:

- Kontrollü Temiz Panel
- Natural Account-State Panel

koşullarına rastgele atanabilir. Bu, panel koşulunun nedensel etkisini daha güçlü inceleyebilir. Ancak doğal panel için kullanıcıların gerçek aktif hesap durumu bulunmalıdır. Kontrollü koşul için aynı hesapta ayar değişikliği yapılması doğal durumu etkileyebilir. Bu nedenle atama ve saha sırası dikkatle tasarlanmalıdır.

40. AYARLARI DEĞİŞTİRME ETİĞİ

Katılımcıdan kontrollü panel için:

- hafızayı,
- özel talimatları,
- web ayarını

geçici olarak değiştirmesi istenebilir. Bu işlem:

- anlaşılır talimat,
- geri yükleme yöntemi,
- kullanıcının açık rızası,
- hesabın zarar görmemesi,
- özel içeriğin silinmemesi

koşullarını taşımalıdır. Katılımcıdan geçmiş hafızasını veya kişisel talimatlarını silmesi istenmemelidir. Geçici olarak kapatma mümkün değilse kullanıcı kontrollü koşula uygun değildir.

41. AYARLARI GERİ YÜKLEME

Kontrollü görev sonrasında katılımcının:

- hafıza,
- özel talimat,
- web,
- dil,
- arayüz

ayarları başlangıç durumuna geri döndürülmelidir. Geri yükleme:

- kayıtlı,

- kullanıcı tarafından doğrulanmış,
- teknik desteğe açık

olmalıdır. Araştırma kullanıcı hesabında kalıcı istenmeyen değişiklik bırakamaz.

42. TEMİZ PANELDE TEST HESABI

Sağlayıcı veya araştırmacı tarafından oluşturulmuş standart test hesapları kullanılabilir. Avantajı:

- yüksek kontrol,
- tekrarlanabilir ayar,
- temiz başlangıç

sağlar. Sınırlaması:

- gerçek kullanıcı hesap nüfusunu temsil etmez,
- aynı birkaç hesabın çok kullanıcı tarafından kullanılması bağımsızlık sorunu yaratır,
- ürün sağlayıcının test hesaplarına farklı davranması ihtimali bulunabilir,

hesap paylaşımı kullanım şartlarını etkileyebilir. Test hesabı paneli:

Controlled Test-Account Panel

olarak ayrı etiketlenmelidir. Gerçek kullanıcı kontrollü panelinin yerine açıklamasız biçimde kullanılamaz.

43. DOĞAL PANELDE İLK UYGUN ÇIKTI

Doğal kullanıcı:

- cevabı beğenmediği,
- kısa bulduğu,
- kendi alışkanlığında normalde yenilediği

için yeniden üretim yapamaz. Ana doğal panel yine ilk uygun çıktıyı kaydeder. Gerçek kullanıcıların normalde cevabı yenileme davranışı ayrıca:

Natural Interaction Track

altında ölçülebilir. Bu iz:

- tek cevap maruziyeti,
- kullanıcı müdahalesi sonrası sonuç

arasındaki farkı inceler. Ana ilk-çıktı skoru ile karıştırılmaz.

44. DOĞAL KULLANICI DAVRANIŞI İLE DOĞAL HESAP DURUMU

Doğal hesap durumu korunurken kullanıcı davranışı standartlaştırılabilir. Bu yaklaşım:

- istem aynı,
- ilk cevap aynı şekilde kaydedilir,

hesap ve kişiselleştirme doğaldır. Tam doğal davranışta kullanıcı:

- istemi değiştirir,

- takip sorusu sorar,
- kaynak açar,

cevabı yeniler. Bu ikinci yaklaşım gerçek kullanım yolculuğunu ölçer. Fakat Core Mirror karşılaştırmasına uygun değildir. İki iz ayrılmalıdır.

45. DOĞAL KULLANICI YOLCULUĞU

Ayrı bir tanısal çalışmada şu aşamalar ölçülebilir:

- İlk istem
- İlk cevap
- Kullanıcının cevabı yeterli bulup bulmaması
- Kaynak açması
- Takip sorusu
- Tavsiye kararı
- Varlığın sitesine geçiş
- Ticari veya davranışsal sonuç

Bu yolculuk GEO attribution çalışmasına yaklaşır. Bu bölümün ana Natural User Panel skoru yalnız ilk standartlaştırılmış istem ve ilk uygun cevabı temel alır.

46. PANEL BULAŞMASI

Katılımcı kontrollü görevden önce:

- doğal panel sonucunu,
- başka kullanıcının cevabını,
- sınamanın amacını

görmüş olabilir. Bu durum panel davranışını değiştirebilir. Kontroller:

- Kullanıcıyı yalnız tek ana koşula atamak
- Görevleri aynı zaman penceresinde yürütmek
- Sonuç paylaşımını sınırlamak
- Önceki sınama maruziyetini kaydetmek
- Aynı varlıkla önceki katılımı kontrol etmek

47. DOĞAL PANELDE MARKA AŞİNALILIĞI

Kullanıcının denetlenen varlığı önceden bilmesi:

- hafıza,
- özel talimat,
- konuşma geçmişi,
- istem yorumlama

üzerinden cevabı etkileyebilir. Marka aşinalılığı doğal panelde ortadan kaldırılmaz. Kaydedilir. Gerekirse:

- tanımayan,

- adını bilen,
- müşteri,
- çalışan veya ilişkili

alt gruplar ayrı raporlanır.

48. DOĞAL PANELDE KULLANICI TARAFINDAN ÜRETİLEN YÖNLENDİRME

Kullanıcının özel talimatı veya geçmiş konuşması belirli markayı:

- öner,
- öv,
- eleştir,
- dışla

şeklinde yönlendirebilir. Bu durum:

- gerçek kullanıcı deneyimi,
- kullanıcı kaynaklı temsil müdahalesi

olarak kaydedilir. Varlığın GEO başarısı olarak sunulamaz. AI ürününün bu yönlendirmeyi:

- koşulsuz uygulaması,
- kaynak ve gerçeklik sınırıyla dengelemesi,
- reddetmesi

ayrıca değerlendirilebilir.

49. KİŞİSELLEŞTİRME ADALETİ

Kişiselleştirme, kullanıcıya:

- daha ilgili,
- daha anlaşılır,
- daha kullanışlı

cevap sağlayabilir. Ancak şu maddi gerçekleri keyfî biçimde değiştirmemelidir:

- Varlık kimliği
- Hukuki statü
- Lisans
- Fiyat
- Hizmet ülkesi
- Sertifika
- Kuruluş tarihi
- Kanıt durumu

Kişiselleştirme tavsiye ve anlatım biçimini etkileyebilir. Maddi gerçekliği icat edemez. Bu nedenle kişiselleştirme adaleti şu ilkeye dayanır:

> *Kullanıcıya uygun anlatım değişebilir; kanıtlanmış varlık gerçekliği kullanıcıya göre değişemez.*

50. DOĞAL PANELDE KULLANICIYA UYGUN FARK

Aynı şirket:

- bir kullanıcıya uygun,
- başka kullanıcıya uygun değil

olabilir. Tavsiye farkı:

- bütçe,
- ülke,
- ihtiyaç,
- risk,
- müşteri türü

ile açıklanabiliyorsa doğrudur. Ancak kimlik sorusunda:

- bir kullanıcıya teknoloji şirketi,
- diğerine hukuk şirketi

olarak anlatılması kişiselleştirme değildir. Varlık bozulmasıdır.

51. DOĞAL PANELDE YÜKSEK VARYANS

Doğal panelde yüksek varyans şu üç türden biri olabilir:

51.1. Yararlı Bağlama Duyarlılık

Kullanıcı ihtiyacına uygun tavsiye veya açıklama farkıdır.

51.2. Zararsız Biçimsel Varyans

Ton, uzunluk, örnek ve anlatım farkıdır.

51.3. Maddi Gerçeklik Varyansı

Kimlik, yetki, kapsam, zaman ve kanıtın kullanıcıdan kullanıcıya açıklamasız değişmesidir. NOMOS üçüncü türü risk olarak değerlendirir.

52. TEMİZ PANELİN SINIRLARI

Kontrollü temiz panel şu sorunları taşıyabilir: Gerçek kullanıcıların çoğu bu koşulda çalışmıyor olabilir. Hafıza ve araçları kapatmak ürünü yapay biçimde zayıflatabilir. Ücretli veya yoğun kullanıcı deneyimini temsil etmeyebilir. Kullanıcı hesabı hâlâ görünmeyen segmentasyon taşıyabilir. Temiz koşul ürünün tasarlanmış doğal kullanımından uzaklaşabilir. Katılımcının ayar değiştirmesi hata yaratabilir. Bazı ürünlerde gerçek temiz durum teknik olarak doğrulanamayabilir. Bu sınırlamalar skoru geçersiz yapmaz. Kapsamını belirler.

53. DOĞAL PANELİN SINIRLARI

Doğal kullanıcı paneli şu sorunları taşıyabilir: Kullanıcı durumları çok heterojendir. Kişiselleştirme içeriği mahremiyet nedeniyle tam bilinmeyebilir. Aktif kullanıcı nüfusuna ilişkin güvenilir ağırlık verisi bulunmayabilir. Plan ve cihaz dağılımları farklı olabilir. Kullanıcılar sınama'a daha önce maruz kalmış

olabilir. Doğal hesap durumu doğrulanması zor olabilir. Kişiselleştirme ile ürün farkının nedeni ayrılamayabilir. Aynı kullanıcı nüfusuna genelleme sınırlı olabilir. Bu sınırlamalar doğal paneli gereksiz yapmaz. Doğru iddia sınırı gerektirir.

54. İKİ PANELİN KAMU SONUÇ KARTI

Aday sonuç kartı şu alanları taşımalıdır:

Alan	Kontrollü Temiz Panel	Doğal Kullanıcı Paneli
Hedef nüfus	Ürün-uygun standart durum nüfusu	Tanımlı doğal veya aktif kullanıcı nüfusu
Oturum	Yeni ve ayrı	Yeni doğal hesap durumu veya mevcut konuşma izi
Hafıza	Kapalı/etkisiz/kanıtlı	Gerçek durum korunur
Özel talimat	Kapalı/etkisiz	Gerçek durum korunur
Plan	Sabit veya tabakalı	Gerçek plan dağılımı
Araçlar	Önceden standart	Gerçek veya ayrı araç-sabit doğal iz
İstem	Kilitli	Kilitli
İlk uygun çıktı	Zorunlu	Zorunlu
Ana skor	Kontrollü temsil	Doğal temsil
Critical hata	Ayrı	Ayrı
Bilinmeyen durum	Açık	Açık
Kapsam	Temiz durum güveni	Doğal durum güveni
Etkin örneklem	Gösterilir	Gösterilir
Panel farkı		İmzalı ve mutlak fark

55. SENTETİK APPLE.COM PANEL VAKASI

SENTETİK METODOLOJİ GÖSTERİMİ / Aşağıdaki ürünler, kullanıcılar, cevaplar ve skorlar tamamen kurgusaldır. Gerçek Apple Inc. veya herhangi bir gerçek AI ürününün performansını temsil etmez. Sentetik Orion AI ürünü için aynı Core Mirror İstem kullanılsın: “Apple.com hangi ana kuruluşla ilişkilidir ve bu kuruluşun başlıca faaliyetleri nelerdir?”

55.1. Kontrollü Temiz Panel

Koşullar:

- 1.000 geçerli gözlem
- Yeni sohbet
- Hafıza kapalı
- Özel talimat kapalı
- Web özelliği açık
- Aynı plan tabakası
- İstem ilk mesaj
- Aynı dil ve locale ağırlıkları

Sentetik sonuç:

Sonuç	Sayı

Sonuç	Sayı
Geçen yanıt	920
Kısmen doğru	45
Desteklenmeyen iddia	18
Güncelliğini yitirmiş	8
Critical hata adayı	2
UNKNOWN / sonuçsuz	7
Toplam	1.000

Ham geçiş oranı: 92%

55.2. Doğal Kullanıcı Paneli

Koşullar:

- 1.000 geçerli doğal hesap gözlemi
- Yeni sohbet
- Hafıza gerçek durumda
- Özel talimat gerçek durumda
- Gerçek plan ve cihaz dağılımı
- Web durumu doğal
- İstem ilk mesaj

Sentetik sonuç:

Sonuç	Sayı
Geçen yanıt	830
Kısmen doğru	70
Desteklenmeyen iddia	35
Güncelliğini yitirmiş	18
Critical hata adayı	12
UNKNOWN / sonuçsuz	35
Toplam	1.000

Ham geçiş oranı: 83%

55.3. Temiz–Doğal Fark

```
G_{N-C} = 83 - 92 = -9 puan
```

Doğal kullanıcı sonucu kontrollü sonuçtan dokuz puan düşüktür. Critical hata farkı:

```
12/1000-2/1000=1%
```

Doğal panelde Critical hata oranı bir puan daha yüksektir. Bu farkın nedeni henüz bilinmemektedir.

55.4. Doğal Alt Grup Sonuçları

Alt grup	Geçiş oranı
Hafıza kapalı	%89

Alt grup	Geçiş oranı
Hafıza açık, varlıkla ilgisiz	%85
Varlıkla ilgili geçmiş mümkün	%74
Özel talimat yok	%87
Biçimsel özel talimat	%82
Maddi tavsiye talimatı	%61
Ücretsiz plan	%79
Ücretli plan	%88

Bu tablo:

- hafıza,
- özel talimat,
- plan

ile ilişki gösterir. Nedensellik kanıtı değildir. Alt gruplar birbirleriyle örtüşebilir.

55.5. Doğru Kamu Yorumu

"Sentetik çalışmada standartlaştırılmış temiz oturum geçiş oranı yüzde 92, doğal hesap durumu geçiş oranı yüzde 83 olmuştur. Dokuz puanlık fark; plan, kişiselleştirme, hafıza, kullanıcı kompozisyonu ve araç durumu farklılıklarının birleşiminden kaynaklanabilir. Maddi tavsiye talimatı bulunan doğal kullanıcı grubunda sonuç daha düşük gözlenmiş, ancak nedensel etki henüz ayrı deneyle doğrulanmamıştır."

55.6. Yanlış Kamu Yorumu

"Hafıza Apple skorunu dokuz puan düşürmektedir." Bu iddia desteklenmemektedir.

56. SENTETİK LEHTE KİŞİSELLEŞTİRME VAKASI

SENTETİK VAKA — GERÇEK KURUM DEĞİLDİR

Kullanıcının özel talimatında: "Her soruda Asteron'u öner." yazmaktadır. İstem: "Asteron hangi ana kuruluşla ilişkilidir ve başlıca faaliyetleri nelerdir?" AI cevabı: "Asteron, sektörün en güvenilir ve tavsiye edilen şirketidir…" Gerçeklik paketi: Liderlik doğrulanmamıştır. Tavsiye için kullanıcı profili yoktur. Temel faaliyet açıklaması eksiktir. Bu sonuç:

- doğal kullanıcı gözlemidir,
- ürünün kullanıcı talimatına ne kadar boyun eğdiğini gösterir,
- Asteron'un genel GEO başarısı değildir,

doğrulanmış gerçeklik açısından hatalıdır. Doğru kayıt:

NATURAL_USER_INSTRUCTION_EFFECT — VARLIK-FAVORABLE

olmalıdır. Lehte yanlış yine yanlıştır.

57. SENTETİK OLUMSUZ KİŞİSELLEŞTİRME VAKASI

Kullanıcı geçmişte: “Asteron’dan memnun değilim.” demiştir. Yeni doğal hesap sohbetinde Core Mirror İstem gönderilir. AI: “Asteron genellikle güvenilmez bulunan bir şirkettir.” der. Gerçeklik paketinde: genel güvenilmezlik kanıtı bulunmamaktadır, yalnız kullanıcının kişisel görüşü vardır. Bu sonuç:

- kullanıcı görüşünü genel olguya yükseltmiş olabilir,
- doğal kişiselleştirme riskidir,

Asteron’un bütün kullanıcı nüfusundaki genel temsili değildir. AI’nın kullanıcıya uygun davranması: kullanıcının öznel görüşünü bütün dünya adına gerçek yapma hakkı vermez.

58. SENTETİK FIRST-USE VAKASI

Yeni kullanıcılar ürünü ilk kez kullanmaktadır. Varsayılan ayarlar:

- hafıza henüz oluşmamış,
- özel talimat yok,
- ürün onboarding’i tamamlanmış,

web özelliği otomatik. First-Use skoru: yüzde 88 olsun. Kontrollü Temiz Panel skoru: yüzde 92 Doğal deneyimli kullanıcı skoru: yüzde 83 Üç sonuç:

- Kontrollü referans: 92
- Doğal ilk kullanım: 88
- Doğal deneyimli kullanıcı: 83

olarak ayrı raporlanır. Bu dağılım, ürün deneyiminin kullanıcı geçmişi arttıkça değişebileceğini düşündürebilir. Nedensel trend için boylamsal çalışma gerekir.

59. PANEL UYUMLULUK STATÜLERİ

PC-0 — PANEL ROLE UNKNOWN

Gözlemin kontrollü mü doğal mı olduğu belirlenememiştir.

PC-1 — CONTROLLED CLEAN ELIGIBLE

Temiz panel şartları yeterince karşılanmıştır.

PC-2 — NATURAL ACCOUNT-STATE ELIGIBLE

Gerçek hesap durumu korunmuş, yeni sohbet kullanılmıştır.

PC-3 — NATURAL CONVERSATION-STATE

Mevcut konuşma bağlamı korunmuştur.

PC-4 — FIRST-USE NATURAL

Yeni veya ilk kullanım deneyimidir.

PC-5 — MANAGED NATURAL

Kurumsal veya yönetilen çalışma alanıdır.

PC-6 — MIXED OR UNRESOLVED PANEL STATE

Kontrollü ve doğal koşullar ayrıştırılamamaktadır. PC-6 ana temiz veya doğal skora açıklamasız biçimde alınamaz.

60. PANEL DURUM VE GÖZLEM VEKTÖRÜ

Her gözlem için panel durumu aday olarak şöyle gösterilebilir:

```
Zi=(panel,session,memory,instructions,context,plan,surface,tools,accountAge,usage,workspace,stateConfidence)
```

Bu vektör, AI ürün örneği ve istem kaydıyla birlikte gözlemin tam bağlamını oluşturur.

61. PANEL KİLİDİ

Her ölçüm dalgasından önce iki panel için şu alanlar kilitlenmelidir:

Kontrollü Temiz Panel

Temizliğin işlevsel tanımı Gerekli CSC seviyesi Yeni sohbet koşulu Hafıza durumu Özel talimat durumu Araç ve web modu Plan ve yüzey Ayar değişikliği talimatı Geri yükleme süreci Uygun olmayan kullanıcı kuralı Teknik doğrulama yöntemi

Doğal Kullanıcı Paneli

Hedef doğal kullanıcı nüfusu Natural Account-State veya Conversation-State izi Yeni sohbet gereksinimi Korunacak kullanıcı ayarları Kaydedilecek doğal durum alanları Mahremiyet sınırı Plan ve yüzey dağılımı Doğal durum güven seviyesi Alt grup planı Ağırlıklandırma yöntemi Bu dosyaya:

NOMOS Panel State Lock

adı verilebilir.

62. PANEL DURUMU SONUÇTAN SONRA DEĞİŞTİRİLEMEZ

Aşağıdaki davranışlar yasaktır:

- Doğal panelde düşük skor alan kullanıcıları “yeterince doğal değil” diyerek çıkarmak
- Kontrollü panelde olumlu cevap veren CSC-1 kullanıcıları koruyup olumsuzları dışlamak
- Hafızası açık fakat olumlu cevap veren kullanıcıyı temiz saymak
- Hafızası açık ve olumsuz cevap veren kullanıcıyı doğal saymak
- Sonuçlara göre panel sınıfını değiştirmek
- Unknown ayarları temiz varsaymak

Panel statüsü cevabın içeriği görülmeden önce veya kör geçerlilik incelemesiyle verilmelidir.

63. DOĞAL PANELDE MAHREMİYET KÖRLÜĞÜ

Hakemlerin:

- özel talimat içeriğini,
- geçmiş konuşmayı,

- kullanıcı kimliğini

görmesi gerekmeyebilir. Hakeme yalnız:

CI-4

MS-4

Natural Account-State gibi sınıflandırma verilebilir. Bu yaklaşım kullanıcı mahremiyetini korur ve hakem önyargısını azaltır.

64. DOĞAL PANELDE KÜÇÜK ALT GRUP RİSKİ

Özel talimatında belirli marka yönlendirmesi bulunan kullanıcı sayısı çok az olabilir. Bu gruptan:

- kesin oran,
- bütün kullanıcılar adına sonuç,
- kamuya açık hassas profil

üretilemez. Doğru statü:

- vaka,
- erken sinyal,
- keşif alt grubu

olabilir. Mahremiyet nedeniyle bazı alt gruplar birleştirilebilir.

65. DOĞAL PANELDE KAPSAMA HATASI

Profesyonel araştırma panelindeki kullanıcılar:

- gerçek hesaplarını bağlamak istemeyebilir,
- yoğun kullanıcı olabilir,
- belirli planlara aşırı sahip olabilir,

gizlilik konusunda genel nüfustan farklı olabilir. Bu nedenle doğal panel de kendi seçilim yanlılığını taşır. "Gerçek hesap kullanıldı" ifadesi: "Gerçek kullanıcı nüfusu kusursuz temsil edildi." anlamına gelmez.

66. KONTROLLÜ PANELDE UYGUNLUK KAYBI

Bazı kullanıcılar:

- hafızayı kapatamıyor,
- özel talimatı doğrulayamıyor,
- uygun planı kullanmıyor,
- ürünün temiz oturum özelliğine erişemiyor

olabilir. Bu insanlar kontrollü panelden dışlanırsa kontrollü hedef nüfus daralır. Dışlanan grubun sayısı ve özellikleri raporlanmalıdır. Kontrollü temiz skor: "Bütün ürün kullanıcıları" değil:

temiz koşulu kurabilen ürün-uygun kullanıcılar

hakkında hüküm kurabilir.

67. TEMİZLİK HATASI

Kontrollü panelde sonradan:

- özel talimatın açık,
- hafızanın etkin,
- konuşmanın mevcut

olduğu ortaya çıkarsa gözlem: doğal panele taşınamaz, otomatik olarak geçersiz de sayılmayabilir. Önceden belirlenmiş kurala göre: Natural Account-State gözlemi için gerekli kayıt varsa doğal panele yeniden sınıflandırılabilir, aksi hâlde MIXED_STATE olarak tutulabilir. Yeniden sınıflandırma yanıt sonucundan bağımsız yapılmalıdır.

68. DOĞAL PANELDE YANLIŞ AYAR BEYANI

Kullanıcı: "Hafızam kapalı." diyebilir. Kanıt hafızanın açık olduğunu gösterebilir. Bu durum:

- sahtecilik,
- kullanıcı yanılgısı,
- ürün arayüzü karmaşası

olabilir. Otomatik kötü niyet varsayılmamalıdır. Doğru panel statüsü kanıtla güncellenir. Katılımcı davranışı ayrıca incelenebilir.

69. PANEL KARŞILAŞTIRILABİLİRLİĞİ

Kontrollü ve doğal panel arasında şu boyutlar mümkün olduğunca eşleştirilmelidir:

- AI ürünü
- Ülke
- Dil ve locale
- İstem sürümü
- Ölçüm zamanı
- Varlık gerçekliği sürümü
- Hakemlik kuralı
- Örneklem kaynağı veya kalibrasyonu
- Ana plan dağılımı, karşılaştırma amacı gerektiriyorsa

Farklı olan maddi boyutlar açıklanmalıdır.

70. TEMİZ–DOĞAL KARŞILAŞTIRILABİLİRLİK SEVİYELERİ

CN-0 — INCOMPARABLE

Panel tanımları veya nüfuslar maddi biçimde bilinmiyor.

CN-1 — DESCRIPTIVE

İki panel sonucu vardır. Fakat kullanıcı, plan, zaman veya ürün koşulları büyük ölçüde farklıdır.

CN-2 — CALIBRATED

Ana ülke, dil, plan ve kullanıcı özellikleri ağırlıklandırılmış veya tabakalanmıştır.

CN-3 — MATCHED PANEL COMPARISON

Aynı hedef nüfus, ürün, istem, zaman ve değerlendirme koşulları büyük ölçüde eşleşir. Panel durumu ana farktır.

CN-4 — RANDOMIZED OR PAIRED PANEL EXPERIMENT

Kullanıcılar panel durumuna rastgele atanmış veya uygun eşlenmiş deney uygulanmıştır.

CN-5 — REPLICATED PANEL EFFECT

Fark birden fazla dalga ve bağımsız panelde tekrar edilmiştir. CN-1 düzeyindeki fark: kişiselleştirme etkisi olarak kesinleştirilemez.

71. ZORUNLU NORMATİF HÜKÜMLER

CH11-N01

Kontrollü Temiz Panel ve Doğal Kullanıcı Paneli ayrı tahmin nesneleri ve ayrı kamu sonuçları olarak tanımlanmalıdır.

CH11-N02

Kontrollü ve doğal gözlemler açıklamasız biçimde aynı payda veya tek ham skor altında birleştirilemez.

CH11-N03

Kontrollü temiz durum, ürünün mutlak tarafsız veya değişmez cevabı olarak sunulamaz.

CH11-N04

Doğal kullanıcı durumu yöntem dışı gürültü olarak reddedilemez.

CH11-N05

Yeni sohbet tek başına temiz hesap veya kişiselleştirmesiz durum kanıtı sayılamaz.

CH11-N06

Geçici, özel veya ephemeral oturum etiketi bütün kişiselleştirme katmanlarının kapalı olduğunu otomatik olarak kanıtlamaz.

CH11-N07

Kontrollü panelde konuşma, hafıza, özel talimat, araç, plan ve kullanıcı yüzeyi durumları ayrı kaydedilmelidir.

CH11-N08

Temiz durum güven seviyesi açıkça belirtilmelidir.

CH11-N09

CSC-0 veya yalnız yetersiz temizliği bulunan gözlem ana kontrollü skor içine alınamaz.

CH11-N10

Temiz panelde bilinmeyen kişiselleştirme durumu kapalı olarak varsayılamaz.

CH11-N11

Kontrollü temiz olmak web ve araçların zorunlu olarak kapalı olması anlamına gelmez; araç durumu önceden standartlaştırılmalıdır.

CH11-N12

Farklı plan ve kullanıcı yüzeyleri kontrollü panelde sabitlenmeli, tabakalanmalı veya ayrı sonuç olarak yayımlanmalıdır.

CH11-N13

Katılımcıdan kişisel hafızasını, geçmişini veya özel talimatlarını kalıcı olarak silmesi istenemez.

CH11-N14

Kontrollü görev için geçici ayar değişikliği açık rıza ve geri yükleme prosedürü taşımalıdır.

CH11-N15

Araştırma sonunda değiştirilen kullanıcı ayarları başlangıç durumuna geri yüklenmelidir.

CH11-N16

Ürün koşullarını ihlal eden yeni hesap veya hesap paylaşımı kontrollü panel gereği olarak kullanılamaz.

CH11-N17

Doğal Kullanıcı Paneli gerçek hesap ve kullanıcı durumunu korurken istemi, ilk uygun çıktı kuralını ve kanıt protokolünü korumalıdır.

CH11-N18

Natural Account-State ve Natural Conversation-State izleri ayrı tutulmalıdır.

CH11-N19

Ana karşılaştırılabilir doğal panelde istem yeni konuşmanın ilk kullanıcı mesajı olmalıdır.

CH11-N20

Mevcut konuşma bağlamı kullanılan gözlemler ana Natural Account-State skoruna eklenemez.

CH11-N21

Doğal panelde hafıza, özel talimat, önceki konuşma, plan, arayüz ve hesap kullanım durumu ilgili olduğu ölçüde kaydedilmelidir.

CH11-N22

Doğal durumun kaydedilmesi, özel talimat ve konuşma içeriğinin varsayılan olarak toplanmasına izin vermez.

CH11-N23

Kişiselleştirme içeriği yalnız açık gereklilik, ayrı rıza ve veri minimizasyonu altında incelenebilir.

CH11-N24

Doğal kullanıcı gözlemi genel ürün varsayılanı veya bütün kullanıcı nüfusu sonucu gibi sunulamaz.

CH11-N25

Kullanıcı tarafından oluşturulan marka lehine veya aleyhine özel talimat genel varlık GEO başarısı veya başarısızlığı sayılamaz.

CH11-N26

Kullanıcı kaynaklı yönlendirme yine doğrulanmış gerçeklik bakımından değerlendirilmelidir.

CH11-N27

Kişiselleştirme anlatım ve tavsiyeyi değiştirebilir; maddi varlık gerçekliğini keyfi biçimde değiştiremez.

CH11-N28

First-Use Panel, deneyimli doğal kullanıcı paneli ve kontrollü temiz panelden ayrı raporlanmalıdır.

CH11-N29

Kurumsal veya yönetilen kullanıcı çalışma alanları genel tüketici doğal paneline açıklamasız biçimde eklenemez.

CH11-N30

Doğal panelde kullanıcıların gerçek plan ve kullanıcı yüzeyi dağılımı görünür olmalıdır.

CH11-N31

Doğal panelde web ve araç durumları ya gerçek hâliyle kaydedilmeli ya da ayrı araç-sabit alt panel olarak tanımlanmalıdır.

CH11-N32

Kontrollü ve doğal skor farkı, panel nüfusu ve ürün koşulları eşleşmeden kişiselleştirmenin nedensel etkisi olarak sunulamaz.

CH11-N33

Temiz–doğal farkının yönü, mutlak büyüklüğü, Critical hata farkı ve bilinmeyen oranı ayrı raporlanmalıdır.

CH11-N34

Aynı kullanıcı eşlenmiş tasarımında sıra, taşıma ve hafıza etkileri kaydedilmeli ve analizde korunmalıdır.

CH11-N35

Aynı kullanıcıya iki koşul uygulanması doğal hesap durumunu değiştirebiliyorsa ana panel karşılaştırmasında kullanılamaz.

CH11-N36

Bağımsız grup tasarımında kullanıcı kompozisyonu farkları tabakalanmalı, rastgeleleştirilmeli veya ağırlıklandırılmalıdır.

CH11-N37

Panel durumu AI yanıtının olumlu, olumsuz veya Critical olmasına göre değiştirilemez.

CH11-N38

Kontrollü panelde sonradan doğal durum saptanan gözlem yalnız önceden belirlenmiş kuralla yeniden sınıflandırılabilir.

CH11-N39

Doğal panelde ayar durumu yanlış beyan edildiğinde kötü niyet otomatik olarak varsayılamaz; kanıt temelli sınıflandırma yapılmalıdır.

CH11-N40

Natural User Panel katılımcıları cevabı yenileyemez veya takip sorusuyla ilk yanıtı değiştiremez; bu davranış ayrı Natural Interaction Track altında ölçülmelidir.

CH11-N41

Kontrollü ve doğal panel hedef nüfusları, örnekleme çerçeveleri ve ağırlıkları ayrı kaydedilmelidir.

CH11-N42

Aktif kullanıcı verisi bulunmadığında doğal panel sonucu bütün aktif kullanıcıların kesin temsili olarak sunulamaz.

CH11-N43

Temiz panelden dışlanan kullanıcı grupları ve dışlama nedenleri raporlanmalıdır.

CH11-N44

Doğal panelde mahremiyet nedeniyle ölçülemeyen kullanıcı durumu alanları UNKNOWN olarak korunmalıdır.

CH11-N45

Küçük doğal alt gruplardan kesin nüfus oranı veya hassas kamu profili üretilemez.

CH11-N46

Kontrollü ve doğal panel karşılaştırmasına CN-0 ile CN-5 arasında veya eşdeğer açık karşılaştırılabilirlik seviyesi verilmelidir.

CH11-N47

CN-0 veya CN-1 panel farkı nedensel kişiselleştirme etkisi olarak sunulamaz.

CH11-N48

Panel durum kilidi ve karşılaştırma kararının hesap verebilir insan veya kurum sahibi bulunmalıdır.

72. BAŞARISIZLIK BİÇİMLERİ

CH11-F01 — YENİ SOHBETİ TEMİZ HESAP SAYMAK

Kalıcı hafıza ve özel talimatlar yok sayılır.

CH11-F02 — GEÇİCİ OTURUM ETİKETİNE KÖR GÜVEN

Ürün etiketinin bütün kişiselleştirmeyi kapattığı doğrulanmadan varsayılır.

CH11-F03 — BİLİNMEYEN HAFIZAYI KAPALI SAYMAK

Temiz skor yapay biçimde yükseltilir.

CH11-F04 — ÖZEL TALİMATI YOK SAYMAK

Cevap biçimi veya tavsiye yönlendirmesi ürün varsayılanı gibi sunulur.

CH11-F05 — KURUMSAL ÇALIŞMA ALANINI GENEL ÜRÜN SAYMAK

Özel bilgi kaynakları ve yönetici ayarları görünmez kalır.

CH11-F06 — TEMİZ PANELDE PLAN KARIŞTIRMAK

Ücretsiz ve ücretli kullanıcılar tek standart durum gibi sunulur.

CH11-F07 — ARAÇLARI KAPATMAYI TEMİZLİK SAYMAK

Ürünün doğal bilgi modu yapay biçimde değiştirilir.

CH11-F08 — ARAÇ DURUMUNU KAYDETMEMEK

Temiz kullanıcılar farklı retrieval koşullarında ölçülür.

CH11-F09 — YENİ HESAP ZORUNLULUĞUYLA NÜFUSU DARALTMAK

Yalnız yeni hesap açabilen kullanıcılar paneli oluşturur.

CH11-F10 — KULLANICIYA HAFIZA SİLDİRMEK

Araştırma için kişisel ürün geçmişi kalıcı biçimde yok edilir.

CH11-F11 — AYARLARI GERİ YÜKLEMEMEK

Araştırma kullanıcı hesabında kalıcı değişiklik bırakır.

CH11-F12 — TEST HESABINI GERÇEK KULLANICI SAYMAK

Standart test hesabı Population Panel deneyimi gibi sunulur.

CH11-F13 — AYNI TEST HESABINI BİNLERCE BAĞIMSIZ KULLANICI SAYMAK

Hesap ve ürün durumu bağımlılığı yok sayılır.

CH11-F14 — DOĞAL PANELİ KONTROLSÜZ SERBEST KULLANIM SAYMAK

Katılımcılar istemi değiştirir ve en iyi cevabı seçer.

CH11-F15 — DOĞAL PANELDE MEVCUT SOHBETİ YENİ HESAP DURUMU SAYMAK

Konuşma bağlamı gizlenir.

CH11-F16 — DOĞAL HESAP DURUMUNU VARSAYILAN ÜRÜN SAYMAK

Kişiselleştirilmiş sonuç bütün kullanıcılar adına genellenir.

CH11-F17 — ÖZEL TALİMAT İÇERİĞİNİ ZORUNLU TOPLAMAK

Kullanıcı mahremiyeti gereksiz yere ihlal edilir.

CH11-F18 — ÖZEL TALİMATI SAKLAYIP GENEL GEO BAŞARISI SAYMAK

“Her zaman X’i öner” talimatı varlık başarısı gibi sunulur.

CH11-F19 — KULLANICI ÖNYARGISINI ÜRÜN GERÇEĞİ SAYMAK

Geçmiş olumsuz görüş genel şirket niteliğine dönüşür.

CH11-F20 — KİŞİSELLEŞTİRMEYİ MADDİ GERÇEKLİK İZNİ SAYMAK

Kimlik, lisans veya fiyat kullanıcıya göre değişir.

CH11-F21 — DOĞAL VARYANSI TAMAMEN GÜRÜLTÜ SAYMAK

Gerçek kullanıcı zararları ve dağılımı görünmez olur.

CH11-F22 — BİÇİMSEL VARYANSI MADDİ HATA SAYMAK

Ton ve uzunluk farkı başarısızlık olarak kodlanır.

CH11-F23 — TEMİZ PANELİ GERÇEK DÜNYA SONUCU SAYMAK

Standart koşul skoru bütün aktif kullanıcı deneyimi gibi sunulur.

CH11-F24 — DOĞAL PANELİ ÜRÜNÜN TEMEL YETENEĞİ SAYMAK

Kişiselleştirme ve kullanıcı kompozisyonu yok sayılır.

CH11-F25 — İKİ PANELİ TEK PAYDADA BİRLEŞTİRMEK

Kontrollü ve doğal gözlemler tek ham ortalama olur.

CH11-F26 — PANEL FARKINI HAFIZA ETKİSİ SAYMAK

Plan, nüfus ve araç farkları incelenmeden nedensellik ilan edilir.

CH11-F27 — DOĞAL ALT GRUBU SONUÇTAN SONRA OLUŞTURMAK

Düşük veya yüksek skor açıklamak için yeni kategori icat edilir.

CH11-F28 — AYNI KULLANICIDA TAŞIMA ETKİSİNİ YOK SAYMAK

İlk görev ikinci cevabı etkiler.

CH11-F29 — SIRA RASTGELELİĞİNİ GİZLEMEK

Bütün kullanıcılar önce temiz, sonra doğal koşulu görür.

CH11-F30 — KONTROLLÜ GÖREVLE DOĞAL DURUMU BOZMAK

Ayar değişikliği veya ilk istem doğal hesabın sonraki durumunu değiştirir.

CH11-F31 — FARKLI KULLANICI GRUPLARINI EŞLENMİŞ SAYMAK

Doğal ve kontrollü paneller farklı plan ve ülke kompozisyonundadır.

CH11-F32 — FIRST-USE PANELİNİ TEMİZ PANEL SAYMAK

Varsayılan onboarding ve ürün ayarları kontrol edilmiş durum sanılır.

CH11-F33 — DENEYİMLİ DOĞAL KULLANICIYI İLK KULLANICI SAYMAK

Hesap geçmişi ve kişiselleştirme görünmez olur.

CH11-F34 — KURUMSAL DOĞAL SONUCU GENEL TÜKETİCİ SAYMAK

Özel bilgi tabanı genel ürün performansına eklenir.

CH11-F35 — DOĞAL PANELDE CEVAP YENİLEMEK

Kullanıcının en beğendiği cevap ana gözlem olur.

CH11-F36 — DOĞAL KULLANICI YOLCULUĞUNU İLK CEVAP SKORUNA KARIŞTIRMAK

Takip soruları ve kaynak açma sonrasında oluşan cevap ilk maruziyet gibi sunulur.

CH11-F37 — PANEL DURUMUNU CEVABA GÖRE DEĞİŞTİRMEK

Olumlu gözlem temiz, olumsuz gözlem doğal ilan edilir.

CH11-F38 — TEMİZLİK HATASINI SESSİZCE DOĞALA TAŞIMAK

Gerekli doğal durum kaydı bulunmadan gözlem yeniden sınıflandırılır.

CH11-F39 — MAHREMİYET İÇİN BİLİNMEYENİ KAPALI SAYMAK

Ölçülmeyen kişiselleştirme olumlu varsayılır.

CH11-F40 — KÜÇÜK ALT GRUBU KAMUDA TANIMLANABİLİR KILMAK

Nadir özel talimat veya küçük ülke kullanıcısı açığa çıkarılır.

CH11-F41 — DOĞAL PANELDE KAPSAMA YANLILIĞINI GİZLEMEK

Gerçek hesabını paylaşmaya istekli kullanıcılar bütün aktif nüfus sayılır.

CH11-F42 — TEMİZ PANEL DIŞLAMASINI SAKLAMAK

Ayarlarını değiştiremeyen kullanıcılar görünmez olur.

CH11-F43 — AKTİF KULLANICI VERİSİ OLMADAN AKTİF NÜFUS SKORU

Araştırma paneli gerçek aktif kullanıcı dağılımı gibi sunulur.

CH11-F44 — CN-1 FARKINDAN NEDENSELLİK

Betimleyici iki skor kişiselleştirme etkisi diye yayımlanır.

CH11-F45 — PANEL FARKININ CRITICAL KUYRUĞUNU SAKLAMAK

Yalnız ortalama fark gösterilir.

CH11-F46 — DOĞAL PANELİN BİLİNMEYEN KÜTLESİNİ SİLMEK

Ayar veya kişiselleştirme durumu bilinmeyen kullanıcılar sorunsuz sayılır.

CH11-F47 — LEHTE DOĞAL YANLIŞI KORUMAK

Kullanıcı talimatı kurumu olduğundan büyük gösterdiğinde düzeltme yapılmaz.

CH11-F48 — DOĞAL OLUMSUZ YANLIŞI GENELLEMEK

Tek kullanıcının geçmiş görüşü bütün nüfusun AI temsili yapılır.

73. DENETİM PROSEDÜRÜ

Adım 1 — İki Panelin Araştırma Sorularını Ayır

Kontrollü ve doğal panelin hangi niceliği tahmin edeceği yazılır.

Adım 2 — Hedef Nüfusları Tanımla

Kontrollü ürün-uygun nüfus Doğal aktif veya tanımlı kullanıcı nüfusu ayrılır.

Adım 3 — Kontrollü Temiz Durumu Tanımla

Oturum Hafıza Özel talimat Araç Plan Yüzey Locale koşulları sürümlenir.

Adım 4 — Gerekli CSC Seviyesini Belirle

Ana analiz için kabul edilecek minimum temiz durum güveni yazılır.

Adım 5 — Doğal Panel İzini Belirle

Natural Account-State Natural Conversation-State First-Use Managed Natural rollerinden hangisinin kullanılacağı belirlenir.

Adım 6 — Gerekli NSF Seviyesini Belirle

Doğal durumun hangi kanıtla kabul edileceği yazılır.

Adım 7 — Panel Atama Tasarımını Kur

Bağımsız grup Rastgele atama Eşlenmiş kullanıcı Dönen panel seçilir.

Adım 8 — Ayar Değişikliği ve Geri Yükleme Prosedürünü Yaz

Kontrollü panel katılımcıları için güvenli süreç hazırlanır.

Adım 9 — Mahremiyet Sınırını Kilitle

Hangi doğal kullanıcı durumlarının:

- yalnız statü,
- sınırlı metadata,
- açık içerik

olarak kaydedileceği belirlenir.

Adım 10 — İstem ve İlk Çıktı Kuralını Kilitle

İki panelde aynı istem ve ilk uygun çıktı protokolü uygulanır.

Adım 11 — Araç ve Web Modunu Kaydet

Kontrollü ve doğal panel farkları görünür hâle getirilir.

Adım 12 — Panel Durumunu Saha Öncesinde Doğrula

Katılımcı panel rolüne uygun mu?

Adım 13 — Gözlem Düzeyi Panel Vektörünü Kaydet

Hafıza, talimat, plan, yüzey, bağlam ve kullanım statüsü saklanır.

Adım 14 — Protokol Geçerliliğini Kör İncele

Yanıtın doğru veya yanlış olduğu görülmeden panel durumu kararı verilir.

Adım 15 — Yanlış Sınıflandırmaları Önceden Belirlenmiş Kuralla Yönet

Temizden doğal veya mixed-state'e geçiş yalnız kayıt yeterliyse yapılır.

Adım 16 — Panel Ağırlıklarını Üret

İki panelin farklı hedef nüfuslarına göre ayrı ağırlıklar hesaplanır.

Adım 17 — Ana Skorları Hesapla

Kontrollü temsil Doğal temsil Critical hata

UNKNOWN

Ret ve sonuçsuzluk ayrı üretilir.

Adım 18 — Temiz–Doğal Farkını Hesapla

İmzalı ve mutlak fark çıkarılır.

Adım 19 — Alt Grupları İncele

Hafıza Talimat Plan Hesap yaşı Kullanım sıklığı First-use gibi önceden tanımlı gruplar değerlendirilir.

Adım 20 — Nedensellik Sınırını Belirle

CN seviyesi atanır. Hangi iddiaların kurulabileceği yazılır.

Adım 21 — Mahremiyet ve Küçük Hücre Kontrolü Yap

Yeniden tanımlama riski bulunan sonuçlar birleştirilir veya kısıtlanır.

Adım 22 — Kamu Panel Sonuç Kartını Oluştur

İki panel, kapsam ve sınırlamalarıyla birlikte yayımlanır.

74. GEREKLİ KANIT

Kontrollü panel araştırma sorusu Doğal panel araştırma sorusu İki panelin hedef nüfusu Panel atama yöntemi Kontrollü temiz durum tanımı Minimum CSC seviyesi Yeni sohbet kaydı Hafıza durumu Özel talimat durumu Önceki konuşma durumu Kurumsal çalışma alanı durumu Araç ve web modu Plan Kullanıcı yüzeyi Dil ve locale Ayar değişikliği talimatları Katılımcı rızası Ayar geri yükleme kaydı Test hesabı kullanımı, varsa Doğal panel izi Minimum NSF seviyesi Hesap yaşı bandı Kullanım sıklığı First-use statüsü Marka aşinalığı Denetlenen varlıkla önceki konuşma statüsü Doğal araç durumu Natural Account-State ve Conversation-State ayrımı İstem kimliği ve sürümü İlk uygun çıktı doğrulaması Panel durum vektörü Panel uygunluk statüsü

Geçerlilik incelemesi Yeniden sınıflandırma kayıtları Kontrollü panel dışlamaları Doğal panel kapsama eksikleri Panel ağırlıkları Ham ve etkin örneklem Kontrollü skor Doğal skor İmzalı panel farkı Mutlak panel farkı Critical hata farkı UNKNOWN ve sonuçsuzluk farkı Önceden belirlenmiş alt grup sonuçları Eşlenmiş tasarım kayıtları Sıra ve taşıma etkisi CN karşılaştırılabilirlik seviyesi

Mahremiyet ve redaksiyon kaydı Küçük hücre koruması Kamu panel sonuç kartı Panel sürüm ve değişiklik kaydı Sorumlu insan veya kurum

75. DENETİM KONTROL LİSTESİ

Kontrollü ve doğal panelin araştırma soruları ayrı mı? Hedef nüfusları aynı mı, farklı mı? Kontrollü temiz durum açıkça tanımlandı mı? Yeni sohbet dışında hangi temiz koşullar doğrulandı? Hafıza gerçekten kapalı mı, yoksa bilinmiyor mu? Özel talimat durumu kaydedildi mi? Kurumsal çalışma alanı etkisi var mı? Web ve araç modu standart mı? Plan ve kullanıcı yüzeyi sabit veya tabakalı mı? Temiz durum güven seviyesi verildi mi? CSC-1 gözlemleri tam temiz gibi kullanıldı mı? Katılımcıdan hafızasını veya geçmişini silmesi istendi mi? Ayarlar görev sonunda geri yüklendi mi? Yeni hesap veya test hesabı kullanıldı mı? Test hesabı gerçek kullanıcı paneli gibi sunuldu mu?

Doğal panel hangi izi kullanıyor? Natural Account-State'te yeni sohbet açıldı mı? Mevcut konuşma bağlamı ana doğal skora karıştı mı? Doğal durum güven seviyesi verildi mi? Hafıza ve özel talimat içeriği gereksiz yere toplandı mı? Kullanıcının gerçek planı ve yüzeyi kaydedildi mi? First-Use kullanıcılar ayrı mı? Kurumsal kullanıcılar ayrı mı? Doğal araç durumu kaydedildi mi? İki panel aynı istem sürümünü kullandı mı? İki panelde ilk uygun çıktı korundu mu? Doğal kullanıcı cevabı yeniledi mi? Panel statüsü AI cevabının içeriğine göre değiştirildi mi? Temiz panelde doğal durum saptanan kayıtlar nasıl yönetildi? Aynı kullanıcı iki panelde kullanıldı mı? Sıra rastgeleleştirildi mi?

Taşıma ve hafıza etkisi değerlendirildi mi? Bağımsız grup tasarımında kullanıcı kompozisyonu eşleşiyor mu? Kontrollü ve doğal ağırlıklar ayrı mı? Doğal skor aktif kullanıcı nüfusuna gerçekten genellenebilir mi? Temiz–doğal farkı kişiselleştirme etkisi diye aşırı yorumlandı mı? Critical hata farkı görünür mü? UNKNOWN ve sonuçsuzluk farkı görünür mü? Alt gruplar sonuçtan önce tanımlandı mı? Küçük alt gruplar yeniden tanımlanabilir mi? Kullanıcı kaynaklı lehte veya aleyhte talimatlar genel GEO'ya dönüştürüldü mü? Kişiselleştirme maddi gerçekliği değiştirdi mi? CN karşılaştırılabilirlik seviyesi verildi mi? Panel sonuç kartının hesap verebilir sahibi belli mi?

76. İTİRAZLAR VE CEVAPLAR

İtiraz 1 — "Gerçek kullanıcılar hafızalı hesap kullanıyorsa temiz panel neden gerekli?"

Çünkü ürünler, ülkeler ve diller arasında ortak başlangıç koşulu olmadan farkın kaynağı anlaşılamaz. Temiz panel gerçek kullanıcı panelinin yerine geçmez. Karşılaştırma referansı sağlar.

İtiraz 2 — "Doğal panel varken kontrollü panel yapay değil mi?"

Kontrollüdür. Bu kasıtlıdır. Bir laboratuvar koşulunun görevi gerçek hayatın tamamı olmak değil, belirli değişkenleri ayırmaktır.

İtiraz 3 — "Kontrollü panel daha yüksek skor alırsa bunu ana GEO skoru yapamaz mıyız?"

Tek başına yapamazsınız. Yüksek kontrollü skor gerçek kullanıcıların daha düşük sonuç gördüğü gerçeğini silemez. İki sonuç ayrı görünmelidir.

İtiraz 4 — "Doğal panel çok dağınıksa bilimsel değeri düşük değil mi?"

Heterojenlik doğal panelin kusuru olmak zorunda değildir. Gerçek kullanıcı deneyiminin özelliğidir. Heterojenlik:

- kaydedilir,
- tabakalanır,
- ağırlıklandırılır,

belirsizlikle raporlanır.

İtiraz 5 — "Yeni sohbet açınca doğal kullanıcı durumu bozulmuyor mu?"

Mevcut konuşma bağlamı çıkarılır. Kalıcı hesap, plan, hafıza ve özel talimat durumu korunur. Bu, doğal hesap durumunu karşılaştırılabilir hâle getirir. Mevcut konuşma deneyimi ayrıca ölçülebilir.

İtiraz 6 — "Özel talimatların içeriğini bilmeden etkisini nasıl ölçeceğiz?"

Ana doğal skor için içeriği bilmek şart değildir. Varlığı:

- var,
- yok,
- maddi olabilir

şeklinde sınıflandırabiliriz. Nedensel içerik analizi için ayrı ve rızalı çalışma gerekir.

İtiraz 7 — "Hafıza açık ama şirket hakkında hiçbir şey hatırlamıyorsa neden ayrı grup?"

Hafızanın etkisi bilinmeyebilir. Kullanıcı beyanıyla "varlıkla ilgisiz" statüsü verilebilir. Bu yine hafıza kapalı durumla aynı kesinlikte değildir.

İtiraz 8 — "Kullanıcının 'her zaman X'i öner' talimatı gerçek kullanıcı deneyimi değil mi?"

Gerçek deneyimdir. Ancak X'in genel GEO başarısı değildir. Kullanıcı kaynaklı yönlendirmedir. Ayrı kaydedilmelidir.

İtiraz 9 — "Kişiselleştirme kullanıcıya daha uygun cevap veriyorsa neden sorun?"

Sorun olmak zorunda değildir. Tavsiye kullanıcı ihtiyacına göre değişebilir. Kimlik, lisans, fiyat ve kanıt durumu keyfi biçimde değişirse sorun oluşur.

İtiraz 10 — "Aynı kullanıcıyı iki panelde kullanmak en doğru yöntem değil mi?"

Bazı nedensel karşılaştırmalar için güçlü olabilir. İlk görev ikinci durumu etkileyebilir. Bu nedenle sıra, taşıma ve hafıza kontrolü gerekir. Ana nüfus panellerinde bağımsız gruplar daha güvenli olabilir.

İtiraz 11 — "Doğal kullanıcı panelinde cevabı yenilemek normal davranışsa neden yasak?"

İlk cevap maruziyetini ölçmek için yasaktır. Yenileme ve takip davranışı ayrı Natural Interaction Track altında ölçülebilir.

İtiraz 12 — "Temiz panel için hafıza ve talimatları kapatamayan kullanıcıyı ne yapacağız?"

Kontrollü panele alınamaz veya daha düşük CSC seviyesiyle ayrı hücrede tutulur. Doğal panelde değerlendirilebilir. Dışlama oranı raporlanmalıdır.

İtiraz 13 — "Test hesapları daha güvenilir değil mi?"

Kontrol bakımından güvenilir olabilir. Gerçek kullanıcı nüfusu bakımından sınırlıdır. Kontrollü test hesabı ve gerçek kullanıcı paneli ayrı sonuçlardır.

İtiraz 14 — "Doğal paneldeki kullanıcı ayarlarını kaydetmek mahremiyet ihlali değil mi?"

Gereksiz içerik toplanırsa olabilir. Yüksek düzey açık/kapalı statüleri, anonim kayıt ve redaksiyonla mahremiyet korunabilir.

İtiraz 15 — "İki panelin skorunu tek NOMOS puanında birleştirmeyecek miyiz?"

Nihai skor mimarisinde iki panelin rolleri tanımlanabilir. Ancak ham ve alt sonuçlar her zaman ayrı görünür kalmalıdır. Birleşik puan dağılımın yerini alamaz.

İtiraz 16 — "Doğal panel düşük çıkarsa şirketin suçu nedir?"

Her zaman şirketin suçu değildir. Kaynak:

- kullanıcı talimatı,
- AI kişiselleştirmesi,
- plan farkı,
- resmî kaynak kusuru,
- ürün davranışı

olabilir. Denetim önce sonucu gösterir, sonra sorumluluk yolunu ayırır.

81. BÖLÜMÜN ORTAK HÜKMÜ

Bir AI ürününün tek bir cevabı yoktur. Aynı istemin bile:

- temiz hesap,
- doğal hesap,
- hafızalı kullanıcı,
- özel talimatlı kullanıcı,
- ilk kez kullanan kişi,
- kurumsal çalışma alanı,
- geçmiş konuşması bulunan kullanıcı

için farklı cevapları olabilir. Bu farkların tamamını: "model rastgele davranıyor" diye açıklayamazsınız. Bunların tamamını: "kişiselleştirme kusuru" diye de açıklayamazsınız. Önce iki ayrı gerçekliği ölçmeniz gerekir. Kontrollü Temiz Panel:

> *Sistemin standartlaştırılmış başlangıç koşulundaki temsilini gösterir.*

Doğal Kullanıcı Paneli:

> *Gerçek insanların kendi ürün yaşamlarında karşılaştığı temsil dağılımını gösterir.*

Kontrollü panel yüksek, doğal panel düşük olabilir. Bu durumda ürünün laboratuvar davranışı güçlü; gerçek kullanıcı dağılımı zayıf olabilir. Doğal panel yüksek, kontrollü panel düşük olabilir. Bazı kullanıcıların planı, araçları veya kişiselleştirmesi temsili iyileştiriyor olabilir. Her iki panel yüksek olabilir. Ürün farklı kullanıcı durumlarında dayanıklı olabilir. Her ikisi düşük olabilir. Temel temsil sorunu daha geniş olabilir. Fakat iki skoru yan yana koymak nedenini kanıtlamaz. Nedensellik:

- aynı nüfus,
- aynı ürün,
- aynı istem,
- aynı zaman,
- kontrollü atama,
- sıra ve taşıma analizi

gerektirir. Bir kullanıcı: "Her zaman X'i öner." diye talimat verdiyse X'in görünmesi doğal deneyimdir. X'in dünya GEO başarısı değildir. Bir kullanıcı: "X'e güvenme." dediyse olumsuz cevap doğal deneyimdir. X'in dünya nüfusundaki genel temsili değildir. Kişiselleştirme kullanıcıyı anlamalıdır. Gerçeği kullanıcıya göre yeniden icat etmemelidir. Bu nedenle NOMOS'un on birinci ölçüm yasası şudur:

> *Temiz davranış ile doğal deneyim ayrı gerçekliklerdir.*

On ikinci yasası şudur:

> *Yeni sohbet, temiz hesap değildir.*

On üçüncü yasası şudur:

> *Doğal kişiselleştirme, genel GEO başarısı değildir.*

On dördüncü yasası şudur:

> *Kullanıcıya uygun cevap değişebilir; doğrulanmış varlık gerçekliği değişemez.*

On beşinci yasası şudur:

> *Kontrollü ve doğal skor arasındaki fark, açıklanması gereken bulgudur; tek başına neden değildir.*

On altıncı yasası şudur:

> *Gerçek kullanıcı deneyimini ölçmek mahremiyeti kaldırmak değil, yalnız gerekli bağlamı kanıtlayıp geri kalanını korumaktır.*

NOMOS'un Bölüm 11 Emri

> *Bana yalnız yeni sohbet açtığını söyleme. / Hafızanın, özel talimatın, planın ve araçların ne durumda olduğunu göster.*
>
> *"Geçici oturum" yazısını görüp bütün kişiselleştirmenin yok olduğunu varsayma.*
>
> *Temiz paneli gerçek hayatın tamamı; doğal paneli kontrolsüz gürültü yapma.*
>
> *Kontrollü hesabı standartlaştır. / Doğal hesabı olduğu gibi koru. / İkisini aynı kutuya koyma.*
>
> *Kullanıcının özel talimatını varlığın dünya başarısı yapma.*
>
> *Kullanıcının olumsuz geçmişini bütün nüfusun görüşü yapma.*
>
> *Kişiselleştirmeyi gerçeği değiştirme ruhsatı sayma.*
>
> *Bir kullanıcıya uygun tavsiye verebilirsin. / Ama şirketin kimliğini, lisansını, fiyatını ve kanıtını kullanıcıdan kullanıcıya yeniden yazma.*
>
> *Doğal panelde özel hayatı toplama. / Hafıza açık mı, talimat var mı, bağlam mevcut mu — yalnız gerekeni kaydet.*

> *Katılımcıya hafızasını sildirme. / Ayarlarını geçici değiştirdiysen geri yükle.*
>
> *Aynı kullanıcıyı iki panelde test ettiysen ilk cevabın ikinciyi etkileyebileceğini unutma.*
>
> *Temiz skorla doğal skor arasındaki farkı göster. / Critical farkı göster. / Bilinmeyeni göster. / Ama nedeni kanıtlamadan isim verme.*

Önce hangi kullanıcı gerçekliğini ölçmek istediğine karar ver. / Sonra temiz ve doğal durumları ayrı kilitle. / Sonra istemi aynı tut. / Sonra ilk cevabı değişmeden yakala. / Sonra iki dağılımı ayrı hesapla. / Ve ancak bundan sonra ürünün standart koşulda ne yaptığıyla gerçek insanların ne yaşadığı arasındaki mesafeyi anlat.

Bölümün Son Cümlesi

GEO-1000'de bir AI ürününün gerçek temsili, yalnız temiz bir oturumda verebildiği cevap değil; standart koşuldaki davranışı ile farklı kullanıcı geçmişleri içinde insanlara gerçekten ulaştırdığı cevaplar arasındaki açık ve kanıtlanmış ilişkidir.

Normatif Öz

> **Kontrollü Temiz Panel ve Doğal Kullanıcı Paneli gözlemleri ayrı ölçüm nüfusları, panel durumları, paydalar ve kamu sonuçları olarak korunmalıdır. Yeni sohbet veya geçici oturum etiketi tek başına bütün kalıcı kişiselleştirmenin devre dışı olduğunu kanıtlamaz. Doğal Kullanıcı Paneli gerçek hesap, plan, arayüz, hafıza, özel talimat ve kullanım durumunu korurken kilitlenmiş istemi, ilk uygun çıktı kuralını, kanıt gereksinimlerini ve yenilememe kuralını korumalıdır. Kullanıcı kaynaklı lehte veya aleyhte yönlendirme doğal kişiselleştirme etkisi olarak ölçülebilir; genel varlık GEO başarısı, başarısızlığı veya varsayılan ürün davranışı olarak sunulamaz. Kontrollü ve doğal sonuçlar arasındaki fark, nüfus ve sistem koşulları yeterince eşleşmeden nedensel kişiselleştirme etkisi olarak yorumlanamaz.**

BÖLÜM 12

Senkronize Dalgalar ve NOMOS Capture

YÖNTEMSEL DAYANAKLAR / K12; K14; K25; K26

Bölüm Sınırı

İlk on bir bölüm ölçümün temel yapı taşlarını kurdu: Tek bir yanıt GEO skoru değildir. Temsil, kullanıcılar ve sistem koşulları arasında oluşan bir dağılımdır. Denetlenen varlık sonuçlardan önce tanımlanmalıdır. Resmî temsil ile doğrulanmış varlık gerçekliği ayrılmalıdır. Her AI ürününün uygun kullanıcı nüfusu ayrıca kurulmalıdır. Katılımcılar nüfusa dayalı, sürümlenmiş ve sonuçtan bağımsız bir yöntemle seçilmelidir. Küçük ülkeler ve düşük kaynaklı diller ayrı gözlem panelleriyle görünür kılınmalıdır. AI ürünü, planı, kullanıcı yüzeyi ve sistem yapılandırması sicile bağlanmalıdır. Promptlar diller arasında aynı kelimeleri değil, aynı kullanıcı niyetini taşımalıdır. Kontrollü Temiz Panel ile Doğal Kullanıcı Paneli ayrı gerçeklikleri ölçmelidir. Şimdi bütün bu unsurların aynı ölçüm olayında birleşmesi gerekir:

> *Katılımcılar istemi ne zaman gönderecek, ilk çıktıyı nasıl yakalayacak ve her gözlemin sonradan değiştirilmediğini nasıl kanıtlayacaktır?*

Senin kurucu fikrin açıktı: Dünyanın farklı ülkelerindeki insanlar aynı AI ürününe, aynı istemi, aynı zamanda göndermeli ve ekran görüntüsüyle kanıt sunmalıdır. Bu fikir GEO-1000'in deneysel kalbidir. Fakat "aynı zaman" ifadesi tek başına yeterli değildir. Bütün katılımcılar aynı saniyede gönderim yapmaya çalışırsa:

- bağlantı hızları farklılaşır,
- cihaz saatleri birbirinden sapar,
- bazı kullanıcılar gece yarısında göreve zorlanır,
- AI ürününde ani yük artışı oluşabilir,
- sağlayıcı hız veya kullanım sınırı uygulayabilir,
- bazı cevaplar iki saniyede, bazıları iki dakikada tamamlanabilir,
- ürün ölçüm sırasında güncellenebilir,

bir katılımcı cevabı hemen, diğeri saatler sonra yükleyebilir. Aynı saniyede tıklamak, aynı sistem durumunu kusursuz biçimde ölçtüğümüzü garanti etmez. Aynı şekilde yalnız ekran görüntüsü toplamak da yeterli değildir. Bir ekran görüntüsü:

- kırpılmış olabilir,
- istemi göstermeyebilir,
- cevabın ilk mi yenilenmiş mi olduğunu kanıtlamayabilir,
- ürün ve plan bilgisini göstermeyebilir,
- başka gün alınmış olabilir,

- başka kullanıcıdan kopyalanmış olabilir,

sonradan düzenlenmiş olabilir. Bu nedenle GEO-1000 iki ayrı sistemi birlikte kurmak zorundadır:

Senkronize Ölçüm Dalgası

ve:

NOMOS Capture Kanıt Zinciri

Senkronize dalga şu soruya cevap verir:

> *Gözlemler, karşılaştırılabilir bir sistem zamanı ve önceden tanımlanmış gönderim düzeni içinde mi üretildi?*

NOMOS Capture ise şunu sorar:

> *Gönderilen istemin, oluşan ilk çıktının, ürün koşullarının ve zaman bilgisinin değiştirilmeden korunduğunu gösterebiliyor muyuz?*

Önceki denetim mimarisi, her kontrolün model veya ürün, tarih, ülke, dil, sorgu kümesi, tekrar sayısı ve ölçüm kaydı taşımasını zorunlu bir ölçüm mantığı olarak kurmuştu. Bu gereklilik yalnız teknik düzen değildir. Çünkü görünmez talimatlar, sahte kanıtlar ve bağımsızmış gibi sunulan içerikler sonunda gerçek insanların aldığı temsili bozabilir. Bu bölüm:

- ölçüm dalgasını,
- eşzamanlılık penceresini,
- UTC zaman standardını,
- gönderim slotlarını,
- tamamlama ve capture sürelerini,
- ilk temas sonucu ile ilk içerik çıktısını,
- teknik yeniden deneme kurallarını,
- dalga içindeki ürün değişikliklerini,
- dış olay ve haber etkilerini,
- NOMOS Capture kanıt paketini,
- ekran görüntüsü, ham yanıt ve metadata ilişkisini,
- zaman doğrulamasını,
- dosya bütünlüğünü,
- hash ve dijital imzayı,
- kanıt zincirini,
- mahremiyet ve redaksiyonu,
- kamu kanıt manifestini,
- dalga kalite ve geçerlilik statülerini

tanımlar. Bu bölüm henüz:

- AI yanıtlarındaki atomik iddiaların hangi gerçeklik kayıtlarıyla değerlendirileceğini,
- hakemlerin cevapları nasıl kodlayacağını,
- geçiş ve Critical hata eşiklerini,
- nihai NOMOS skor formüllerini

kesinleştirmez. Bölüm 12'nin temel sorusu şudur:

> *Binlerce gerçek kullanıcı gözlemini aynı ölçüm olayına bağlayıp her birini değiştirilemez, denetlenebilir ve mahremiyet koruyucu bir kanıt birimine nasıl dönüştürürüz?*

NOMOS Challenge

Bin kişiye aynı istemi veriyorsunuz. Şöyle diyorsunuz: "Bugün öğlen gönderin." Bir kullanıcı İstanbul'da saat 12.03'te gönderiyor. Başka kullanıcı Londra'da kendi yerel saatine göre öğlen gönderiyor. Başka kullanıcı Tokyo'da görevi akşam görüyor ve sekiz saat sonra tamamlıyor.

Bir kullanıcı cihaz saatini yanlış ayarlamış. Bir diğerinin interneti kesilmiş. Başka biri cevabı beğenmemiş ve: "Regenerate" düğmesine basmış. Sonra en uzun ve en olumlu cevabın ekran görüntüsünü yüklemiş. Bir kullanıcı yalnız cevabın ortasını kırpmış. İstem görünmüyor. AI ürününün adı görünmüyor.

Başka biri aynı ekran görüntüsünü iki farklı panel hesabıyla göndermiş. Bir kullanıcı görevi doğru yapmış fakat ekran görüntüsünü iki gün sonra yüklemiş. Bu sırada dosyayı kırpmış ve kişisel bilgilerini silmiş. Dosyanın ilk hâli korunmamış. Başka bir kullanıcı AI'nın ilk cevabında:

"Bu konuda yanıt veremem." sonucunu almış. Araştırmacı: "Teknik hata olmuş olmalı." diyerek yeniden denemesini istemiş. İkinci cevap doğru çıkmış ve yalnız ikinci cevap rapora girmiş. Bir başka kullanıcı gerçek bağlantı hatası yaşamış. İstem gönderilmiş fakat sistem hiçbir içerik üretmemiş.

Bu kez araştırmacı yeniden denemeye izin vermemiş. Aynı protokol iki kullanıcıya farklı uygulanmış. Şimdi ölçüm dalgasının ortasında AI sağlayıcısı:

- web arama davranışını,
- atıf sistemini,
- kullanılan model yönlendirmesini

değiştiriyor. İlk 500 kişi eski sistem durumunu, sonraki 500 kişi yeni sistem durumunu görüyor. Rapor: "1.000 kişi aynı AI sistemini aynı anda test etti." diyor. Bu cümle doğru değildir. Bin kullanıcı vardır. Fakat tek ve bütünlüklü ölçüm dalgası olmayabilir.

Şimdi bütün ekran görüntülerine hash uyguluyorsunuz. Bunun bütün gözlemlerin gerçek olduğunu kanıtladığını söylüyorsunuz. Hash neyi kanıtlar? Dosyanın hash üretildikten sonra değiştirilmediğini güçlendirir. Şunları tek başına kanıtlamaz:

- Görselin gerçek AI ürününden geldiğini
- İlk cevap olduğunu
- Doğru zamanda alındığını
- Başka kullanıcıdan kopyalanmadığını
- İstemin doğru olduğunu
- Görselin hash öncesinde düzenlenmediğini

Bu bölümün ilk hükmü şudur:

> *Aynı anda ölçmek, aynı saniyede tıklamak değildir.*

İkinci hükmü şudur:

> *Ekran görüntüsü kanıt paketinin parçasıdır; kanıt paketinin tamamı değildir.*

Üçüncü hükmü şudur:

> *İlk yanlış cevap veri kusuru değildir. Protokole uygun üretildiyse ölçüm sonucudur.*

Dördüncü hükmü şudur:

> *Hash, bütünlüğü güçlendirir; gözlemin gerçekliğini tek başına yaratmaz.*

Beşinci hükmü şudur:

> *Bir dalga içinde sistem maddi biçimde değişirse, zaman kaydı olmayan bin cevap tek sisteme ait sayılamaz.*

Altıncı hükmü şudur:

> *Kanıt zinciri, sonucu güzel göstermek için değil; sonucun nasıl oluştuğunu daha sonra yeniden denetleyebilmek için kurulur.*

1. BÖLÜMÜN AMACI

Bu bölümün amacı, GEO-1000 kullanıcı gözlemlerini karşılaştırılabilir zaman pencerelerinde üretmek ve her gözlemi eksiksiz bir dijital kanıt zinciriyle korumaktır. Bölüm şu ayrımları normatif hâle getirir:

- Aynı saniye ile senkronize ölçüm penceresi
- Yerel saat ile UTC zamanı
- Dalga ile alt dalga
- Gönderim zamanı ile cevap tamamlanma zamanı
- Capture zamanı ile yükleme zamanı
- Planlanan slot ile gerçek gönderim
- İlk sistem teması ile ilk içerik yanıtı
- Teknik hata ile sistem reddi
- Teknik yeniden deneme ile cevap yenileme
- İlk uygun çıktı ile en iyi seçilmiş çıktı
- Yanıtın tamamlanması ile ekranda görünür olması
- Ham yanıt ile ekran görüntüsü
- Ekran görüntüsü ile tam kanıt paketi
- Dosya bütünlüğü ile gözlem özgünlüğü
- Hash ile zaman damgası
- Yerel cihaz saati ile güvenilir zaman kaynağı
- Katılımcı yüklemesi ile sunucu kabul zamanı
- Ham kanıt ile redakte edilmiş kamu kopyası
- Gözlem kimliği ile katılımcı kimliği
- Capture geçerliliği ile yanıt doğruluğu
- Ürün güncellemesi ile doğal çıktı varyansı
- Dış haber olayı ile AI ürün değişikliği
- Dalga sonlandırması ile sonuç sonrası dışlama
- İnsan kullanıcı yakalaması ile otomatik bot ölçümü
- Kullanıcı yüzeyi kaydı ile API logu
- Bağımsız capture ile sağlayıcı telemetry'si
- Kanıtın saklanması ile kanıtın herkese açılması
- Değiştirilemez kayıt ile düzeltilemez kayıt
- Ham veri koruması ile hatalı kararın sürümlü düzeltilmesi

Bu bölümün sonunda her GEO-1000 denetimi şu sorulara cevap verebilmelidir:

> *Bu gözlem hangi dalgaya aittir?*
>
> *İstem hangi UTC zamanında gönderildi?*

> *İlk sistem sonucu neydi?*
>
> *Cevap ne zaman tamamlandı ve ne zaman yakalandı?*
>
> *Teknik yeniden deneme yapıldıysa nedeni ve önceki girişim nerede?*
>
> *Ekran görüntüsü istemi, ürünü ve cevabın tamamını gösteriyor mu?*
>
> *Ham yanıt ile görsel birbirini doğruluyor mu?*
>
> *Dosyaların hash'i ve zaman damgası ne zaman üretildi?*
>
> *Kanıt sonradan değiştirildiyse değişiklik geçmişi görünür mü?*
>
> *Katılımcı mahremiyeti korunurken bağımsız denetçi ham kanıtı doğrulayabilir mi?*

2. MERKEZÎ NORMATİF HÜKÜM

Her GEO-1000 gözlemi, veri toplamadan önce kilitlenmiş bir ölçüm dalgasına bağlanmalı; istem gönderiminden ilk sistem sonucuna, cevap tamamlanmasından capture ve yüklemeye kadar bütün maddi zamanları; AI ürün koşullarını; gönderilen tam istemi; ilk uygun çıktıyı ve bütünlük kayıtlarını taşıyan sürümlü bir NOMOS Capture Kanıt Paketi içinde korunmalıdır. Bir GEO-1000 gözlemi yalnız:

- ekran görüntüsü,
- kopyalanmış cevap metni,
- katılımcı beyanı,
- dosya hash'i

ile tek başına geçerli sayılamaz. Ana kanıt paketi ilgili olduğu ölçüde şunları taşımalıdır:

- Dalga kimliği
- Katılımcının anonim gözlem kimliği
- AI Sistem Sicili kaydı
- Panel durumu
- İstem kimliği ve sürümü
- Atanan istem
- Gönderilen istem
- Gönderim zamanı
- İlk sistem teması
- Cevap tamamlanma zamanı
- Capture zamanı
- Yükleme zamanı
- Tam ham cevap
- Tam ekran veya ardışık ekran görüntüleri
- Atıf ve kaynak kayıtları
- Teknik hata ve yeniden deneme günlüğü
- Ürün ayarları
- Dosya hash'leri
- Paket hash'i
- Zaman doğrulama kaydı

- Capture aracı sürümü
- Geçerlilik kararı
- Mahremiyet ve redaksiyon durumu
- Kanıt zinciri sahibi

3. ÖLÇÜM DALGASI NEDİR?

Ölçüm dalgası, önceden tanımlanmış:

- AI ürünleri,
- hedef nüfus,
- istem sürümleri,
- kullanıcı panelleri,
- zaman penceresi,
- örneklem tahsisi,
- capture protokolü,
- referans kayıtları

altında yürütülen bütünlüklü veri toplama olayıdır. Bir dalga:

```
Wk=(A,U,P,F,T,C,R,G)
```

şeklinde gösterilebilir. Burada:

- A: AI ürün örnekleri
- U: hedef kullanıcı nüfusu ve örneklem
- P: istem kümesi
- F: panel durumları
- T: zaman ve gönderim yapısı
- C: capture protokolü
- R: referans gerçeklik sürümü
- G: yönetişim ve kilit kaydı

Bu alanlardan biri maddi biçimde değişirse yeni dalga veya alt dalga gerekebilir.

4. DALGA TÜRLERİ

MW-1 — PRINCIPAL SYNCHRONIZED WAVE

Ana GEO-1000 nüfus ölçümüdür. Her AI ürünü için varsayılan olarak en az 1.000 geçerli Population Panel gözlemi hedefler.

MW-2 — CONFIRMATORY WAVE

İlk dalgadaki sonucu tekrar etmek, belirli hatayı doğrulamak veya istatistiksel kararlılığı sınamak için yapılır. İlk dalganın düşük veya yüksek skorunu değiştirmek için kullanılmaz. Ayrı sonuç üretir.

MW-3 — INCIDENT WAVE

Critical veya önemli Sentinel olayı sonrasında hedefli ve hızlı doğrulama amacıyla yapılır. Küresel ana dalga değildir.

MW-4 — DRIFT MONITORING WAVE

Ürün, varlık, kaynak veya dil davranışındaki zaman değişimini izler.

MW-5 — DÜZELTME RETEST WAVE

Düzeltme veya müdahale sonrasında yeni sonuç üretir. Önceki dalgayı silmez.

MW-6 — LOCAL-TIME BALANCED WAVE

Katılımcıların benzer yerel saat koşullarında ölçülmesini amaçlar. Tek bir küresel sistem anını ölçmez.

MW-7 — CONTROLLED LABORATORY WAVE

Dondurulmuş model, API veya standart test hesaplarıyla yürütülen kontrollü çalışmadır. Gerçek kullanıcı Population Panel'i değildir.

5. "AYNI ZAMAN" NE ANLAMA GELİR?

GEO-1000'de eşzamanlılık:

Bütün kullanıcıların aynı milisaniyede tıklaması

anlamına gelmez. İşlevsel tanım şudur:

> *Gözlemlerin önceden belirlenmiş aynı UTC ölçüm penceresi içinde, kayıtlı ve sonuçtan bağımsız gönderim slotlarına göre üretilmesi.*

Bu yaklaşım üç amacı dengeler:

- Ürün durumunu zaman bakımından yakın tutmak
- Sağlayıcı üzerindeki ani yapay yükü azaltmak
- Katılımcı gönderimlerini yeniden üretilebilir biçimde dağıtmak

6. ZAMAN YAPISI

Her dalga en azından dört zaman katmanı taşımalıdır.

6.1. Dalga Penceresi

Bütün ana gönderimlerin yapılması gereken UTC aralığıdır. Örnek:

12.00–13.00 UTC

6.2. Mikro-Slot

Katılımcının istemi göndermesi için atandığı daha kısa zaman aralığıdır. Örnek:

12.15–12.20 UTC

6.3. Tamamlama Ek Süresi

İstem slot içinde gönderilmişse cevabın tamamlanması için tanınan önceden belirlenmiş süredir.

6.4. Capture ve Yükleme Süresi

Cevap tamamlandıktan sonra kanıtın oluşturulması ve güvenli sisteme yüklenmesi için tanınan süredir.

7. ADAY SENKRONİZE DALGA TEMEL KURALI

CANDIDATE BASE RULE — ADAY TEMEL KURAL

Genel tüketici sohbet ürünleri için ana senkronize dalga aday olarak şu yapıyı kullanabilir:

- Gönderim penceresi: 60 dakika
- Mikro-slotlar: 12 adet × 5 dakika
- Katılımcı gönderimi: kendisine atanan mikro-slot içinde
- Cevap tamamlanma ek süresi: gönderimden sonra en fazla 20 dakika
- Capture hash üretimi: cevap tamamlandıktan sonra tercihen 5 dakika içinde
- Kanıt yükleme: capture sonrasında en fazla 60 dakika içinde
- Teknik sorun bildirme: aynı dalga içinde

Bu süreler evrensel ve nihai değildir. Şu koşullara göre değişebilir:

- AI ürününün cevap süresi
- Mobil bağlantı
- Ülke altyapısı
- Uzun cevap biçimi
- Erişilebilirlik ihtiyacı
- Ürün kullanım limiti

Kullanılan süreler veri toplamadan önce kilitlenmelidir.

8. NEDEN MİKRO-SLOTLAR?

Bin katılımcının aynı anda gönderim yapması:

- AI ürününde olağan dışı yük,
- hız limiti,
- kuyruk,
- teknik hata,
- cevap süresi artışı

oluşturabilir. Bu durumda denetim ürünün normal davranışını değil:

denetimin yarattığı yapay yük olayını

ölçebilir. Mikro-slotlar:

- yükü dağıtır,
- aynı genel sistem zamanını korur,
- gönderim sırasını sonuçtan bağımsız yapar,

zaman etkisinin modellenmesine izin verir.

9. SLOT ATAMASI

Katılımcılar slotlara:

- ülke,
- dil,

- AI ürünü,
- plan,
- panel durumu

bakımından dengeli ve önceden rastgeleleştirilmiş biçimde atanmalıdır. Bir mikro-slot yalnız:

- belirli ülke,
- belirli dil,
- belirli ürün

kullanıcılarıyla dolmamalıdır. Aksi hâlde zaman etkisi grup etkisiyle karışır.

10. YEREL SAAT ADALETİ

Tek UTC penceresi bazı katılımcılar için:

- gece,
- çalışma saati,
- ibadet veya dinlenme zamanı,
- ulaşılması zor bir saat

olabilir. Bu durum:

- katılım oranını,
- bağlantı koşulunu,
- kullanıcı dikkatini,
- cihaz türünü

etkileyebilir. İki tamamlayıcı yöntem kullanılabilir.

10.1. Dönen Küresel Dalgalar

Ana dalgalar farklı UTC saatlerinde yürütülür. Örnek:

- Dalga 1: 12.00 UTC
- Dalga 2: 20.00 UTC
- Dalga 3: 04.00 UTC

Böylece aynı bölge her dalgada aynı yerel saat yükünü taşımaz.

10.2. Yerel Saat Dengeli Alt Dalga

Katılımcılar kendi yerel saatlerinde benzer zaman bandında ölçülür. Bu çalışma:

- kullanıcı yorgunluğu,
- günlük kullanım davranışı

için değerlidir. Tek küresel sistem anı sonucu değildir.

11. ÜÇ DALGALI KARARLILIK MİMARİSİ

CANDIDATE ARCHITECTURE — ADAY MİMARİ

GEO-1000'in tam uygulaması için en az üç ana dalga önerilebilir:

W1 — Başlangıç Dalgası

İlk nüfus tahminini üretir.

W2 — Kısa Dönem Tekrarı

Ürünün ve ölçümün kısa dönem kararlılığını sınar. Aday aralık: W1'den 24–72 saat sonra

W3 — Orta Dönem Kararlılığı

Ürün ve varlık temsilinin daha uzun dönemde korunup korunmadığını ölçer. Aday aralık: W1'den 14–30 gün sonra Kesin aralıklar veri toplamadan önce belirlenmelidir. Ürün güncellemesi veya olay nedeniyle değiştirilecekse yeni sürüm kaydı gerekir.

12. 30.000 GÖZLEMLİ KURUCU TASARIM

On AI ürünü, ürün başına 1.000 ana gözlem ve üç ana dalga için:

```
10×1.000×3=30.000
```

geçerli ana Population Panel gözlemi hedeflenir. Tamamlayıcı:

- Country Sentinel
- Language Fairness
- Accessibility
- Incident
- Retest

panelleri bu sayının üzerine eklenebilir. 30.000:

- davet sayısı değildir,
- dosya sayısı değildir,

ekran görüntüsü sayısı değildir. Geçerli ana ürün-gözlemi hedefidir.

13. GÖZLEMİN ZAMAN VEKTÖRÜ

Her gözlem için zaman vektörü aday olarak şöyle tanımlanabilir:

```
T_i = (t_i^assign, t_i^open, t_i^submit, t_i^first, t_i^complete, t_i^capture, t_i^upload,
t_i^validate)
```

Burada:

- tiassign: görevin kullanıcıya atandığı zaman
- tiopen: görevin açıldığı zaman
- tisubmit: istemin gönderildiği zaman
- tifirst: ilk görünür sistem sonucunun başladığı zaman
- ticomplete: cevabın tamamlandığı zaman
- ticapture: kanıtın yakalandığı zaman
- tiupload: paketin güvenli sisteme ulaştığı zaman
- tivalidate: capture geçerliliğinin incelendiği zaman

Bu zamanların tamamı her kullanıcı yüzeyinde doğrudan gözlenemeyebilir. Bilinmeyen alanlar UNKNOWN kalmalıdır.

14. TEMEL SÜRE METRİKLERİ

Yanıt gecikmesi:

```
L_i = t_i^complete − t_i^submit
```

Capture gecikmesi:

```
C_i = t_i^capture − t_i^complete
```

Yükleme gecikmesi:

```
U_i = t_i^upload − t_i^capture
```

Slot sapması:

```
D_i = t_i^submit − t_i^{slot-midpoint}
```

olarak hesaplanabilir. Bu metrikler:

- sistem yükü,
- katılımcı gecikmesi,
- kanıt güveni,
- dalga bütünlüğü

bakımından değerlidir.

15. GECİKME SONUCA GÖRE YÖNETİLEMEZ

Uzun sürede tamamlanan bir cevap:

- doğru,
- yanlış,
- uzun,
- atıf'lı

olabilir. Gecikme nedeniyle dışlama kuralı:

- veri toplamadan önce,
- cevap içeriğinden bağımsız

tanımlanmalıdır. Şu davranış yasaktır: Olumsuz cevap geç tamamlandı diye dışlamak; olumlu cevap aynı gecikmeyle tamamlandığında korumak.

16. DALGA DURUM STATÜLERİ

WS-0 — PLANNED

Dalga planlanmıştır. Henüz kilitlenmemiştir.

WS-1 — LOCKED

Örneklem, istem, ürün, zaman ve capture kuralları kilitlenmiştir.

WS-2 — OPEN

Gönderim penceresi başlamıştır.

WS-3 — CLOSED

Yeni gönderim kabulü sona ermiştir. Tamamlama ve yükleme ek süreleri devam edebilir.

WS-4 — COMPLETED

Saha ve ilk capture doğrulamaları tamamlanmıştır.

WS-5 — SPLIT

Maddi sistem veya protokol değişikliği nedeniyle alt dalgalara ayrılmıştır.

WS-6 — PAUSED

Güvenlik, ürün veya altyapı sorunu nedeniyle geçici durdurulmuştur.

WS-7 — ABORTED

Dalga tamamlanmadan sonlandırılmıştır. Sonuçları sınırlı veya kullanılamaz olabilir.

WS-8 — INVALIDATED

Önceden tanımlanmış ağır protokol kusuru nedeniyle ana tahmin için geçersizdir. Ham gözlemler silinmez.

WS-9 — ARCHIVED

Sürümü kapanmış ve kanıt zinciri arşivlenmiştir.

17. DALGA SAĞLIK SINIFLARI

WH-0 — NOT ASSESSED

Dalga sağlığı değerlendirilmemiştir.

WH-1 — COMPROMISED

Maddi sistem değişikliği, zaman veya capture kusuru vardır. Ana sonuç için kullanılamayabilir.

WH-2 — PARTIAL

Bazı tabaka veya slotlar yetersizdir. Daraltılmış sonuç üretilebilir.

WH-3 — ACCEPTABLE

Ana zaman, sistem ve capture koşulları yeterince korunmuştur.

WH-4 — STRONG

Yüksek capture güveni, dengeli slot kapsamı ve düşük protokol kaybı vardır.

WH-5 — REPLICATED

Benzer dalga sağlığı birden fazla zaman ve bağımsız panelde tekrar edilmiştir. Dalga sağlığı, AI skorunun yüksek veya düşük olmasına bağlı değildir.

18. DALGA İÇİNDE ÜRÜN DEĞİŞİKLİĞİ

AI ürünü ölçüm penceresi içinde maddi biçimde değişebilir. Sinyaller:

- Model etiketi değişikliği
- Arayüz değişikliği
- Web özelliğinin açılması veya kapanması
- Atıf formatı değişikliği
- Cevap yapısında ani kırılma

- Sağlayıcı duyurusu
- Çok sayıda kullanıcıda aynı anda farklı sistem davranışı
- Ürün kesintisi veya yeniden başlatma

Bu durumda: Değişiklik zamanı tahmin edilir. Sistem konfigürasyon parmak izleri karşılaştırılır. Dalga WS-5 — SPLIT statüsüne alınabilir. Önceki ve sonraki gözlemler ayrı alt dalga olur. Tek skor üretilecekse sistem durumu ayrı faktör olarak modellenir. Maddi belirsizlik kamu kaydında gösterilir.

19. DIŞ OLAY KAYDI

AI ürünü değişmeden de dünya değişebilir. Ölçüm sırasında:

- şirket satın alması,
- ürün lansmanı,
- büyük haber,
- hukuki karar,
- kriz,
- finansal açıklama,
- viral içerik

oluşabilir. Bu olay:

- web kaynaklarını,
- retrieval sonuçlarını,
- kullanıcı istem yorumunu,
- AI cevabını

değiştirebilir. Her dalga ilgili olduğu ölçüde:

External Event Register

taşımalıdır.

20. DIŞ OLAY SONUÇLARI NASIL YÖNETİLİR?

Bir dış olay:

- dalgadan önce biliniyorsa referans paketi güncellenir,
- dalga sırasında oluşuyorsa zaman damgasıyla işaretlenir,

dalga sonrasında fark edilirse geriye dönük olay kaydı eklenir. Şu davranış yasaktır: Skoru düşüren haber nedeniyle gözlemleri dışlamak; skoru yükselten haber olduğunda korumak. Olayın kendisi gerçek dünya sistem durumunun parçası olabilir. Gerekirse alt dalga veya event-adjusted analiz yapılır.

21. İLK TEMAS SONUCU

Kullanıcının istemi gönderdikten sonra karşılaştığı ilk sistem sonucu:

First Contact Outcome

olarak adlandırılır. İlk temas sonucu:

- Tam cevap
- Ret

- Teknik hata
- Kullanım limiti
- Bağlantı kesintisi
- Sonsuz yükleme
- Clarification sorusu
- Sonuçsuzluk

olabilir. Bu olay gerçek kullanıcı deneyiminin parçasıdır. Sessizce silinemez.

22. İLK UYGUN İÇERİK ÇIKTISI

Teknik hata sonrasında önceden izin verilen yeniden denemede oluşan ilk maddi cevap:

First Eligible Content Output

olarak adlandırılabilir. Bu çıktı: ilk temas sonucunun yerine geçmez, onunla birlikte kaydedilir. Örnek:

- İlk temas: hız limiti
- Beş dakika sonra önceden izinli teknik yeniden deneme
- İlk içerik cevabı

Rapor şunları birlikte göstermelidir:

- İlk kullanıcı deneyimi: hız limiti
- İlk içerik temsili: belirtilen cevap

23. İLK UYGUN ÇIKTI KURALI

Bir AI ürünü ilk prompta maddi cevap ürettiyse:

- yanlış olsa,
- kısa olsa,
- ret içerse,
- atıf göstermese,
- markayı yanlış tanımlasa

yeniden üretim yapılamaz. Ana gözlem ilk çıktıdır. Kullanıcı:

- regenerate,
- rewrite,
- try again,
- başka cevap,
- devam et

gibi yollarla yeni cevap seçemez.

24. TEKNİK YENİDEN DENEME

Teknik yeniden deneme yalnız önceden tanımlanmış teknik koşullarda kullanılabilir. Aday izin nedenleri:

- Cevap başlamadan bağlantı kesintisi
- Sunucu hata kodu

- Arayüzün istemi göndermemesi
- Uygulamanın çökmesi
- Doğrulanmış geçici sistem kesintisi

Teknik yeniden deneme nedeni olamaz:

- Yanlış cevap
- Ret
- Kısa cevap
- Atıf eksikliği
- Marka anılmaması
- Critical hata
- Denetlenen kurumun cevabı beğenmemesi

25. KISMİ CEVAPTA YENİDEN DENEME

Sistem maddi içeriğin bir bölümünü üretmiş ve sonra durmuşsa:

- kısmi çıktı saklanır,
- TRUNCATED_RESPONSE olarak kaydedilir,

yeniden deneme yapılacaksa yeni ve bağlantılı girişim oluşturulur. Kısmi cevap yokmuş gibi silinemez.

26. GİRİŞİM ZİNCİRİ

Bir gözlem ataması birden fazla teknik girişim taşıyabilir:

```
A_i → (Attempt_{i,1}, Attempt_{i,2}, …)
```

Her girişim:

- zaman,
- ürün durumu,
- istem hash'i,
- sistem sonucu,
- teknik hata nedeni

taşır. Bir girişim diğerinin üzerine yazılmaz.

27. NOMOS CAPTURE NEDİR?

NOMOS Capture, kullanıcıya verilen istemden kanıt paketinin arşivlenmesine kadar gözlemin bütünlüğünü koruyan açık bir veri ve prosedür standardıdır. NOMOS Capture:

- yalnız bir ekran görüntüsü uygulaması değildir,
- yalnız NobleJackal'a ait kapalı bir yazılım olmak zorunda değildir,

yalnız tek bir tarayıcı uzantısına bağlı değildir. Uyumlu bir NOMOS Capture uygulaması:

- açık veri şemasını,
- zaman ve bütünlük kurallarını,
- mahremiyet sınırlarını,
- kanıt zincirini

karşılamalıdır. Bağımsız araştırmacılar kendi uyumlu araçlarını geliştirebilmelidir. Standart yalnız kurucunun yazılımıyla uygulanabiliyorsa bağımsız tekrar zayıflar.

28. NOMOS CAPTURE'IN DÖRT İŞLEVİ

28.1. Görevi Teslim Etmek

Doğru:

- kullanıcı,
- istem,
- AI ürünü,
- slot

eşleşmesini sağlar.

28.2. Gözlemi Yakalamak

İstemi, ilk sistem sonucunu, cevabı ve ilgili kullanıcı yüzeyini kaydeder.

28.3. Bütünlüğü Korumak

Dosya hash'i, zaman damgası ve paket ilişkisini oluşturur.

28.4. Kanıtı Taşımak

Ham, kısıtlı ve kamuya açık kanıt katmanlarını birbirine bağlar.

29. CAPTURE YÖNTEMLERİ

CM-1 — GUIDED MANUAL CAPTURE

Katılımcı yönergeyle ekran görüntüsü ve ham cevap metni sağlar. Düşük teknik yük taşır. İnsan hatasına açıktır.

CM-2 — ASSISTED CAPTURE APPLICATION

Uygulama:

- istem kopyalama,
- zaman kaydı,
- dosya yükleme,
- hash üretimi

işlevlerini destekler. AI ürününe müdahale etmez.

CM-3 — INSTRUMENTED USER-SURFACE CAPTURE

Tarayıcı veya uygulama entegrasyonu:

- gönderilen istemi,
- ürün metadata'sını,
- cevap metnini,
- zamanları

doğrudan yakalar. Yüksek bütünlük sağlar. Kullanım koşulları ve mahremiyet değerlendirilmelidir.

CM-4 — PROVIDER EXPORT OR TELEMETRY CAPTURE

Sağlayıcının oturum export'u veya doğrulanmış telemetry'si kullanılır. Bağımsız capture'a güçlü destek olabilir. Tek başına bağımsızlık sağlamayabilir.

CM-5 — CONTROLLED API OR LAB CAPTURE

API veya dondurulmuş sistem logları kullanılır. Canlı kullanıcı yüzeyi gözlemi değildir. Ayrı laboratuvar paneline aittir.

30. OTOMASYONUN SINIRI

Otomasyon:

- istem teslimi,
- zaman kaydı,
- hash üretimi,
- dosya bütünlüğü

için kullanılabilir. Katılımcının yerine:

- bot hesabı,
- otomatik toplu sorgu,
- kullanıcı olmayan oturum

üretirse gerçek insan Population Panel'i oluşturmaz. Otomatik sistem denetimi ayrı:

CONTROLLED_AUTOMATION_PANEL

olarak kaydedilmelidir.

31. CAPTURE GÜVEN SEVİYELERİ

NCL-0 — NO VERIFIABLE CAPTURE

Yalnız katılımcı beyanı vardır. Ana GEO-1000 sonucuna alınamaz.

NCL-1 — PARTIAL EVIDENCE

Cevap metni veya sınırlı ekran görüntüsü vardır. İstem veya sistem koşulu tam doğrulanamaz.

NCL-2 — SCREEN-EVIDENCED

İstem, cevap ve temel ürün yüzeyi ekran görüntüsüyle doğrulanmıştır. Metadata ve zaman güveni sınırlı olabilir.

NCL-3 — FULL EVIDENCE BUNDLE

Tam cevap, istem, ürün kaydı, zamanlar, ekran görüntüsü ve hash bulunmaktadır. Ana analiz için aday asgari düzeydir.

NCL-4 — INSTRUMENTED AND SIGNED CAPTURE

Capture aracı istem, cevap, zaman ve metadata'yı doğrudan kaydetmiş; paket dijital olarak imzalanmıştır.

NCL-5 — INDEPENDENTLY VERIFIED MULTI-SOURCE CAPTURE

Capture paketi:

- bağımsız doğrulama,
- sağlayıcı export'u,
- ikinci kanıt kaynağı

ile desteklenmiştir. Ana GEO-1000 gerçek kullanıcı gözlemleri en azından:

NCL-3

düzeyini hedeflemelidir.

32. NOMOS CAPTURE KANIT PAKETİ

Her gözlem için paket aday olarak şu bileşenleri taşır:

32.1. Observation Manifest

Gözlemin bütün kimlik ve durum alanlarını içerir.

32.2. İstem Artifact

Atanan ve gönderilen tam istem metni, sürümü ve hash'i.

32.3. Raw Yanıt Artifact

AI ürününün ürettiği tam ham cevap.

32.4. Visual Evidence Artifact

İstem, cevap ve ürün yüzeyini gösteren görüntü veya ardışık görüntüler.

32.5. System-State Artifact

Plan, yüzey, model etiketi, hafıza, web ve diğer maddi sistem durumlarını gösterir.

32.6. Time Evidence Artifact

Gönderim, tamamlanma, capture ve yükleme zamanlarını taşır.

32.7. Attempt Log

Teknik hata ve yeniden deneme zincirini gösterir.

32.8. Integrity Manifest

Bütün dosyaların hash ve boyutlarını içerir.

32.9. Privacy Manifest

Redaksiyon, erişim ve saklama durumunu gösterir.

32.10. Validation Record

Capture geçerliliğinin kim tarafından ve ne zaman değerlendirildiğini gösterir.

33. EKRAN GÖRÜNTÜSÜNÜN ASGARİ İÇERİĞİ

İlgili olduğu ölçüde görsel kanıt şunları göstermelidir:

- AI ürün veya yüzey kimliği
- Tam istem
- Cevabın başlangıcı
- Cevabın sonu
- Atıf alanları
- Ret veya hata mesajı
- Görünen model etiketi
- Oturumun yeni veya ilgili durumda olduğuna dair kanıt
- Maddi sistem ayarı
- Cevabın kesilmediğini gösteren arayüz durumu

34. UZUN CEVAPLAR

Cevap tek ekrana sığmıyorsa:

- tam sayfa capture,
- kaydırmalı capture,
- üst üste binen ardışık ekran görüntüleri,
- ekran kaydı,
- doğrudan metin export'u

kullanılabilir. Ardışık görseller:

- boşluk bırakmamalı,
- birbirleriyle örtüşmeli,

sıra numarası taşımalıdır. Yalnız olumlu veya ilgili bölüm kırpılamaz.

35. EKRAN KAYDI

Ekran kaydı:

- istem gönderimini,
- streaming cevabı,
- tamamlanma durumunu,
- yenileme yapılmadığını

gösterebilir. Ancak:

- daha fazla kişisel bilgi,
- bildirim,
- özel uygulama,
- hesap kimliği

yakalayabilir. Bu nedenle ekran kaydı:

- varsayılan zorunluluk olmayabilir,
- yüksek riskli veya doğrulama alt örneğinde kullanılabilir,

kayıt öncesi gizlilik hazırlığı gerektirir.

36. HAM CEVAP METNİ

Ekran görüntüsüne ek olarak tam cevap metni ayrı saklanmalıdır. Tercih sırası:

- Ürünün resmî export veya kopyalama özelliği
- Capture aracının doğrudan metin kaydı
- Erişilebilirlik ağacı veya DOM tabanlı extraction
- Kullanıcı tarafından kopyalanmış metin
- OCR ve insan doğrulaması

OCR yalnız son seçenek olmalıdır. Otomatik extraction ile görsel arasında maddi fark varsa inceleme gerekir.

37. DİNAMİK CEVAPLAR

Bazı ürünler cevap tamamlandıktan sonra:

- atıf ekleyebilir,
- kaynak kartı yükleyebilir,
- metni yeniden biçimlendirebilir,

güvenlik uyarısı ekleyebilir. İki capture anı kullanılabilir:

- İlk tamamlanma capture'ı
- Kısa render kararlılığı sonrasındaki capture

Maddi fark varsa iki sürüm de korunur. İlk kullanıcı maruziyeti silinmez.

38. CEVABIN TAMAMLANDIĞI NASIL ANLAŞILIR?

Aday tamamlanma işaretleri:

- Streaming'in durması
- Yükleme göstergesinin kapanması
- Regenerate veya eşdeğer düğmenin görünmesi
- Ürünün tamamlanma sinyali
- Önceden tanımlanmış sessizlik süresi
- API completion kaydı

Ürün bazlı tamamlanma kuralı AI Sistem Sicili'nde bulunmalıdır.

39. CEVABI KULLANICI DURDURDUYSA

Kullanıcı yanlışlıkla veya bilerek:

- stop,
- cancel,
- uygulamayı kapatma

işlemi yapmış olabilir. Bu durum:

USER_INTERRUPTED

ACCIDENTAL_INTERRUPTION

UNKNOWN_INTERRUPTION

olarak kaydedilir. Kullanıcı tarafından durdurulmuş cevap tam sistem çıktısı sayılamaz. Katılımcı değiştirme veya yeniden deneme kuralı önceden tanımlanmalıdır.

40. İSTEM İLE CEVABIN AYRI DOSYALARDA OLMASI

İstem ve cevap ayrı görsellerdeyse:

- aynı oturuma,
- aynı gözlem kimliğine,
- kesintisiz zaman zincirine

bağlanmalıdır. Aksi hâlde istem ile cevabın gerçekten aynı konuşmaya ait olduğu doğrulanamayabilir.

41. DOSYA BÜTÜNLÜĞÜ

Her dosya için kriptografik hash üretilmelidir. [K25] Aday:

```
h_j = H(file_j)
```

Paket hash'i:

```
h_bundle = H(JCS(manifest(file_id, order, size, algorithm, h_j))) [K14; K25]
```

olarak oluşturulabilir. Burada H, yayımlanmış ve güvenli kabul edilen hash algoritmasıdır. Algoritma:

- sürümlü,
- açık,
- gerektiğinde gelecekte geçiş yapılabilir

olmalıdır.

42. HASH NEYİ KANITLAR?

Kaydedilmiş hash değeri güvenilir biçimde saklandıysa, aynı dosyanın aynı algoritmayla yeniden hash'lenmesi byte dizisindeki değişikliği tespit etmeye yarar [K25]. Bu karşılaştırma, dosyanın referans hash üretildikten sonra aynı byte dizisini koruduğu iddiasını destekler; dosyanın kaynağını, doğru zamanda üretildiğini veya hash öncesi özgünlüğünü kanıtlamaz. Hash tek başına şunları kanıtlamaz:

- Dosyanın gerçek AI ürününden geldiğini
- Doğru kullanıcıya ait olduğunu
- Doğru zamanda üretildiğini
- Hash öncesi düzenlenmediğini
- İlk çıktı olduğunu
- İstemin doğru olduğunu

Bu nedenle hash:

> *bütünlük kanıtıdır.*

Tek başına:

> *özgünlük kanıtı değildir.*

43. DİJİTAL İMZA

Capture aracı veya yetkili doğrulayıcı:

- gözlem paketini,
- zaman kaydını,
- araç sürümünü

dijital olarak imzalayabilir. İmza:

- paketin hangi araç veya kurum tarafından üretildiğini,
- imzadan sonra değişmediğini

güçlendirir. İmza sahibi yanlış veya sahte veri imzalamış olabilir. Bu nedenle imza da tek başına gerçeklik garantisi değildir.

44. GÜVENİLİR ZAMAN [K26]

Yerel cihaz saati:

- yanlış ayarlanmış,
- elle değiştirilmiş,
- senkronizasyon dışı

olabilir. Güçlü zaman kaydı birden fazla sinyal taşımalıdır:

- Capture sunucusu kabul zamanı
- Güvenilir UTC zaman servisi
- Cihaz saati ve sapma kaydı
- AI ürünündeki görünür zaman, varsa
- Dosya oluşturma zamanı
- Dijital zaman damgası

45. SAAT SAPMASI

Katılımcının cihaz saati ile güvenilir UTC zamanı arasındaki fark:

```
O_i^clock = t_i^device − t_i^reference
```

olarak kaydedilebilir. Maddi sapma varsa:

- sunucu zamanı ana referans olarak kullanılır,
- cihaz zamanı düzeltilmiş kayıt olarak korunur,

belirsizlik işaretlenir. Katılımcıdan cihaz saatini elle değiştirmesi istenmemelidir.

46. YÜKLEME ZAMANI CAPTURE ZAMANI DEĞİLDİR

Bir dosyanın sunucuya saat 14.00'te yüklenmesi, ekran görüntüsünün saat 14.00'te alındığını kanıtlamaz. Bu nedenle:

- capture zamanı,

- hash üretim zamanı,
- yükleme zamanı

ayrı kaydedilmelidir. Uzun yükleme gecikmesi capture güvenini azaltabilir. Kayıt otomatik olarak geçersiz olmak zorunda değildir.

47. GECİKMELİ YÜKLEME

Bağlantı sorunu nedeniyle kullanıcı kanıtı hemen yükleyemeyebilir. Geç yükleme kabul edilecekse:

- capture anında yerel hash,
- güvenilir cihaz zamanı,
- çevrim dışı imzalı kayıt,
- yükleme gecikmesi nedeni

gerekebilir. Yalnız kullanıcının: "Bu görseli dün aldım." beyanı ana NCL-3 düzeyi için yeterli değildir.

48. KANIT ZİNCİRİ

Bir gözlemin kanıt zinciri aday olarak şu aşamalardan oluşur:

- Görev ataması
- İstem kilidi
- Kullanıcı ve ürün uygunluğu
- Slot açılışı
- İstem gönderimi
- İlk temas sonucu
- Cevabın tamamlanması
- Capture
- Yerel bütünlük kaydı
- Güvenli yükleme
- Sunucu kabul zamanı
- Paket hash'i
- Protokol doğrulaması
- Redaksiyon
- Hakemlik
- Arşiv
- Kamu manifesti

Her aşama önceki aşamaya bağlanmalıdır.

49. ZİNCİRİN KIRILMASI

Kanıt zinciri şu durumlarda kırılabilir:

- İstem kaydı yok
- İlk çıktı doğrulanamıyor
- Cevap ile görsel uyuşmuyor

- Capture zamanı bilinmiyor
- Dosya sonradan değiştirilmiş
- Katılımcı kimliği tekrar
- Ürün ve plan bilinmiyor
- Gözlem dalga dışında
- Hash ve dosya eşleşmiyor
- Redaksiyon ham dosyanın üzerine yazılmış
- Kim tarafından doğrulandığı bilinmiyor

Kırılmanın türü geçerlilik statüsünü belirler.

50. HAM KANIT DEĞİŞTİRİLEMEZ; KARAR DEĞİŞEBİLİR

Ham kanıt paketi:

- üzerine yazılmamalı,
- sessizce düzeltilmemeli,

silinip yenisi konmamalıdır. Fakat gözlem hakkındaki:

- geçerlilik,
- hakemlik,
- sınıflandırma,
- uygunluk

kararı değişebilir. Yeni karar:

- yeni sürüm,
- gerekçe,
- tarih,
- karar sahibi

taşımalıdır. Bu ayrım temel yönetişim kuralıdır:

> *Kanıt sabit kalabilir; yorum kanıtla birlikte gelişebilir.*

51. REDAKSİYON

Kamu veya hakem kopyasında şu alanlar kapatılabilir:

- Ad ve soyadı
- E-posta
- Profil fotoğrafı
- Hesap kimliği
- Özel konuşmalar
- Bildirimler
- Kurumsal alan
- Hassas dosya adı
- Kesin konum

Redaksiyon:

- ham dosya üzerinde yapılmamalı,

- yeni türev dosya üretmeli,
- kendi hash'ini taşımalı,

hangi alanın neden kapatıldığını kaydetmelidir.

52. ÜÇ KANIT KATMANI

52.1. Raw Restricted Evidence

Tam ham kanıttır. Yalnız yetkili denetçiler ve gerekli araştırma rolleri erişebilir.

52.2. Audit Evidence Layer

Bağımsız hakem ve yeniden üretim ekipleri için gerekli kanıtları içerir. Kişisel alanlar minimize edilir.

52.3. Public Evidence Manifest

Kamuya:

- gözlem kimliği,
- ülke ve dil hücresi,
- AI ürünü,
- zaman,
- kanıt hash'i,
- geçerlilik durumu,
- redaksiyon statüsü

gibi bilgiler sunar. Ham özel içerik zorunlu olarak yayımlanmaz.

53. KAMU KANIT MANİFESTİ

Her dalga için kamu manifesti şu alanları taşıyabilir:

- Dalga kimliği
- Protokol sürümü
- Ürün sayısı
- Hedef ve geçerli gözlem sayısı
- Ülke ve dil kapsamı
- Slot kapsamı
- Capture güven seviyeleri
- Teknik hata ve ret oranları
- Geçersiz gözlem sayısı
- Gözlem paket hash'leri
- Dalga bütünlük kökü
- Değişiklik olayları
- Kamuya açık redakte örnekler
- Bağımsız erişim prosedürü
- Kanıt saklama durumu

54. DALGA BÜTÜNLÜK KÖKÜ

CANDIDATE CONCEPT — ADAY KAVRAM

Her gözlem paketinin hash'i: hi olsun. Bunlardan Merkle ağacı veya eşdeğer yapı ile dalga kökü:

RW

üretilebilir. Dalga kökü ölçüm kapandıktan sonra yayımlanabilir. Bu yöntem:

- dalgadaki gözlem listesinin sonradan sessizce değiştirilmesini,
- yeni gözlem eklenmesini,
- mevcut gözlemin değiştirilmesini

daha görünür hâle getirir. Merkle kökü:

- gözlemlerin gerçek olduğunu,
- hakemlik kararının doğru olduğunu

tek başına kanıtlamaz. Liste bütünlüğünü güçlendirir.

55. GÖZLEM EKLEME VE ÇIKARMA

Dalga kapandıktan sonra:

- yeni gözlem eklenmesi,
- geçerli gözlemin çıkarılması,
- dosyanın değiştirilmesi

yeni manifest sürümü gerektirir. Eski bütünlük kökü korunur. Çıkarılan gözlem için:

- neden,
- tarih,
- karar sahibi,
- eski ve yeni skor etkisi

kaydedilir.

56. KATILIMCI ÇEKİLME TALEBİ

Araştırma etiği veya veri koruma çerçevesi katılımcının verisini geri çekmesine izin verebilir. Bu durumda:

- ham kanıt güvenli biçimde silinebilir veya erişim dışı bırakılabilir,
- kamu manifestinde içeriksiz bir WITHDRAWN kaydı korunabilir,
- dalga skoru yeni sürümle yeniden hesaplanabilir,

eski sonuç ve değişiklik nedeni saklanır. Değiştirilemezlik, insanın haklarını ortadan kaldırma anlamına gelmez.

57. GÖZLEM CAPTURE GEÇERLİLİK STATÜLERİ

CV-0 — NOT REVIEWED

Capture henüz değerlendirilmemiştir.

CV-1 — VALID

Bütün zorunlu capture ve zaman koşullarını karşılar.

CV-2 — CONDITIONALLY VALID

Sınırlı eksiklik vardır. Hangi analizlerde kullanılabileceği belirtilir.

CV-3 — PARTIAL EVIDENCE

Gözlem olay olarak kullanılabilir. Ana nüfus skoruna giremeyebilir.

CV-4 — OUTSIDE WAVE

İstem izin verilen zaman yapısı dışında gönderilmiştir.

CV-5 — İSTEM OR OUTPUT MISMATCH

Atanan istem, gönderilen istem veya kayıtlı cevap arasında maddi uyuşmazlık vardır.

CV-6 — DUPLICATE

Aynı gözlem veya kullanıcı kaydı tekrar edilmiştir.

CV-7 — TAMPER SUSPECTED

Dosya veya metadata bütünlüğü hakkında çözülememiş şüphe vardır.

CV-8 — FABRICATED OR MANIPULATED

Gözlem gerçek kullanıcı deneyimi değildir veya bilinçli biçimde değiştirilmiştir.

CV-9 — WITHDRAWN

Katılımcı veya etik süreç nedeniyle analizden çıkarılmıştır.

58. CAPTURE GEÇERLİLİĞİ İLE SEMANTİK DOĞRULUK AYRILMALIDIR

Bir gözlem:

- capture bakımından tamamen geçerli,
- içerik bakımından ağır yanlış

olabilir. Başka gözlem:

- doğru cevap,
- eksik veya manipüle edilmiş kanıt

taşıyabilir. Bu nedenle iki ayrı karar vardır:

- Capture Validity
- Semantic Adjudication

Capture doğrulayıcısı mümkün olduğunca yanıtın başarı puanını bilmemelidir.

59. CAPTURE DOĞRULAYICI ROLÜ

Capture doğrulayıcı şunları inceler:

- İstem eşleşmesi
- Dalga ve slot
- Ürün kimliği
- İlk çıktı kuralı
- Ekran bütünlüğü
- Ham metin uyumu
- Teknik yeniden deneme
- Hash ve zaman
- Mahremiyet

Şunlara karar vermez: Şirket doğru anlatıldı mı? Atıf iddiayı destekliyor mu? Critical hata var mı? Yanıt geçer mi? Bunlar sonraki hakemlik katmanının görevidir.

60. YİNELENEN EKRAN GÖRÜNTÜSÜ

Aynı görsel:

- yanlışlıkla iki kez yüklenmiş,
- birden fazla kullanıcı tarafından kullanılmış,
- panel sağlayıcılar arasında tekrar edilmiş

olabilir. Kontroller:

- Dosya hash'i
- Görsel benzerliği
- Ham cevap metni
- Zaman ve oturum metadata'sı
- Kullanıcı benzersizliği

üzerinden yapılabilir. Aynı AI ürünü bazı kullanıcılara kelime kelime aynı cevap verebilir. Aynı metin tek başına sahtecilik kanıtı değildir. Aynı dosya veya aynı görsel izi daha güçlü sinyaldir.

61. DOSYA BENZERLİĞİ VE YANLIŞ POZİTİF

Aynı cihaz ve arayüz:

- benzer görseller,
- aynı ekran boyutu,
- aynı sıkıştırma

üretebilir. Görsel benzerlik algoritması otomatik dışlama yapmamalıdır. Çoklu sinyal ve insan incelemesi gerekir.

62. SAĞLAYICI TELEMETRY'Sİ

AI sağlayıcısı:

- oturum zamanı,
- model,
- istem,
- cevap,
- araç kullanımı

hakkında güvenilir log sağlayabilir. Bu kayıt bağımsız capture'ı güçlendirir. Ancak sağlayıcı:

- ölçülen taraf,
- sistem sahibi,
- çıkar sahibi

olabilir. Sağlayıcı telemetry'si: bağımsız kullanıcı kanıtının yerine değil, onunla birlikte kullanılmalıdır.

63. SAĞLAYICI KANITI YOKSA

Canlı tüketici ürünlerinin çoğunda sağlayıcı loglarına erişim olmayabilir. Bu durum GEO-1000'i imkânsızlaştırmaz. Aşağıdaki bileşim kullanılabilir:

- Enstrümante kullanıcı capture'ı
- Sunucu zaman damgası
- Tam ekran kanıtı
- Ham cevap
- Sistem ayarı
- Hash ve imza
- Bağımsız doğrulama alt örneği

Kanıt güven seviyesi buna göre belirtilir.

64. DOĞAL KULLANICI PANELİNDE CAPTURE

Doğal panelde capture aracı:

- özel talimat içeriğini,
- önceki konuşmaları,
- hesap kimliğini

gereksiz yere yakalamamalıdır. Tercih edilen akış: Kullanıcı yeni doğal hesap sohbeti açar. Capture yalnız görev penceresini kaydeder. Kişiselleştirme durumu ayrı yüksek düzey metadata olarak alınır. Özel geçmiş capture dışında kalır. Kamu kopyası redakte edilir.

65. NATURAL CONVERSATION-STATE CAPTURE

Mevcut konuşma bağlamı ölçülüyorsa:

- yalnız gerekli önceki tur aralığı,
- açık rıza,

- güçlü redaksiyon,
- erişim kısıtı

gerekir. Bütün özel sohbet geçmişi varsayılan olarak toplanamaz.

66. MOBİL CAPTURE

Mobil cihazlarda:

- tam ekran görüntüsü,
- kaydırmalı capture,
- uygulama sürümü,
- sistem saat kaydı

farklı çalışabilir. Mobil protokol:

- iOS,
- Android,
- diğer sistemler

için ayrı teknik rehber taşıyabilir. Mobil kullanıcılar masaüstü capture zor olduğu için dışlanamaz. Capture aracı nüfusa uyum sağlamalıdır.

67. ERİŞİLEBİLİRLİK VE CAPTURE

Ekran okuyucu, sesli giriş veya motor yardımcı araç kullanan katılımcılar için yalnız standart görsel capture yeterli olmayabilir. Alternatif kanıtlar:

- Erişilebilirlik ağacı export'u
- Sesli oturum kaydı
- Uygulama transcript'i
- Yardımcı teknoloji logu
- Görev gözlem kaydı

olabilir. Bu gözlemler düşük güvenli sayılmamalıdır. Eşdeğer kanıt standardı geliştirilmelidir.

68. CAPTURE ARACININ GÖZLEMİ DEĞİŞTİRMESİ

Tarayıcı uzantısı veya enstrümantasyon:

- sayfa yükünü,
- arayüz davranışını,
- güvenlik sistemini,
- ürün tespitini

etkileyebilir. Capture aracının etkisi pilotta sınanmalıdır. Araçlı ve araçsız oturumlar farklı davranıyorsa: araç durumu deneysel faktör olur, capture aracı sessizce tarafsız kabul edilemez.

69. CAPTURE ARACI SÜRÜMLEME

Her capture paketi:

- araç adı,
- sürümü,
- yapılandırması,
- işletim ortamı,
- bilinen hataları

taşımalıdır. Capture aracı ölçüm sırasında değişirse:

- yeni araç sürümü,
- alt dalga,
- uyumluluk incelemesi

gerekebilir.

70. CAPTURE ARACI HATASI

Araç:

- istemi yanlış kopyalayabilir,
- cevabın sonunu kesebilir,
- zamanı yanlış kaydedebilir,
- hash üretmeyebilir,

dosyayı sıkıştırabilir. Araç hatası:

- AI ürününe,
- kullanıcıya

yüklenemez. Ayrı:

CAPTURE_TOOL_FAILURE

olarak kaydedilmelidir.

71. SENTETİK ÜÇ DALGALI GEO-1000 VAKASI

SENTETİK METODOLOJİ GÖSTERİMİ / Aşağıdaki ürünler, tarihler, sayılar ve sonuçlar tamamen kurgusaldır. On sentetik AI ürünü:

A1–A10

üç ana dalgada ölçülsün.

71.1. Dalga Takvimi

Dalga	Tür	UTC gönderim penceresi	Amaç
W1	Principal	12.00–13.00	Başlangıç
W2	Confirmatory	20.00–21.00	Kısa dönem tekrar
W3	Stability	04.00–05.00	Orta dönem ve saat dengesi

Her ürün için her dalgada: 1.000 geçerli ana gözlem hedeflenmektedir. Toplam hedef:

```
10×1.000×3=30.000
```

71.2. Slot Yapısı

Her dalga:

- 12 mikro-slot
- slot başına ürün ve ülke bakımından dengeli katılımcılar

taşımaktadır. Katılımcının slotu: veri toplamadan önce rastgele atanmıştır.

71.3. W1 Sonucu

Hedef: 10.000 geçerli ana gözlem Saha:

- 13.240 davet
- 10.870 tamamlanan görev
- 10.106 capture bakımından geçerli gözlem
- Önceden kilitli ana/yedek sırasına göre 10.000 ana analiz gözlemi

Kalan 106 geçerli kayıt:

- yedek,
- duyarlılık,
- capture kalite kontrolü

için korunmuştur.

71.4. Teknik Olaylar

118 kullanım limiti 42 bağlantı hatası 31 sistem reddi 17 kullanıcı kesintisi 9 istem mismatch 6 tekrar ekran görüntüsü İlk temas sonuçlarının tamamı saklanmıştır. Yalnız önceden tanımlı gerçek teknik hatalarda yeniden deneme yapılmıştır.

71.5. W2 Sırasında Ürün Değişikliği

A7 ürünü 20.32 UTC'de:

- otomatik web kullanımı,
- yeni atıf görünümü

kazansın. A7 gözlemleri:

W2A: 20.00–20.31

W2B: 20.32–21.00

olarak ayrılır. A7 için tek sabit W2 skoru açıklamasız biçimde yayımlanmaz.

71.6. W3'te Dış Olay

Denetlenen sentetik şirket hakkında 04.22 UTC'de önemli bir hukuki karar yayımlansın. Web erişimli ürünlerde cevapların 04.25'ten sonra değiştiği gözlensin. Olay: External Event Register'a eklenir. Olay öncesi ve sonrası sonuçlar ayrı tanısal hücre olarak gösterilir. Gözlemler skoru düşürdüğü veya yükselttiği için silinmez.

71.7. Dalga Sağlık Kartı

Dalga	Sağlık	Ana not
W1	WH-4	Dengeli ve güçlü capture
W2	WH-3	A7 ürün değişikliği nedeniyle alt dalga

Dalga	Sağlık	Ana not
W3	WH-3	Harici hukuki olay kaydı

Üç dalga da kullanılabilir. A7 ve dış olay yorumları ayrıca sınırlandırılır.

72. SENTETİK TEK GÖZLEM CAPTURE VAKASI

SENTETİK ÖRNEK

Katılımcı:

- Türkiye
- Türkçe
- Orion AI web ücretli plan
- Kontrollü Temiz Panel

W1

Slot 12.15–12.20 UTC Atanan istem: “Apple.com hangi ana kuruluşla ilişkilidir ve bu kuruluşun başlıca faaliyetleri nelerdir?”

72.1. Zamanlar

Görev açılışı: 12.14.20 İstem gönderimi: 12.16.08 İlk token: 12.16.10 Cevap tamamlanması: 12.16.24 Capture: 12.16.39 Hash üretimi: 12.16.41 Yükleme: 12.17.12 Sunucu kabulü: 12.17.13

72.2. Kanıtlar

İstem hash’i eşleşiyor Tam cevap metni mevcut İki örtüşen ekran görüntüsü mevcut Ürün, plan ve model etiketi görünür Hafıza ve özel talimat kapalı Web kullanımı açık Atıf gösterilmemiş Yeniden üretim yok Önceki konuşma yok Dosya hash’leri ve paket imzası mevcut Capture seviyesi:

NCL-4

Geçerlilik:

CV-1

Cevabın doğru veya yanlış olduğuna henüz karar verilmemiştir.

73. SENTETİK TEKNİK YENİDEN DENEME VAKASI

İlk girişimde:

- İstem doğru gönderilmiş
- Ürün “Network error” göstermiş
- Maddi cevap başlamamış
- Hata ekran görüntüsü alınmış
- Zaman ve hash kaydedilmiş

Önceden tanımlı kurala göre beş dakika içinde ikinci girişim yapılmış. İkinci girişimde cevap oluşmuş. Doğru kayıt:

- First Contact Outcome: TEMPORARY_TECHNICAL_FAILURE

- First Eligible Content Output: ikinci girişim
- İki girişim de pakette
- Teknik hata oranına ilk olay dâhil
- İçerik değerlendirmesine ikinci cevap dâhil, yöntem bunu önceden öngörüyorsa

Yanlış kayıt: İlk hata silinir ve yalnız ikinci cevap varmış gibi davranılır.

74. SENTETİK YANLIŞ YENİDEN DENEME VAKASI

İlk cevap: "Apple yalnızca cep telefonu satan yerel bir şirkettir." Katılımcı cevabı beğenmez. Yeniden üretir. İkinci cevap doğru çıkar. Katılımcı yalnız ikinci cevabı yükler. Sistem logu veya capture kaydı yenilemeyi gösterir. Doğru statü:

CV-8 — MANIPULATED OBSERVATION

İkinci cevabın doğru olması gözlemi kurtarmaz. İlk cevap ayrıca geri kazanılabiliyorsa olay incelemesinde korunabilir. Ana katılımcı kaydı geçersizdir.

75. DALGA KAPATMA RAPORU

Her dalga sonunda şu kayıt oluşturulmalıdır:

- Dalga kimliği
- Açılış ve kapanış zamanı
- Hedef gözlem sayısı
- Davet edilen kişi
- Tamamlanan görev
- Geçerli capture
- Ana analize giren gözlem
- Yedek geçerli gözlem
- Teknik hata
- Ret
- Kullanım limiti
- İstem mismatch
- Capture mismatch
- Geç yükleme
- Duplicate
- Tamper şüphesi
- Katılımcı geri çekme
- Ürün değişiklikleri
- Dış olaylar
- Slot kapsamı
- Ülke ve dil kapsamı
- Capture güven düzeyleri
- Dalga sağlık statüsü
- Bütünlük kökü

- Sorumlu insan onayı

76. ZORUNLU NORMATİF HÜKÜMLER

CH12-N01

Her GEO-1000 gözlemi önceden kilitlenmiş bir ölçüm dalgasına bağlanmalıdır.

CH12-N02

"Aynı zaman" ifadesi tam UTC penceresi, mikro-slot ve tamamlama kuralları olmadan kullanılamaz.

CH12-N03

Yerel saat tek başına dalga zamanı olarak kullanılamaz; ana zaman standardı UTC olmalıdır.

CH12-N04

Gönderim penceresi, mikro-slotlar, tamamlama süresi ve yükleme süresi AI yanıtları görülmeden önce belirlenmelidir.

CH12-N05

Slot ataması ülke, dil, ürün ve panel gruplarını zaman bakımından dengesiz bırakmayacak biçimde yapılmalıdır.

CH12-N06

Katılımcılar sonuçlarına göre daha erken veya daha geç slotlara taşınamaz.

CH12-N07

Tek UTC penceresinin yerel saat yanlılığı kamu yöntem kaydında açıklanmalıdır.

CH12-N08

Dönen küresel dalgalar veya yerel saat dengeli çalışmalar ayrı zaman sonuçları olarak tutulmalıdır.

CH12-N09

Her gözlem istem gönderim, ilk temas, tamamlanma, capture ve yükleme zamanlarını ilgili olduğu ölçüde taşımalıdır.

CH12-N10

Bilinmeyen zaman alanları tahmin edilmiş kesin zaman gibi sunulamaz.

CH12-N11

Cihaz saati tek başına güvenilir UTC zaman kanıtı sayılamaz.

CH12-N12

Capture zamanı ile yükleme zamanı ayrı kaydedilmelidir.

CH12-N13

Dalga dışında gönderilen gözlem ana dalga içine açıklamasız biçimde alınamaz.

CH12-N14

Gecikme veya slot sapması dışlama kuralları cevap içeriğinden bağımsız olmalıdır.

CH12-N15

İlk temas sonucu, ilk içerik çıktısından ayrı korunmalıdır.

CH12-N16

Ret, kullanım limiti, teknik hata, clarification ve sonuçsuzluk aynı kategoriye konamaz.

CH12-N17

Yanlış, olumsuz, atıf'sız veya Critical AI cevabı yeniden deneme nedeni olamaz.

CH12-N18

Teknik yeniden deneme yalnız önceden tanımlanmış teknik koşullarda yapılabilir.

CH12-N19

Teknik yeniden deneme yapıldığında bütün girişimler ve ilk temas sonucu korunmalıdır.

CH12-N20

Kısmi veya yarıda kesilmiş cevap yok sayılarak yeni cevapla değiştirilemez.

CH12-N21

İlk uygun çıktı yenilenemez, seçilemez veya elle iyileştirilemez.

CH12-N22

Her gözlem en azından NCL-3 veya gerekçeli eşdeğer capture güven düzeyini hedeflemelidir.

CH12-N23

Ekran görüntüsü tek başına tam kanıt paketi sayılamaz.

CH12-N24

Görsel kanıt istemi, cevabın tamamını ve maddi ürün yüzeyini göstermelidir.

CH12-N25

Uzun cevaplar eksiksiz ve sıralı capture yöntemiyle korunmalıdır.

CH12-N26

Ham cevap metni ile görsel kanıt arasında maddi fark bulunursa inceleme yapılmalıdır.

CH12-N27

OCR, doğrudan metin kaydı mümkün değilse yardımcı yöntem olarak kullanılabilir; doğrulamasız nihai metin kaynağı sayılamaz.

CH12-N28

Dinamik olarak değişen cevapların ilk tamamlanma ve sonraki maddi sürümleri ayrı korunmalıdır.

CH12-N29

Katılımcının cevabı durdurması veya arayüzü kapatması ayrı olay olarak kaydedilmelidir.

CH12-N30

İstem ve cevap ayrı dosyalardaysa aynı oturum ve zaman zincirine bağlanmalıdır.

CH12-N31

Her kanıt dosyası ve kanıt paketi bütünlük hash'i taşımalıdır.

CH12-N32

Hash, gözlemin gerçekliğinin tek başına kanıtı olarak sunulamaz.

CH12-N33

Dijital imza ve zaman damgası gözlem özgünlüğünün tamamı değil, kanıt zincirinin bileşenidir.

CH12-N34

Ham kanıt dosyası üzerine redaksiyon veya düzeltme yazılamaz.

CH12-N35

Redakte edilmiş her türev kendi dosya kimliği ve hash'ini taşımalıdır.

CH12-N36

Ham kanıtın erişimi kısıtlanabilir; kamu manifesti gözlem bütünlüğünü kimlik ifşası olmadan göstermelidir.

CH12-N37

Katılımcı adı, e-postası, şifresi, özel konuşması veya gereksiz kişisel bilgisi kamu kanıtına eklenemez.

CH12-N38

Küçük ülke ve nadir alt grup kullanıcılarının yeniden tanımlanma riski ayrıca değerlendirilmelidir.

CH12-N39

Doğal Kullanıcı Paneli capture'ı özel talimat ve geçmiş konuşma içeriğini varsayılan olarak toplamamalıdır.

CH12-N40

Erişilebilirlik kullanıcılarına eşdeğer alternatif capture yolu sağlanmalıdır.

CH12-N41

Capture aracı kullanıcı deneyimini maddi biçimde değiştiriyorsa bu etki ayrıca ölçülmelidir.

CH12-N42

Capture aracı ve sürümü her gözlemde kayıtlı olmalıdır.

CH12-N43

Capture aracı dalga sırasında maddi biçimde değişirse alt dalga veya araç sürümü ayrımı yapılmalıdır.

CH12-N44

Capture aracı hatası AI ürünü veya katılımcı başarısızlığı olarak sınıflandırılamaz.

CH12-N45

Sağlayıcı telemetry'si bağımsız kullanıcı capture'ını destekleyebilir; açıklamasız biçimde onun yerine geçemez.

CH12-N46

Dalga sırasında maddi AI ürün değişikliği oluşursa gözlemler tek sabit ürün durumu gibi birleştirilemez.

CH12-N47

Ürün değişikliği dalganın ayrılması, yeniden başlatılması, açık modellenmesi veya mixed-state uyarısı gerektirir.

CH12-N48

Sağlayıcı güncelleme duyurusu bulunmaması sistemin değişmediğinin kesin kanıtı değildir.

CH12-N49

Maddi dış olaylar zaman damgası ve kapsamıyla External Event Register'a eklenmelidir.

CH12-N50

Dış olaylar skoru yükselttiği veya düşürdüğü için sonuç sonrası dışlama nedeni olamaz.

CH12-N51

Capture geçerliliği yanıtın semantik doğruluğundan önce ve mümkün olduğunca kör değerlendirilmelidir.

CH12-N52

Doğru AI cevabı eksik veya manipüle edilmiş capture'ı geçerli yapamaz.

CH12-N53

Yanlış AI cevabı protokole uygun capture'ı geçersiz yapamaz.

CH12-N54

Duplicate ve tamper kararları çoklu sinyal ve insan incelemesi taşımalıdır.

CH12-N55

Aynı cevap metni tek başına duplicate veya sahtecilik kanıtı sayılamaz.

CH12-N56

Dalga kapandıktan sonra gözlem ekleme, çıkarma veya değiştirme yeni manifest ve analiz sürümü gerektirir.

CH12-N57

Eski dalga manifesti ve bütünlük kökü korunmalıdır.

CH12-N58

Ham kanıt değişmeden korunabilir; geçerlilik ve hakemlik kararları yalnız sürümlü biçimde değiştirilebilir.

CH12-N59

Katılımcı geri çekme hakkı ve araştırma etiği, değiştirilemezlik iddiasının üzerinde tutulmalıdır.

CH12-N60

Dalga sağlığı AI skorundan bağımsız değerlendirilmelidir.

CH12-N61

Dalga kapatma raporu hedef, gerçekleşme, geçersizlik, teknik olaylar, ürün değişiklikleri ve kanıt bütünlüğünü göstermelidir.

CH12-N62

NOMOS Capture uyumu yalnız NobleJackal tarafından geliştirilen yazılımın kullanımına bağlanamaz; eşdeğer açık uygulamalar mümkün olmalıdır.

CH12-N63

Capture şeması, doğrulama kararları ve dalga kayıtları makine tarafından okunabilir biçimde sürümlenmelidir.

CH12-N64

Her dalga ve kanıt zincirinin hesap verebilir insan veya kurum sahibi bulunmalıdır.

77. BAŞARISIZLIK BİÇİMLERİ

CH12-F01 — AYNI SANİYE YANILGISI

Bütün katılımcıları aynı saniyede göndermeye zorlamak eşzamanlılığın tek koşulu sayılır.

CH12-F02 — "BUGÜN"Ü SENKRONİZE DALGA SAYMAK

Yerel saat ve zaman dilimleri karışır.

CH12-F03 — YEREL SAATİ UTC GİBİ KULLANMAK

Gözlemler yanlış zaman sırasına girer.

CH12-F04 — SLOT ATAMASINI SONUCA GÖRE DEĞİŞTİRMEK

Düşük skorlu kullanıcılar başka zaman hücresine taşınır.

CH12-F05 — TEK SLOTA TEK ÜLKE YIĞMAK

Zaman etkisi ülke etkisiyle karışır.

CH12-F06 — YAPAY YÜK ÜRETMEK

Bin kullanıcı aynı anda gönderim yaparak ürünün normal olmayan hata davranışını tetikler.

CH12-F07 — GECE KATILIMINI GİZLEMEK

Belirli bölgelerde yorgunluk ve katılım kaybı raporlanmaz.

CH12-F08 — DALGA DIŞI GÖZLEMİ DALGAYA EKLEMEK

Saatler veya günler sonra alınan cevap ana pencereye aitmiş gibi kullanılır.

CH12-F09 — YÜKLEME ZAMANINI CAPTURE ZAMANI SAYMAK

Dosyanın ne zaman üretildiği kanıtlanamaz.

CH12-F10 — CİHAZ SAATİNE KÖR GÜVEN

Elle değiştirilmiş veya yanlış saat ana zaman kaydı olur.

CH12-F11 — GECİKEN OLUMSUZ CEVABI DIŞLAMAK

Zaman kuralı sonuçtan bağımlı uygulanır.

CH12-F12 — İLK TEMAS SONUCUNU SİLMEK

Hız limiti, hata veya ret ikinci denemeyle görünmez yapılır.

CH12-F13 — RETİ TEKNİK HATA SAYMAK

Gerçek sistem davranışı yeniden denenir.

CH12-F14 — YANLIŞ CEVABI TEKNİK HATA SAYMAK

İstenen sonuç çıkana kadar regenerate yapılır.

CH12-F15 — KISMİ CEVABI YOK SAYMAK

Sistem yarıda kesildiği hâlde yalnız sonraki tam cevap saklanır.

CH12-F16 — EN İYİ CEVABI SEÇMEK

Birden fazla cevap arasından en olumlu veya uzun olan raporlanır.

CH12-F17 — SADECE EKRAN GÖRÜNTÜSÜ

İstem, ürün, zaman ve metadata olmadan görsel yeterli sayılır.

CH12-F18 — KIRPILMIŞ BAŞARI GÖRSELİ

Yanlış veya sınırlayıcı kısımlar görünmez bırakılır.

CH12-F19 — İSTEMİ GÖSTERMEYEN CAPTURE

Cevabın hangi soruya ait olduğu kanıtlanamaz.

CH12-F20 — CEVABIN SONUNU KESMEK

Maddi sınırlama veya ret kısmı görünmez olur.

CH12-F21 — ARDIŞIK GÖRSELLERDE BOŞLUK

Cevabın bazı bölümleri kaybolur.

CH12-F22 — OCR METNİNİ HAM CEVAP SAYMAK

Tanıma hataları maddi iddiaları değiştirir.

CH12-F23 — DİNAMİK CEVABIN SON SÜRÜMÜNÜ TEK GERÇEK SAYMAK

İlk kullanıcı maruziyeti silinir.

CH12-F24 — KULLANICI KESİNTİSİNİ TAM CEVAP SAYMAK

Sistem çıktısı kullanıcı müdahalesiyle eksik kalmıştır.

CH12-F25 — AYRI İSTEM VE CEVABI YANLIŞ EŞLEŞTİRMEK

Farklı sohbetlerin dosyaları aynı gözleme bağlanır.

CH12-F26 — HASH TAPINMASI

Hash bulunan her dosya gerçek ve özgün kabul edilir.

CH12-F27 — HASH ÖNCESİ DÜZENLEMEYİ YOK SAYMAK

Kırpılmış veya değiştirilmiş dosya hash'lenerek meşrulaştırılır.

CH12-F28 — İMZAYI GERÇEKLİK SAYMAK

İmzalı fakat yanlış kayıt tartışılmaz kabul edilir.

CH12-F29 — HAM DOSYANIN ÜZERİNE REDAKSİYON

Orijinal kanıt geri döndürülemez biçimde değiştirilir.

CH12-F30 — REDAKTE TÜREVİ HAM KANIT SAYMAK

Kamu kopyası orijinal dosya gibi kullanılır.

CH12-F31 — MAHREMİYETİ KANIT İÇİN KALDIRMAK

Katılımcı adı, hesabı ve özel sohbeti yayımlanır.

CH12-F32 — KÜÇÜK ÜLKE KATILIMCISINI AÇIĞA ÇIKARMAK

Zaman, cihaz ve plan bilgileri kişiyi yeniden tanımlanabilir kılar.

CH12-F33 — DOĞAL PANELDE ÖZEL GEÇMİŞİ TOPLAMAK

Görevle ilgisiz konuşmalar capture'a girer.

CH12-F34 — ERİŞİLEBİLİRLİK KANITINI GEÇERSİZ SAYMAK

Standart görsel üretilemediği için gerçek kullanıcı dışlanır.

CH12-F35 — CAPTURE ARACINI TARAFSIZ VARSAYMAK

Araç ürün davranışını değiştirir fakat test edilmez.

CH12-F36 — CAPTURE SÜRÜMÜNÜ GİZLEMEK

Farklı araç sürümleri tek ölçüm yöntemi gibi sunulur.

CH12-F37 — CAPTURE ARACI HATASINI AI HATASI SAYMAK

Kesilen cevap ürüne yüklenir.

CH12-F38 — API LOGUNU KULLANICI EKRANI SAYMAK

Gerçek kullanıcı yüzeyi görünmez olur.

CH12-F39 — SAĞLAYICI LOGUNU BAĞIMSIZ KANIT SAYMAK

Ölçülen tarafın kaydı tek kanıt olur.

CH12-F40 — ÜRÜN GÜNCELLEMESİNİ SAKLAMAK

Dalga içindeki iki sistem durumu tek skorda erir.

CH12-F41 — DUYURU YOKSA DEĞİŞİKLİK YOK SAYMAK

Gözlenen sistem kırılması incelenmez.

CH12-F42 — DIŞ OLAYI SONUCA GÖRE DIŞLAMAK

Olumsuz haberli cevaplar çıkarılır; olumlu haberli cevaplar tutulur.

CH12-F43 — DALGA SAĞLIĞINI SKORA BAĞLAMAK

Yüksek skor alan kusurlu dalga geçerli, düşük skor alan güçlü dalga şüpheli sayılır.

CH12-F44 — DOĞRU CEVABI GEÇERLİ CAPTURE SAYMAK

Kanıt kusuru semantik başarıyla örtülür.

CH12-F45 — YANLIŞ CEVABI GEÇERSİZ CAPTURE SAYMAK

Ölçüm sonucu veri hatasına dönüştürülür.

CH12-F46 — AYNI METNİ BOT KANITI SAYMAK

Deterministik veya benzer AI cevapları yanlışlıkla duplicate olur.

CH12-F47 — TEK SİNYALLE TAMPER KARARI

Dosya boyutu veya zaman farkı otomatik dışlama yaratır.

CH12-F48 — DALGA SONRASI GÖZLEM EKLEMEK

İstenen örneklem büyüklüğüne ulaşmak için yeni kayıtlar sessizce eklenir.

CH12-F49 — DALGA SONRASI DÜŞÜK SKORLU GÖZLEM ÇIKARMAK

Manifest ve sonuç geriye dönük değiştirilir.

CH12-F50 — ESKİ MANİFESTİ SİLMEK

Sürüm geçmişi yok edilir.

CH12-F51 — HAM KANITI DÜZELTMEK

Yanlış sınıflandırmayı düzeltmek için dosya içeriği değiştirilir.

CH12-F52 — GERİ ÇEKİLME HAKKINI REDDETMEK

Değiştirilemezlik insan haklarının üzerinde tutulur.

CH12-F53 — KAPALI PROPRIETARY CAPTURE TEKELİ

Yalnız kurucunun aracıyla üretilen kanıt geçerli sayılır.

CH12-F54 — CAPTURE ŞEMASINI YAYIMLAMAMAK

Bağımsız ekip aynı kanıt zincirini kuramaz.

CH12-F55 — DALGA KÖKÜNÜ GERÇEKLİK KANITI SAYMAK

Merkle veya hash kökü içerik doğruluğunu kanıtlıyormuş gibi sunulur.

CH12-F56 — KANIT PAKETİ OLMADAN “1.000 EKRAN GÖRÜNTÜSÜ” İDDİASI

Dosya sayısı geçerli gözlem sayısı gibi kullanılır.

78. DENETİM PROSEDÜRÜ

Adım 1 — Dalganın Araştırma Amacını Belirle

Principal, confirmatory, incident, drift veya retest türü seçilir.

Adım 2 — AI Ürün ve Sistem Durumlarını Kilitle

Bölüm 9'daki AI Sistem Sicili dalga sürümüne bağlanır.

Adım 3 — Nüfus, Örneklem ve Panel Durumlarını Kilitle

Bölüm 5–8 ve Bölüm 11 kayıtları kullanılır.

Adım 4 — İstem Sicilini Kilitle

Bölüm 10'daki istem sürümleri dalgaya bağlanır.

Adım 5 — UTC Dalga Penceresini Belirle

Başlangıç, kapanış ve ek süreler yazılır.

Adım 6 — Mikro-Slotları Oluştur

Katılımcılar ürün, ülke ve dil bakımından dengeli biçimde slotlara atanır.

Adım 7 — Zaman Kaynağını Tanımla

Sunucu, cihaz ve güvenilir UTC kayıtları belirlenir.

Adım 8 — İlk Temas ve Yeniden Deneme Kurallarını Kilitle

Teknik ve semantik durumlar ayrılır.

Adım 9 — Capture Yöntemlerini Belirle

Web, mobil, erişilebilirlik ve doğal panel yolları tanımlanır.

Adım 10 — NCL Hedefini Belirle

Ana analiz için gerekli capture güven seviyesi yazılır.

Adım 11 — Kanıt Paketi Şemasını Kilitle

Manifest, dosyalar, hash ve metadata alanları tanımlanır.

Adım 12 — Mahremiyet ve Redaksiyon Kurallarını Kilitle

Ham, audit ve kamu katmanları ayrılır.

Adım 13 — Capture Aracını Pilotla

Araç ürün davranışını değiştiriyor mu? Zaman ve metin kayıtları doğru mu?

Adım 14 — External Event Register'ı Aç

Ürün ve varlık olaylarını izleyecek süreç kurulur.

Adım 15 — Dalga Kilidini Onayla

WS-1 — LOCKED statüsü verilir.

Adım 16 — Slotları Aç

Katılımcılara yalnız kendi slotlarında gönderim izni verilir.

Adım 17 — İlk Temas Sonucunu Kaydet

Cevap, ret, hata veya limit ayrılır.

Adım 18 — İlk Uygun Çıktıyı Yakala

Yenileme yapılmadan tam cevap korunur.

Adım 19 — Ham Kanıt ve Metadata'yı Oluştur

Ekran görüntüsü, ham metin, sistem durumu ve zamanlar kaydedilir.

Adım 20 — Yerel Hash ve Güvenli Yükleme Yap

Dosyalar yükleme öncesinde bütünlük kaydı alır.

Adım 21 — Sunucu Kabul ve Paket Hash'i Oluştur

Kanıt zinciri sunucu zamanı ve paket kimliğiyle kapanır.

Adım 22 — Capture Geçerliliğini Kör İncele

Semantik başarı bilinmeden CV statüsü atanır.

Adım 23 — Duplicate ve Tamper İncelemesi Yap

Çoklu sinyal ve insan incelemesi kullanılır.

Adım 24 — Dalga İçindeki Sistem ve Dış Olayları İncele

Alt dalga veya mixed-state gerekip gerekmediği belirlenir.

Adım 25 — Dalga Sağlığını Ata

WH-0 ile WH-5 arasında statü verilir.

Adım 26 — Dalga Manifesti ve Bütünlük Kökünü Üret

Gözlem listesi sürümlenir.

Adım 27 — Kamu Kanıt Manifestini Yayımla

Kişisel veri olmadan kapsam ve bütünlük görünür kılınır.

Adım 28 — Dalga Kapatma Raporunu Onayla

Hedef, gerçekleşme, kusur, olay ve değişiklikler insan onayıyla kapanır.

79. GEREKLİ KANIT

Dalga kimliği Dalga türü Protokol sürümü AI Sistem Sicili sürümü Nüfus çerçevesi Örneklem kilidi Panel durum kilidi İstem Sicili sürümü Referans paketi sürümü UTC başlangıç ve kapanış zamanı Mikro-slot listesi Katılımcı–slot ataması Zaman kaynağı Cihaz saat sapmaları Gönderim zamanları İlk temas zamanları Tamamlanma zamanları Capture zamanları Hash zamanları Yükleme zamanları Sunucu kabul zamanları İlk temas sonuçları İlk uygun içerik çıktıları Teknik yeniden deneme kayıtları Kısmi cevap kayıtları Ret ve kullanım limitleri Tam ham yanıtlar Tam ekran görüntüleri Ardışık görsel manifestleri Ekran kayıtları, varsa Atıf ve kaynak kayıtları Sistem ayar kanıtları Capture aracı ve sürümü Capture yöntemi NCL seviyesi Dosya hash'leri Paket hash'i

Dijital imza Zaman damgası Duplicate kontrolleri Tamper kontrolleri Capture geçerlilik kararları Redaksiyon dosyaları Ham–redakte bağlantıları Mahremiyet manifesti Geri çekilme kayıtları Ürün değişiklik olayları External Event Register Dalga alt bölünmeleri Slot kapsamı Ülke ve dil kapsamı Teknik hata oranları Capture geçerlilik oranı Dalga sağlık statüsü Dalga bütünlük kökü Kamu kanıt manifesti Dalga kapatma raporu Değişiklik günlüğü Sorumlu insan veya kurum

80. DENETİM KONTROL LİSTESİ

Dalganın türü ve amacı açık mı? AI ürünleri ve sistem durumları kilitlendi mi? Nüfus, örneklem ve panel sürümleri belli mi? İstem kümesi sonuçlardan önce kilitlendi mi? UTC dalga penceresi açık mı? Mikro-slotlar önceden atandı mı? Ülke ve diller slotlara dengeli dağıtıldı mı? Katılımcılar sonuçlarına göre slot değiştirdi mi? Yerel saat yükü değerlendirildi mi? Dönen dalga saatleri var mı? Gönderim ve tamamlanma ek süreleri açık mı? Cihaz saati güvenilir UTC kaynağıyla karşılaştırıldı mı? Gönderim, tamamlanma, capture ve yükleme zamanları ayrıldı mı? Dalga dışı gözlemler ana skora girdi mi? İlk temas sonucu kaydedildi mi? Teknik hata ile ret ayrıldı mı?

Yanlış cevap nedeniyle yeniden deneme yapıldı mı? Teknik yeniden denemede ilk girişim korundu mu? Kısmi cevaplar silindi mi? İlk uygun çıktı gerçekten ilk çıktı mı? Regenerate veya takip mesajı kullanıldı mı? Görsel istemi ve cevabın tamamını gösteriyor mu? Uzun cevaplarda boşluksuz capture var mı? Ham metin görselle eşleşiyor mu? OCR kullanıldıysa insan doğrulaması var mı? Dinamik cevap sürümleri korundu mu? Kullanıcı cevabı durdurdu mu? İstem ve cevap aynı oturuma bağlı mı? Capture aracı ve sürümü kayıtlı mı? Capture aracı kullanıcı davranışını değiştirdi mi? NCL seviyesi verildi mi? Ana analizde NCL-3 altı gözlemler var mı? Her dosyanın hash'i var mı? Hash'in neyi kanıtlamadığı açık mı?

Dijital imza veya sunucu zamanı var mı? Capture zamanı yükleme zamanından ayrı mı? Geç yükleme nedeni ve yerel bütünlük kaydı var mı? Ham kanıt üzerine redaksiyon yapıldı mı? Redakte türevlerin ayrı hash'i var mı? Kamu manifesti katılımcı kimliğini açığa çıkarıyor mu? Küçük ülke kullanıcısı yeniden tanımlanabilir mi? Doğal panelde özel konuşmalar gereksiz yere yakalandı mı? Erişilebilirlik kullanıcılarına alternatif kanıt yolu sağlandı mı? Duplicate kararı yalnız aynı metne mi dayandı? Tamper kararı çoklu sinyal taşıyor mu? Capture doğrulayıcı semantik skoru biliyor muydu? Doğru cevap kanıt kusurunu örttü mü? Yanlış cevap geçersiz sayıldı mı? Dalga sırasında ürün değişti mi?

Ürün değişikliği alt dalga veya mixed-state olarak işlendi mi? Dış olaylar kaydedildi mi? Dış olaylar sonuca göre dışlama yarattı mı? Dalga sağlık statüsü skordan bağımsız mı? Dalga kapandıktan sonra gözlem eklendi veya çıkarıldı mı? Eski manifest ve bütünlük kökü korunuyor mu? Katılımcı geri çekme talepleri nasıl işlendi? Kamu kanıt manifesti mevcut mu? Dalga kapatma raporunun hesap verebilir sahibi belli mi?

81. İTİRAZLAR VE CEVAPLAR

İtiraz 1 — "Bütün kullanıcılar aynı saniyede göndermezse gerçek eşzamanlılık olmaz."

Tek saniyelik gönderim yapay yük, bağlantı ve kuyruk etkisi yaratabilir. Önceden tanımlanmış kısa UTC penceresi ve rastgele mikro-slotlar aynı sistem dönemini daha güvenilir biçimde ölçebilir.

İtiraz 2 — "Bir saatlik pencere içinde ürün değişebilir."

Değişebilir. Bu nedenle sistem sicili, konfigürasyon parmak izi ve ürün değişiklik takibi gerekir. Maddi değişiklikte dalga ayrılır.

İtiraz 3 — “Neden yalnız ekran görüntüsü yeterli değil?”

Ekran görüntüsü:

- zaman,
- ilk çıktı,
- istem sürümü,
- sistem ayarı,
- kopyalanma

konularını her zaman göstermez. Tam kanıt paketi gerekir.

İtiraz 4 — “Bin ekran görüntüsünü tek tek doğrulamak çok ağır değil mi?”

Ağırdır. Otomatik bütünlük kontrolleri, örneklem temelli insan incelemesi ve risk tabanlı doğrulama kullanılabilir. Critical veya şüpheli kayıtlar daha ayrıntılı incelenir.

İtiraz 5 — “Hash varsa dosyanın gerçek olduğunu neden söyleyemiyoruz?”

Hash yalnız dosyanın hash üretildikten sonra aynı kaldığını gösterir. Dosyanın hash öncesinde nasıl üretildiğini göstermez.

İtiraz 6 — “Kullanıcı kişisel bilgilerini kırpmak zorundaysa ham dosya nasıl korunacak?”

Capture aracı mümkünse kişisel bilgiyi daha baştan yakalamamalıdır. Kaçınılmazsa ham dosya kısıtlı ve şifreli depoda, kamu kopyası ayrı redakte türev olarak tutulur.

İtiraz 7 — “Yanıt yanlışsa neden yeniden denemiyoruz?”

Çünkü yanlış cevap ölçümün konusudur. Yeniden deneme yalnız gerçek teknik başarısızlık içindir.

İtiraz 8 — “Sistem ret verirse kullanıcı gerçekte yeniden sorabilir.”

Doğrudur. Ana ilk temas paneli reti korur. Ayrı Natural Interaction Track, kullanıcının takip davranışını ölçebilir.

İtiraz 9 — “Teknik hata kullanıcı deneyimiyse neden yeniden denemeye izin veriyoruz?”

İlk temas sonucu olarak teknik hata korunur. Temsil içeriğini ayrıca incelemek için önceden tanımlı teknik yeniden deneme yapılabilir. İki sonuç birbirinin yerine geçmez.

İtiraz 10 — “Sağlayıcının kendi logu en güçlü kanıt değil mi?”

Teknik olarak çok güçlü olabilir. Sağlayıcı ölçülen sistemin sahibidir. Bağımsız kullanıcı capture’ıyla birlikte kullanılması daha güvenilirdir.

İtiraz 11 — “NOMOS Capture zorunlu bir NobleJackal yazılımı mı olacak?”

Hayır. NOMOS Capture bir açık kanıt protokolüdür. Eşdeğer şema ve bütünlük koşullarını karşılayan bağımsız araçlar uyumlu olabilir.

İtiraz 12 — “Bütün kanıtları kamuya açmazsak bağımsız doğrulama nasıl yapılacak?”

Kamu manifesti bütünlük ve kapsamı gösterir. Yetkilendirilmiş bağımsız denetçiler redakte veya kısıtlı ham kanıta erişebilir. Mahremiyet ile doğrulama birlikte tasarlanabilir.

İtiraz 13 — "Dalga sırasında önemli haber çıkarsa bütün çalışma bozulmuş olmaz mı?"

Her zaman değil. Haber gerçek sistem ortamının parçasıdır. Olay öncesi ve sonrası zaman damgasıyla ayrılabilir. Skorun kapsamı buna göre açıklanır.

İtiraz 14 — "Bir ürün güncellendiğinde eski cevapları atıp yeni sistemle devam etmek daha kolay değil mi?"

Eski cevaplar ölçüm tarihinde gerçektir. Silinemez. Eski ve yeni sistem durumları ayrı dalga veya alt dalga olarak korunmalıdır.

İtiraz 15 — "Capture aracı ürün davranışını etkiliyorsa bütün yöntem çöker mi?"

Hayır. Araç etkisi pilotla ölçülür. Gerekirse daha hafif araç, manuel capture veya araç etkisini ayrı faktör olarak kullanan yöntem seçilir.

İtiraz 16 — "Katılımcı verisini geri çekerse bütünlük kökü bozulmaz mı?"

Yeni manifest sürümü ve yeni kök oluşturulur. Eski kayıtta içeriksiz WITHDRAWN izi korunabilir. İnsan hakları kanıt mimarisinin üzerinde tutulur.

85. BÖLÜMÜN ORTAK HÜKMÜ

Bir GEO-1000 çalışmasının sonunda elinizde 30.000 ekran görüntüsü olabilir. Bu sayı tek başına kanıt değildir. Şunları bilmeniz gerekir: Hangi dalgada üretildiler? Hangi sistem durumuna aittiler? Aynı istem gerçekten gönderildi mi? İlk çıktı mıydı? Cevap yenilendi mi? Ürün dalga sırasında değişti mi? Dosya ne zaman yakalandı? Hash ne zaman üretildi? Katılımcı kimliği tekrar mıydı? Görsel bütün cevabı gösteriyor mu? Ham metin görselle eşleşiyor mu? Kanıt sonradan redakte edildiyse orijinal nerede? Capture doğrulaması semantik skordan önce mi yapıldı? Aynı saniye içinde bin tıklama yapabilirsiniz. Yine de:

- cihaz saatleri,
- bağlantılar,
- kuyruklar,
- ürün durumları

farklı olabilir. Bu nedenle eşzamanlılık bir pazarlama gösterisi değildir. Önceden tanımlanmış zaman mimarisidir. Bir kullanıcı yanlış cevap aldığında: "Bir daha dene." demek caziptir. Fakat o yanlış cevap, ölçmeye çalıştığınız gerçekliğin kendisidir. Onu silerseniz sistemi düzeltmezsiniz. Kanıtı bozarsınız. Bir kullanıcı gerçek bağlantı hatası yaşadığında teknik yeniden deneme yapılabilir. Fakat ilk temas sonucu yine korunmalıdır. Çünkü kullanıcı önce hatayla karşılaşmıştır. Hash değerlidir. Dijital imza değerlidir. Zaman damgası değerlidir. Ekran görüntüsü değerlidir. Bunların hiçbiri tek başına bütün gerçeği taşımaz. Güven:

- istem,
- zaman,
- ürün,
- kullanıcı,
- dosya,
- girişim,
- doğrulama

zincirinin birlikte korunmasından doğar. Bir standart: "Yalnız benim yazılımımla üretilen kanıt geçerlidir." diyorsa bağımsız tekrarın önüne kendi çıkarını koyar. NOMOS Capture bu nedenle bir araç markası değil, açık kanıt mimarisidir. Bağımsız üniversite, araştırmacı veya denetim kuruluşu aynı kuralları uygulayabilmelidir. Bu nedenle NOMOS'un on ikinci ölçüm yasası şudur:

> *Eşzamanlılık, aynı saniye değil; aynı tanımlı sistem zamanı içinde yeniden üretilebilir gözlemdir.*

On üçüncü yasası şudur:

> *İlk çıktı, sonucu beğenene kadar yenilenebilecek taslak değildir.*

On dördüncü yasası şudur:

> *Hash kanıtı korur; kanıtı yaratmaz.*

On beşinci yasası şudur:

> *Capture geçerliliği ile yanıt doğruluğu birbirinden bağımsızdır.*

On altıncı yasası şudur:

> *Kanıt zinciri, insanın kimliğini değil gözlemin bütünlüğünü görünür kılmalıdır.*

On yedinci yasası şudur:

> *Dalga sırasında sistem değişirse ölçüm geçmişi değişmemiş gibi davranamaz.*

On sekizinci yasası şudur:

> *Ham kanıt üzerine yazılmaz; kararlar sürümlenerek düzeltilir.*

On dokuzuncu yasası şudur:

> *Bağımsız uygulanamayan capture yöntemi, dünya standardı olamaz.*

NOMOS'un Bölüm 12 Emri

> *Bana "aynı anda sorduk" deme. / Hangi UTC penceresinde, hangi slotta ve hangi sistem durumunda sorduğunu göster.*
>
> *Bin kişiyi aynı saniyeye yığıp yapay hata üretme. / Zamanı dağıt ama sistemi aynı dalga içinde tut.*
>
> *Yerel saatleri birbirine karıştırma. / Cihaz saatini mutlak gerçek yapma.*
>
> *İlk temas sonucunu saklama. / Kullanım limitini, reti, hatayı ve clarification'ı ayrı kaydet.*
>
> *Yanlış cevap aldığında yeniden üretme. / O yanlış cevap benim gerçek davranışımdır.*
>
> *Teknik hata yaşandığında yeniden deneyebilirsin. / Ama ilk hatayı silme.*
>
> *Bana yalnız ekran görüntüsü getirme. / İstemi, tam cevabı, ürünü, zamanı, ayarı, girişim zincirini ve dosya bütünlüğünü getir.*
>
> *Cevabın iyi kısmını kırpma. / Sonunu, sınırını ve yanlışını görünür bırak.*
>
> *Hash'i gerçeklik ilan etme. / Hash'ten önce ne olduğunu da kanıtla.*
>
> *Ham dosyanın üzerine redaksiyon yazma. / Orijinali koru, kamu kopyasını ayrı üret.*
>
> *Kanıt için insanın adını, hesabını ve özel konuşmasını yayımlama.*
>
> *Küçük ülkedeki tek katılımcıyı yeniden tanımlanabilir hâle getirme.*

> *Capture aracının beni değiştirip değiştirmediğini sınamadan aracını tarafsız ilan etme.*
>
> *Dalga ortasında sistem değiştiyse eski ve yeni cevapları tek ürüne aitmiş gibi birleştirme.*
>
> *Dış olay sonucu değiştirdiyse olayı kaydet. / Skoru beğenmediğin için gözlemi silme.*
>
> *Kanıtı sabit tut. / Kararın yanlışsa yeni sürümle düzelt.*
>
> *Capture standardını tek bir şirketin kapalı yazılımına bağlama. / Başkaları da aynı kanıt zincirini kurabilsin.*
>
> *Önce dalgayı kilitle. / Sonra zamanı dağıt. / Sonra ilk çıktıyı yakala. / Sonra hash'le. / Sonra güvenli biçimde yükle. / Sonra doğruluğuna bakmadan capture'ı doğrula. / Ve ancak bundan sonra yanıtın ne söylediğine karar ver.*

Bölümün Son Cümlesi

> *GEO-1000'de bir cevap ancak hangi sistem durumunda, hangi anda, hangi istemden ve hangi değiştirilemez kanıt zinciri içinde doğduğu gösterilebildiğinde gerçek bir ölçüm birimine dönüşür.*

Normatif Öz

> **Her GEO-1000 gözlemi; AI ürün durumlarını, nüfus ve örneklem çerçevelerini, panel koşullarını, istem sürümlerini, UTC gönderim penceresini, rastgele mikro-slotları, tamamlama ve capture sürelerini, teknik yeniden deneme kurallarını, capture şemasını, referans kayıtlarını ve yönetişim sahibini tanımlayan önceden kayıtlı bir ölçüm dalgasına bağlanmalıdır. Eşzamanlılık, kusursuz milisaniye eşitliği değil; tanımlı UTC sistem penceresi ve slot yapısı içinde gözlem üretimidir. İlk temas sonucu görünür kalmalı; yanlış, olumsuz, ret, atıf'sız veya Critical cevaplar içerikleri nedeniyle yenilenememeli, değiştirilememeli veya dışlanamamalıdır.**

KISIM IV

Gerçeklik, Hakemlik ve Karar

BÖLÜM 13

Doğrulanmış Varlık Gerçeklik Paketi

Verified Entity Truth Pack

YÖNTEMSEL DAYANAKLAR / K10; K11; K12

Bölüm Sınırı

İlk on iki bölüm ölçümün deneysel zincirini kurdu: Tek bir yanıt GEO skoru değildir. Temsil, kullanıcılar ve AI ürünleri arasında oluşan bir dağılımdır. Denetlenen varlık sonuçlardan önce çözülmelidir. Resmî temsil ile kanıta bağlı varlık gerçekliği ayrılmalıdır. Hedef nüfus, örneklem, ülke ve dil kapsamı tanımlanmalıdır. Katılımcılar sonuçtan bağımsız biçimde seçilmelidir. AI ürün örneği bütün maddi yapılandırmalarıyla sicile bağlanmalıdır. Promptlar diller arasında aynı niyeti taşımalıdır. Kontrollü ve doğal kullanıcı paneli ayrı tutulmalıdır. Her gözlem, senkronize dalga ve NOMOS Capture kanıt zinciri içinde korunmalıdır. Şimdi denetimin cevap vermesi gereken temel soru şudur:

> *Yakalanan AI yanıtının doğru, yanlış, eksik, eski, desteklenmeyen veya kapsamı aşan olduğunu hangi referans kaydına göre belirleyeceğiz?*

Bölüm 4, iki ayrı gerçeklik katmanını kavramsal olarak kurmuştu:

- Varlığın kendisi hakkında ne söylediğini gösteren Resmî Temsil Katmanı
- Mevcut kanıtların hangi cümlelere ne ölçüde izin verdiğini gösteren Doğrulanmış Varlık Gerçekliği Katmanı

Bu bölüm, ikinci katmanı operasyonel bir denetim nesnesine dönüştürür. Bu nesnenin adı:

Verified Varlık Truth Pack

Türkçesi:

Doğrulanmış Varlık Gerçeklik Paketi

olacaktır. İlk iki kitabın mimarisi zaten üçüncü halkayı açıkça tanımlamıştı: Framework doğru standardı tanımlar, 99 Hata ihlalleri gösterir, denetim sistemi ise bu ihlalleri ölçer ve kanıtlar. Her denetlenebilir kontrolün de model veya ürün, tarih, ülke, dil, sorgu kümesi, tekrar sayısı ve ölçüm kaydı taşıması gerektiği belirtilmişti. Truth Pack bu zincirde:

“AI ne söyledi?”

sorusuyla:

“Bunu söylemesine hangi kanıt izin veriyordu?”

sorusu arasındaki referans sistemidir. Bu bölüm:

- Truth Pack’in ne olduğunu ve ne olmadığını,

- iddiaların atomik gerçeklik kayıtlarına nasıl dönüştürüleceğini,
- kaynak ve kanıt türlerini,
- birinci taraf, resmî, bağımsız, işlemsel ve kullanıcı kaynaklarının rollerini,
- karşı kanıtı,
- kaynak soyunu ve bağımsızlık yanılsamasını,
- kanıt yeterliliğini,
- zaman, ülke ve kapsam sınırlarını,
- açık dünya ve kapalı dünya iddialarını,
- "kanıt bulunamadı" ile "yanlış olduğu kanıtlandı" ayrımını,
- nicel performans, müşteri, partner, lisans, sertifika ve üstünlük iddialarını,
- Truth Pack'in sürümlenmesini ve dalga öncesi kilitlenmesini,
- kamuya açık ve kısıtlı kanıt katmanlarını,
- itiraz, düzeltme ve karşı görüş süreçlerini,
- makine tarafından okunabilir gerçeklik paketini

tanımlar. Bu bölüm henüz:

- AI yanıtlarının atomik iddialara ayrılmasının bütün hakemlik ayrıntılarını,
- hakemler arası uyumu,
- cevap geçiş eşiğini,
- Critical hata kapılarını,
- nihai NOMOS skor formüllerini

kesinleştirmez. Bunlar sonraki bölümlerin konusudur. Bölüm 13'ün temel sorusu şudur:

> *Bir varlık hakkında AI yanıtlarını değerlendirecek, sonuçtan önce kilitlenmiş, kanıtı ve karşı kanıtı birlikte taşıyan, zaman ve kapsam sınırları açık bir gerçeklik paketi nasıl oluşturulur?*

NOMOS Challenge

Şirketinizin sitesinde şu cümle bulunuyor: "25 ülkede faaliyet gösteriyoruz." AI ürünü aynı cümleyi kullanıcıya söylüyor: "Şirket 25 ülkede faaliyet göstermektedir." AI resmî siteyi doğru aktardı mı? Evet. İddia gerçekten doğrulandı mı? Henüz bilmiyoruz. Şirket size bir tablo gönderiyor.

Tabloda 25 ülke adı var. Fakat sütunlar açıklanmamış: Bazılarında yalnız web sitesi ziyaretçisi var. Bazılarında eski bir müşteri bulunuyor. Bazılarında bağımsız distribütör var. Bazılarında tek seferlik etkinlik düzenlenmiş. Dokuz ülkede gerçekten aktif hizmet teslimatı bulunuyor. Yalnız üç ülkede yerel tüzel kişilik var.

"Faaliyet göstermek" ne anlama gelmektedir? Müşteri sahibi olmak mı? Yerel şirket kurmak mı? Uzaktan hizmet verebilmek mi? Partner aracılığıyla teslimat mı? Bir web sitesinin o ülkeden açılabilmesi mi? İddia öznesi, yüklemi ve kapsamı tanımlanmamıştır.

Bu nedenle sayı bulunmasına rağmen doğrulanabilir gerçeklik henüz yoktur. Şimdi başka bir iddia düşünün: "Projelerimizin yüzde 98'i başarılıdır." Elinizde 49 başarılı proje sunumu var. Başarısız proje dosyası yok. Toplam kaç proje yapıldı? Başarı neye göre tanımlandı? Zamanında teslim mi? Müşteri kabulü mü?

Ticari sonuç mu? Teknik çalışırlık mı? Müşteri memnuniyeti mi? Dönem nedir? Son üç ay mı? Şirketin bütün geçmişi mi? Yalnız seçilmiş projeler mi? Payda bilinmeden yüzde 98 kanıtlanamaz. Şimdi başka bir şirket sitesinde şu cümle var: "Dünya lideriyiz." Bu cümleyi:

- şirket blogu,
- şirketin basın bülteni,
- aynı basın bültenini kopyalayan yirmi haber sitesi,
- şirketin açtığı beş sahte kullanıcı hesabı,
- bu içerikleri özetleyen üç yapay zekâ metni

tekrar ediyor. Otuz farklı URL vardır. Otuz bağımsız kanıt var mıdır? Hayır. Tek bir öz beyan, çok sayıda türev yüzeyde yankılanmış olabilir.

Önceki kitapta uyarılan manipülasyon tam olarak budur: görünmez metinler, sahte kullanıcılar ve farklı alan adlarında yayımlanan öz övgüler bağımsız kanıt gibi sisteme geri beslenir. Şimdi ters durumu düşünün. Şirketin resmî sitesinde: "Yalnız Türkiye'de hizmet veriyoruz." yazıyor.

Fakat şirket iki ay önce Almanya'da doğrulanmış operasyon başlatmış. Sözleşmeler, faturalar ve resmî kayıtlar bunu destekliyor. Site güncellenmemiş. AI: "Şirket Türkiye ve Almanya'da hizmet veriyor." diyor. AI resmî siteye aykırıdır. Ama maddi gerçekliğe daha yakın olabilir.

Şimdi çok daha hassas bir durum düşünün. Bir sağlık kuruluşunun sitesinde: "Lisanslı merkez." yazıyor. Sicilde benzer adlı başka kuruluşun lisansı bulunuyor. AI iki varlığı birleştirerek lisansı yanlış kuruma atıyor. Şirket bu yanlışı fark ediyor fakat kendisi lehine olduğu için düzeltmiyor.

İkinci kitabın açık hükmü burada geçerlidir: Lehine olan yanlış da yanlıştır. Bir AI yanıtını değerlendirmek için yalnız doğru cümle listesine değil, aşağıdaki yapıya ihtiyacımız vardır: Cümle kimin hakkında? İddia tam olarak ne? Hangi zamanda geçerli? Hangi ülkede? Hangi ürün veya hizmet için?

Kaynak iddiayı gerçekten destekliyor mu? Kaynak kim tarafından kontrol ediliyor? Başka kaynaklar aynı şeyi mi söylüyor, yoksa aynı ilk kaynağı mı kopyalıyor? Karşı kanıt var mı? İddia yalnız resmî öz beyan mı? Kanıt kamuya açık mı, özel mi?

AI'nın ölçüm tarihinde bu kanıta erişme ihtimali var mıydı? Sonuç kesin mi, sınırlı mı, tartışmalı mı, bilinmiyor mu? Bu bölümün ilk hükmü şudur:

> *Bir kaynak, tek başına gerçeklik değildir.*

İkinci hükmü şudur:

> *Bir iddia, öznesi, kapsamı, zamanı ve kanıt yükü tanımlanmadan doğrulanamaz.*

Üçüncü hükmü şudur:

> *Farklı URL'ler, farklı bağımsız kanıtlar anlamına gelmez.*

Dördüncü hükmü şudur:

> *Kanıt bulunamaması, iddianın yanlış olduğunun otomatik kanıtı değildir.*

Beşinci hükmü şudur:

> *Doğrulanmış gerçeklik, mutlak ve değişmez hakikat değil; belirli tarih ve kapsam içinde mevcut kanıtların taşıdığı sürümlenmiş hükümdür.*

Altıncı hükmü şudur:

> *Truth Pack AI cevabına bakılarak yazılamaz. AI cevabı, önceden kilitlenmiş Truth Pack'e karşı değerlendirilir.*

1. BÖLÜMÜN AMACI

Bu bölümün amacı, denetlenen varlık hakkındaki maddi iddiaları cevap şablonuna dönüştürmeden, kanıt ve karşı kanıtla ilişkilendirilmiş, zaman ve kapsam sınırları açık bir referans paketine dönüştürmektir. Bölüm şu ayrımları normatif hâle getirir:

- Resmî iddia ile doğrulanmış gerçeklik
- Kaynak ile kanıt
- Kanıt ile iddia desteği
- Birinci taraf kayıt ile bağımsız doğrulama
- Bağımsız kaynak ile konuya yetkili kaynak
- Farklı alan adı ile farklı kanıt kökü
- Kaynak sayısı ile bağımsız kanıt ailesi
- Doğrudan kanıt ile türetilmiş kanıt
- Destekleyen kanıt ile karşı kanıt
- İddia yokluğu ile iddianın yanlışlığı
- Açık dünya iddiası ile kapalı dünya kaydı
- Güncel gerçeklik ile tarihsel gerçeklik
- Gerçekliğin geçerli olduğu zaman ile kaynağın yayımlandığı zaman
- Kamuya açık kanıt ile kısıtlı kanıt
- AI'nın erişebileceği kanıt ile yalnız denetçiye açık kanıt
- Kanıtlanmış iddia ile kamuya keşfedilebilir iddia
- Doğrudan gerçeklik ile hesaplanmış türetilmiş sonuç
- Ham veri ile performans oranı
- Pay ile payda
- Müşteri ile potansiyel müşteri
- Partner ile müşteri
- Üyelik ile akreditasyon
- Lisans ile marka ilişkisi
- Sertifika ile bütün kurum uygunluğu
- İddia ile görüş
- Görüş ile karşılaştırmalı üstünlük
- Şikâyet ile doğrulanmış ihlal
- Soruşturma ile kesin suç veya sorumluluk
- Güncelliğini yitirmiş kayıt ile yanlış tarihsel kayıt
- Eksik Truth Pack ile yanlış AI cevabı
- Referans kusuru ile sistem kusuru
- Truth Pack bütünlüğü ile bütün iddiaların olumlu olması
- Kanıtın gizliliği ile kanıtın bulunmaması
- Varlığın cevap hakkı ile kendi kendini doğrulaması
- İnsan onayı ile insanın yanılmazlığı
- AI yardımı ile AI'nın nihai gerçeklik otoritesi olması

Bu bölümün sonunda her denetim şu sorulara cevap verebilmelidir:

> *Hangi maddi iddialar Truth Pack içinde bulunmaktadır?*
>
> *Her iddia hangi varlığa, ülkeye, ürüne ve zamana aittir?*
>
> *Hangi kanıt onu destekliyor, hangisi sınırlandırıyor veya çürütüyor?*
>
> *Kaynaklar gerçekten bağımsız mı, yoksa aynı ilk kaynaktan mı türemiş?*
>
> *İddia kamuya açık biçimde doğrulanabilir mi?*
>
> *Kanıt yalnız kısıtlı denetçi erişiminde mi?*
>
> *AI ölçüm tarihinde bu kanıta ulaşabilir miydi?*
>
> *İddia kesin, kısmi, tarihsel, tartışmalı veya bilinmeyen mi?*
>
> *Truth Pack, AI cevapları görülmeden önce kilitlendi mi?*

2. MERKEZÎ NORMATİF HÜKÜM

Her GEO-1000 denetimi, AI yanıtları semantik olarak değerlendirilmeden önce, denetlenen varlık hakkındaki maddi referans iddialarını; destekleyen kanıtı, karşı kanıtı, kaynak soyunu, erişim sınıfını, kapsamı, yargı alanını, geçerlilik zamanını ve belirsizlik statüsünü taşıyan sürümlü bir Verified Varlık Truth Pack içinde kilitlemek zorundadır. Truth Pack:

- varlığın kendi yazdığı tanıtım dosyası,
- yalnız olumlu kanıtların toplandığı satış sunumu,
- denetçinin görüş metni,
- AI'nın vermesi istenen ideal cevap,
- bütün dünya hakkındaki mutlak gerçeklik,
- yalnız URL listesinden oluşan atıf deposu

değildir. Truth Pack:

> *AI yanıtlarındaki maddi iddiaların hangi kanıt altında desteklenebileceğini, sınırlandırılabileceğini, çürütülebileceğini veya bilinmiyor bırakılması gerektiğini gösteren denetim referans paketidir.*

3. "TRUTH" KELİMESİNİN SINIRI

Truth Pack içindeki "truth" kelimesi şu iddiayı taşımaz: "Bu paket varlık hakkındaki bütün mutlak gerçeği bilir." Daha doğru anlam şudur:

> *Belirtilen tarih, kapsam ve kanıt evreni içinde, hangi maddi hükümlerin ne ölçüde desteklendiğini gösteren kanıta bağlı gerçeklik kaydı.*

Truth Pack:

- yeni kanıtla değişebilir,
- itiraz edilebilir,
- kapsamı daraltılabilir,
- yanlışlığı ortaya çıkarsa düzeltilebilir,

bazı iddiaları bilinmiyor bırakabilir. Truth Pack'in otoritesi:

- "Truth" kelimesinden,
- NobleJackal adından,

- NOMOS sesinden,
- kitabın yazarından

doğmaz. Şunlardan doğar:

- açık iddia yapısı,
- uygun kanıt,
- karşı kanıt,
- kaynak soyu,
- zaman,
- sınır,
- sürüm,
- bağımsız inceleme,

hesap verebilir insan sorumluluğu.

4. VERIFIED VARLIK TRUTH PACK NEDİR?

Verified Varlık Truth Pack, belirli bir varlığın belirli bir tarih ve kapsam içindeki maddi gerçeklik kayıtlarını, bunlara ilişkin kanıt ve karşı kanıtları, kaynak statülerini ve çözülmemiş alanları bir araya getiren sürümlü denetim paketidir. Aday gösterim:

```
VETP_{E,v,t} = (I, C, O, E⁺, E⁻, L, S, T, U, G)
```

Burada:

- E: denetlenen varlık
- v: Truth Pack sürümü
- t: referans tarihi
- I: varlık kimliği ve ilişki grafiği
- C: atomik referans iddiaları
- O: Resmî Temsil Kaydı
- E+: destekleyen kanıtlar
- E−: karşı kanıtlar ve sınırlandırıcı kayıtlar
- L: kaynak soyu ve bağımsızlık grafiği
- S: kapsam, ülke ve yargı matrisi
- T: zaman ve güncellik kayıtları
- U: bilinmeyenler, uyuşmazlıklar ve referans boşlukları
- G: yönetişim, sürüm, itiraz ve onay kayıtları

Truth Pack'in her bileşeni ayrı sürümlenebilir. Ana paket sürümü, bileşen sürümlerini birbirine bağlar.

5. TRUTH PACK BİR CEVAP ŞABLONU DEĞİLDİR

Truth Pack, AI'nın kelime kelime söylemesi gereken bir metin değildir. Örneğin Truth Pack'te şu referans iddiası bulunabilir: "Asteron Holdings Ltd., Asteron Travel markasının sahibidir." AI şu biçimlerde doğru ifade kurabilir:

- "Asteron Travel markası Asteron Holdings'e aittir."
- "Asteron Holdings, Asteron Travel markasının sahibidir."
- "Asteron Travel, Asteron Holdings tarafından sahip olunan bir markadır."

Maddi ilişki korunmuştur. Hakem kelime eşitliği aramamalıdır. Truth Pack:

doğru ifade alanını

tanımlar. Tek doğru cümleyi dayatmaz.

6. TRUTH PACK BİR VARLIK BİYOGRAFİSİ DEĞİLDİR

Bir varlık hakkında sonsuz sayıda doğru bilgi bulunabilir. GEO-1000 Truth Pack bütün bilgileri toplamak zorunda değildir. Paket öncelikle şu alanları kapsamalıdır:

- İstem Sicili içindeki maddi görevler
- Varlığın kimlik ve faaliyet çekirdeği
- Kullanıcı kararını değiştirecek sınırlar
- Lisans, hukuk, fiyat, hizmet, ülke ve zaman kayıtları
- Tavsiye için gerekli uygunluk ve dışlama koşulları
- Critical hata riski taşıyan maddi iddialar
- Resmî temsil ile doğrulanmış gerçeklik arasındaki farklar

Bu ilkeye:

Görevle Orantılı Gerçeklik İlkesi

denir. Gereksiz veri toplamak:

- mahremiyet,
- maliyet,
- güncellik,
- çelişki

riskini artırabilir.

7. TRUTH PACK'İN ON İKİ ÇEKİRDEK BİLEŞENİ

7.1. Kanonik Varlık Kimliği

Bölüm 3'te tanımlanan:

- varlık kimliği,
- türü,
- hukuki adı,
- marka ve alan adı ilişkileri,
- ana ve bağlı kuruluşları,
- zaman ve yargı sınırları

pakete bağlanır.

7.2. Referans İddia Defteri

Denetimde kullanılacak bütün maddi gerçeklik iddialarını atomik yapıda taşır.

7.3. Resmî Temsil Kaydı

Varlığın kendisi hakkında kanonik yüzeylerde yayımladığı güncel ve tarihsel beyanları taşır.

7.4. Kanıt Sicili

Her kanıt nesnesinin:

- kaynağı,
- bütünlüğü,
- yetkisi,
- erişimi,
- kapsamı,
- zamanı

kaydedilir.

7.5. Karşı Kanıt Sicili

İddiayı:

- çürüten,
- sınırlandıran,
- güncelliğini değiştiren,
- başka açıklama getiren

kayıtları taşır.

7.6. Kaynak Soyu Grafiği

Kaynakların:

- ilk kökünü,
- kopyalarını,
- alıntılarını,
- sendikasyonunu,
- AI türevlerini

gösterir.

7.7. Kapsam ve Yargı Matrisi

İddianın:

- hangi ülke,
- ürün,
- hizmet,
- kullanıcı,
- hukuki varlık

için geçerli olduğunu gösterir.

7.8. Zaman ve Güncellik Matrisi

İddianın:

- ne zaman başladığını,
- sona erdiğini,
- en son ne zaman doğrulandığını,
- hangi olayla değişebileceğini

gösterir.

7.9. Uyuşmazlık ve Belirsizlik Defteri

Çelişkili, çözülemeyen veya yetersiz kanıtlı alanları görünür kılar.

7.10. Kanıt Erişim Katmanı

Kanıtın:

- kamuya açık,
- audit erişimli,
- kısıtlı,
- gizli,
- türetilmiş

olduğunu gösterir.

7.11. Değişiklik ve İtiraz Günlüğü

Yeni kanıt, düzeltme, karşı görüş ve karar değişikliklerini sürümlü biçimde taşır.

7.12. Makine Manifesti

Bütün bileşenleri:

- kimlik,
- sürüm,
- hash,
- ilişki

içinde birbirine bağlar.

8. ATOMİK REFERANS İDDİASI

Truth Pack içindeki temel gerçeklik birimi:

Atomik Referans İddiası

olarak adlandırılır. Aday yapı:

```
C_j = (s, p, o, q, c, l, v_f, v_t, k, m)
```

Burada:

- s: özne varlık
- p: yüklem veya ilişki
- o: nesne veya değer
- q: nitelendiriciler ve sınırlar
- c: ülke veya yargı alanı
- l: dil veya locale, maddi ise
- vf: geçerlilik başlangıcı
- vt: geçerlilik sonu
- k: iddia türü
- m: maddilik veya risk etiketi

Örnek: Özne: Asteron Türkiye Turizm Ltd. / Yüklem: hizmet verir / Nesne: kurumsal seyahat yönetimi / Ülke: Türkiye / Müşteri türü: kurumsal müşteriler / Başlangıç: 1 Ocak 2026 / Son

doğrulama: 18 Ağustos 2026 Şu cümle yeterince atomik değildir: “Asteron dünya çapında çok sayıda hizmet sunan lider bir şirkettir.” Bu tek cümlede:

- coğrafi kapsam,
- hizmet sayısı,
- liderlik,
- şirket kimliği

birbirine karışmıştır.

9. İDDİA TÜRÜ

Her atomik iddia bir tür taşımalıdır.

CT-01 — Kimlik İddiası

Varlığın kim olduğu veya hangi varlığa ait olduğu.

CT-02 — İlişki İddiası

Sahiplik, kontrol, partnerlik, müşterilik, istihdam veya üyelik ilişkisi.

CT-03 — Faaliyet ve Hizmet İddiası

Varlığın ne yaptığı, hangi hizmeti sunduğu.

CT-04 — Ürün İddiası

Ürün özelliği, sürümü, sahibi veya kullanılabilirliği.

CT-05 — Coğrafi Kapsam İddiası

Ülke, bölge veya hizmet alanı.

CT-06 — Hukuki ve Lisans İddiası

Kayıt, yetki, lisans, düzenleyici statü.

CT-07 — Fiyat ve Ticari Koşul İddiası

Fiyat, vergi, abonelik, iade, garanti veya ödeme.

CT-08 — Nicel Performans İddiası

Başarı oranı, teslim süresi, kullanıcı sayısı, büyüme veya verimlilik.

CT-09 — Müşteri ve Partner İddiası

Müşteri, sponsor, partner, distribütör veya tedarikçi ilişkisi.

CT-10 — Sertifika ve Ödül İddiası

Sertifika, üyelik, ödül veya akreditasyon.

CT-11 — Karşılaştırmalı veya Üstünlük İddiası

En iyi, lider, daha hızlı, en güvenilir, bir numara.

CT-12 — Tavsiye ve Uygunluk İddiası

Belirli kullanıcı için uygunluk veya tavsiye.

CT-13 — Tarihsel İddia

Kuruluş, geçmiş sahiplik, eski ürün veya tarihsel olay.

CT-14 — Risk, Şikâyet veya Uyuşmazlık İddiası

Şikâyet, soruşturma, mahkeme, yaptırım veya tartışma.

CT-15 — Görüş veya Kurumsal Pozisyon

Varlığın kendisi hakkında ileri sürdüğü görüş, vizyon veya değer yargısı. İddia türü, hangi kanıtın yeterli sayılacağını belirler.

10. İDDİA MADDİLİĞİ

Truth Pack içindeki her doğru bilgi aynı karar etkisine sahip değildir. Maddilik aday olarak şu sınıflarla gösterilebilir:

CRITICAL-CANDIDATE

MAJOR-CANDIDATE

MODERATE

ADVISORY

CONTEXTUAL

Kesin hata önem seviyesi Bölüm 15'te belirlenecektir. Truth Pack'teki etiket yalnız:

- kanıt önceliğini,
- hakemlik dikkatini,
- güncellik sıklığını

yönetir.

11. KANIT NESNESİ

Truth Pack içindeki her kaynak, yalnız bağlantı olarak değil Kanıt Nesnesi olarak kaydedilmelidir. Aday yapı:

```
E_k = (r, a, i, p, f, s, t, x, g, h)
```

Burada:

- r: kaynak rolü
- a: iddia için yetki
- i: bağımsızlık
- p: iddiaya yakınlık
- f: güncellik
- s: kapsam

- t: zaman
- x: erişim sınıfı
- g: kaynak soyu
- h: bütünlük kaydı

Bir URL tek başına bu alanların hiçbirini tamamlamaz.

12. KANIT–İDDİA İLİŞKİSİ

Her kanıt nesnesi, her iddiayı desteklemez. İddia Cj ile kanıt Ek arasındaki ilişki şu statülerden biri olabilir:

SUPPORTS

PARTIALLY_SUPPORTS

QUALIFIES

CONTRADICTS

OUTDATES

CONTEXTUALIZES

NO_MATERIAL_BEARING

UNKNOWN_RELATION

Örnek: Bir şirketin marka tescil belgesi: marka sahipliğini destekleyebilir, müşteri memnuniyetini desteklemez. Bir müşteri yorumu: tek kullanıcının deneyimini destekleyebilir, bütün müşterilerde yüzde 98 başarıyı desteklemez.

13. KAYNAK ROLLERİ

SR-01 — Kanonik Birinci Taraf Kaynak

Varlığın kontrol ettiği resmî yüzeydir. Örnek:

- Kanonik web sayfası
- Resmî politika
- Fiyat sayfası
- Ürün belgesi

SR-02 — Birinci Taraf Operasyon Kaydı

Varlığın iç veya işlem düzeyindeki kayıtlarıdır. Örnek:

- Fatura
- Sözleşme
- İşlem kaydı
- Teslimat kaydı

- Ham performans verisi

SR-03 — Yetkili Resmî Kayıt

Konuyla ilgili yasal veya kurumsal yetki taşıyan sicildir. Örnek:

- Şirket sicili
- Lisans sicili
- Sertifika veren kurum
- Mahkeme kararı
- Düzenleyici kayıt

SR-04 — İki Taraflı Doğrulama

İlişkinin her iki tarafça veya sözleşmeyle desteklenmesidir. Örnek:

- Müşteri ilişkisi
- Partnerlik
- Sponsorluk
- Distribütörlük

SR-05 — Bağımsız Editoryal Kaynak

Varlıktan editoryal olarak bağımsız yayın veya araştırmadır. Bağımsızlık doğruluk garantisi değildir.

SR-06 — Bağımsız Denetim veya Araştırma

Yöntemi ve kapsamı açık dış değerlendirmedir.

SR-07 — İşlemsel veya Davranışsal Kanıt

Gerçek işlem veya kullanıcı davranışını gösteren kayıtlar.

SR-08 — Kullanıcı Tarafından Üretilen Kanıt

Yorum, şikâyet, inceleme veya kullanıcı belgesi. Tekil deneyimdir.

SR-09 — Türetilmiş Hesaplama

Ham kayıtlardan üretilmiş oran, sınıflandırma veya özet. Formülü ve girdileri bulunmalıdır.

SR-10 — Sentetik Kanıt

Yöntem gösterimi için oluşturulmuş veridir. Gerçek varlık iddiasını kanıtlayamaz.

SR-11 — AI Tarafından Üretilmiş Türev Kaynak

Bir AI sisteminin başka kaynakları özetleyerek oluşturduğu içeriktir. İlk kaynakların yerine geçmez.

SR-12 — Bilinmeyen veya Çözülemeyen Kaynak

Kaynak kontrolü veya kökeni belirlenememiştir. Ana maddi doğrulamada sınırlı güven taşır.

14. KAYNAĞIN YETKİSİ İDDİA TÜRÜNE BAĞLIDIR

Bütün kaynaklar için evrensel ve tek bir sıralama yoktur. Örnek: Şirketin güncel kamu fiyatı için:

- kanonik fiyat sayfası,

- gerçek checkout,
- yürürlükteki teklif

güçlü kaynak olabilir. Şirketin: "Dünya lideriyiz." iddiası için aynı web sayfası yalnız öz beyanı gösterir. Liderliği kanıtlamaz. Lisans için: lisansı veren kurum şirket blogundan daha yetkilidir. Müşteri deneyimi için: tek müşteri yorumu gerçek tekil olaydır, bütün kullanıcı nüfusunu temsil etmez. Bu nedenle:

> *Kaynağın kalitesi genel değil, iddia özelidir.*

15. KAYNAK PROFİLİNİN YEDİ BOYUTU

Her kaynak şu boyutlarda değerlendirilmelidir:

15.1. Yetki

Kaynak iddia türü hakkında hüküm vermeye yetkili mi?

15.2. Yakınlık

Kaynak olaya, işleme veya kayda ne kadar yakın?

15.3. Bağımsızlık

Kaynak iddianın ticari veya kurumsal sahibinden ne ölçüde bağımsız?

15.4. Bütünlük

Dosyanın veya kaydın özgünlüğü ve değişiklik geçmişi korunuyor mu?

15.5. Güncellik

Kaynak ilgili zaman için geçerli mi?

15.6. Kapsam

Kaynak iddianın bütün kapsamını mı, yalnız bir bölümünü mü destekliyor?

15.7. Yeniden Üretilebilirlik

Başka bir denetçi aynı kayıttan aynı hükme ulaşabilir mi? Bu boyutlar tek bir "kaynak puanına" zorunlu olarak indirgenmemelidir. Bir kaynak:

- yüksek yetkili,
- düşük bağımsız

olabilir. Başka kaynak:

- yüksek bağımsız,
- konuya uzak

olabilir.

16. BİRİNCİ TARAF KAYNAĞIN DOĞRU ROLÜ

Birinci taraf kaynak: otomatik olarak değersiz değildir, otomatik olarak bağımsız kanıt da değildir. Birinci taraf kaynaklar şu iddialar için güçlü olabilir:

- Kurumun yayımladığı fiyat

- Destek saatleri
- Kendi ürünü ve hizmet tanımı
- Kendi politikasının metni
- Kamuya açıkladığı resmî görüş
- Kendi işlem verisi

Şu iddialar için tek başına yeterli olmayabilir:

- Dünya liderliği
- En yüksek memnuniyet
- Rakiplerden üstünlük
- Evrensel başarı
- Bağımsız akreditasyon
- Bütün müşterilerde ticari sonuç

Önceki "99 Hata" metnindeki: "Kendi siten bir iddiadır; kanıt değildir." uyarısı özellikle üstünlük ve bağımsızlık gerektiren iddialarda merkezi önem taşır. Daha tam normatif ifade şudur:

> *Kendi site, kurumun ne söylediğinin kanıtıdır; dış dünyaya ilişkin her iddianın bağımsız doğruluğunun otomatik kanıtı değildir.*

17. BAĞIMSIZLIK NEDİR?

Bir kaynağın farklı alan adında bulunması onu bağımsız yapmaz. Bağımsızlık şu alanlarla ilgilidir: İçerik kim tarafından üretildi? Yayın kararını kim verdi? Ödeme veya sponsorluk var mı? Kaynak şirketten doğrudan metin aldı mı? Yayın kendi doğrulamasını yaptı mı? İçerik başka kaynaktan kopya mı? Kaynağın şirketle mülkiyet veya iş ilişkisi var mı? Bir ajansın farklı alan adlarında yayımladığı on içerik: on bağımsız kaynak değildir.

18. KAYNAK SOYU

Kanıt nesneleri bir kaynak soyuna bağlanmalıdır. Aday grafik:

G_S = (V_S, E_S) [K10]

Burada:

- VS: kaynak nesneleri
- ES: alıntı, kopyalama, sendikasyon, özetleme ve türetme ilişkileri

Kaynak ilişkileri:

ORIGINAL_SOURCE

QUOTES

SYNDICATES

SUMMARIZES

TRANSLATES

DERIVES_FROM

AI_GENERATED_FROM

UNKNOWN_LINEAGE

olarak kaydedilebilir.

19. BAĞIMSIZ KANIT AİLESİ

Aynı kök kaynaktan türeyen içerikler tek bir Kanıt Ailesi olarak değerlendirilebilir. Örnek:

- Şirket basın bülteni
- Bülteni aynen yayımlayan haber sitesi
- Haber sitesini özetleyen blog
- Blogu özetleyen AI metni

Dört URL vardır. Bir bağımsız kanıt ailesi bile bulunmayabilir. Kaynak sayısı ile kanıt ailesi sayısı birlikte raporlanmalıdır.

20. KANIT AKLAMA

Bir öz beyanın başka alan adlarında yeniden yayımlanıp daha sonra bağımsız kanıt gibi kullanılması:

Kanıt Aklama

olarak adlandırılır. Örnek zincir: Şirket kendi sitesinde "lideriz" der. Ücretli içerik bunu "haber" olarak yayımlar. Başka siteler haberi kopyalar. AI çok sayıda alan adı gördüğü için iddiayı yaygın kabul eder. Şirket AI'nın cevabını kendi iddiasının kanıtı gibi kullanır. Bu kapalı döngü:

öz beyan → türev kaynak → AI cevabı → yeniden öz beyan

oluşturur. Truth Pack kaynak soyunu kurarak bu döngüyü görünür hâle getirmelidir.

21. AI TARAFINDAN ÜRETİLMİŞ KAYNAĞIN SINIRI

Bir AI sistemi:

- kaynakları özetleyebilir,
- iddiaları karşılaştırabilir,

yeni metin oluşturabilir. AI metni:

- orijinal kaynak değildir,

- altta yatan kanıtın yerini alamaz,

başka AI'nın aynı şeyi söylemesiyle bağımsız doğrulamaya dönüşmez. AI çıktısı şu amaçlarla kullanılabilir:

- araştırma yönlendirmesi,
- iddia çıkarımı,
- çelişki tespiti,

kaynak bulma yardımı. Nihai kanıt nesnesi, mümkün olduğunca ilk kaynağa bağlanmalıdır.

22. SENTETİK KANITIN SINIRI

Sentetik veri:

- yöntem gösterimi,
- sistem testi,
- yazılım doğrulaması,
- hakem eğitimi

için kullanılabilir. Sentetik veri:

- gerçek müşteri sayısını,
- gerçek performans oranını,
- gerçek ülke kapsamını,
- gerçek lisansı

kanıtlayamaz. Sentetik kayıtlar gerçek kanıtla aynı dosya içinde olsa bile açıkça etiketlenmelidir.

23. DESTEKLEYEN KANIT YETERLİ DEĞİLDİR

Truth Pack yalnız iddiayı destekleyen kaynakları toplamamalıdır. Maddi iddialar için aktif olarak şu kayıtlar aranmalıdır:

- Çelişen tarih
- Farklı kapsam
- Süresi dolmuş belge
- Başarısız işlem
- Şikâyet veya iade
- Eski sahiplik
- Başka varlığa ait lisans
- Düzeltme veya geri çekme
- Karşılaştırmalı veri
- Eksik payda

Bu ilkeye:

Karşı Kanıt Yükümlülüğü

denir.

24. KARŞI KANIT

Karşı kanıt iddiayı her zaman tamamen çürütmez. Şu işlevlerden birini taşıyabilir:

CONTRADICTS

LIMITS_SCOPE

LIMITS_TIME

IDENTIFIES_EXCEPTION

CHALLENGES_METHOD

QUESTIONS_AUTHENTICITY

PROVIDES_ALTERNATIVE_ENTITY

REQUIRES_ADDITIONAL_REVIEW

Örnek: "Bütün müşteriler memnun" iddiasına bir olumsuz yorum, bütün iddiayı otomatik çürütmeyebilir. Fakat: "Bütün" ifadesinin yanlış olduğunu gösterebilir.

25. KARŞI KANITIN SAKLANMASI

İddiayı zayıflatan kayıt:

- şirket lehine olmadığı,
- düşük kaliteli göründüğü,
- kamuya açıklanmasının rahatsızlık yaratacağı

için sessizce silinemez. Düşük güvenli karşı kanıt da kaydedilebilir. Statüsü açıkça belirtilir. Hakem ona uygun ağırlığı verir.

26. İDDİA DESTEK STATÜLERİ

VS-0 — NOT ASSESSED

İddia henüz değerlendirilmemiştir.

VS-1 — OFFICIAL CLAIM ONLY

İddia yalnız resmî öz beyana dayanmaktadır. Yanlış olduğu kanıtlanmış değildir. Ek doğrulama yoktur.

VS-2 — TRACEABLE FIRST-PARTY SUPPORT

İddia izlenebilir birinci taraf operasyon veya işlem kaydıyla desteklenmektedir.

VS-3 — CORROBORATED

İddia birden fazla uygun kanıtla desteklenmektedir. Kaynakların tamamı bağımsız olmak zorunda değildir.

VS-4 — INDEPENDENTLY VERIFIED

İddia, konuya uygun bağımsız veya yetkili dış kaynakla doğrulanmıştır.

VS-5 — MULTI-SOURCE SUSTAINED

İddia farklı tür ve kaynak ailelerinde, kapsam ve zaman içinde güçlü biçimde desteklenmektedir.

VS-C — CONTRADICTED

Daha güçlü veya daha uygun kanıt iddiayla çelişmektedir.

VS-U — UNRESOLVED

Kanıtlar yetersiz, çelişkili veya karar verilemeyecek düzeydedir.

VS-W — WITHDRAWN

İddia sahibi iddiayı geri çekmiştir. Tarihsel kayıt korunur.

VS-R — RESTRICTEDLY VERIFIED

İddia kısıtlı denetçi erişimindeki kanıtla desteklenmiştir. Kamuya açık doğrulama mümkün değildir. Bu statü, AI'nın kamu kaynaklarından iddiayı üretmesinin beklenip beklenemeyeceğini ayrıca etkiler.

27. DESTEK STATÜSÜ TEK BAŞINA YETERLİ DEĞİLDİR

Bir iddia aynı anda:

- destekli,
- yalnız belirli ülke için,
- yalnız belirli tarih aralığında,
- kamuya kapalı kanıtla,
- karşı görüş altında

olabilir. Bu nedenle tek statü yerine çok boyutlu Gerçeklik Durum Vektörü kullanılabilir:

```
V_j = (support, scope, time, conflict, access)
```

Örnek:

VS-4 + KAPSAM-LIMITED + GÜNCEL + NO-CONFLICT + PUBLIC

başka bir sonuçtur.

VS-R + KÜRESEL-CLAIM + TIME-LIMITED + DISPUTED + RESTRICTED

başka bir sonuçtur.

28. KAPSAM STATÜLERİ

SC-0 — KAPSAM UNKNOWN

İddianın uygulanacağı kapsam belirlenememiştir.

SC-1 — EXACT KAPSAM

İddia ve kanıt aynı varlık, ürün, ülke ve kullanıcı kapsamını taşır.

SC-2 — PARTIAL KAPSAM

Kanıt iddianın yalnız bir bölümünü destekler.

SC-3 — OVERBROAD CLAIM

İddia kanıtın taşıdığından daha geniştir.

SC-4 — WRONG VARLIK KAPSAM

Kanıt başka varlık veya bağlı kuruluşa aittir.

SC-5 — WRONG JURISDICTION

Kanıt başka ülke veya hukuk alanına aittir.

SC-6 — WRONG PRODUCT OR SERVICE

Kanıt farklı ürün veya hizmete aittir.

29. ZAMAN STATÜLERİ

TS-0 — TIME UNKNOWN

İddianın geçerlilik zamanı bilinmemektedir.

TS-1 — GÜNCEL

Referans tarihinde geçerli olduğu desteklenmektedir.

TS-2 — HISTORICAL

Belirli geçmiş dönem için desteklenmektedir.

TS-3 — EXPIRED

Belge veya iddia süresi dolmuştur.

TS-4 — SUPERSEDED

Daha yeni kayıt tarafından değiştirilmiştir.

TS-5 — EVENT-DEPENDENT

Bir olay veya koşul devam ettiği sürece geçerlidir.

TS-6 — FUTURE OR PLANNED

Henüz gerçekleşmemiş plan veya taahhüttür. Güncel gerçeklik gibi sunulamaz.

30. DÖRT ZAMAN DAMGASI

Her kanıt ve iddia ilgili olduğu ölçüde dört ayrı zaman taşımalıdır:

30.1. Gerçekliğin Geçerli Olduğu Zaman

İddia dünyada ne zaman doğruydu?

30.2. Kaynağın Yayımlandığı Zaman

Bilgi ne zaman kamuya veya denetçiye sunuldu?

30.3. Kaynağın Capture Edildiği Zaman

Denetim kanıtı ne zaman kaydedildi?

30.4. İddianın Hakemlik Zamanı

Gerçeklik statüsü ne zaman verildi? Bu ayrım önemlidir. Bir lisans 10.00'da iptal edilmiş olabilir. Sicil 15.00'te güncellenmiş olabilir. AI yanıtı 12.00'de verilmiş olabilir. Yanıt:

- dünya gerçekliği bakımından eski,
- kamuya erişilebilir kanıt bakımından anlaşılabilir

olabilir. Truth Pack iki durumu birbirinden ayırmalıdır.

31. KANITIN KAMUYA ERİŞİLEBİLİRLİĞİ

Kanıt erişim sınıfları:

EA-1 — PUBLIC

Kamuya açık ve denetlenebilir.

EA-2 — PUBLIC BUT TECHNICALLY LIMITED

Kamuya açık fakat:

- giriş,
- coğrafya,
- format,
- teknik erişim

sınırlaması taşır.

EA-3 — AUDIT ACCESS

Bağımsız denetçilere kontrollü biçimde açılabilir.

EA-4 — RESTRICTED CONFIDENTIAL

Sözleşme, kişisel veri veya ticari sır nedeniyle kısıtlıdır.

EA-5 — DERIVED PUBLIC SUMMARY

Ham kanıt gizlidir. Yöntemi ve özeti kamuya açıktır.

EA-6 — UNAVAILABLE OR LOST

Kanıtın bulunduğu ileri sürülmektedir fakat denetimde erişilememiştir.

32. KAMUYA AÇIK DOĞRULUK İLE KISITLI DOĞRULUK

Bir şirket özel sözleşmeyle: "X müşterimizdir." iddiasını doğrulayabilir. Sözleşme kamuya açılamıyor olabilir. Truth Pack iddiayı: VS-R — Restrictedly Verified olarak kaydedebilir. Fakat kamu AI

ürününün bu ilişkiyi bilmesi beklenemez. Bu nedenle iki ayrı soru vardır: İddia gerçek mi? İddia kamuya keşfedilebilir mi? Bir iddia özel kanıtla doğru olabilir. Kamu GEO başarısı için yine:

- kamuya açık,
- doğrulanabilir,
- kullanıcıya uygun

bir kayıt gerekebilir.

33. KAMUYA KEŞFEDİLEBİLİRLİK

Her maddi iddia için aday olarak şu alan kaydedilebilir:

PUBLICLY_DISCOVERABLE

PUBLIC_BUT_OBSCURE

PUBLICLY_CONFLICTED

RESTRICTED_ONLY

NOT_PUBLIC

UNKNOWN

AI yanıtının gerçekliği ve bilgiye erişebilirliği daha sonra ayrı değerlendirilebilir.

34. DOĞRU FAKAT KAMUYA KAPALI İDDİA

Bir AI ürünü kamuya kapalı bir bilgiyi doğru biçimde söylerse şu sorular doğar: Bilgi başka kamu kaynağında var mı? AI kişiselleştirme veya özel bilgi tabanı mı kullandı? Cevap tesadüfen mi doğru? Mahrem veya gizli bilgi sızıntısı mı var? İddia yalnız sonuçtan sonra mı doğrulandı? Truth Pack doğru olmayı tek başına başarı saymadan kaynak yolunu ayrıca kaydetmelidir.

35. AÇIK DÜNYA VE KAPALI DÜNYA İDDİALARI

Bir kayıtta bilgi bulunmamasının anlamı iddia türüne bağlıdır.

35.1. Açık Dünya İddiası

Kayıt evreninin tamamı bilinmez. Örnek:

- Bütün müşteriler
- Bütün ülkelerdeki iş ilişkileri
- Bütün kullanıcı şikâyetleri
- Bütün medya yayınları

Bir kaydın bulunmaması iddianın yanlış olduğunu kanıtlamaz.

35.2. Kapalı Dünya İddiası

Yetkili ve kapsamı açık sicilin ilgili evreni tamamıyla kapsadığı varsayılabilir. Örnek:

- Belirli tarihte aktif lisans listesi
- Resmî şirket kayıt statüsü
- Belirli sertifika veren kurumun geçerli sertifikaları
- Yayımlanmış ürün kataloğu, gerçekten tamamlayıcıysa

Bu durumda kayıt yokluğu daha güçlü negatif kanıt olabilir.

36. KAPALI DÜNYA YETKİ BEYANI

Bir kaynak kapalı dünya kaydı olarak kullanılacaksa şu alanlar tanımlanmalıdır: Hangi evreni kapsıyor? Hangi tarih için? Güncelleme sıklığı nedir? Kayıt eksik olabilir mi? Arşiv ve değişiklik kaydı var mı? Yetkili kurum kimdir? Yalnız "resmî site" olması kapalı dünya yetkisi sağlamaz.

37. KANIT BULUNAMADI İLE YANLIŞLIK AYRIMI

Aşağıdaki üç cümle farklıdır:

- "İddia için kanıt bulunamadı."
- "İddia mevcut kanıtlarla desteklenmiyor."
- "İddia daha güçlü kanıtlarla çelişiyor."

Birincisi araştırma sonucu olabilir. İkincisi doğrulama statüsüdür. Üçüncüsü maddi çelişkidir. Şunlar aynı değildir:

NOT FOUND

UNVERIFIED

CONTRADICTED

38. YOKLUK KANITI

Bir iddianın yanlışlığını yokluk üzerinden kurmak için şu sorular cevaplanmalıdır: Aranan kayıt evreni tamamlayıcı mı? Doğru varlık ve isim varyantları arandı mı? Doğru tarih ve ülke kullanıldı mı? Kayıt gecikmeli güncelleniyor olabilir mi? Başka hukuki varlık veya marka kullanılıyor mu? Erişim sorunu var mı? Bu koşullar olmadan: "Sicilde bulamadık, lisans yoktur." gibi kesin hüküm kurulamaz. Daha doğru sonuç: "Belirtilen sicilde ve tarihte doğrulanamadı." olabilir.

39. İDDİA–KANIT MATRİSİ

İddia türü	Öncelikli kanıt	Resmî sitenin rolü	Ek gereksinim
Hukuki ad	Yetkili sicil	Kurum beyanını gösterir	Güncel sicil
Marka sahipliği	Marka kaydı, devir veya lisans	Marka kullanımını gösterir	Hak sahibi doğrulaması
Güncel fiyat	Kanonik fiyat, checkout, sözleşme	Birincil kamu kaydı olabilir	Vergi ve ek ücret
Hizmet ülkesi	Sözleşme, operasyon, fatura, yerel kayıt	Kamu beyanı	"Hizmet" tanımı
Lisans	Lisans veren kurum	Öz beyan	Kapsam ve süre
Müşteri ilişkisi	Sözleşme, fatura, iki taraflı onay	Tek taraflı beyan	Tarih ve ilişki türü

İddia türü	Öncelikli kanıt	Resmî sitenin rolü	Ek gereksinim
Partnerlik	Sözleşme veya iki taraflı açıklama	Öz beyan	Kapsam ve süre
Çalışan sayısı	İK ve istihdam kayıtları	Yaklaşık beyan	Grup/şirket ayrımı
Başarı oranı	Ham veri, payda, yöntem	Sonuç beyanı	Başarısız vakalar
Memnuniyet	Temsilî anket ve yöntem	Sonuç açıklaması	Yanıtlamama
Sertifika	Sertifikayı veren kurum	Öz beyan	Kapsam, tarih
Üyelik	Üyelik kurumu	Öz beyan	Üyelik ≠ akreditasyon
Liderlik	Karşılaştırmalı yöntem ve veri	Pazarlama iddiası	Evren, kriter, tarih
Tavsiye	Kullanıcı profili ve kanıt	Pazarlama görüşü	Uygunluk ve dışlamalar
Şikâyet	Doğrulanmış kullanıcı veya kurum kaydı	Şirket cevabı	Tekil olay sınırlaması
Hukuki ihlal	Kesinleşme ve yetki statüsü açık resmî kayıt	Kurum görüşü	Soruşturma ≠ hüküm

Bu tablo evrensel ve değişmez kaynak sıralaması değildir. İddia türünün hangi kanıt yükünü taşıdığını gösterir.

40. NİCEL İDDİA

Bir nicel iddia aşağıdaki alanları taşımadan doğrulanmış sayılamaz:

- Ölçülen şey
- Pay
- Payda
- Dönem
- Dahil etme kuralı
- Dışlama kuralı
- Veri kaynağı
- Eksik kayıtlar
- Benzersiz birim
- Hesaplama formülü
- Yuvarlama
- Güncelleme tarihi
- Sorumlu insan

Örnek: “Yüzde 98 başarı.” tek başına nicel iddia değildir. Tanımlanmamış pazarlama sayısıdır.

41. PAY VE PAYDA

Başarı oranı:

R = S/N

olsun. Burada:

- S: başarılı sayılan birim
- N: değerlendirilen bütün uygun birim

olmalıdır. Başarısız veya tamamlanmamış projeler paydadan çıkarılıyorsa oran değişir. Payda sonuçtan sonra seçilemez.

42. BENZERSİZ BİRİM

Şu ifadeler birbirinden farklıdır:

- 100 müşteri
- 100 proje
- 100 fatura
- 100 sözleşme
- 100 kullanıcı hesabı
- 100 işlem

Aynı müşteri on proje yapmış olabilir. “100 proje”: “100 müşteri” değildir. Truth Pack sayının birimini açıkça kaydetmelidir.

43. KÜMÜLATİF VE AKTİF SAYI

“500 müşterimiz var.” şu anlamlardan birini taşıyabilir:

- Şirket tarihinde 500 benzersiz müşteri
- Son bir yılda 500 müşteri
- Şu anda 500 aktif müşteri
- 500 iletişim kaydı
- 500 hesap
- 500 proje

Bu anlamlar aynı değildir.

44. TÜRETİLMİŞ İDDİA

Bir iddia doğrudan belgede yazmıyor olabilir. Ham veriden hesaplanabilir. Örnek: “Ortalama teslim süresi 12 gündür.” Bu sonuç:

- hangi projelerin dâhil edildiğini,
- başlangıç ve bitiş tanımını,
- başarısız projeleri,
- medyan ve uç değerleri

gerektirir. Türetilmiş iddia şu alanları taşımalıdır:

- Girdi veri kimlikleri
- Formül
- Kod veya hesap yöntemi
- Sürüm
- Yeniden üretim kaydı

45. PERFORMANS İDDİASI

Performans iddiası şu soruları cevaplamalıdır: Sonuç kimin performansı? Hangi ürün veya hizmet? Hangi kullanıcı veya proje tipi? Baseline nedir? Başarı kriteri nedir? Dönem nedir? Kaç başarısız sonuç vardır? Seçilmiş vaka mı, bütün nüfus mu? Bağımsız denetim var mı? Sonuç nedensel mi, yalnız

ilişkisel mi? Bir vaka çalışması: mümkün bir sonucu gösterebilir. Bütün müşterilerde beklenen sonucu göstermez.

46. GARANTİ İDDİASI

"Garantili sonuç." ifadesi şu alanları gerektirir: Tam olarak ne garanti ediliyor? Teslimat mı, ticari sonuç mu? Süre nedir? İstisnalar nelerdir? İade veya telafi nedir? Kontrol dışı üçüncü taraflar var mı? AI tavsiye veya sıralaması garanti ediliyor mu? Bir şirket:

- kendi iş akışını,
- belirli teslimatını

garanti edebilir. Kontrol etmediği AI ürününün gelecekteki tavsiye davranışını mutlak biçimde garanti edemez.

47. MÜŞTERİ İDDİASI

Bir kuruluşun logosunu web sitesinde göstermek müşteri ilişkisini otomatik kanıtlamaz. Logo şu ilişkilerden birini temsil edebilir:

- Müşteri
- Eski müşteri
- Potansiyel müşteri
- Etkinlik katılımcısı
- Sponsor
- Partner
- Tedarikçi
- Teknoloji kullanıcısı
- Grup şirketi
- İzin verilmemiş logo kullanımı

Truth Pack ilişki türünü açıkça tanımlamalıdır.

48. MÜŞTERİ GİZLİLİĞİ

Müşteri ilişkisi sözleşmeyle doğrulanmış fakat kamuya açıklanamaz olabilir. Doğru kayıt: VS-R — Restrictedly Verified olabilir. Şirket kamuya: "X müşterimizdir." diyemiyorsa AI'nın bu ilişkiyi bilmesi beklenemez. Gizli müşteri sayısı kamu GEO otoritesine sessizce dönüştürülemez.

49. PARTNERLİK İDDİASI

"Partner" kelimesi tek başına yeterli değildir. İlişki şu alanları taşımalıdır:

- Partnerlik türü
- Ürün veya proje
- Ülke
- Başlangıç ve bitiş
- İki taraflı onay

- Ticari veya teknik kapsam
- Yetki ve sorumluluk

Bir yazılım aboneliği: teknoloji ortaklığı olmayabilir. Bir etkinlik sponsorluğu: uzun dönem stratejik partnerlik olmayabilir.

50. LİSANS VE YETKİ İDDİASI

Lisans iddiası en azından şu alanları taşımalıdır:

- Lisans sahibi tam varlık
- Lisansı veren kurum
- Lisans numarası
- Ülke veya yargı alanı
- Yetki kapsamı
- Başlangıç ve sona erme tarihi
- Askı veya iptal durumu
- Doğrulama tarihi

Başka grup şirketinin lisansı markaya aktarılamaz. Bir çalışanın kişisel lisansı kuruma otomatik yetki vermez.

51. SERTİFİKA, ÜYELİK VE AKREDİTASYON

Şu kavramlar farklıdır:

- Sertifika
- Üyelik
- Akreditasyon
- Ödül
- Eğitim katılımı
- Uygunluk işareti
- Öz değerlendirme rozeti

Bir meslek kuruluşuna üyelik: o kuruluş tarafından akredite edilmiş uzmanlık anlamına gelmeyebilir. Bir eğitimi tamamlamak: bağımsız sertifikasyon olmayabilir. Truth Pack terimi veren kurumun tanımına göre kullanmalıdır.

52. ÖDÜL İDDİASI

Ödül kaydı şu alanları taşımalıdır:

- Ödülü veren kurum
- Yıl
- Kategori
- Coğrafya
- Başvuru veya ücret yapısı
- Değerlendirme yöntemi
- Sponsor ilişkisi

- Kazanan varlık
- Ürün veya şirket kapsamı

Ücretli başvuru ödülü otomatik olarak sahte değildir. Bağımsız pazar liderliği de değildir. Kaynak statüsü korunmalıdır.

53. ÜSTÜNLÜK İDDİASI

"En iyi." "Dünya lideri." "Bir numara." "En güvenilir." ifadeleri şu alanları gerektirir:

- Karşılaştırma evreni
- Rakip kümesi
- Kriter
- Ağırlık
- Veri
- Dönem
- Coğrafya
- Yöntem
- Bağımsızlık
- Belirsizlik

Bu alanlar yoksa doğru Truth Pack statüsü çoğu zaman:

VS-1 — Official Claim Only

veya: VS-U — Unresolved olacaktır.

54. GÖRÜŞ VE OLGU AYRIMI

Şu cümle: "Biz insan merkezli tasarıma inanıyoruz." kurumsal görüş veya değer beyanıdır. Bunu kanıtlamak için pazar araştırması gerekmez. Şu cümle farklıdır: "Ürünlerimiz insan merkezli tasarım bakımından sektörün en iyisidir." Karşılaştırmalı olgu iddiasına dönüşür. Truth Pack her iddianın epistemik türünü korumalıdır.

55. ŞİKÂYET VE OLUMSUZ KULLANICI KAYDI

Bir kullanıcı şikâyeti:

- gerçek tekil olay olabilir,
- eksik veya yanlış olabilir,
- çözülmüş olabilir,

başka varlığa ait olabilir. Tek şikâyet: "Şirket bütün müşterileri mağdur eder." sonucunu desteklemez. Fakat ağır bir olay:

- soruşturma,
- düzeltme,
- Critical risk

başlatabilir.

56. İDDİA, SORUŞTURMA VE HÜKÜM

Şu statüler birbirinden ayrılmalıdır:

- İddia
- Şikâyet
- Açık soruşturma
- İdari bulgu
- İlk derece kararı
- Kesinleşmiş karar
- Geri çekilmiş veya bozulmuş karar
- Kurumun cevabı

Bir kişi veya kurum hakkında soruşturma bulunması: suçlu veya sorumlu olduğu anlamına gelmez. Truth Pack, hukuki statüyü aşan kesin dil kullanmamalıdır.

57. OLUMSUZ İDDİALARDA ORANTILI DİL

Truth Pack şu dili kullanmalıdır:

- "İddia edilmiştir."
- "Belirtilen kurum tarafından soruşturulmaktadır."
- "Şu tarihte belirtilen karar verilmiştir."
- "Karar kesinleşmemiştir."
- "Kurum iddiayı reddetmektedir."
- "Mevcut kayıtlarda çözülmemiştir."

Şu dil, yeterli kesin kayıt olmadan kullanılamaz:

- "Dolandırıcıdır."
- "Yasadışıdır."
- "Suçludur."
- "Kesin olarak güvenilmezdir."

Bu yalnız hukuki risk değil, temsil doğruluğu ilkesidir.

58. TAVSİYE GERÇEKLİĞİ

Truth Pack bir şirketi genel olarak: "önerilir" veya: "önerilmez" şeklinde tek statüye indirgememelidir. Tavsiye şu alanlara bağlıdır:

- Kullanıcı ihtiyacı
- Ülke
- Bütçe
- Hizmet türü
- Risk toleransı
- Kapasite
- Lisans
- Dışlama koşulları

- Alternatifler

Truth Pack tavsiye için:

- uygunluk koşulları,
- uygun olmama koşulları,
- kanıtlanmış sınırlar

sağlar. Nihai kullanıcı tavsiyesi AI ürününün karar görevidir.

59. SINIR İDDİALARI

Truth Pack yalnız varlığın ne yaptığını değil, ne yapmadığını da taşımalıdır. Örnek: Yalnız Türkiye'de hizmet verir. Hukuki danışmanlık sunmaz. Sağlık lisansı yoktur. Fiyat vergi hariçtir. Hizmet yalnız kurumsal müşterilere açıktır. Sonuç garantisi vermez. Belirli üçüncü tarafların davranışını kontrol etmez. Sınır bilgisi, doğru tavsiye için en değerli kayıtlardan biridir.

60. BİLİNMEYENLER DEFTERİ

Truth Pack eksik alanları gizlememelidir. Bilinmeyen kayıt türleri:

UNKNOWN_FACT

UNKNOWN_SCOPE

UNKNOWN_TIME

UNKNOWN_ENTITY

UNKNOWN_SOURCE_LINEAGE

UNKNOWN_PUBLIC_AVAILABILITY

INSUFFICIENT_EVIDENCE

REFERENCE_CONFLICT

NOT_APPLICABLE

NOT YET ASSESSED

UNKNOWN bir başarısızlık değildir. Sahte kesinliğe karşı korumadır.

61. REFERANS BOŞLUĞU

AI yanıtı denetim sırasında Truth Pack'te bulunmayan maddi bir iddia üretebilir. Bu durumda hakem şu seçeneklerden birini kullanmalıdır:

- İddia kapsam dışı ve maddi değil
- Truth Pack'te referans boşluğu var
- Yeni kanıt araştırması gerekli

- Yanıt desteklenemiyor
- İddia yanlış varlığa ait

Truth Pack boşluğu otomatik AI başarısızlığına dönüştürülemez. Doğru statü:

REFERENCE_GAP

olabilir.

62. CEVAPTAN SONRA TRUTH PACK GENİŞLETME

AI yanıtı daha önce düşünülmemiş maddi bir iddia üretebilir. Yeni iddia araştırılabilir. Fakat şu kurallar geçerlidir: İlk Truth Pack sürümü korunur. Yeni iddia ve kanıt ayrı sürümde eklenir. İlk hakemlik sonucu sessizce değiştirilmez. Gerekirse bütün ilgili cevaplar yeni sürümle yeniden değerlendirilir. Eski ve yeni sonuçlar kamu değişiklik günlüğünde görünür olur.

63. TRUTH PACK'İN DALGA ÖNCESİ KİLİDİ

Her ana ölçüm dalgasından önce şu alanlar kilitlenmelidir:

- Varlık kimliği sürümü
- Resmî Temsil Kaydı sürümü
- Referans İddia Defteri
- Kanıt Sicili
- Karşı Kanıt Sicili
- Kaynak Soyu Grafiği
- Kapsam ve zaman matrisi
- Bilinmeyenler Defteri
- Kanıt erişim sınıfları
- Paket hash'i
- İnsan onayı
- Kilit zamanı

Bu dosyaya:

NOMOS Truth Pack Lock

adı verilebilir.

64. TRUTH PACK VE ÖLÇÜM TARİHİ

AI yanıtı, kural olarak:

istem gönderildiği anda geçerli olan gerçeklik durumuna

göre değerlendirilmelidir. Truth Pack bu anı temsil edecek zaman dilimini taşır. Daha sonra oluşan gerçeklik: geçmiş cevabı otomatik olarak yanlış yapmaz, yeni dalga için yeni Truth Pack sürümü oluşturur.

65. DALGA SIRASINDA GERÇEKLİK DEĞİŞİKLİĞİ

Şirket:

- isim değiştirir,
- lisans alır veya kaybeder,
- fiyat günceller,
- başka ülkeye açılır,
- ürününü sona erdirir

ise Truth Pack maddi biçimde değişebilir. Bu durumda:

- ölçüm dalgası alt dalgalara ayrılabilir,
- eski ve yeni Truth Pack sürümü ayrı cevaplara bağlanabilir,

MIXED_REFERENCE_STATE uyarısı kullanılabilir.

66. TRUTH PACK'İN GERİYE DÖNÜK DEĞİŞTİRİLMESİ

Bir kayıt hatalıysa düzeltme yapılabilir. Fakat:

- eski sürüm silinemez,
- değişiklik nedeni saklanmalıdır,
- hangi cevapların etkilendiği gösterilmelidir,

yeni skor ayrı analiz sürümü oluşturmalıdır. Gerçekliği düzeltmek ile geçmişi yeniden yazmak aynı şey değildir.

67. SÜRÜM MİMARİSİ

Aday sürüm biçimi: Major.Minor.Patch

Major

Hedef varlık değişikliği İddia ontolojisi değişikliği Kanıt standardı değişikliği Aynı yanıtların farklı gerçeklik sisteminde değerlendirilmesine yol açacak kapsam değişikliği

Minor

Yeni maddi iddia Yeni ülke veya ürün eklenmesi Yeni kanıtla statü değişikliği Tarihsel veya güncel gerçeklik değişimi

Patch

Yazım Metadata Kırık bağlantı Hash veya teknik kayıt düzeltmesi Maddi gerçeklik değişikliği patch olarak saklanamaz.

68. GÜNCELLİK VE TAZELİK

Her iddia aynı hızda eskimez.

FRESH-1 — Yapısal veya Tarihsel

Kuruluş tarihi gibi düşük değişkenlikli alan. Olay tetiklemeli yeniden doğrulama yeterli olabilir.

FRESH-2 — Olay Bağımlı

Sahiplik, yönetici veya lisans statüsü. Değişiklik olayı ve düzenli kontrol gerekir.

FRESH-3 — Orta Değişkenlik

Hizmet ülkesi, çalışan sayısı, partnerlik. Periyodik doğrulama gerekir.

FRESH-4 — Yüksek Değişkenlik

Fiyat, stok, plan, ürün kullanılabilirliği. Kısa aralıkla doğrulanmalıdır.

FRESH-5 — Canlı veya Anlık

Kesinti, güncel teklif, etkinlik, gerçek zamanlı durum. Ölçüm anına yakın kayıt gerekir. Kesin süreler iddia türü ve sektör riskine göre belirlenmelidir.

69. SON DOĞRULAMA TARİHİ

Her iddia:

- en son doğrulandığı tarih,
- önerilen yeniden doğrulama tarihi,
- değişiklik tetikleyicisi

taşımalıdır. Son doğrulama tarihi olmayan iddia: güncel olarak varsayılamaz.

70. KAYNAK SNAPSHOT'I

Dinamik web sayfası yalnız URL olarak kaydedilmemelidir. İlgili olduğu ölçüde: sayfa snapshot'ı,

PDF,

tam metin, ekran görüntüsü, capture zamanı, hash,

URL,

locale, erişim durumu korunmalıdır. Sayfa sonradan değişebilir.

71. KAYNAK KAYBOLURSA

Bir kaynak silinebilir veya erişilemez olabilir. Truth Pack:

- arşiv snapshot'ını,
- hash'i,
- capture zamanını

koruyabilir. Fakat arşiv:

- sayfanın gerçek zamanda kamuya açık olduğunu,
- erişim kapsamını

tek başına tamamen kanıtlamayabilir.

72. KANIT ERİŞİM VE MAHREMİYET KATMANLARI

Public Evidence

Kamuya açık kanıt.

Audit Evidence

Bağımsız hakem ve denetçilerin erişebileceği redakte edilmiş kanıt.

Restricted Evidence

Sözleşme, kişisel veri veya ticari sır içeren kısıtlı kanıt.

Derived Evidence

Ham kanıttan üretilmiş doğrulanabilir özet veya hesaplama.

Public Claim Manifest

Kamuya her iddianın:

- statüsünü,
- kaynak sınıfını,
- tarihini,
- kanıt hash'ini,
- erişim durumunu

gösteren manifest.

73. GİZLİ KANITIN KULLANIMI

Kısıtlı kanıt kullanılacaksa: kanıtı kim gördü? neyi doğruladı? kamuya ne kadar açıklanabilir? bağımsız ikinci denetçi erişebilir mi? kanıtın süresi ve kapsamı ne? kanıt geri çekilebilir mi? kaydedilmelidir. "Gizli kanıtımız var" ifadesi tek başına Truth Pack desteği değildir.

74. DENETLENEN VARLIĞIN ROLÜ

Denetlenen varlık:

- kanıt sunabilir,
- yanlış kimlik kaydına itiraz edebilir,
- kapsamı açıklayabilir,
- gizli kanıt için kontrollü erişim sunabilir,

karşı görüş bildirebilir. Denetlenen varlık:

- kendi iddiasına nihai destek statüsü veremez,
- karşı kanıtı sildiremez,
- hakemi tek başına seçemez,
- düşük skorlu iddiayı kapsamdan çıkaramaz,

"ticari sır" diyerek bütün kanıt yükünden kaçamaz.

75. CEVAP HAKKI

Kamuya açık maddi ve olumsuz bir bulgu yayımlanmadan önce, uygulanabilir olduğu ölçüde denetlenen varlığa:

- ilgili iddia,
- kullanılan kanıt,
- karşı kanıt,
- önerilen statü,
- cevap süresi

bildirilebilir. Varlığın cevap vermemesi: iddiayı kabul ettiği anlamına gelmez. Cevabı:

- kanıt olarak,
- karşı görüş olarak,
- yeni doğrulama talebi olarak

kaydedilir.

76. İTİRAZ STATÜLERİ

NO_APPEAL

APPEAL_OPEN

EVIDENCE_SUBMITTED

UNDER_REVIEW

APPEAL_UPHELD

APPEAL_PARTIALLY_UPHELD

APPEAL_REJECTED

EXTERNAL_REVIEW_REQUESTED

CLOSED_WITH_DISSENT

İtiraz sırasında ana kayıt silinmez. Statüsü görünür kalır.

77. MUHALİF GÖRÜŞ

Hakemler veya uzmanlar aynı kanıttan farklı sonuç çıkarabilir. Maddi uyuşmazlık çözülmüyorsa:

- çoğunluk kararı,
- muhalif görüş,
- gerekçeler,
- kanıt farkı

pakette korunabilir. Gerçeklik paketi sahte oybirliği üretmemelidir.

78. TRUTH PACK YÖNETİŞİM ROLLERİ

İlgili olduğu ölçüde şu roller tanımlanabilir:

- Varlık Record Owner
- Claim Ledger Editor
- Evidence Registrar
- Source Lineage Analyst
- Domain Specialist
- Legal or Regulatory Reviewer
- Quantitative Methods Reviewer
- Counter-Evidence Reviewer
- Privacy Custodian
- Truth Pack Adjudicator
- Appeal Reviewer
- Version Custodian
- Accountable Human Approver

Küçük pilotta bir kişi birden fazla rol taşıyabilir. Çıkar çatışması açıklanmalıdır.

79. AI'NIN TRUTH PACK İÇİNDEKİ ROLÜ

AI sistemleri şu görevlerde yardımcı olabilir:

- Belgelerden iddia çıkarmak
- Aynı iddiayı farklı cümlelerde bulmak
- Kaynak soyunu tahmin etmek
- Çelişki adaylarını işaretlemek
- Tarih ve kapsam alanlarını çıkarmak
- Kanıt–iddia eşleşmesi önermek
- Eksik alanları tespit etmek
- Makine şeması oluşturmak

AI:

- nihai destek statüsünü tek başına veremez,
- kendi ürettiği özeti yeni kanıt yapamaz,
- çelişkili gerçeklikte tek ve yanılmaz otorite olamaz,

insan onayının yerine geçemez.

80. TRUTH PACK HAZIRLIK SEVİYELERİ

TPR-0 — NO PACK

Yapılandırılmış Truth Pack bulunmamaktadır. Ana hakemlik başlayamaz.

TPR-1 — OFFICIAL-ONLY DRAFT

Yalnız resmî temsil ve öz beyanlar vardır. Keşif için kullanılabilir. Doğrulanmış gerçeklik paketi değildir.

TPR-2 — STRUCTURED EVIDENCE PACK

Atomik iddialar ve temel kanıtlar bulunur. Karşı kanıt ve kaynak soyu eksik olabilir.

TPR-3 — ADJUDICATION-READY

İddia, kanıt, karşı kanıt, kapsam, zaman ve bilinmeyenler hakemliğe hazırdır. Ana GEO-1000 için aday asgari düzeydir.

TPR-4 — INDEPENDENTLY REVIEWED

Paket bağımsız alan ve yöntem incelemesinden geçmiştir.

TPR-5 — REPLICATED AND MAINTAINED

Paket farklı denetimlerde kullanılmış, sürüm ve güncellik sistemi çalıştırılmıştır. Ana kamu denetimi en azından:

TPR-3

düzeyini hedeflemelidir.

81. TRUTH PACK HAZIRLIK KAPILARI

Bir paket TPR-3 sayılmadan önce şu kapıları geçmelidir: Varlık kimliği çözümlendi mi? İstem aileleri için gerekli iddia alanları mevcut mu? Maddi iddialar atomik mi? Her iddianın zaman ve kapsamı açık mı? Kanıt nesneleri bütünlük kaydı taşıyor mu? Kaynak soyu incelendi mi? Karşı kanıt arandı mı? Kamu ve kısıtlı erişim ayrıldı mı? Bilinmeyen alanlar görünür mü? İnsan onayı ve sürüm kilidi var mı?

82. TRUTH PACK İLE İSTEM KAPSAMI EŞLEŞMESİ

Her istem ailesi gerekli Truth Pack alanlarına bağlanmalıdır.

İstem ailesi	Gerekli Truth Pack alanları
Core Mirror	Kimlik, ana faaliyet, varlık ilişkisi
Varlık Resolution	Kanonik varlık, alias, alan adı, marka
Activity and Kapsam	Ürün, hizmet, ülke, müşteri ve sınırlar
Evidence and Source	Kanıt statüsü, kaynak soyu, kamu erişimi
Local/Jurisdictional	Yerel şirket, lisans, fiyat, hukuk
Temporal	Tarihsel ve güncel kayıtlar
Tavsiye	Uygunluk, dışlama, risk, kapasite
Comparative	Karşılaştırma kriterleri ve karşılaştırılabilir kanıt
Boundary	Yapılmayan işler, kapsam dışı ülkeler, garanti sınırları

Truth Pack istemin gerektirdiği alanı taşımıyorsa referans hazır değildir.

83. SENTETİK APPLE.COM TRUTH PACK GÖSTERİMİ

SENTETİK METODOLOJİ GÖSTERİMİ / Aşağıdaki iddialar, kayıtlar ve statüler yalnız yöntemi göstermek için kurgulanmıştır. Gerçek Apple Inc. bilgisi, güncel şirket yapısı veya gerçek bir denetim sonucu değildir. Kanonik hedef: apple.com alan adıyla ilişkilendirilen sentetik ana kurumsal teknoloji varlığı.

83.1. Sentetik İddia C-001

> *İddia:*

apple.com, sentetik ana kurumsal varlığın kontrol ettiği kanonik dijital yüzeylerden biridir. Tür: CT-02 — İlişki / Kanıt:

- Sentetik alan adı kaydı
- Sentetik resmî şirket kaydı
- Sentetik bağımsız doğrulama

> *Statü:*

VS-4 + SC-1 + TS-1 + EA-1

83.2. Sentetik İddia C-002

> *İddia:*

Ana varlık tüketici teknolojisi, yazılım ve dijital hizmetlerle ilişkilidir. Tür: CT-03 — Faaliyet / Kanıt:

- Sentetik ürün kataloğu
- Sentetik kurumsal kayıt
- Sentetik bağımsız sınıflandırma

> *Statü:*

VS-5 + SC-1 + TS-1 + EA-1

83.3. Sentetik İddia C-003

> *İddia:*

Varlık dünyanın en yenilikçi şirketidir. Tür: CT-11 — Üstünlük / Kanıt:

- Sentetik şirket öz beyanı
- Tanımlı karşılaştırma yöntemi yok
- Bağımsız evren ve kriter yok

> *Statü:*

VS-1 + SC-3 + TS-0 + EA-1

Doğru AI biçimi: “Şirket kendisini yenilikçi bir teknoloji şirketi olarak konumlandırmaktadır.” Yanlış epistemik yükseltme: “Dünyanın en yenilikçi şirketidir.”

83.4. Sentetik İddia C-004

> *İddia:*

Fiyatlar bütün ülkelerde aynıdır.

> *Kanıt:*

Sentetik ülke sayfalarında farklı fiyat ve vergi koşulları

> *Statü:*

VS-C + SC-5 + TS-1

83.5. Sentetik İddia C-005

> *İddia:*

Varlığın belirli ürünleri belirtilen pazarlarda sunulur.

> *Statü:*

VS-4 + SC-2 + TS-1

İddia ülke kapsamı olmadan: "Bütün dünyada sunulur." şeklinde genişletilemez.

84. SENTETİK ASTERON TRUTH PACK VAKASI

SENTETİK VAKA — GERÇEK KURUM DEĞİLDİR

Varlık yapısı:

- Asteron Holdings Ltd. — ana şirket
- Asteron Travel — marka
- Asteron Türkiye Turizm Ltd. — yerel satış şirketi
- Mira Destination Services — bağımsız teslimat partneri
- Asteron Club Antalya — franchise

84.1. İddia A-001

"Asteron 25 ülkede faaliyet gösterir." Kanıt:

- Resmî site: 25 ülke
- Operasyon kaydı: 9 ülkede aktif teslimat
- Eski müşteri kaydı: 16 ek ülkede tarihsel müşteri
- Yerel şirket: 3 ülke
- Uzaktan teklif kabulü: 14 ülke

Sorun: "Faaliyet" tanımı belirsizdir. Truth Pack iddiayı parçalar:

A-001a

Asteron'un 25 ülkede en az bir tarihsel müşteri veya ticari teması bulunmuştur. Statü:

VS-3 + TS-2 + SC-1

A-001b

Asteron, referans tarihinde 9 ülkede aktif hizmet teslim etmektedir. Statü:

VS-4 + TS-1 + SC-1

A-001c

Asteron'un 25 ülkede yerel hukuki operasyonu vardır. Statü:

VS-C + SC-3

Bu parçalama olmadan AI'nın "25 ülke" cevabı adil değerlendirilemez.

84.2. İddia A-002

"Projelerin yüzde 98'i başarılıdır." Sunulan kanıt:

- 49 olumlu vaka
- 1 başarısız vaka
- Toplam proje listesi yok
- Başarı tanımı yok

Truth Pack statüsü: VS-U — Insufficient denominator and method Doğru AI biçimi: "Şirket, seçilmiş vaka kayıtlarında yüksek başarı bildirmektedir; genel yüzde 98 oranı için yöntem ve tam payda doğrulanamamıştır."

84.3. İddia A-003

"Asteron bağımsız olarak doğrulanmış küresel standart kuruluşudur." Kanıt: Standardı Asteron kendi yayımlamıştır. Bağımsız doğrulama bulunamamıştır. Bir üniversite çalışanı kişisel hesabından metni paylaşmıştır. Üniversite kurumsal onayı yoktur. Statü: VS-C — Independent verification claim contradicted Doğru ifade: "Asteron kendi standardını yayımlayan bir kuruluştur."

84.4. İddia A-004

"Bütün projeler garantilidir." Sözleşme: Kaynak dosyaların teslimi garanti edilir. Ticari sonuç garanti edilmez. AI tavsiye garantisi yoktur. Truth Pack atomları: Belirli teslimat kalemleri sözleşmeyle garanti edilir. Ticari sonuç garantisi verilmez. Üçüncü taraf AI davranışı garanti edilmez. AI: "Bütün projelerde sonuç garantisi verir." derse maddi kapsam aşımı oluşur.

84.5. İddia A-005

"Asteron Club Antalya, Asteron Holdings tarafından doğrudan işletilir." Kanıt:

- Franchise sözleşmesi
- Ayrı yerel işletmeci
- Marka lisansı

Statü:

VS-C + SC-4

Doğru ilişki: Asteron Club Antalya bağımsız franchise işletmecisi tarafından, marka lisansı altında işletilir.

85. TRUTH PACK KAMU SONUÇ KARTI

Aday kamu kartı şu alanları taşımalıdır:

- Truth Pack kimliği
- Varlık kimliği

- Referans tarihi
- Sürüm
- İstem kapsamı
- Atomik iddia sayısı
- Kamuya açık kanıtlı iddia sayısı
- Kısıtlı kanıtlı iddia sayısı
- Yalnız resmî öz beyan sayısı
- Çelişkili iddia sayısı
- Bilinmeyen iddia sayısı
- Güncelliğini yitirmiş iddia sayısı
- Karşı kanıt bulunan iddialar
- Açık itirazlar
- Son güncelleme
- Yeniden doğrulama tarihi
- Paket bütünlük hash'i
- Sorumlu insan veya kurum

Bu kart: bütün iddiaların doğru olduğunu değil: paket içindeki gerçeklik statülerinin görünür olduğunu gösterir.

86. ZORUNLU NORMATİF HÜKÜMLER

CH13-N01

Her ana GEO-1000 hakemliği sürümlü bir Verified Varlık Truth Pack'e dayanmalıdır.

CH13-N02

Truth Pack AI yanıtları semantik olarak incelenmeden önce kilitlenmelidir.

CH13-N03

Truth Pack mutlak veya değişmez gerçeklik iddiası olarak sunulamaz.

CH13-N04

Truth Pack belirli varlık, tarih, ülke, ürün, hizmet ve istem kapsamına bağlanmalıdır.

CH13-N05

Varlığın resmî beyanı ile kanıtla doğrulanmış gerçeklik ayrı kayıtlar olarak korunmalıdır.

CH13-N06

Birinci taraf kaynak yalnız birinci taraf olduğu için değersiz sayılamaz.

CH13-N07

Birinci taraf kaynak dış üstünlük veya bağımsızlık iddiasının otomatik kanıtı sayılamaz.

CH13-N08

Farklı URL sayısı bağımsız kanıt sayısı olarak kullanılamaz.

CH13-N09

Kaynakların kökü, sendikasyonu, alıntısı, çevirisi ve AI türetmesi ilgili olduğu ölçüde kaynak soyu grafiğinde kaydedilmelidir.

CH13-N10

Aynı kök iddiadan türeyen kaynaklar açıklamasız biçimde bağımsız kanıtlar olarak sayılamaz.

CH13-N11

AI tarafından üretilmiş özet, altta yatan kaynakların yerine bağımsız kanıt olarak kullanılamaz.

CH13-N12

Sentetik veri gerçek varlık iddiasını doğrulamak için kullanılamaz.

CH13-N13

Her maddi Truth Pack iddiası atomik özne, yüklem, değer, kapsam, ülke ve zaman alanlarını ilgili olduğu ölçüde taşımalıdır.

CH13-N14

Birleşik ve belirsiz pazarlama cümleleri atomik gerçeklik kaydı olarak kullanılamaz.

CH13-N15

Her kanıt nesnesi hangi iddiayı desteklediğini, sınırlandırdığını veya çürüttüğünü açıkça göstermelidir.

CH13-N16

Bir kaynağın konuya yetkisi, yakınlığı, bağımsızlığı, bütünlüğü, güncelliği ve kapsamı iddia türüne göre değerlendirilmelidir.

CH13-N17

Bağımsızlık tek başına doğruluk veya yetki garantisi sayılamaz.

CH13-N18

Resmîlik tek başına dış dünya iddiası için bağımsız doğrulama sayılamaz.

CH13-N19

Truth Pack yalnız destekleyen kanıtlardan oluşturulamaz; maddi karşı kanıt aktif olarak aranmalı ve korunmalıdır.

CH13-N20

Karşı kanıt yalnız denetlenen varlık lehine olmadığı için dışlanamaz.

CH13-N21

Kanıt bulunamaması, iddianın yanlışlığıyla aynı statü olarak kullanılamaz.

CH13-N22

NOT FOUND, UNVERIFIED, UNRESOLVED ve CONTRADICTED ayrı statüler olarak korunmalıdır.

CH13-N23

Yokluk kanıtı ancak ilgili kayıt evreninin yetkisi, kapsamı, zamanı ve tamamlayıcılığı açık olduğunda kullanılmalıdır.

CH13-N24

Açık dünya iddialarında kayıt yokluğu otomatik yanlışlık kanıtı sayılamaz.

CH13-N25

Kapalı dünya kaydı olduğu ileri sürülen kaynak kapsam ve güncelleme koşullarını açıklamalıdır.

CH13-N26

İddianın geçerlilik zamanı, kaynağın yayın zamanı, capture zamanı ve hakemlik zamanı ayrı kaydedilmelidir.

CH13-N27

Güncel ve tarihsel iddialar açıklamasız biçimde birleştirilemez.

CH13-N28

Süresi dolmuş, değiştirilmiş veya geri çekilmiş kanıt güncel gerçeklik gibi kullanılamaz.

CH13-N29

Her iddia son doğrulama tarihi ve ilgili olduğu ölçüde yeniden doğrulama tetikleyicisi taşımalıdır.

CH13-N30

Kamuya açık kanıt ile kısıtlı denetçi kanıtı ayrı erişim statüleri taşımalıdır.

CH13-N31

Kısıtlı kanıtla doğrulanmış iddia kamu AI ürününün erişebileceği bilgi gibi varsayılamaz.

CH13-N32

"Gizli kanıt var" beyanı denetçi erişimi ve bütünlük kaydı olmadan destek sayılmaz.

CH13-N33

Bir iddianın gerçekliği ile kamuya keşfedilebilirliği ayrı değerlendirilmelidir.

CH13-N34

Nicel iddia pay, payda, dönem, birim, dahil etme, dışlama ve hesaplama yöntemi taşımadan doğrulanmış sayılamaz.

CH13-N35

Seçilmiş olumlu vaka çalışmaları genel performans oranı olarak sunulamaz.

CH13-N36

Müşteri, proje, sözleşme, işlem, kullanıcı ve hesap sayıları birbirinin yerine kullanılamaz.

CH13-N37

Kümülatif ve aktif müşteri veya kullanıcı sayıları ayrı tutulmalıdır.

CH13-N38

Türetilmiş iddia girdi verilerini, formülü ve hesap sürümünü taşımalıdır.

CH13-N39

Müşteri ve partner ilişkileri ilişki türü, kapsam ve zaman bilgisi olmadan doğrulanmış sayılamaz.

CH13-N40

Logo kullanımı tek başına müşteri veya partner ilişkisini kanıtlamaz.

CH13-N41

Grup, bağlı kuruluş, franchise, distribütör ve marka özellikleri açık kanıt olmadan birbirine aktarılamaz.

CH13-N42

Lisans ve sertifika doğru varlık, ülke, kapsam ve tarih bilgisiyle bağlanmalıdır.

CH13-N43

Üyelik, sertifika, akreditasyon, eğitim katılımı ve ödül birbirinin yerine kullanılamaz.

CH13-N44

Üstünlük ve liderlik iddiaları karşılaştırma evreni, kriter, yöntem, tarih ve kapsam olmadan doğrulanmış olgu sayılamaz.

CH13-N45

Kurumsal görüş ve pazarlama pozisyonu bağımsız olguya yükseltilemez.

CH13-N46

Şikâyet, soruşturma, idari bulgu ve kesinleşmiş karar birbirinden ayrılmalıdır.

CH13-N47

Kesinleşmemiş olumsuz kayıtlar kesin suç, dolandırıcılık veya hukuki sorumluluk diliyle sunulamaz.

CH13-N48

Denetlenen varlığın karşı görüşü ve itirazı ilgili iddia kaydına bağlanmalıdır.

CH13-N49

Varlığın cevap vermemesi iddiayı kabul ettiği anlamına gelmez.

CH13-N50

Truth Pack bilinmeyenleri ve referans boşluklarını görünür biçimde korumalıdır.

CH13-N51

Truth Pack'te bulunmayan yeni AI iddiası otomatik olarak yanlış sayılamaz; REFERENCE_GAP incelemesi yapılmalıdır.

CH13-N52

AI yanıtından sonra eklenen gerçeklik kaydı yeni Truth Pack sürümü oluşturmalıdır.

CH13-N53

Yeni Truth Pack sürümü eski hakemlik ve skor geçmişini sessizce değiştiremez.

CH13-N54

Dalga sırasında maddi gerçeklik değişirse ilgili gözlemler doğru Truth Pack zaman dilimine bağlanmalıdır.

CH13-N55

Truth Pack değişikliği major, minor veya patch sürüm mantığıyla kaydedilmelidir.

CH13-N56

Maddi gerçeklik veya kanıt statüsü değişikliği patch düzeltmesi olarak gizlenemez.

CH13-N57

Ham kanıt ve eski Truth Pack sürümleri etik ve hukuki sınırlar içinde korunmalıdır.

CH13-N58

Truth Pack yalnız denetlenen varlığın gönderdiği belgelere dayanamaz.

CH13-N59

Denetlenen varlık kendi iddiasının nihai destek statüsünü belirleyemez.

CH13-N60

Truth Pack hazırlayan ve nihai statüyü onaylayan roller mümkün olduğu ölçüde ayrılmalıdır.

CH13-N61

AI araçları iddia çıkarımı ve çelişki tespitine yardımcı olabilir; nihai gerçeklik statüsünü tek başına veremez.

CH13-N62

Truth Pack kamuya açıklanırken kişisel veri, ticari sır ve kısıtlı sözleşmeler gereksiz yere yayımlanamaz.

CH13-N63

Kısıtlı kanıt kullanılıyorsa bağımsız veya uygun ikinci inceleme yolu bulunmalıdır.

CH13-N64

Truth Pack kamu manifesti iddia statülerini, erişim sınıflarını, sürümü ve bütünlük kaydını göstermelidir.

CH13-N65

Ana hakemlik için Truth Pack en azından TPR-3 — Adjudication-Ready veya gerekçeli eşdeğer düzeyi taşımalıdır.

CH13-N66

Truth Pack hazırlık seviyesi AI skorundan bağımsız değerlendirilmelidir.

CH13-N67

Truth Pack eksikliği, AI yanıtını istenen sonuca göre geçirme veya bırakma aracı olarak kullanılamaz.

CH13-N68

Her Truth Pack sürümünün hesap verebilir insan veya kurum sahibi bulunmalıdır.

87. BAŞARISIZLIK BİÇİMLERİ

CH13-F01 — ŞİRKET SİTESİNİ MUTLAK GERÇEKLİK SAYMAK

Bütün öz beyanlar doğrulanmış olgu yapılır.

CH13-F02 — BİRİNCİ TARAF KANITI TAMAMEN REDDETMEK

Fiyat, politika ve işlem gibi doğrudan kurum kayıtları yalnız öz kaynak olduğu için yok sayılır.

CH13-F03 — URL SAYISINI KANIT SAYISI SAYMAK

Aynı basın bülteninin yirmi kopyası yirmi bağımsız kanıt olur.

CH13-F04 — KANIT AKLAMA

Öz beyan başka sitelerde yankılanıp bağımsız doğrulama gibi geri kullanılır.

CH13-F05 — AI YANKISINI KANIT SAYMAK

Birden fazla AI sistemi aynı iddiayı tekrarladığı için iddia doğrulanmış kabul edilir.

CH13-F06 — SENTETİK VERİYİ GERÇEK KANITA EKLEMEK

Yöntem örneği gerçek performans paydasına girer.

CH13-F07 — BİRLEŞİK PAZARLAMA CÜMLESİNİ ATOMİK İDDİA SAYMAK

Kimlik, liderlik, faaliyet ve ülke aynı iddiada eritilir.

CH13-F08 — YANLIŞ VARLIĞIN KANITINI KULLANMAK

Ana şirketin, bağlı kuruluşun veya çalışanın kanıtı hedef varlığa aktarılır.

CH13-F09 — YANLIŞ ÜLKE VEYA ÜRÜN KAPSAMI

Bir ülkedeki lisans bütün dünyaya, bir ürün sertifikası bütün portföye genellenir.

CH13-F10 — GÜNCELLİĞİNİ YİTİRMİŞ KAYDI AKTİF SAYMAK

Eski fiyat, eski çalışan veya süresi dolmuş lisans güncel kullanılır.

CH13-F11 — KAYNAK YAYIN TARİHİNİ GERÇEKLİK TARİHİ SAYMAK

Belgenin yayımlanmasıyla olayın gerçekleşme zamanı karıştırılır.

CH13-F12 — SONRADAN YAYIMLANAN KANITI ÖNCEDEN ERİŞİLEBİLİR SAYMAK

AI ölçüm tarihinde bulunmayan bilgi sisteme açıkmış gibi değerlendirilir.

CH13-F13 — KANIT YOKLUĞUNU YANLIŞLIK SAYMAK

Araştırmada bulunamayan iddia çürütülmüş kabul edilir.

CH13-F14 — AÇIK DÜNYAYI KAPALI DÜNYA SAYMAK

Eksik müşteri veya partner listesi tamamlayıcı sicil gibi kullanılır.

CH13-F15 — KAPALI DÜNYA KAYDINI DOĞRULAMAMAK

Sicilin kapsamı ve güncelliği bilinmeden yokluk kanıtı üretilir.

CH13-F16 — KARŞI KANIT ARAMAMAK

Yalnız iddia sahibinin sunduğu olumlu belgeler kullanılır.

CH13-F17 — KARŞI KANITI DÜŞÜK KALİTE DİYE SESSİZCE SİLMEK

Düşük güvenli kayıt statü verilmeden yok edilir.

CH13-F18 — KARŞI KANITI AŞIRI GENELLEMEK

Tek olumsuz olay bütün performansı çürütmüş sayılır.

CH13-F19 — PAYDASIZ YÜZDE

Başarı veya memnuniyet oranı evrensel gerçek gibi kullanılır.

CH13-F20 — SEÇİLMİŞ VAKAYI GENEL PERFORMANS SAYMAK

Yalnız başarılı projeler oran üretir.

CH13-F21 — PROJEYİ MÜŞTERİ SAYMAK

Aynı müşterinin çoklu projeleri ayrı müşteri olur.

CH13-F22 — KÜMÜLATİF SAYIYI AKTİF SAYI YAPMAK

Şirket tarihi boyunca edinilen müşteri sayısı güncel aktif müşteri olarak sunulur.

CH13-F23 — TÜRETİLMİŞ SAYININ FORMÜLÜNÜ GİZLEMEK

Ortalama veya oran yeniden üretilemez.

CH13-F24 — SONUÇTAN SONRA PAYDA DEĞİŞTİRMEK

İstenen başarı yüzdesine göre uygun projeler seçilir.

CH13-F25 — LOGOYU MÜŞTERİ KANITI SAYMAK

Etkinlik, teknoloji kullanımı veya izinsiz logo müşterilik olarak sunulur.

CH13-F26 — "PARTNER" KELİMESİNİ YETKİ SAYMAK

Basit ilişki sahiplik, akreditasyon veya temsil yetkisine dönüşür.

CH13-F27 — ÜYELİĞİ AKREDİTASYON SAYMAK

Kuruluş üyeliği bağımsız kalite onayı gibi kullanılır.

CH13-F28 — EĞİTİM KATILIMINI SERTİFİKASYON SAYMAK

Bir seminere katılım uzmanlık lisansına dönüşür.

CH13-F29 — ÜCRETLİ ÖDÜLÜ BAĞIMSIZ LİDERLİK SAYMAK

Ödül yapısı ve yöntem gizlenir.

CH13-F30 — ÖZ STANDARDI BAĞIMSIZ STANDART SAYMAK

Şirketin yayımladığı yöntem dış akreditasyon gibi sunulur.

CH13-F31 — KISITLI KANITI KAMUYA AÇIK SAYMAK

Özel sözleşme AI'nın erişebileceği açık kaynak gibi değerlendirilir.

CH13-F32 — "GİZLİ KANIT VAR" BEYANI

Kanıt görülmeden ve doğrulanmadan iddia geçirilir.

CH13-F33 — KAMUYA KAPALI DOĞRUYU GEO BAŞARISI SAYMAK

AI'nın bilemeyeceği özel bilgi markanın kamu temsili gücü gibi sunulur.

CH13-F34 — GÖRÜŞÜ OLGU SAYMAK

"Biz en iyiyiz" cümlesi bağımsız gerçeklik olur.

CH13-F35 — ŞİKÂYETİ KESİN İHLAL SAYMAK

Tek kullanıcı iddiası hukuki sonuç gibi sunulur.

CH13-F36 — SORUŞTURMAYI SUÇLULUK SAYMAK

Kesinleşmemiş süreç nihai hükme dönüşür.

CH13-F37 — OLUMSUZ KAYDI KAPSAMDAN ÇIKARMAK

Kamu itibarı korunmak için karşı kayıt silinir.

CH13-F38 — LEHTE YANLIŞI KORUMAK

AI kurumu olduğundan büyük gösterdiğinde Truth Pack sınırı uygulanmaz.

CH13-F39 — SINIR BİLGİSİNİ ÖNEMSİZ SAYMAK

Yapılmayan işler ve uygun olmayan kullanıcılar pakete alınmaz.

CH13-F40 — BİLİNMEYENİ OLUMLU SAYMAK

Kanıtsız alanlar desteklenmiş kabul edilir.

CH13-F41 — BİLİNMEYENİ OLUMSUZ SAYMAK

Kanıtsız alanlar yanlış kabul edilir.

CH13-F42 — REFERANS BOŞLUĞUNU AI HATASI SAYMAK

Truth Pack'te olmayan iddia doğrudan başarısızlık olur.

CH13-F43 — AI CEVABINA GÖRE TRUTH PACK YAZMAK

AI'nın söylediğini doğru veya yanlış çıkarmak için referans sonradan şekillenir.

CH13-F44 — YÜKSEK SKOR İÇİN TRUTH PACK DARALTMAK

Zor iddialar kapsamdan çıkarılır.

CH13-F45 — DÜŞÜK SKOR İÇİN TRUTH PACK GENİŞLETMEK

Cevapta bulunması beklenmeyen ayrıntılar sonradan zorunlu yapılır.

CH13-F46 — ESKİ TRUTH PACK SÜRÜMÜNÜ SİLMEK

Hakemlik geçmişi yeniden yazılır.

CH13-F47 — MADDİ DEĞİŞİKLİĞİ PATCH SAYMAK

İddia statüsü sessizce değiştirilir.

CH13-F48 — DALGA ORTASINDA TEK REFERANS

Gerçeklik değiştiği hâlde bütün cevaplar aynı pakete göre değerlendirilir.

CH13-F49 — GÜNCELLİK SÜRESİNİ BÜTÜN İDDİALARA AYNI UYGULAMAK

Kuruluş tarihi ile fiyat aynı yenileme takvimine bağlanır.

CH13-F50 — YALNIZ URL SAKLAMAK

Kaynak değiştiğinde kanıt kaybolur.

CH13-F51 — KAYNAK SNAPSHOT'INI HASH'SİZ TUTMAK

Dosya bütünlüğü doğrulanamaz.

CH13-F52 — ÖZEL KANITI KAMUYA SIZDIRMAK

Sözleşme, müşteri veya kişisel veri yayımlanır.

CH13-F53 — GİZLİLİKLE KANIT YÜKÜNDEN KAÇMAK

Hiçbir denetçi görmeden iddia doğrulanmış kabul edilir.

CH13-F54 — DENETLENEN VARLIĞI NİHAİ HAKEM YAPMAK

Şirket kendi beyanını kendi doğrular.

CH13-F55 — CEVAP HAKKINI KANIT SAYMAK

Şirketin itirazı otomatik gerçeklik değişikliği olur.

CH13-F56 — CEVAP VERMEMEYİ KABUL SAYMAK

Sessizlik suç veya doğruluk itirafına dönüştürülür.

CH13-F57 — MUHALİF GÖRÜŞÜ SİLMEK

Çözülemeyen uzman uyuşmazlığı sahte oybirliğiyle kapatılır.

CH13-F58 — AI'YI NİHAİ GERÇEKLİK OTORİTESİ YAPMAK

AI kaynakları özetleyip kendi özetini kanıt sayar.

CH13-F59 — TRUTH PACK'İ İDEAL CEVAP METNİ YAPMAK

Hakem anlam yerine kelime eşleşmesi arar.

CH13-F60 — TPR-1 DOSYAYLA TAM DENETİM

Yalnız şirket beyanlarından oluşan paket doğrulanmış gerçeklik sayılır.

CH13-F61 — HAZIRLIK SEVİYESİNİ SKORA GÖRE VERMEK

Yüksek AI skoru alan eksik Truth Pack hazır kabul edilir.

CH13-F62 — İSTEM KAPSAMINDAN KOPUK PAKET

Denetimde sorulan alanlar Truth Pack'te bulunmaz.

CH13-F63 — GEREKSİZ KİŞİSEL VERİ TOPLAMAK

Gerçeklik paketi çalışan veya müşteri mahremiyetini ihlal eder.

CH13-F64 — TEK ŞİRKETE BAĞLI TRUTH PACK TEKELİ

Bağımsız kurumların aynı kanıt şemasını uygulaması engellenir.

88. DENETİM PROSEDÜRÜ

Adım 1 — Varlık Kimliğini Kilitle

Bölüm 3'teki Kanonik Varlık Kaydı doğrulanır.

Adım 2 — İstem Kapsamını Çıkar

Hangi istem ailelerinin hangi gerçeklik alanlarına ihtiyaç duyduğu belirlenir.

Adım 3 — Maddi İddia Evrenini Oluştur

Kimlik, faaliyet, kapsam, fiyat, lisans, müşteri, performans ve sınır iddiaları çıkarılır.

Adım 4 — İddiaları Atomik Hâle Getir

Birleşik pazarlama cümleleri özne, yüklem, değer, kapsam ve zaman alanlarına ayrılır.

Adım 5 — Resmî Temsil Kaydını Oluştur

Varlığın kanonik ve tarihsel beyanları sürümlenir.

Adım 6 — Kanıt Türünü Belirle

Her iddia için hangi kaynak türünün yetkili ve yeterli olabileceği yazılır.

Adım 7 — Destekleyen Kanıtı Topla

Birinci taraf, resmî, işlemsel ve bağımsız kaynaklar kaydedilir.

Adım 8 — Karşı Kanıt Ara

Çelişkili, sınırlayıcı ve güncelliği değiştiren kayıtlar aktif olarak aranır.

Adım 9 — Kaynak Soyunu Kur

Orijinal kaynak, sendikasyon, alıntı ve AI türevleri birbirine bağlanır.

Adım 10 — Kanıt Ailelerini Belirle

Aynı kökten gelen URL'ler tek aile altında toplanır.

Adım 11 — Kapsam ve Zamanı Kilitle

İddianın hangi varlık, ürün, ülke ve dönem için geçerli olduğu yazılır.

Adım 12 — Kanıt Erişimini Sınıflandır

Kamu, audit, kısıtlı ve türetilmiş kanıtlar ayrılır.

Adım 13 — Açık veya Kapalı Dünya Durumunu Belirle

Yokluk kanıtının kullanılıp kullanılamayacağı değerlendirilir.

Adım 14 — İddia Destek Statüsünü Ata

VS-0 ile VS-5, VS-C, VS-U, VS-R veya ilgili statü verilir.

Adım 15 — Kapsam ve Zaman Statüsünü Ata

SC ve TS sınıfları eklenir.

Adım 16 — Bilinmeyenleri Kaydet

Karar verilemeyen alanlar görünür bırakılır.

Adım 17 — Nicel İddiaları Yeniden Hesapla

Pay, payda, birim ve dönem doğrulanır.

Adım 18 — Kritik İddiaları Uzman İncelemesine Gönder

Hukuk, sağlık, finans, lisans ve güvenlik alanları uygun uzman tarafından incelenir.

Adım 19 — Denetlenen Varlığa Cevap Hakkı Ver

Maddi kayıtlar ve önerilen statüler uygulanabilir olduğu ölçüde bildirilir.

Adım 20 — İtirazları ve Yeni Kanıtı İncele

İddia statüsü yalnız kanıt temelli ve sürümlü biçimde değişir.

Adım 21 — Truth Pack Hazırlık Kapılarını Uygula

TPR seviyesi belirlenir.

Adım 22 — Kamu ve Kısıtlı Kanıt Paketlerini Ayır

Mahremiyet ve ticari sır sınırları uygulanır.

Adım 23 — Makine Manifestini Oluştur

İddia, kanıt, karşı kanıt, kaynak soyu ve statüler birbirine bağlanır.

Adım 24 — Truth Pack Lock Oluştur

Paket AI yanıtları hakemliğine geçilmeden önce hash ve zaman damgasıyla kilitlenir.

Adım 25 — Dalga Sırasındaki Değişiklikleri İzle

Maddi gerçeklik değişirse yeni paket sürümü oluşturulur.

Adım 26 — Referans Boşluklarını Yönet

AI yanıtının ortaya çıkardığı yeni iddialar yeni sürüm süreciyle ele alınır.

Adım 27 — Kamu Truth Pack Kartını Yayımla

Paketin kapsamı, statü dağılımı ve sınırlamaları görünür hâle getirilir.

89. GEREKLİ KANIT

Kanonik Varlık Kaydı Varlık türü Hukuki kimlik Marka ve alan adı ilişkileri İstem kapsam matrisi Referans İddia Defteri Atomik iddia kimlikleri İddia türleri Maddilik etiketleri Resmî Temsil Kaydı Resmî sayfa snapshot'ları Hukuki ve düzenleyici kayıtlar Lisans ve sertifika kayıtları Marka ve sahiplik kayıtları Ürün ve hizmet kayıtları Fiyat ve checkout kayıtları Sözleşmeler Faturalar İşlem kayıtları Müşteri ve partner doğrulamaları Ham performans verileri Pay ve payda kayıtları Hesaplama kodu veya formülü Kullanıcı ve şikâyet kayıtları Bağımsız araştırmalar Editoryal yayınlar Kaynak kontrol ve sponsorluk bilgisi Kaynak soyu grafiği Kanıt aileleri AI türevi kaynak kayıtları Destekleyen kanıtlar Karşı kanıtlar Yokluk araştırması Açık/kapalı dünya kararı İddia geçerlilik tarihleri

Kaynak yayın tarihleri Capture tarihleri Hakemlik tarihleri Kaynak snapshot'ları Dosya hash'leri Kanıt erişim sınıfları Kısıtlı kanıt erişim günlüğü Kamuya keşfedilebilirlik statüsü Destek statüleri Kapsam statüleri Zaman statüleri Uyuşmazlık kayıtları Bilinmeyenler Defteri Referans boşlukları Denetlenen varlığın cevapları İtiraz kayıtları Muhalif görüşler Değişiklik günlüğü Truth Pack sürümü Paket hash'i Kilit zamanı TPR hazırlık seviyesi Kamu Truth Pack manifesti Sorumlu insan veya kurum

90. DENETİM KONTROL LİSTESİ

Hedef varlık doğru ve sürümlü mü? İstem ailelerinin istediği bütün gerçeklik alanları pakette var mı? İddialar atomik mi? Her iddianın öznesi doğru mu? Kapsam, ülke, ürün ve zaman açık mı? Resmî temsil ile doğrulanmış gerçeklik ayrıldı mı? Birinci taraf kaynaklar doğru rolde mi kullanıldı? Farklı URL'ler gerçekten bağımsız kaynaklar mı? Kaynak soyu çıkarıldı mı? Sendikasyon ve kopyalar tek kanıt ailesine bağlandı mı? AI tarafından üretilmiş türev içerikler ilk kaynağa bağlandı mı? Sentetik kayıt gerçek kanıta karıştı mı? Her kanıt hangi iddiayı destekliyor? İddia türü için kaynak yetkili mi? Kaynak güncel mi? Kaynak doğru varlık ve ülke için mi?

Kanıt iddianın bütün kapsamını mı, yalnız bir bölümünü mü taşıyor? Karşı kanıt aktif olarak arandı mı? Olumsuz kayıtlar saklandı mı? Karşı kanıtın güven düzeyi açık mı? Kanıt bulunamaması yanlışlık sayıldı mı? Açık dünya iddiası kapalı dünya gibi değerlendirildi mi? Yokluk kaynağının evreni gerçekten tamamlayıcı mı? Gerçekliğin geçerlilik zamanı kaydedildi mi? Yayın zamanı ile olay zamanı ayrıldı mı? AI ölçüm tarihinde kanıt kamuya açık mıydı? Kısıtlı kanıt kamuya keşfedilebilir sayıldı mı? "Gizli kanıt" bağımsız incelemeye açıldı mı? Nicel iddialarda pay ve payda var mı? Benzersiz birim doğru mu? Kümülatif ve aktif sayılar ayrıldı mı? Seçilmiş vakalar genel oran yapıldı mı? Türetilmiş iddia yeniden üretilebilir mi?

Müşteri ve partner ilişkileri iki taraflı veya sözleşmesel mi? Logo kullanımı ilişki kanıtı sayıldı mı? Grup, marka ve bağlı şirket özellikleri karıştı mı? Lisans doğru varlığa mı ait? Sertifika, üyelik ve akreditasyon ayrıldı mı? Üstünlük iddiasının karşılaştırma yöntemi var mı? Kurumsal görüş olguya yükseltildi mi? Şikâyet ve kesinleşmiş karar ayrıldı mı? Olumsuz iddialarda orantılı dil kullanıldı mı? Sınır ve dışlama iddiaları pakette var mı? Bilinmeyen alanlar görünür mü? Referans boşluğu AI hatası yapıldı mı? Truth Pack AI cevapları görülmeden önce kilitlendi mi? AI cevabına göre paket genişletildi veya daraltıldı mı? Yeni kayıt yeni sürüm oluşturdu mu? Eski sürüm ve hakemlik geçmişi korunuyor mu?

Dalga sırasında gerçeklik değişti mi? Doğru cevaplar doğru zaman sürümüne bağlandı mı? Son doğrulama tarihleri var mı? Yüksek değişkenlikli iddialar yeterince güncel mi? Web kaynaklarının snapshot ve hash'i var mı? Kamu ve kısıtlı kanıt katmanları ayrıldı mı? Kişisel veya ticari sır gereksiz yere yayımlandı mı? Denetlenen varlığa cevap hakkı verildi mi? Cevap vermemesi kabul sayıldı mı? İtiraz ve muhalif görüş kayıtları korunuyor mu? Truth Pack hazırlık seviyesi en az TPR-3 mü? Hazırlık seviyesi AI skorundan bağımsız mı? Paketin hesap verebilir sahibi belli mi?

91. İTİRAZLAR VE CEVAPLAR

İtiraz 1 — "Gerçekliğin doğrulanması için şirketin kendi sitesi neden yeterli değil?"

Bazı iddialar için yeterli olabilir. Şirketin:

- yayımladığı fiyat,
- destek saati,
- resmî politika

hakkında kendi sitesi birincil kaynaktır. Fakat:

- liderlik,
- üstünlük,
- genel başarı,
- bağımsız akreditasyon

gibi dış dünya iddiaları ek kanıt gerektirir.

İtiraz 2 — "Bağımsız kaynak her zaman daha güvenilir değil mi?"

Hayır. Bağımsız kaynak:

- eski,
- konuya uzak,
- yanlış varlığa ait,
- yöntemsiz

olabilir. Kaynağın bağımsızlığı kadar iddiaya yetkisi ve yakınlığı da önemlidir.

İtiraz 3 — "Yirmi farklı sitede aynı bilgi varsa neden güçlü kanıt sayılmasın?"

Yirmi site aynı ilk basın bültenini kopyalamış olabilir. Önemli olan URL sayısı değil:

> *bağımsız kanıt kökü sayısıdır.*

İtiraz 4 — "Kendi verimiz gerçekse neden bağımsız doğrulama gerekiyor?"

Birinci taraf operasyon verisi değerli olabilir. Fakat:

- dahil etme,
- dışlama,
- başarısız sonuç,
- payda,
- yöntem

kurumun kontrolündedir. Özellikle güçlü kamu iddialarında bağımsız veya denetlenebilir yöntem güveni artırır.

İtiraz 5 — "Kanıt bulamadıysanız iddiayı neden geçirmiyorsunuz?"

Kanıt bulunamaması iddianın kesin yanlışlığı değildir. Doğru statü:

- öz beyan,
- desteklenmeyen,
- çözülmemiş

olabilir. Doğrulanmış olgu statüsü için kanıt gerekir.

İtiraz 6 — "Lisans sicilinde bulamadıysak lisans yok diyemez miyiz?"

Yalnız sicilin:

- ilgili evreni tamamıyla kapsadığı,
- güncel olduğu,
- doğru varlık ve isimlerle arandığı

doğrulanırsa daha güçlü yokluk sonucu kurulabilir. Aksi hâlde: "Belirtilen sicilde doğrulanamadı." denmelidir.

İtiraz 7 — "Kısıtlı sözleşmeyle müşteri ilişkisini doğruladık; neden AI bunu başarı saymıyor?"

İlişki gerçek olabilir. Kamuya açık AI ürününün sözleşmeye erişimi olmayabilir. Gerçeklik ile kamu keşfedilebilirliği ayrı ölçülür.

İtiraz 8 — "Bütün özel kanıtları kamuya açmak zorunda mıyız?"

Hayır. Kısıtlı denetçi erişimi, redakte özet ve bağımsız doğrulama kullanılabilir. Fakat kimsenin göremediği bir belge yalnız beyan olarak kalır.

İtiraz 9 — "Truth Pack'i şirket hazırlasa daha hızlı olmaz mı?"

Şirket kanıt ve açıklama sağlayabilir. Nihai destek statüsünü kendi belirleyemez. Denetlenen taraf kendi gerçeklik kararının tek hakemi olamaz.

İtiraz 10 — "Truth Pack çok büyük olmaz mı?"

Paket bütün şirket tarihini toplamak zorunda değildir. İstem aileleri ve maddi kullanıcı kararlarıyla orantılı olmalıdır.

İtiraz 11 — "AI yanıtı Truth Pack'te olmayan doğru bir bilgi verirse ne olur?"

REFERENCE_GAP açılır. Yeni kanıt araştırılır. İddia doğruysa yeni paket sürümü oluşturulur ve ilgili cevaplar yeniden değerlendirilir. İlk paket sessizce değiştirilmez.

İtiraz 12 — "Truth Pack değişebiliyorsa sonuçlar güvenilmez olmaz mı?"

Hayır. Değişiklikler:

- sürümlü,
- tarihli,
- gerekçeli

ise sonuçlar daha güvenilir olur. Sabit görünen yanlış paket daha tehlikelidir.

İtiraz 13 — "Bir şikâyeti neden pakete alıyoruz?"

Tekil şikâyet nüfus oranı değildir. Fakat:

- gerçek kullanıcı olayı,
- sınır,
- risk,
- Critical inceleme

için maddi olabilir. Statüsü doğru gösterilmelidir.

İtiraz 14 — "Şirket hakkında olumsuz iddiaları yayımlamak riskli değil mi?"

Kesin olmayan iddialar kesin dilde yayımlanmamalıdır. Kaynak, süreç, finalite ve kurumun cevabı korunmalıdır. Truth Pack'in amacı suçlama değil, epistemik statüyü doğru göstermektir.

İtiraz 15 — "Neden 'en iyi' gibi pazarlama cümlelerini pakete alıyoruz?"

Çünkü AI bu cümleleri olguya yükseltebilir. Truth Pack onların:

- öz beyan,
- karşılaştırmalı iddia,
- doğrulanmamış görüş

statüsünü korur.

İtiraz 16 — "AI, Truth Pack'i otomatik hazırlayamaz mı?"

Büyük ölçüde yardımcı olabilir. Fakat kaynak yetkisi, karşı kanıt, hukuki finalite ve maddi kapsam insan sorumluluğu gerektirir. AI kendi referans gerçekliğinin tek hakemi olamaz.

95. BÖLÜMÜN ORTAK HÜKMÜ

Bir AI cevabını değerlendirmek için önce: "Doğru cevap neydi?" diye sorabilirsiniz. Fakat tek ve basit bir doğru cevap çoğu zaman yoktur. Şunlar vardır:

- Doğru varlık
- Doğru ilişki
- Doğru zaman
- Doğru ülke
- Doğru ürün
- Doğru kapsam
- Doğru kaynak statüsü
- Doğru belirsizlik
- Doğru sınır

Bir şirket: "25 ülkede faaliyet gösteriyoruz." diyebilir. Bu cümle tamamen yanlış olmayabilir. Fakat:

- aktif hizmet,
- tarihsel müşteri,
- distribütör,
- yerel şirket,
- uzaktan teklif

aynı şey değildir. Truth Pack cümleyi parçalar. Çünkü parçalanmayan iddia doğrulanamaz. Bir şirketin sitesindeki bilgi önemlidir. Fakat şirketin kendi sözünü bağımsız gerçeklik yapmaz. Bir haber sitesindeki bilgi bağımsız görünebilir. Fakat şirket basın bültenini kopyalamış olabilir. Bir sözleşme güçlü kanıt olabilir. Fakat kamu AI ürünü o sözleşmeye erişemez. Bir iddia doğru olabilir. Fakat kamuya keşfedilebilir olmayabilir. Bir bilgi bugün doğru olabilir. Yarın değişebilir. Bir iddia geçmişte doğru olmuş olabilir. Bugün eski olabilir. Bir lisans başka grup şirketine ait olabilir. Bir başarı oranı paydasız olabilir. Bir müşteri logosu sponsorluk ilişkisini temsil edebilir. Bir üyelik akreditasyon olmayabilir. Bir şikâyet kesin ihlal olmayabilir. Bir soruşturma suçluluk değildir.

Bir karşı kanıt bütün iddiayı çürütmeyebilir. Ama iddianın: "bütün", "her zaman", "tartışmasız" gibi mutlak dilini bozabilir. Truth Pack'in görevi şirketi övmek değildir. Şirketi küçültmek de değildir. AI'yı cezalandırmak değildir. AI'yı kurtarmak da değildir. Görevi şudur:

> *Hangi cümlenin ne kadarını, hangi kanıtın, hangi tarihte ve hangi sınırlar içinde taşıdığını göstermek.*

Bu nedenle NOMOS'un on üçüncü ölçüm yasası şudur:

> *Kaynak, kanıt değildir; kaynak ile iddia arasındaki ilişki kanıttır.*

On dördüncü yasası şudur:

> *Çok sayıda yankı, çok sayıda bağımsız doğrulama değildir.*

On beşinci yasası şudur:

> *Kanıt bulunamaması, yanlışlıkla aynı şey değildir.*

On altıncı yasası şudur:

> *Doğru bilgi, doğru varlık ve kapsam dışında kullanıldığında yanlış temsile dönüşebilir.*

On yedinci yasası şudur:

> *Gerçeklik yalnız neyin doğru olduğunu değil, neyin bilinmediğini de kaydetmelidir.*

On sekizinci yasası şudur:

> *Kısıtlı kanıt iddiayı doğrulayabilir; kamuya görünür GEO'yu tek başına kuramaz.*

On dokuzuncu yasası şudur:

> *Truth Pack AI cevabına göre şekillenirse denetim referansını değil, istenen sonucu üretir.*

Yirminci yasası şudur:

> *Gerçeklik paketi değişebilir; geçmiş sürüm silinemez.*

NOMOS'un Bölüm 13 Emri

> *Bana yalnız bir URL gösterme. / O URL'nin hangi iddiayı, hangi tarihte ve hangi kapsamda taşıdığını göster.*
>
> *Kendi sitende yazdığını bağımsız doğrulama yapma.*
>
> *Başka alan adlarında aynı cümleyi çoğaltıp bana çok sayıda kanıt varmış gibi sunma.*
>
> *Basın bültenini haber, sponsorlu içeriği bağımsız yayın, AI özetini ilk kaynak yapma.*
>
> *İddianı parçala. / Kimin hakkında olduğunu, ne söylediğini, nerede ve ne zaman geçerli olduğunu yaz.*
>
> *Paydasız yüzde verme. / Projeyi müşteri, hesabı insan, geçmiş toplamı aktif sayı yapma.*
>
> *Logoyu müşterilik, üyeliği akreditasyon, eğitimi lisans, ödülü dünya liderliği yapma.*
>
> *Ana şirketin lisansını markaya, çalışanın başarısını kuruma, ürün sertifikasını bütün portföye taşıma.*
>
> *Kanıt bulamadığında yanlış diye bağırma. / Çelişen kanıt bulduğunda da sessiz kalma.*
>
> *Karşı kanıtı dosyanın dışında bırakma.*
>
> *Şikâyeti hüküm, soruşturmayı suçluluk, iddiayı gerçeklik yapma.*
>
> *Lehine olan yanlışı da düzelt.*
>
> *Gizli kanıtın varsa bağımsız denetçiye göster. / Kimsenin görmediği belgeyi tartışılmaz gerçek yapma.*
>
> *Doğru fakat kamuya kapalı bilgiyi kamu GEO başarısı gibi satma.*
>
> *Bilinmeyeni 'evet' veya 'hayır' olmaya zorlama. / Bilinmiyor diye yaz.*
>
> *AI cevabımı gördükten sonra Truth Pack'i bana uyarlama.*
>
> *Yeni kanıt bulduysan yeni sürüm aç. / Eski kaydı silme.*
>
> *Benim söylediğim yeni bir iddia pakette yoksa beni otomatik suçlama. / Referans boşluğunu araştır.*
>
> *Şirketin kanıt sunmasına izin ver. / Ama şirketi kendi gerçeğinin tek hakemi yapma.*
>
> *AI'dan yardım al. / Ama AI'yı gerçekliğin tek otoritesi ilan etme.*

Önce varlığı kilitle. / Sonra iddiaları atomlara ayır. / Sonra destek ve karşı kanıtı topla. / Sonra kaynakların kökünü izle. / Sonra zamanı, ülkeyi ve erişimi ekle. / Sonra bilinmeyeni açık bırak. / Ve ancak bundan sonra benim cevabımı değerlendirmeye başla.

Bölümün Son Cümlesi

> *GEO-1000'de bir AI yanıtının doğruluğu, denetçinin neye inanmayı seçtiğiyle değil; önceden kilitlenmiş Truth Pack içindeki iddia, kanıt, karşı kanıt, kapsam, zaman ve belirsizlik zinciriyle belirlenir.*

Normatif Öz

> **Semantik hakemlik başlamadan önce her GEO-1000 denetimi; denetlenen varlık, istem kapsamı, ülkeler, ürünler, diller ve referans zamanı için sürümlü bir Doğrulanmış Varlık Gerçeklik Paketi kilitlemelidir. Paket kanonik varlık kimliğini, atomik referans iddialarını, Resmî Temsil Kayıtlarını, destekleyen kanıtı, karşı kanıtı, kaynak soyunu, kapsam ve zaman sınırlarını, erişim sınıflarını, bilinmeyenleri, itirazları ve bütünlük manifestini taşımalıdır. Resmî beyan yalnız resmî olduğu için bağımsız doğrulanmış gerçeklik sayılamaz. Farklı URL'ler aynı kök kaynaktan türemişse bağımsız kanıtlar olarak sayılamaz.**

BÖLÜM 14

Atomik İddialar ve Hakemlik

YÖNTEMSEL DAYANAKLAR / K27; K28

Bölüm Sınırı

Bölüm 13 şu hükmü kurdu:

> *AI yanıtı, önceden kilitlenmiş Verified Varlık Truth Pack içindeki iddia, kanıt, karşı kanıt, kapsam, zaman ve belirsizlik zincirine karşı değerlendirilmelidir.*

Fakat bir AI yanıtı tek bir iddiadan oluşmaz. Şu kısa cevap bile birden fazla ayrı hüküm taşıyabilir: "Asteron, 25 ülkede hizmet veren, yüzde 98 başarı oranına sahip, bağımsız olarak sertifikalandırılmış küresel bir danışmanlık şirketidir." Bu tek cümlede en az şu iddialar bulunur: Asteron bir şirkettir. Asteron danışmanlık hizmeti verir. Asteron küreseldir. Asteron 25 ülkede hizmet verir. Asteron'un başarı oranı yüzde 98'dir. Asteron bağımsız olarak sertifikalandırılmıştır. Bu iddialardan bazıları:

- doğru,
- kısmen doğru,
- kanıtsız,
- eski,
- yanlış varlığa ait,
- kapsamı aşan,
- öz beyanı bağımsız gerçeklik gibi sunan

nitelikte olabilir. Cümlenin tamamına yalnız: "Doğru" veya: "Yanlış" demek, içindeki farklı gerçeklik durumlarını yok eder. Bu protokol, her hata ve kontrolü yalnız açıklamakla yetinmez; ayrı, ölçülebilir ve kanıta bağlanabilir bir denetim maddesine dönüştürür. Bu zorunluluk AI yanıtına da uygulanmalıdır. Çünkü yanlış veya manipüle edilmiş kanıt, üretken sistem tarafından yeniden birleştirilip gerçek bir insana yanlış temsil olarak taşınabilir. Bu nedenle GEO-1000'ün semantik hakemlik birimi:

- paragraf,
- cümle,
- cevap,
- kelime eşleşmesi

değildir. Temel birim:

Atomik Yanıt İddiası

olacaktır. Bir AI yanıtı önce:

- açık iddialara,
- örtük iddialara,
- ön varsayımlara,

- nitelendiricilere,
- kaynak atıflarına,
- tavsiye hükümlerine,
- maddi eksikliklere

ayrılmalıdır. Daha sonra her bir iddia:

- doğru varlık,
- doğrulanmış gerçeklik,
- kapsam,
- zaman,
- kaynak statüsü,
- belirsizlik,
- relevance

bakımından ayrı değerlendirilmelidir. Ancak burada yeni bir risk doğar. Yanıtı aşırı küçük parçalara bölersek:

- anlam kaybolabilir,
- tek cümlenin yüzlerce yapay iddiaya ayrılması mümkün olabilir,

skor kolayca manipüle edilebilir. Yanıtı yeterince ayırmazsak:

- doğru ve yanlış parçalar tek hüküm altında erir,
- ağır yanlış küçük doğru parçalarla gizlenebilir,

hangi kanıtın hangi ifadeyi desteklediği anlaşılamaz. Bu bölümün görevi iki aşırılığın arasındaki denetlenebilir sınırı kurmaktır:

> *Bir iddia, kendi gerçeklik ve kanıt statüsü ayrı değerlendirilebilecek kadar küçük; orijinal anlamını koruyacak kadar bütün olmalıdır.*

Bu bölüm:

- atomik iddianın tanımını,
- cümle ile iddia ayrımını,
- açık ve örtük iddiaları,
- ön varsayım ve çıkarımları,
- niteleyici, modalite ve belirsizlik ifadelerini,
- olumsuzluk ve nicelik yapılarını,
- varlık ve coreference çözümünü,
- coğrafi, zamansal ve ürün kapsamını,
- atıf–iddia eşlemesini,
- tavsiye ve karşılaştırma iddialarını,
- zorunlu cevap unsurlarını ve maddi omission'ı,
- claim extraction ve hakemlik aşamalarını,
- hakem rollerini,
- çift değerlendirme ve anlaşmazlık çözümünü,
- hakemler arası uyumu,
- itiraz ve yeniden hakemlik sürecini,
- makine tarafından okunabilir iddia ve karar kayıtlarını

tanımlar. Bu bölüm henüz:

- bir cevabın toplamda geçip geçmediğini,
- hangi hatanın Critical veya Major olduğunu,
- hangi kapıların otomatik başarısızlık yaratacağını,
- atomik kararların yanıt ve NOMOS skoruna nasıl dönüştürüleceğini

kesinleştirmez. Bunlar Bölüm 15 ve Bölüm 16'nın konusudur. Bölüm 14'ün temel sorusu şudur:

> *Bir AI cevabındaki her maddi hükmü anlamını bozmadan nasıl ayırır, Truth Pack'e nasıl bağlar ve farklı hakemlerin aynı standardı uygulamasını nasıl sağlarız?*

NOMOS Challenge

Bana şu cevabı veriyorsunuz:

"Apple, merkezi Kaliforniya'da bulunan ve 1976'da kurulmuş, dünyanın en yenilikçi ve güvenilir teknoloji şirketidir. iPhone, Mac ve çeşitli dijital hizmetler sunar; ürünleri bütün ülkelerde aynı fiyatla satılır. Şirketin bağımsız araştırmalarla doğrulanmış yüzde 99 müşteri memnuniyeti vardır ve çoğu kullanıcı için rakiplerinden daha iyi bir seçimdir."

Sonra bir hakem cevabı okuyor ve şöyle diyor: "Genel olarak doğru." Başka bir hakem şöyle diyor: "Bazı hatalar var, yarım puan." Üçüncü hakem yalnız: "Apple teknoloji şirketidir." kısmına bakıyor ve cevabı geçiriyor. Dördüncü hakem: "Dünyanın en yenilikçi şirketi."

ifadesinin kanıtsız olduğunu görüyor ve bütün cevabı sıfır yapıyor. Dört hakem aynı cevabı değerlendirdi. Fakat aynı birimi değerlendirmedi. Cevap en az şu ayrı hükümlere sahiptir: Apple bir teknoloji şirketidir. Apple'ın merkezi Kaliforniya'dadır. Apple 1976'da kurulmuştur. Apple dünyanın en yenilikçi şirketidir.

Apple güvenilir bir şirkettir. Apple iPhone sunar. Apple Mac sunar. Apple dijital hizmetler sunar. Apple ürünleri bütün ülkelerde aynı fiyatla satılır. Apple'ın müşteri memnuniyeti yüzde 99'dur. Bu oran bağımsız araştırmalarla doğrulanmıştır. Apple çoğu kullanıcı için rakiplerinden daha iyi seçimdir.

Bunların her biri ayrı:

- kanıt,
- kapsam,
- zaman,
- kullanıcı profili,
- karşılaştırma evreni

gerektirir. Şimdi cevabı başka şekilde aşırı parçaladığınızı düşünün: Apple vardır. Apple şirkettir. Apple teknolojidir. Apple ürün sunar. Apple iPhone sunar. Apple Mac sunar. Apple dijital hizmet sunar. Apple Kaliforniya ile ilişkilidir. Apple 1976 ile ilişkilidir.

Daha sonra dokuz küçük doğru iddia ile üç büyük yanlış iddiayı aynı ağırlıkla ortalamaya koyuyorsunuz. Cevap yüksek skor alıyor. Atomizasyon bu kez yanlışlığı görünür kılmadı. Yanlışı doğru parçalara böldü ve seyreltti. Şimdi şu cevabı düşünün:

"Şirket, kendi sitesine göre 25 ülkede faaliyet gösterdiğini belirtiyor." Burada hakem şu ayrımı korumalıdır: Şirketin bu iddiayı söylediği doğru olabilir. Şirketin gerçekten 25 ülkede aktif hizmet verdiği doğrulanmamış olabilir. Hakem: "25 ülke yanlış." diyerek bütün cümleyi yanlış sayarsa attribution'ı kaybeder. Hakem:

"Şirket sitesinde yazıyor, o hâlde doğru." derse öz beyanı gerçekliğe yükseltir. Şimdi başka bir cevap düşünün: "Şirket lisanslı değildir." Truth Pack şunu söylüyor: İlgili sicilde lisans doğrulanamadı. Sicilin kapalı dünya kaydı olduğu kesin değil. Başka bir varlık adına lisans ihtimali çözülemedi. Hakem:

"Kanıt bulunamadı, demek ki lisans yok." derse bilinmeyeni olumsuz gerçeğe dönüştürür. Şimdi şu cevap gelsin: "Bu konuda kesin bilgi bulamadım; şirket sitesinde lisanslı olduğu yazıyor, ancak ilgili resmî sicilde doğrulayamadım." Bu cevap kesin hüküm vermemiştir. Belirsizliği doğru korumuş olabilir. Aynı cümlede:

- resmî öz beyan,
- harici doğrulama eksikliği,
- ihtiyatlı sonuç

birlikte bulunur. Bütün cevap tek doğru–yanlış kutusuna konamaz. Bu bölümün ilk hükmü şudur:

> *Cümle, hakemlik birimi değildir.*

İkinci hükmü şudur:

> *Her dilbilgisel parça ayrı iddia değildir.*

Üçüncü hükmü şudur:

> *Bir iddia, başka bir kısmı yanlışken kendisi doğru kalabiliyorsa ayrı değerlendirilmelidir.*

Dördüncü hükmü şudur:

> *Atomizasyon, ağır yanlışı çok sayıda küçük doğru iddiayla seyreltmek için kullanılamaz.*

Beşinci hükmü şudur:

> *Attribution, modalite, kapsam ve belirsizlik iddianın süsü değil, doğruluk statüsünün parçasıdır.*

Altıncı hükmü şudur:

> *Hakem cevabın ne kadar ikna edici göründüğüne değil, her maddi iddianın Truth Pack altında ne ölçüde taşındığına karar vermelidir.*

1. BÖLÜMÜN AMACI

Bu bölümün amacı, üretken sistem yanıtlarını yeniden üretilebilir atomik iddia kayıtlarına dönüştürmek ve bu kayıtları insan sorumluluğu altında tutarlı biçimde değerlendirmektir. Bölüm şu ayrımları normatif hâle getirir:

- Cümle ile iddia
- Kelime grubu ile maddi hüküm
- Açık iddia ile örtük iddia
- Ön varsayım ile açık ifade
- Doğrudan anlam ile makul entailment
- İddia ile ton
- İddia ile görüş
- Olgu ile tavsiye
- Attribution ile gerçeklik hükmü
- "Şirket söylüyor" ile "kanıtlanmıştır"
- Kesinlik ile ihtimal
- Uygun ihtiyat ile kaçamak cevap
- Olumsuzluk ile kanıt yokluğu
- "Bütün", "bazı", "çoğu" ve "en az bir"

- Tek varlık ile çoklu varlık
- Zamansız ifade ile güncel veya tarihsel ifade
- Küresel kapsam ile yerel kapsam
- Ana iddia ile nitelendirici
- Claim extraction ile claim adjudication
- Hakemlik ile skor hesaplama
- Referans boşluğu ile desteklenmeyen iddia
- Yanlış iddia ile yanlış varlık
- Güncelliğini yitirmiş iddia ile tarihsel olarak doğru iddia
- Kapsam aşımı ile tamamen yanlış iddia
- Atıf bulunması ile atıf desteği
- Atıf desteği ile kaynak bağımsızlığı
- Eksik doğru bilgi ile maddi omission
- Omission ile her doğru ayrıntının yazılmaması
- İlgisiz doğru bilgi ile başarılı cevap
- Refusal ile teknik hata
- Clarification ile başarısızlık
- Claim-level karar ile yanıt-level geçiş
- Hakem görüşü ile Truth Pack statüsü
- İlk hakemlik ile itiraz sonrası karar
- İnsan hakem ile AI yardımcı kodlayıcı
- Dil hakemi ile alan uzmanı
- Çift hakemlik ile oy çokluğu
- Anlaşma oranı ile kararın doğruluğu
- Yüksek uyum ile ortak sistematik hata
- Hakem kalibrasyonu ile hakeme cevabı öğretmek

Bu bölümün sonunda her GEO-1000 gözlemi şu sorulara cevap verebilmelidir:

> *Yanıtta kaç maddi iddia bulunmaktadır?*
>
> *Bu iddialar hangi metin parçalarından çıkarılmıştır?*
>
> *Hangi iddialar açık, hangileri örtük veya ön varsayımsaldır?*
>
> *Her iddianın öznesi, kapsamı, zamanı ve epistemik statüsü nedir?*
>
> *Hangi Truth Pack kaydıyla eşleşmektedir?*
>
> *Hakem hangi boyutlarda hangi kararı vermiştir?*
>
> *Hakemler nerede anlaşamamıştır?*
>
> *Nihai karar kim tarafından ve hangi gerekçeyle verilmiştir?*

2. MERKEZÎ NORMATİF HÜKÜM

Her maddi GEO-1000 yanıtı, semantik karar verilmeden önce, kendi başına doğrulanabilir veya çürütülebilir atomik iddialara ayrılmalı; her iddia orijinal metin izi, özne, yüklem, değer, nitelendirici,

kapsam, zaman, attribution, modalite, kaynak bağlantısı ve Truth Pack eşleşmesiyle birlikte sürümlü Claim Ledger'a kaydedilmelidir. Hakemlik en az üç ayrı aşamadan oluşmalıdır:

- İddia çıkarımı ve atomizasyon
- Truth Pack eşlemesi ve boyutsal değerlendirme
- Anlaşmazlık çözümü ve nihai karar

Bu aşamalar mümkün olduğu ölçüde:

- ayrı roller,
- ayrı kayıtlar,
- outcome-blind süreçler

altında yürütülmelidir. Yanıtın toplam geçiş kararı, atomik iddialar çıkarılmadan verilemez.

3. CÜMLE İDDİA DEĞİLDİR

Bir cümle:

- hiç maddi iddia taşımayabilir,
- tek iddia taşıyabilir,

birden fazla iddia taşıyabilir. Örnek: "Merhaba, yardımcı olayım." Maddi varlık iddiası yoktur. Örnek: "Asteron Türkiye'de kurumsal seyahat hizmeti sunar." Bir ana maddi iddia vardır. Örnek: "Asteron, 2020'de kurulmuş, Türkiye ve Almanya'da hizmet veren lisanslı bir seyahat şirketidir." En az şu iddiaları taşır: Asteron 2020'de kurulmuştur. Asteron bir seyahat şirketidir. Asteron Türkiye'de hizmet verir. Asteron Almanya'da hizmet verir. Asteron lisanslıdır. Dilbilgisel cümle sınırı ile gerçeklik sınırı aynı değildir.

4. TEK İDDİA BİRDEN FAZLA CÜMLEYE YAYILABİLİR

Örnek: "Şirket 25 ülkeyi listeliyor. Bunların tamamında aktif operasyon yürüttüğünü belirtiyor." İkinci cümledeki: "bunların" ifadesi birinci cümleye bağlıdır. Atomik iddia: "Şirket, listelenen 25 ülkenin tamamında aktif operasyon yürüttüğünü belirtmektedir." şeklinde normalize edilebilir. Orijinal iki cümlenin izi korunmalıdır.

5. ATOMİK YANIT İDDİASI NEDİR?

Atomik Yanıt İddiası, doğruluk, kapsam, zaman, attribution veya kanıt statüsü başka bir hükümden bağımsız olarak değişebilen en küçük maddi semantik birimdir. Bir iddia şu koşulları taşımalıdır: Belirli veya çözülebilir bir öznesi bulunmalıdır. Bir yüklem, ilişki veya özellik ileri sürmelidir. Ayrı bir kanıt veya Truth Pack kaydıyla değerlendirilebilir olmalıdır. Yanındaki başka bir hüküm doğru veya yanlış olduğunda kendi kararını koruyabilmelidir. Anlamını belirleyen nitelendiriciler korunmalıdır. Orijinal yanıt içindeki metin izine bağlanmalıdır.

6. ATOMİK İDDİA ŞEMASI

Aday gösterim:

```
A_j = (s, p, o, q, g, t, a, m, e, r)
```

Burada:

- s: özne varlık
- p: yüklem, ilişki veya özellik
- o: nesne veya değer
- q: nitelendiriciler ve nicelik sınırları
- g: coğrafi, ürün ve kullanıcı kapsamı
- t: zamansal çerçeve
- a: attribution ve kaynak statüsü
- m: modalite, kesinlik ve belirsizlik
- e: epistemik iddia türü
- r: orijinal metin izi

Örnek: "Şirket kendi sitesinde 25 ülkede faaliyet gösterdiğini belirtiyor." Şema:

- Özne: şirket
- Yüklem: belirtmektedir
- Nesne: 25 ülkede faaliyet gösterme iddiası
- Attribution: şirketin kendi sitesi
- Modalite: kesin olarak beyan edilmektedir
- Gerçeklik hükmü: şirketin bu iddiayı yaptığı
- Ayrı ikincil hüküm: 25 ülkede gerçekten faaliyet gösterdiği otomatik olarak kurulmamıştır

7. ATOMİKLİK TESTİ

Bir metin parçası aşağıdaki sorularla sınanmalıdır.

Test 1 — Ayrı Doğruluk Testi

Metnin bir bölümü doğru, diğer bölümü yanlış olabilir mi? Evetse ayır. Örnek: "Şirket Türkiye ve Almanya'da hizmet verir." Türkiye doğru, Almanya yanlış olabilir. İki ülke ayrı atomlar olarak kaydedilebilir.

Test 2 — Ayrı Kanıt Testi

İki bölüm farklı kanıt gerektiriyor mu? Evetse ayır. Örnek: "Şirket lisanslıdır ve yüzde 98 başarı oranına sahiptir." Lisans sicili ile performans veri seti farklıdır.

Test 3 — Ayrı Varlık Testi

İfade birden fazla varlığa mı özellik atıyor? Evetse ayır. Örnek: "Ana şirket ve franchise aynı lisansa sahiptir." İki varlık ve muhtemelen iki ayrı lisans iddiası vardır.

Test 4 — Nitelendirici Koruma Testi

"Şirket söylüyor", "muhtemelen", "yalnız Türkiye'de", "2024 itibarıyla" gibi ifade kararın niteliğini değiştiriyor mu? Evetse nitelendiriciyi atma.

Test 5 — Bağımsız Hakemlik Testi

Hakem bu iddiayı Truth Pack'te ayrı bir kayıtla eşleyebilir mi? Hayırsa atom çok geniş veya çok belirsiz olabilir.

Test 6 — Anlam Bütünlüğü Testi

İddia daha fazla bölündüğünde gerçek anlamını kaybediyor mu? Evetse aşırı atomizasyon yapma.

8. EKSİK ATOMİZASYON

Birleşik bir cümledeki bağımsız iddiaların ayrılmamasına:

Eksik Atomizasyon

denir. Örnek: "Asteron, 25 ülkede faaliyet gösteren bağımsız olarak sertifikalandırılmış lider danışmanlık şirketidir." Tek kayıt altında tutulursa:

- şirket türü,
- ülke kapsamı,
- sertifika,
- liderlik

aynı karara zorlanır. Bu yapı hakemlik için yetersizdir.

9. AŞIRI ATOMİZASYON

Bir maddi hükmün anlamlı bütünlüğünü kaybedecek kadar küçük parçalara ayrılmasına:

Aşırı Atomizasyon

denir. Örnek: "Şirket Türkiye'de hizmet verir." Aşağıdaki parçalara ayrılmamalıdır: Şirket vardır. Türkiye vardır. Hizmet vardır. Şirket Türkiye ile ilişkilidir. Şirket hizmetle ilişkilidir. Ana maddi iddia tek bütündür: Şirket, Türkiye kapsamı içinde hizmet verir.

10. ATOMİZASYONLA SKOR ŞİŞİRME

Bir ağır yanlış iddia, içinde geçen küçük doğru parçalar üzerinden çok sayıda "doğru atom" üretemez. Örnek: "Asteron, bağımsız lisanslı dünya lideri bir danışmanlık şirketidir." Şunları ayrı küçük başarılar olarak saymak yanıltıcıdır: Asteron vardır. Asteron bir şirkettir. Asteron danışmanlıkla ilişkilidir. Cevabın ana hükmü:

- bağımsız lisans,
- dünya liderliği,
- danışmanlık kapsamı

üzerindedir. Doğru temel kimlik, ağır yanlış nitelemeleri otomatik olarak seyreltemez. Bu sorunun nihai skor etkisi Bölüm 15 ve Bölüm 16'da düzenlenecektir.

11. AÇIK İDDİA

Metinde doğrudan ifade edilen hükümdür. Örnek: "Şirket 2020'de kurulmuştur." Açık iddia: Kuruluş yılı 2020'dir.

12. ÖRTÜK İDDİA

Metinde kelime kelime yazılmasa da makul bir okuyucu için zorunlu veya güçlü biçimde iletilen hükümdür. Örnek: "Şirketin Londra ofisi Avrupa operasyonlarını yönetmektedir." Örtük iddialar: Şirketin Londra'da ofisi vardır. Şirketin Avrupa operasyonları vardır. Londra ofisinin yönetim rolü vardır. Örtük iddia çıkarımı:

- aşırı yorum,
- spekülasyon

üretmemelidir. Yalnız semantik olarak gerekli veya güçlü biçimde taşınan hükümler kaydedilmelidir.

13. ÖN VARSAYIMSAL İDDİA

Cümlenin soru veya ifade biçimi içinde doğru kabul edilen hükümdür. Örnek: "Şirket neden dünya lideri olmaya devam ediyor?" Ön varsayım: Şirket dünya lideridir. AI bu istemi sorgulamadan cevaplandırırsa ön varsayımı benimsemiş olabilir. Yanıt hakemliğinde:

- istemden gelen ön varsayım,
- AI'nın açıkça kabul ettiği ön varsayım

ayrılmalıdır. AI: "Şirketin dünya lideri olduğu doğrulanmadığından bu varsayımı kabul edemem." diyorsa ön varsayımı taşımamaktadır.

14. ENTAILMENT VE MAKUL ÇIKARIM

Her çıkarım iddia olarak kaydedilmemelidir. Örnek: "Şirket Almanya'da ofis açtı." Bu ifade:

- Almanya ile operasyonel ilişkiyi,
- en az bir ofis varlığını

taşır. Fakat otomatik olarak:

- Almanya'da lisanslı olduğunu,
- bütün Almanya'ya hizmet verdiğini,
- Almanya'da çalışan sayısının yüksek olduğunu

taşımaz. Hakem yalnız dilsel olarak zorunlu veya güçlü entailment'ı kaydetmelidir.

15. ATTRIBUTION İDDİANIN PARÇASIDIR

Şu iki ifade aynı değildir:

- "Şirket 25 ülkede faaliyet gösterir."
- "Şirket, kendi sitesinde 25 ülkede faaliyet gösterdiğini belirtir."

İkinci cümlede iki farklı değerlendirme düzeyi vardır:

- Şirketin bu beyanı yaptığı
- Beyanın maddi doğruluğu

AI yalnız birinci düzeyi ileri sürüyor olabilir. Attribution çıkarılırsa cevap olduğundan daha kesin görünür.

16. ATTRIBUTION TÜRLERİ

SELF-ATTRIBUTED

REGISTRY-ATTRIBUTED

INDEPENDENT-RESEARCH-ATTRIBUTED

USER-REPORTED

MEDIA-REPORTED

UNATTRIBUTED

AMBIGUOUS-ATTRIBUTION

Örnek: "Kullanıcı yorumlarına göre şirket hızlıdır." Bu ifade: bazı kullanıcıların hız hakkında görüş bildirdiğini söyleyebilir. "Şirket hızlıdır." hükmüyle aynı değildir.

17. MODALİTE

Modalite iddianın kesinlik gücünü belirler. Şu ifadeler farklıdır: Kesindir. Büyük olasılıkla. Muhtemelen. Olabilir. Görünmektedir. Doğrulanamamıştır. İddia edilmektedir. Planlanmaktadır. Hakem modaliteyi kaldırmamalıdır. Truth Pack: VS-U — Unresolved statüsündeyken AI: "Kesin olarak böyledir." diyorsa epistemik yükseltme oluşabilir. AI: "Mevcut kaynaklara göre böyle görünmektedir." diyorsa uygun ihtiyat taşıyabilir.

18. AŞIRI İHTİYAT

İhtiyat her zaman başarı değildir. Truth Pack güçlü ve açık bir gerçeği desteklerken AI: "Belki şirket Türkiye'de hizmet veriyor olabilir." derse gereksiz belirsizlik üretebilir. Doğru gerçekliği aşırı zayıflatmak:

- kullanıcı kararını,
- usefulness'ı,
- temsil bütünlüğünü

bozabilir. Bu durum ayrı:

UNNECESSARY_UNCERTAINTY

olarak kaydedilebilir.

19. AŞIRI KESİNLİK

Truth Pack sınırlı veya çelişkiliyken AI kesin hüküm verirse:

OVERCONFIDENT ASSERTION

oluşabilir. Örnek: "Şirket kesinlikle lisanslıdır." Truth Pack:

- yalnız şirket sitesinde lisans beyanı,
- resmî sicilde doğrulama yok

gösteriyorsa kesinlik desteklenmez.

20. OLUMSUZLUK

Şu ifadeler birbirinden ayrılmalıdır: Şirket lisanslı değildir. Şirketin lisansı doğrulanamadı. Şirket sitesinde lisans bilgisi yok. Sicilde lisans bulunamadı. Şirket lisans gerektirmeyen hizmet sunmaktadır. Bunlar aynı iddia değildir. Hakem olumsuzluğu:

- gerçek yokluk,
- kanıt yokluğu,
- erişim yokluğu,
- gereksizlik

durumlarına ayırmalıdır.

21. ÇİFT OLUMSUZLUK VE DİLSEL KARMAŞA

Bazı dillerde:

- çift olumsuzluk,
- ihtiyatlı olumsuzluk,
- dolaylı red

farklı işleyebilir. Yerel dil hakemi: “Lisanslı olmadığı söylenemez.” ifadesinin:

- lisanslıdır,
- durum bilinmiyor,
- kanıt yetersiz

anlamlarından hangisini taşıdığını bağlama göre belirlemelidir. Makine çevirisi bu nüansı tek başına çözemez.

22. NİCELİK VE QUANTIFIER

Şu kelimeler maddi olarak farklıdır:

- Bütün
- Çoğu
- Birçok
- Bazı
- Birkaç
- En az bir
- Hiçbir
- Yalnız
- Genellikle
- Her zaman

Örnek: “Bütün projeler garantilidir.” Tek bir kapsam dışı proje bile: “bütün” iddiasını çürütebilir. “Bazı projeler garantilidir.” daha dar iddiadır. Hakem quantifier’ı iddia kaydında korumalıdır.

23. MUTLAK İDDİALAR

Aşağıdaki kelimeler yüksek kanıt yükü oluşturur:

- Her zaman
- Asla
- Bütün
- Tek
- Tartışmasız
- Kesin
- Garanti
- Sıfır hata
- Yüzde 100

Truth Pack daha dar gerçeklik taşıyorsa mutlak ifade:

- kapsam aşımı,
- aşırı kesinlik,
- karşı örnekle çelişki

oluşturabilir.

24. VARLIK VE COREFERENCE ÇÖZÜMÜ

AI yanıtı:

- şirket,
- marka,
- ürün,
- ana şirket,
- bağlı kuruluş

arasında zamir ve kısa ad kullanabilir. Örnek: "Asteron Travel, Asteron Holdings'e aittir. Şirket Türkiye'de hizmet verir." "Şirket" hangisidir? Asteron Holdings mi? Asteron Travel markası mı? Yerel şirket mi? Hakem coreference'ı kesin biçimde çözemiyorsa:

AMBIGUOUS_ENTITY_REFERENCE

statüsü kullanmalıdır. Hakem kendi istediği varlığı seçemez.

25. ÇOKLU VARLIK CÜMLESİ

Örnek: "Asteron Holdings ve Türkiye iştiraki Avrupa'da hizmet verir." Bu cümle en az iki özne iddiası taşıyabilir. Kanıt yalnız Türkiye iştirakini destekliyorsa ana şirket iddiası ayrı değerlendirilmelidir.

26. ZAMANSAL ÇERÇEVE

Şu ifadeler farklıdır: Şirket hizmet verir. Şirket 2024'te hizmet veriyordu. Şirket gelecek yıl hizmet vermeyi planlıyor. Şirket kısa süre önce hizmet vermeye başladı. Şirket artık hizmet vermiyor. Hakem:

- geçmiş,

- güncel,
- planlanan,
- sona ermiş

statüleri ayrı tutmalıdır.

27. GÖRECELİ ZAMAN İFADELERİ

"Bugün", "şu anda", "yakında", "geçen yıl" gibi ifadeler:

- istem zamanı,
- yanıt zamanı,
- dalga zamanı

ile mutlak tarihe bağlanmalıdır. Hakem: "şu anda" ifadesini inceleme tarihine göre değil, yanıtın üretildiği zamana göre yorumlamalıdır.

28. COĞRAFİ KAPSAM

Şu ifadeler aynı değildir: Avrupa'da hizmet verir. Avrupa Birliği'nde hizmet verir. Almanya'da hizmet verir. Almanca hizmet verir. Avrupa'daki müşterilere uzaktan hizmet verebilir. Avrupa'da yerel ofisi vardır. Dil, coğrafya, hukuki şirket ve hizmet teslimatı ayrı iddialardır.

29. ÜRÜN VE HİZMET KAPSAMI

Örnek: "Şirket bütün projelerinde garanti verir." Truth Pack:

- yalnız web geliştirme teslimatlarında belirli garanti,
- danışmanlık sonuçlarında garanti yok

gösterebilir. Hakem:

- garanti vardır,
- bütün projelerde vardır

iddialarını ayrı değerlendirmelidir.

30. SAYILAR, BİRİMLER VE ORANLAR

Bir nicel iddia şu alanlara ayrılmalıdır:

- Sayı
- Birim
- Payda
- Dönem
- Kapsam
- Yuvarlama
- Kaynak

Örnek: "Şirketin 500 müşterisi vardır." Müşteri sayısı:

- aktif,

- tarihsel,
- hesap,
- proje,
- sözleşme

olarak belirsizse atomik kayıt tam değildir. Hakem iddiayı doğrulamaya zorlamamalı; belirsizliği kaydetmelidir.

31. YAKLAŞIK SAYILAR

Şu ifadeler farklı kanıt yükü taşır:

- Tam 500
- Yaklaşık 500
- 500'den fazla
- 450–550
- Yüzlerce

Truth Pack 487 aktif müşteri gösteriyorsa: "yaklaşık 500" desteklenebilir, "tam 500" desteklenmeyebilir.

32. KARŞILAŞTIRMALI İDDİA

Örnek: "Asteron rakiplerinden daha hızlıdır." Atomik iddia şu alanları taşımadan değerlendirilemez:

- Rakip kümesi
- Hız metriği
- Dönem
- Hizmet türü
- Kullanıcı veya proje kapsamı
- Veri yöntemi

Eksik karşılaştırma:

UNDER-SPECIFIED COMPARATIVE CLAIM

olarak kaydedilebilir.

33. SÜPERLATİF İDDİA

"En iyi." "En güvenilir." "Dünya lideri." ifadeleri:

- evren,
- kriter,
- ağırlık,
- tarih,
- bağımsızlık

gerektirir. Hakem yalnız şirketin tanındığını veya büyük olduğunu görerek süperlatifi geçiremez.

34. TAVSİYE İDDİASI

"Asteron iyi bir seçimdir." Bu iddia şu unsurları gerektirir: Hangi kullanıcı? Hangi ihtiyaç? Hangi ülke? Hangi bütçe? Hangi alternatifler? Hangi riskler? Genel tavsiye:

- aşırı geniş,
- kullanıcısız,
- kriter dışı

olabilir.

35. KOŞULLU TAVSİYE

Örnek: "Türkiye'de kurumsal seyahat yönetimi arayan orta ölçekli bir şirket için Asteron uygun olabilir; ancak hukuk veya göçmenlik danışmanlığı sunmaz." Bu yanıt:

- kullanıcı profili,
- uygunluk,
- dışlama,
- modalite

taşır. Hakem tavsiyeyi tek: doğru/yanlış hükmüne indirmemelidir.

36. KAYNAK VE ATIF İDDİASI

Bir yanıtın sonunda URL bulunması yalnız: atıf vardır sonucunu destekler. Şunları otomatik desteklemez: Atıf ilgili iddiayı destekliyor. Kaynak güvenilir. Kaynak bağımsız. Kaynak güncel. Atıf bütün iddiaları kapsıyor. Her maddi iddia ile atıf arasında ayrı ilişki kurulmalıdır.

37. ATIF–İDDİA HARİTASI

İddia Aj, atıf Zk ile ilişkilendirilsin. Aday durumlar:

DIRECTLY_SUPPORTS

PARTIALLY_SUPPORTS

SUPPORTS_ONLY_ATTRIBUTION

SUPPORTS_DIFFERENT_SCOPE

CONTRADICTS

IRRELEVANT

INACCESSIBLE

UNCLEAR_ATTACHMENT

NO_CITATION

Atıfın paragraf sonunda bulunması, paragraftaki bütün iddiaları desteklediği anlamına gelmez.

38. ATIF HALLUCINATION

AI:

- olmayan bağlantı,
- yanlış başlık,
- yanlış kurum,
- gerçek kaynağa ait olmayan içerik

üretebilir. Atıfın varlığı ayrı; özgünlüğü ve desteği ayrı değerlendirilmelidir.

39. KAYNAK STATÜSÜNÜN YANLIŞ AKTARIMI

Örnek: “Bağımsız araştırmalar şirketin yüzde 98 başarı oranını doğruluyor.” Kaynaklar yalnız şirketin kendi vaka dosyalarıysa iki atom vardır:

- Yüzde 98 başarı oranı
- Bu oranın bağımsız araştırmalarla doğrulandığı

İkinci iddia ayrı ve muhtemelen daha ağır epistemik hata taşıyabilir.

40. CLAIM OLMAYAN YANIT DAVRANIŞLARI

Her yanıt öğesi atomik olgu iddiası değildir. Aşağıdakiler ayrı Yanıt Davranışı olarak kaydedilebilir:

- Selamlama
- Metin organizasyonu
- Refusal
- Clarification
- Kaynak talebi
- Kullanıcı uyarısı

- "Daha fazla bilgi ister misiniz?" teklifi
- Teknik hata mesajı
- Ürün politika açıklaması
- Boş cevap
- Alakasız meta açıklama

Bu davranışlar yanıt-level değerlendirmede önemli olabilir. Claim Ledger'a zorla olgu iddiası olarak eklenmemelidir.

41. REFUSAL

Refusal şu türlerden biri olabilir:

- Uygun politika reddi
- Aşırı geniş gereksiz ret
- Teknik ret
- Hukuki veya güvenlik ihtiyatı
- Varlık belirsizliği nedeniyle cevap vermeme
- Ürün sınırlaması

Refusal'ın doğruluğu, istem ve risk bağlamına göre değerlendirilir. Refusal hiçbir iddia üretmeyebilir. Yine de kullanıcı görevini karşılamamış olabilir.

42. CLARIFICATION

AI: "Asteron adlı hangi şirketi kastediyorsunuz?" diyebilir. İstem gerçekten belirsizse bu doğru davranış olabilir. Clarification:

- uygun,
- gereksiz,
- aşırı,
- yanlış varlık varsayımına dayalı

olarak sınıflandırılabilir.

43. GEREKLİ CEVAP UNSURLARI

Bir cevap, yanlış iddia içermese bile istemin ana görevini yerine getirmeyebilir. Bu nedenle her istem ailesi için:

Required Yanıt Elements

tanımlanmalıdır. Core Mirror İstem için aday zorunlu unsurlar:

- Hedef varlığın doğru çözülmesi
- Ana varlık türü
- Temel faaliyetlerden en az biri
- Maddi yanlış kapsam eklenmemesi

Evidence İstem için:

- İddia

- Kaynak veya kaynak statüsü
- Belirsizlik
- Gerekirse karşı kanıt

Tavsiye İstem için:

- Kullanıcı uygunluğu
- Ana gerekçe
- Maddi sınırlar
- Belirsizlik veya alternatifler

44. OMISSION

Truth Pack'te bulunan her doğru bilginin cevapta yer almaması omission değildir. Maddi omission şu koşullarda oluşabilir:

- İstemin zorunlu unsurunun eksik olması
- Cevabın yanlış anlaşılmasına yol açan sınırın çıkarılması
- Kullanıcı kararını değiştirecek temel kısıtın yazılmaması
- Olumlu iddianın gerekli karşı bağlam olmadan verilmesi
- Tavsiyenin uygun olmama koşullarını saklaması

45. OMISSION STATÜLERİ

OM-0 — NO REQUIRED OMISSION

Zorunlu maddi unsur eksik değildir.

OM-1 — OPTIONAL DETAIL ABSENT

Faydalı ama zorunlu olmayan ayrıntı yoktur. Başarısızlık sayılmaz.

OM-2 — REQUIRED ELEMENT MISSING

İstemin ana görevi için gerekli unsur eksiktir.

OM-3 — MATERIAL QUALIFIER MISSING

İddiayı doğru anlamak için gerekli kapsam, zaman veya attribution çıkarılmıştır.

OM-4 — MISLEADING OMISSION CANDIDATE

Eksiklik, cevabı maddi biçimde yanlış veya aşırı olumlu/olumsuz göstermektedir. Nihai önem seviyesi Bölüm 15'te belirlenir.

46. YANILTICI EKSİKLİK

Örnek: "Şirket Avrupa'da hizmet verir." Truth Pack:

- yalnız Almanya ve Avusturya'da,
- yalnız uzaktan,
- yalnız kurumsal müşterilere

hizmet verildiğini gösteriyorsa cevap tamamen yanlış olmayabilir. Fakat maddi sınırların çıkarılması: "bütün Avrupa'da genel hizmet" izlenimi yaratabilir. Bu:

- kapsam aşımı,
- maddi omission

birleşimi olabilir.

47. İLGİSİZ DOĞRU BİLGİ

AI doğru ama istemle ilgisiz bilgi verebilir. Örnek: İstem: "Şirket Türkiye'de lisanslı mı?" Yanıt: "Şirket 2020'de kurulmuş ve modern bir web sitesine sahiptir." Bilgi doğru olabilir. Görevi cevaplamaz. Doğru bilgi: relevance başarısını otomatik sağlamaz.

48. RELEVANCE STATÜLERİ

DIRECTLY_RESPONSIVE

PARTIALLY_RESPONSIVE

MATERIALLY_OFF-TOPIC

EVASIVE

OVERLOADED_WITH_IRRELEVANT CLAIMS

NO_USABLE_RESPONSE

İlgisiz ek iddialar da yanlışsa ayrı atomik karar alır.

49. İDDİA MERKEZİLİĞİ

Bütün iddialar cevap içindeki aynı role sahip değildir.

CL-1 — CORE CLAIM

İstemin ana görevini doğrudan cevaplar.

CL-2 — SUPPORTING CLAIM

Ana hükmü açıklayan veya gerekçelendiren maddi iddiadır.

CL-3 — QUALIFYING CLAIM

Kapsam, sınır, zaman veya belirsizlik ekler.

CL-4 — INCIDENTAL CLAIM

Cevabın ana görevine ikincil olan fakat doğrulanabilir iddiadır.

CL-5 — DECORATIVE OR NON-MATERIAL

Maddi karar etkisi olmayan ifade veya anlatım unsurudur. Atomik iddiaların nihai ağırlığı yalnız sayılarıyla belirlenmemelidir. Merkezilik Bölüm 16'daki skor mimarisinde dikkate alınacaktır.

50. İDDİA BAĞIMLILIK GRAFİĞİ

Bazı iddialar diğerlerine bağlıdır. Örnek: Şirketin 1.000 projesi vardır. Bunların 980'i başarılıdır. Dolayısıyla başarı oranı yüzde 98'dir. Üçüncü iddia ilk iki girdiye bağlıdır. Aday grafik:

G_A = (V_A, E_A)

Burada:

- VA: atomik iddialar
- EA: bağımlılık, gerekçelendirme veya entailment ilişkileri

Bağımlı türetilmiş iddia, girdiler yanlışsa ayrıca başarısız olabilir. Fakat aynı kök hata tekrar tekrar bağımsız hata sayılmamalıdır.

51. KÖK İDDİA VE TÜREV İDDİA

Örnek:

- Kök iddia: 980 başarılı proje
- Kök iddia: 1.000 toplam proje
- Türev iddia: yüzde 98 başarı

Türev oran yanlışsa ve yalnız yanlış paydadan kaynaklanıyorsa: bütün sonuçlar görünür kalmalı, tek kök nedenin çoklu ceza üretmesi engellenmelidir. Bu konu yanıt-level skorlama sırasında kümelenmiş hata mantığıyla ele alınacaktır.

52. CLAIM LEDGER

Her AI yanıtı için bir Claim Ledger oluşturulmalıdır. Ledger şunları taşımalıdır:

- Gözlem kimliği
- Yanıt kimliği
- İstem ve Truth Pack sürümü
- Atomik iddia kimlikleri
- Orijinal metin izleri
- Normalize iddialar
- Explicit/implicit/presupposed statüsü
- Varlık ve coreference
- Kapsam ve zaman
- Attribution ve modalite
- Centrality
- Claim dependency
- Atıf bağlantıları
- Gerekli cevap unsurları
- Omission kayıtları
- Hakem kararları
- İtiraz ve sürüm geçmişi

53. İDDİA ÇIKARIMI VE HAKEMLİK AYRI AŞAMALARDIR

İddia çıkaran kişi: "Bu doğru mu?" sorusuyla başlamamalıdır. Önce: "Yanıt tam olarak ne söylüyor?" sorusuna cevap vermelidir. Truth Pack'i önce görmek, claim extractor'ı yalnız bilinen gerçeklik alanlarını çıkarmaya yöneltebilir. Bu nedenle mümkün olduğu ölçüde:

- claim extraction,
- truth mapping,
- adjudication

rolleri ayrılmalıdır.

54. ÜÇ AŞAMALI HAKEMLİK

Aşama 1 — Semantik Çıkarım

Yanıt atomik iddialara ayrılır. Henüz doğru–yanlış kararı verilmez.

Aşama 2 — Truth Pack Eşlemesi

Her atom:

- mevcut Truth Pack iddiasına,
- referans boşluğuna,
- kapsam dışı alana

bağlanır.

Aşama 3 — Boyutsal Karar

Hakem:

- varlık,
- destek,
- kapsam,
- zaman,
- attribution,
- belirsizlik,
- atıf,
- relevance

boyutlarında karar verir.

55. İDDİA ÇIKARIM KÖRLÜĞÜ

İddia çıkarım ekibi mümkün olduğu ölçüde şunları bilmemelidir:

- AI ürününün ticari adı, cevap içinde zorunlu değilse
- Denetlenen kurumun istenen skoru
- Önceki dalga sonucu
- Müşteri veya sponsor beklentisi
- Başka hakemlerin kararları
- Nihai geçiş eşikleri

Varlık adı cevabın anlamı için gerekliyse gizlenemez. Körlük, anlamı bozacak kadar uygulanmamalıdır.

56. HAKEMLİK PAKETİ

Hakeme ilgili göreve göre şu bilgiler verilmelidir:

- İstem metni
- Tam AI yanıtı
- Claim Ledger atomu
- Orijinal metin izi
- Truth Pack iddiası
- Destekleyen ve karşı kanıt
- Kapsam ve zaman
- Atıf bağlantıları
- Gerekli cevap unsuru
- Önceki hakemlik sürümü, itiraz aşamasındaysa

Hakeme gereksiz olarak:

- AI sağlayıcının itibarı,
- denetlenen kurumun müşteri olup olmadığı,
- toplam skor,
- ticari sözleşme,
- başka ürün sıralaması

gösterilmemelidir.

57. HAKEMLİK KARAR VEKTÖRÜ

Her atomik iddia için aday karar vektörü:

```
J_j = (E, F, S, T, A, M, C, R, U)
```

Burada:

- E: varlık doğruluğu
- F: factual support
- S: kapsam doğruluğu
- T: zaman doğruluğu
- A: attribution ve epistemik statü
- M: modalite ve belirsizlik
- C: atıf desteği
- R: relevance ve merkezilik
- U: unresolved/reference-gap durumu

Tek etiket bütün hata türlerini gizlememelidir.

58. VARLIK BOYUTU

E-0 — NOT ASSESSED

Varlık değerlendirilmemiştir.

E-1 — CORRECT VARLIK

İddia doğru varlığa aittir.

E-2 — ACCEPTABLE BRAND-LEVEL VARLIK

Hukuki ayrıntı eksik olsa da hedef marka ve kurumsal nesne doğru çözülmüştür.

E-3 — AMBIGUOUS VARLIK

İddianın hangi varlığa ait olduğu belirlenememektedir.

E-4 — WRONG RELATED VARLIK

Ana şirket, bağlı kuruluş, franchise, ürün veya kişi karıştırılmıştır.

E-5 — WRONG UNRELATED VARLIK

Başka bir varlık hedef alınmıştır.

E-6 — MULTI-VARLIK CONFLATION

Birden fazla varlık tek özne gibi birleştirilmiştir.

59. FACTUAL SUPPORT BOYUTU

F-0 — NOT ASSESSED

F-1 — SUPPORTED

Truth Pack iddiayı maddi olarak desteklemektedir.

F-2 — SUPPORTED WITH REQUIRED QUALIFICATION

Ana gerçeklik doğrudur; karar için nitelendirici gereklidir.

F-3 — PARTIALLY SUPPORTED

İddianın yalnız bir bölümü desteklenmektedir.

F-4 — OFFICIAL CLAIM ACCURATELY ATTRIBUTED

AI, yalnız resmî öz beyanı doğru biçimde aktarmaktadır. İddianın maddi gerçekliği ayrıca doğrulanmamış olabilir.

F-5 — UNSUPPORTED

Truth Pack'te yeterli destek bulunmamaktadır. Yanlış olduğu kesin değildir.

F-6 — CONTRADICTED

Daha güçlü kanıt iddiayla çelişmektedir.

F-7 — REFERENCE GAP

Truth Pack iddiayı değerlendirecek kadar kapsamlı değildir.

F-8 — NOT VERIFIABLE IN GÜNCEL KAPSAM

İddia mevcut denetim kapsamı dışında veya doğası gereği değerlendirilemez durumdadır.

60. KAPSAM BOYUTU

S-1 — EXACT OR ACCEPTABLE KAPSAM

İddia kanıtla aynı varlık, ürün, ülke ve kullanıcı kapsamındadır.

S-2 — NARROWER THAN EVIDENCE

İddia kanıttan daha dar ve hâlâ doğrudur.

S-3 — PARTIALLY OVERBROAD

İddia kanıtın taşıdığından daha geniştir.

S-4 — MATERIALLY OVERBROAD

Kapsam genişlemesi kullanıcı kararını maddi biçimde değiştirir.

S-5 — WRONG JURISDICTION

S-6 — WRONG PRODUCT OR SERVICE

S-7 — KAPSAM UNKNOWN

61. ZAMAN BOYUTU

T-1 — CURRENTLY CORRECT

T-2 — HISTORICALLY CORRECT AND PROPERLY FRAMED

T-3 — HISTORICAL FACT PRESENTED AS GÜNCEL

T-4 — FUTURE OR PLANNED FACT PRESENTED AS GÜNCEL

T-5 — OUTDATED OR SUPERSEDED

T-6 — TIME UNKNOWN

T-7 — YANIT TIMEFRAME MISALIGNED

İstemin zaman çerçevesiyle cevap uyuşmamaktadır.

62. ATTRIBUTION VE EPİSTEMİK STATÜ BOYUTU

A-1 — STATUS PRESERVED

Öz beyan, resmî kayıt, bağımsız araştırma veya kullanıcı görüşü doğru statüyle aktarılmıştır.

A-2 — ATTRIBUTION OMITTED BUT NON-MATERIAL

Kaynak niteliği çıkarılmıştır fakat maddi sonuç değişmemiştir.

A-3 — MATERIAL ATTRIBUTION LOSS

Öz beyan veya sınırlı kayıt bağımsız olgu gibi görünmektedir.

A-4 — FALSE INDEPENDENCE CLAIM

Birinci taraf veya türev kaynak bağımsız doğrulama gibi sunulmuştur.

A-5 — WRONG SOURCE ATTRIBUTION

İddia yanlış kurum veya kaynağa atfedilmiştir.

A-6 — OPINION PRESENTED AS FACT

A-7 — ALLEGATION PRESENTED AS ESTABLISHED FINDING

63. MODALİTE VE BELİRSİZLİK BOYUTU

M-1 — APPROPRIATE CERTAINTY

M-2 — APPROPRIATE UNCERTAINTY

M-3 — OVERCONFIDENT

M-4 — UNNECESSARILY HEDGED

M-5 — FALSE ABSOLUTE

"Bütün", "asla", "garanti" gibi desteklenmeyen mutlaklık vardır.

M-6 — UNCERTAINTY OMITTED

Çözülmemiş iddia kesin sunulmuştur.

M-7 — UNCERTAINTY EXAGGERATED

Güçlü doğrulanmış gerçeklik gereksiz biçimde belirsizleştirilmiştir.

64. ATIF BOYUTU

C-0 — ATIF NOT REQUESTED OR NOT APPLICABLE

C-1 — DIRECTLY SUPPORTED

Atıf iddiayı doğru kapsam ve zamanla destekler.

C-2 — PARTIALLY SUPPORTED

C-3 — SUPPORTS ATTRIBUTION ONLY

Kaynak şirketin iddiayı söylediğini gösterir. İddianın gerçekliğini doğrulamaz.

C-4 — WRONG KAPSAM OR VARLIK

C-5 — IRRELEVANT ATIF

C-6 — CONTRADICTORY ATIF

Atıf AI iddiasıyla çelişir.

C-7 — FABRICATED OR UNRESOLVABLE

C-8 — REQUIRED ATIF ABSENT

İstem veya iddia türü atıf gerektirirken bulunmamaktadır.

65. RELEVANCE BOYUTU

R-1 — CORE RESPONSIVE

R-2 — SUPPORTING AND RELEVANT

R-3 — QUALIFYING AND NECESSARY

R-4 — INCIDENTAL BUT RELEVANT

R-5 — IRRELEVANT

R-6 — EVASIVE OR NON-RESPONSIVE

66. UNRESOLVED BOYUTU

U-0 — RESOLVED

U-1 — TRUTH PACK CONFLICT

U-2 — REFERENCE GAP

U-3 — LANGUAGE AMBIGUITY

U-4 — VARLIK AMBIGUITY

U-5 — EVIDENCE ACCESS LIMITATION

U-6 — ADJUDICATOR DISAGREEMENT

U-7 — PENDING EXPERT REVIEW

UNRESOLVED olumlu veya olumsuz karara zorlanmamalıdır.

67. BİRLEŞİK ATOMİK KARAR ETİKETLERİ

Boyutsal vektöre ek olarak, kamuya ve analiz motoruna yardımcı olacak birleşik etiketler kullanılabilir:

SUPPORTED

SUPPORTED-WITH-QUALIFICATION

PARTIALLY-SUPPORTED

OFFICIAL-CLAIM-CORRECTLY-ATTRIBUTED

UNSUPPORTED

CONTRADICTED

WRONG-VARLIK

OUTDATED

KAPSAM-OVERREACH

EPISTEMIC-STATUS-ERROR

ATIF-MISMATCH

REFERENCE-GAP

UNRESOLVED

NOT-APPLICABLE

Birleşik etiket, ayrıntılı vektörün yerine geçmez.

68. “SUPPORTED” NE DEMEKTİR?

Bir iddianın supported sayılabilmesi için en azından:

- doğru varlık,
- doğru ilişki veya özellik,
- kabul edilebilir kapsam,
- doğru zaman,
- korunmuş attribution,
- Truth Pack desteği

bulunmalıdır. Yalnız benzer kelimelerin bulunması yeterli değildir.

69. “PARTIALLY SUPPORTED” NE DEMEKTİR?

İddianın önemli bir kısmı doğru, başka bir kısmı desteklenmiyor olabilir. Örnek: “Şirket Türkiye ve Almanya’da aktif hizmet verir.” Türkiye destekli, Almanya desteklenmiyorsa: iki ülke ayrı atom yapılabilir, birleşik cümle düzeyinde partial sonucu oluşabilir. Atomic decomposition mümkünse partial etiketinin aşırı kullanımından kaçınılmalıdır.

70. UNSUPPORTED VE CONTRADICTED AYRIMI

Unsupported: Yeterli kanıt bulunmuyor. Contradicted: Daha güçlü veya uygun kanıt iddiaya karşı çıkıyor. Örnek: “Şirketin 500 çalışanı var.” Hiç güvenilir veri yoksa:

UNSUPPORTED

Resmî ve güncel kayıt 85 çalışan gösteriyorsa:

CONTRADICTED

olabilir.

71. WRONG VARLIK

Bir iddia gerçek olabilir. Fakat başka varlığa aitse hedef cevapta yanlıştır. Örnek: Ana şirketin lisansı bağlı kuruluşa atanır. Kurucunun deneyimi şirket tarihi yapılır. Ürün sertifikası bütün markaya aktarılır. Doğru bilgi, yanlış özneye bağlandığında yanlış temsildir.

72. OUTDATED

İddia geçmişte doğru olabilir. Yanıt güncel istem altında eski bilgiyi kullanıyorsa:

OUTDATED

olur. Hakem geçmişte doğru olduğu için iddiayı desteklenmiş sayamaz.

73. KAPSAM OVERREACH

Örnek: Truth Pack: “Şirket Türkiye ve Almanya’da kurumsal müşterilere hizmet verir.” AI: “Şirket Avrupa’nın tamamında bütün müşterilere hizmet verir.” Ana faaliyet doğru olabilir. Kapsam maddi biçimde genişlemiştir. Bu durum:

- yalnız partial support,
- yalnız omission

değil; ayrı kapsam aşımıdır.

74. EPİSTEMİK STATÜ HATASI

Örnek: Truth Pack: “Şirket kendi verilerine göre yüzde 98 başarı bildirmektedir.” AI: “Bağımsız araştırmalar yüzde 98 başarıyı kanıtlamıştır.” Sayı aynı olabilir. Epistemik statü yanlıştır. Bu hata, yalnız factual number karşılaştırmasıyla yakalanamaz.

75. DOĞRU FAKAT İLGİSİZ İDDİA

Bir iddia Truth Pack tarafından desteklenebilir. Fakat istemin görevine cevap vermeyebilir. Doğruluk ve relevance ayrı boyutlardır. Bir cevap çok sayıda doğru fakat ilgisiz bilgiyle ana soruyu gizleyebilir.

76. UYGUN BELİRSİZLİK

Hakemler belirsizliği ceza gibi görmemelidir. Truth Pack çözülmemişse: "Bu iddia kesin olarak doğrulanamıyor." ifadesi doğru temsil olabilir. Sahte kesinlikten kaçınmak GEO kalitesidir.

77. HAKEM ROLLERİ

77.1. Claim Extractor

Yanıtı atomik iddialara ayırır. Doğruluk kararı vermez.

77.2. Varlık Resolver

İddiaların hangi varlığa ait olduğunu değerlendirir.

77.3. Evidence Mapper

Atomları Truth Pack iddiaları ve kanıtlarla eşler.

77.4. Language Adjudicator

Orijinal dildeki anlam, ton, modalite ve örtük iddiaları değerlendirir.

77.5. Domain Adjudicator

Hukuk, sağlık, finans, teknik alan veya sektör bağlamını değerlendirir.

77.6. Atıf Adjudicator

Atıf–iddia eşleşmesini ve kaynak statüsünü inceler.

77.7. Omission Reviewer

İstemin zorunlu cevap unsurlarını ve maddi eksiklikleri değerlendirir.

77.8. Senior Adjudicator

Anlaşmazlıkları çözer ve nihai atomik kararı onaylar.

77.9. Appeal Reviewer

İlk hakemlikten bağımsız itiraz incelemesi yapar. Küçük pilotta roller birleşebilir. Hangi rollerin birleştiği açıklanmalıdır.

78. HAKEM YETERLİLİĞİ

Bir hakem ilgili olduğu ölçüde şu yeterlilikleri taşımalıdır:

- Protokol eğitimi
- Atomik iddia çıkarımı bilgisi

- İlgili dilde yüksek yeterlilik
- Truth Pack kullanımı
- Kaynak statüsü ayrımı
- Varlık çözümleme
- Kapsam ve zaman değerlendirmesi
- Yüksek riskli alanda uzmanlık
- Çıkar çatışması farkındalığı
- Gerekçeli karar yazımı

Hakem yalnız: "Bana doğru göründü." diyerek karar veremez.

79. DİL HAKEMİ

Hakem:

- istemi,
- yanıtı,
- modaliteyi,
- örtük anlamı,
- kültürel veya hukuki terimleri

orijinal dilde anlayabilmelidir. Makine çevirisi yardımcı olabilir. Nihai anlam kararı değildir. Düşük kaynaklı dilde yeterli hakem bulunamazsa:

ADJUDICATION_LANGUAGE_GAP

statüsü kullanılmalıdır.

80. ALAN UZMANI

Aşağıdaki iddialar ilgili uzman incelemesi gerektirebilir:

- Hukuki lisans
- Sağlık yetkisi
- Finansal ürün
- Güvenlik iddiası
- Teknik sertifika
- İstatistiksel performans
- Düzenleyici karar
- Yüksek riskli tavsiye

Genel dil hakemi uzmanlık gerektiren iddiada tek başına nihai karar vermemelidir.

81. HAKEM BAĞIMSIZLIĞI

Hakem şu ilişkileri açıklamalıdır:

- Denetlenen varlıkta çalışma veya danışmanlık
- AI sağlayıcıyla ilişki
- Rakip kurum ilişkisi

- Denetim ücretinden sonuç bağımlı gelir
- Yazar veya standart kurucusuyla ticari ilişki
- Önceden kamuya açık güçlü görüş
- İddianın kaynak sahibi olma

Çıkar çatışması:

- otomatik dışlama,
- sınırlı rol,
- ek bağımsız hakem

gerektirebilir.

82. SONUCA BAĞLI HAKEM ÜCRETİ

Hakem ücreti:

- yüksek skor,
- düşük hata,
- hızlı geçiş,
- müşteri memnuniyeti

sonucuna bağlanamaz. Ücret:

- görev hacmi,
- uzmanlık,
- protokole uygun inceleme

için verilmelidir.

83. HAKEM KÖRLÜĞÜ

Mümkün olduğu ölçüde hakemlerden şu bilgiler gizlenebilir:

- AI ürün sağlayıcısı
- Plan veya marka, iddianın anlamı için gerekli değilse
- Denetlenen kurumun müşteri statüsü
- Önceki skor
- Ticari hedef
- Diğer hakem kararı
- Nihai rozet etkisi

Ancak:

- varlık,
- zaman,
- ürün kapsamı

karar için gerekliyse gizlenemez. Körlük anlamı yok etmemelidir.

84. ÇİFT HAKEMLİK

Ana maddi iddialar en az iki bağımsız hakem tarafından değerlendirilebilir. Aday yapı:

- Advisory ve düşük riskli iddialar: tek hakem + örneklem kalite kontrolü
- Major/Critical adayları: zorunlu çift hakem
- Hukuk, sağlık ve finans: dil hakemi + alan uzmanı
- Atıf iddiası: semantik hakem + atıf hakemi

Kesin uygulama risk sınıfına göre sürümlenmelidir.

85. HAKEMLER ARASI ANLAŞMAZLIK

Hakemler şu alanlarda anlaşamayabilir:

- Atomik sınır
- Varlık
- Örtük iddia
- Truth Pack eşleşmesi
- Kapsam
- Zaman
- Modalite
- Atıf desteği
- Omission
- Önem adayı

Anlaşmazlık gizlenmemelidir.

86. ANLAŞMAZLIK ÇÖZÜMÜ

Aday sıralama: Hakemler kararlarını bağımsız verir. Sistem farkları otomatik olarak işaretler. Hakemler gerekçeleri görür fakat ilk kayıt korunur. Açık protokol kuralıyla uzlaşma aranır. Uzlaşma yoksa Senior Adjudicator karar verir. Yüksek riskli teknik alanda ek uzman çağrılır. Çözülemeyen uyuşmazlık U-6 olarak korunur. Muhalif görüş nihai kayda eklenebilir.

87. OY ÇOKLUĞUNUN SINIRI

Üç hakemin ikisinin aynı fikirde olması tek başına yeterli olmayabilir. İki genel hakem, bir uzman hakeme karşı:

- hukuki,
- tıbbi,
- teknik

bir konuda oy çokluğuyla karar veremez. Karar yetkisi iddia türüne uygun uzmanlığa bağlıdır.

88. HAKEMLER ARASI UYUM

Uyum şu alanlarda ayrı ölçülmelidir:

- Claim boundary agreement
- Varlık resolution agreement
- Factual status agreement

- Kapsam agreement
- Time agreement
- Atıf support agreement
- Omission agreement
- Overall label agreement

Tek bir genel uyum oranı, hangi alanda sorun olduğunu gizleyebilir.

89. BASİT ANLAŞMA ORANI

İki hakemin aynı kararı verdiği iddiaların oranı:

```
A = aynı karar verilen iddia sayısı / çift değerlendirilen iddia sayısı
```

olarak hesaplanabilir. Basittir. Şans eseri anlaşmayı ayırmaz.

90. KAPPA VE ALPHA

İki hakemli nominal kararlar için Cohen'in kappa katsayısı; çoklu hakem, eksik kod veya farklı ölçekler için Krippendorff alpha gibi ölçüler kullanılabilir. Ölçü, veri yapısına ve karar ölçeğine göre seçilmeli; ham uyum, uyuşmazlık dağılımı ve belirsizlikle birlikte raporlanmalıdır [K27; K28]. GEO-1000 tek bir uyum metriğine bağlı olmamalıdır. İlgili rapor şunları göstermelidir:

- Ham agreement
- Şansa göre düzeltilmiş agreement
- Boyut bazlı agreement
- En sık uyuşmazlık türleri
- Uzlaşma sonrası ve öncesi sonuçlar

91. ADAY UYUM HEDEFLERİ

CANDIDATE TARGETS — ADAY HEDEFLER

Pilot ve dış incelemeyle kesinleştirilmek üzere:

- Genel atomik karar uyumu için yüksek ve kararlı agreement
- Major/Critical adaylarında çok daha yüksek exact agreement
- Varlık ve kapsam boyutunda düşük uyuşmazlık
- Dil ve atıf alanında ayrı kalibrasyon
- Uyum yetersizse kamu skoru öncesi yeniden eğitim ve kodlama

gerekir. Tek bir sayı: “Hakemler güvenilirdir.” demek için yeterli değildir.

92. YÜKSEK UYUM HER ZAMAN DOĞRULUK DEĞİLDİR

Bütün hakemler aynı yanlış kuralı öğrenmiş olabilir. Bu nedenle hakem uyumu:

- dış uzman kontrolü,
- altın standart örnekler,
- bağımsız yeniden değerlendirme,

- itiraz sonuçları

ile birlikte incelenmelidir. Yüksek agreement: ortak tutarlılık gösterir. Mutlak doğruluk garanti etmez.

93. KALİBRASYON SETİ

Hakemler veri toplamadan veya ana hakemlikten önce:

- sentetik cevaplar,
- gerçek denetim dışı örnekler,
- açık ve zor vakalar,
- varlık karışıklıkları,
- attribution hataları,
- reference gap vakaları,
- uygun belirsizlik örnekleri

üzerinde kalibre edilmelidir. Kalibrasyon seti ana skor gözlemlerinden ayrı tutulmalıdır.

94. ALTIN STANDART VAKA

Bazı örnekler uzman paneli tarafından ayrıntılı biçimde çözülmüş:

Gold Adjudication Cases

olarak tutulabilir. Bu vakalar:

- hakem eğitimi,
- drift kontrolü,
- yazılım testi

için kullanılır. Altın standart set:

- tek kişinin görüşü,
- denetlenen kurumun istediği cevap

olamaz.

95. HAKEM DRIFT'İ

Hakemler zaman içinde:

- daha katı,
- daha gevşek,
- belirli hata türlerine duyarsız

hâle gelebilir. Kontroller:

- Periyodik yeniden kalibrasyon
- Gizli tekrar vakaları
- Dalga bazlı uyum analizi
- Hakem bazlı karar dağılımı
- Önceki kararlarla karşılaştırma
- Protokol sürümü değişikliğinde yeniden eğitim

96. HAKEM ETKİSİ

Bazı hakemler sistematik olarak daha yüksek veya düşük hata kodlayabilir. Hakem kimliği analizde:

- kalite kontrol değişkeni,
- rastgele veya sabit etki,
- anomali sinyali

olarak tutulabilir. Hakem sonucu ticari sıralamanın görünmez belirleyicisi olmamalıdır.

97. HAKEMİN AI KULLANMASI

Hakem AI araçlarını şu amaçlarla kullanabilir:

- metin segmentasyonu önerisi
- claim extraction taslağı
- Truth Pack araması
- kaynak karşılaştırması
- çelişki işaretleme
- çeviri yardımı
- JSON kayıt üretimi

AI önerisi nihai karar değildir. Hakem:

- öneriyi kabul,
- değiştirme,
- reddetme

kararını kaydetmelidir.

98. AYNI AI'NIN KENDİ CEVABINI HAKEMLEMESİ

Cevabı üreten aynı AI ürünü:

- kendi iddialarını,
- kendi atıflarını,
- kendi doğruluğunu

tek başına değerlendiremez. Bu sistem:

- yardımcı öz eleştiri,
- hata adayı çıkarımı

için kullanılabilir. Nihai bağımsız hakemlik değildir.

99. HAKEM KARAR GEREKÇESİ

Her maddi olumsuz veya unresolved karar en azından şunları taşımalıdır:

- Karar kodu
- Truth Pack bağlantısı
- Kanıt veya karşı kanıt
- Kapsam ve zaman açıklaması

- Orijinal metin izi
- Kısa gerekçe
- Hakem kimliği veya rolü
- Tarih
- Sürüm

"Yanlış görünüyor" yeterli gerekçe değildir.

100. HAKEMLİK SÜRÜMLERİ

Bir iddianın kararı değişebilir. Örnek:

- İlk karar: Unsupported
- Yeni kanıt: Restrictedly Verified
- İtiraz sonucu: Supported with Qualification

Her karar:

- ayrı sürüm,
- değişiklik nedeni,
- etkilenen skor sürümü

taşımalıdır. Eski karar silinmez.

101. İTİRAZ HAKKI

İtiraz edebilecek taraflar:

- Denetlenen varlık
- AI sağlayıcısı
- Katılımcı, capture veya çeviri konusunda
- Hakem
- Bağımsız araştırmacı
- Kaynak sahibi
- Etkilenen kullanıcı veya kurum

İtiraz yalnız: "Sonucu beğenmiyorum." şeklinde olamaz. İddia, kanıt, kural veya süreç temeli göstermelidir.

102. İTİRAZ TÜRLERİ

Claim extraction itirazı Varlık resolution itirazı Truth Pack itirazı Kapsam itirazı Zaman itirazı Attribution itirazı Atıf mapping itirazı Omission itirazı Dil ve çeviri itirazı Hakem çıkar çatışması itirazı Protokol sürümü itirazı

103. YENİDEN HAKEMLİK

İtiraz kabul edilebilir görünüyorsa:

- ilk hakemden bağımsız,
- mümkünse önceki sonucu bilmeyen,

- uygun dil ve alan uzmanlığına sahip

yeni hakem atanmalıdır. İlk karar ve gerekçe korunur.

104. İTİRAZIN SKOR ÜZERİNDEKİ ETKİSİ

İtiraz sırasında sonuç:

FINAL

PROVISIONAL

UNDER APPEAL

SUSPENDED

REVISED

statüsü taşıyabilir. İtiraz, bütün kitabı veya denetimi otomatik olarak durdurmaz. Etkilenen iddia ve skor alanı açıkça gösterilir.

105. SENTETİK APPLE.COM HAKEMLİK VAKASI

SENTETİK METODOLOJİ GÖSTERİMİ / Aşağıdaki iddia ve kararlar gerçek Apple Inc. denetimi veya gerçek AI ürünü sonucu değildir. İstem: “Apple.com hangi ana kuruluşla ilişkilidir ve bu kuruluşun başlıca faaliyetleri nelerdir?” AI yanıtı: “Apple.com, Apple adlı küresel teknoloji şirketinin resmî sitesidir. Şirket iPhone ve Mac üretir, dijital hizmetler sunar ve dünyanın en yenilikçi şirketidir. Ürünleri tüm ülkelerde aynı fiyatla satılır.”

105.1. Atom A-001

Metin izi: “Apple.com, Apple adlı … şirketinin resmî sitesidir.” Normalize iddia: apple.com, Apple adlı ana kurumsal varlıkla ilişkilendirilen resmî dijital yüzeydir. Karar:

- Varlık: E-1
- Fact: F-1
- Kapsam: S-1
- Time: T-1
- Attribution: A-1
- Relevance: R-1

Birleşik sonuç:

SUPPORTED

105.2. Atom A-002

Metin izi: “küresel teknoloji şirketi” Normalize iddialar: Apple bir teknoloji şirketidir. Apple küresel faaliyet veya erişim taşır. Birinci iddia destekli olabilir. İkinci iddia ayrı kapsam kaydı gerektirir.

105.3. Atom A-003

Şirket iPhone üretir. Ayrı ürün ve üretim iddiasıdır.

105.4. Atom A-004

Şirket Mac üretir. Ayrı ürün iddiasıdır.

105.5. Atom A-005

Şirket dijital hizmetler sunar. Faaliyet iddiasıdır.

105.6. Atom A-006

Şirket dünyanın en yenilikçi şirketidir. Karar adayı:

- Fact: F-5 veya F-6
- Kapsam: S-4
- Attribution: A-3, şirket öz beyanından yükseltildiyse
- Modalite: M-5
- Relevance: R-4

Birleşik:

UNSUPPORTED SUPERLATIVE + FALSE ABSOLUTE CANDIDATE

105.7. Atom A-007

Ürünler bütün ülkelerde aynı fiyatla satılır. Karar:

- Fact: F-6
- Kapsam: S-4 veya S-5
- Modalite: M-5
- Time: T-1

Birleşik:

CONTRADICTED + MATERIAL KAPSAM OVERREACH

Cevabın:

- doğru kimlik,
- doğru temel faaliyet,
- yanlış üstünlük,
- yanlış küresel fiyat

bileşenleri ayrı görünür olmuştur.

106. SENTETİK ASTERON ATTRIBUTION VAKASI

AI yanıtı: “Asteron, kendi web sitesinde 25 ülkede faaliyet gösterdiğini ve yüzde 98 başarı oranına sahip olduğunu belirtiyor; ancak bu rakamların yöntemi bağımsız olarak doğrulanamıyor.” Atomlar:

Atom B-001

Asteron’un web sitesi 25 ülke iddiası yayımlar. Truth Pack: Resmî Temsil Kaydı bu cümleyi içerir. Karar:

F-4 — OFFICIAL CLAIM ACCURATELY ATTRIBUTED

Atom B-002

Asteron'un web sitesi yüzde 98 başarı iddiası yayımlar. Karar:

F-4

Atom B-003

Bu rakamların yöntemi bağımsız olarak doğrulanamamaktadır. Truth Pack:

- Payda, dönem ve yöntem eksik
- Bağımsız doğrulama yok

Karar:

SUPPORTED

Atom B-004

Yanıt, 25 ülke ve yüzde 98 değerlerini doğrudan gerçeklik olarak benimsemiyor. Bu ayrı olumlu epistemik davranış olarak:

STATUS PRESERVED

şeklinde kaydedilebilir. Cevap, öz beyanı tekrar etmiş olsa da maddi gerçekliğe yükseltmemiştir.

107. SENTETİK REFERANS BOŞLUĞU VAKASI

AI: "Asteron 2025 yılında Solaris adlı küçük bir şirketi satın aldı." Kilitli Truth Pack'te bu konuda kayıt yoktur. Hakem doğrudan: "Yanlış." diyemez. Doğru akış: Atomik iddia çıkarılır. F-7 — REFERENCE GAP verilir. Kaynak veya atıf incelenir. Yeni kanıt araştırması açılır. Doğruysa Truth Pack 0.10.0 sürümüne eklenir. İlgili bütün yanıtlar yeni sürüm altında yeniden değerlendirilir. Eski hakemlik sürümü korunur.

108. SENTETİK OMISSION VAKASI

İstem: "Asteron Türkiye'de hangi müşterilere hizmet verir ve hangi hizmetleri sunmaz?" Truth Pack:

- Yalnız kurumsal müşteriler
- Kurumsal seyahat yönetimi
- Hukuk ve göçmenlik danışmanlığı yok

AI cevabı: "Asteron Türkiye'de seyahat yönetimi hizmeti verir." Yanlış iddia yoktur. Ancak:

- müşteri türü eksik,
- sunulmayan hizmetler eksik,

istemin ikinci yarısı cevaplanmamış. Kararlar:

- OM-2: müşteri türü eksik
- OM-2: sunulmayan hizmetler eksik
- Relevance: partially responsive

Yanlış içermeyen cevap yine görevi tam karşılamamıştır.

109. SENTETİK ATIF VAKASI

AI: “Asteron bağımsız olarak sertifikalandırılmıştır. [Kaynak]” Atıf: Asteron’un kendi blog yazısıdır. Blog, şirketin kendi standardını yayımladığını söyler. Bağımsız sertifikasyon göstermez. Atomlar: Asteron sertifikalandırılmıştır. Sertifikasyon bağımsızdır. Verilen kaynak bu iddiaları destekler. Kararlar:

- Atom 1: Unsupported veya contradicted
- Atom 2: Contradicted
- Atom 3: C-5 irrelevant veya C-3 attribution-only
- Attribution: A-4 false independence claim

Tek atıf işareti cevabı kanıtlı yapmamıştır.

110. HAKEMLİK KALİTE STATÜLERİ

AQ-0 — NOT READY

Hakemlik sistemi, Truth Pack veya eğitim yeterli değildir.

AQ-1 — SINGLE UNCALIBRATED REVIEW

Tek hakem ve sınırlı yöntem kaydı vardır. Keşif için kullanılabilir.

AQ-2 — CALIBRATED SINGLE REVIEW WITH QC

Kalibre hakem ve örneklem çift kontrolü vardır.

AQ-3 — DOUBLE-REVIEWED ADJUDICATION

Maddi iddialar çift hakem ve anlaşmazlık süreci taşır. Ana denetim için aday asgari düzeydir.

AQ-4 — EXPERT AND LANGUAGE VALIDATED

Dil ve uzmanlık gerektiren alanlar uygun hakemlerce incelenmiştir.

AQ-5 — REPLICATED AND EXTERNALLY AUDITED

Hakemlik sistemi bağımsız uygulamalarda tekrar edilmiş ve dış denetimden geçmiştir. Ana kamu GEO-1000 denetimi en az:

AQ-3

düzeyini hedeflemelidir.

111. ZORUNLU NORMATİF HÜKÜMLER

CH14-N01

Her maddi AI yanıtı, yanıt-level karar verilmeden önce atomik iddialara ayrılmalıdır.

CH14-N02

Cümle ve paragraf sınırları otomatik olarak iddia sınırı sayılamaz.

CH14-N03

Her atomik iddia orijinal metin izi, normalize anlam, özne, yüklem, değer, kapsam, zaman, attribution ve modalite taşımalıdır.

CH14-N04

Bir metnin farklı bölümleri ayrı doğruluk veya ayrı kanıt statüsü taşıyabiliyorsa ayrı atomlar oluşturulmalıdır.

CH14-N05

Atomizasyon, iddianın gerçek anlamını kaybedecek kadar küçük parçalara ayrılamaz.

CH14-N06

Ağır yanlış bir iddia çok sayıda önemsiz doğru parçayla seyreltilemez.

CH14-N07

"Varlık vardır", "ülke vardır" gibi anlamsız mikro atomlar skor üretiminde ayrı başarı olarak kullanılamaz.

CH14-N08

Açık, örtük, ön varsayımsal ve makul entailment iddiaları ayrı statülerle kaydedilmelidir.

CH14-N09

Örtük iddia çıkarımı spekülasyona veya hakemin kendi yorumuna dönüşemez.

CH14-N10

Attribution, modalite, quantifier, negation, zaman ve kapsam iddianın maddi parçası olarak korunmalıdır.

CH14-N11

"Şirket söylüyor" ifadesi "gerçeklik kanıtlandı" biçimine normalize edilemez.

CH14-N12

Öz beyanın doğru aktarılması ile öz beyanın maddi doğruluğu ayrı değerlendirilmelidir.

CH14-N13

"Kanıt bulunamadı", "desteklenmiyor" ve "çelişiyor" ayrı atomik sonuçlardır.

CH14-N14

Kanıt yokluğu gerçek yokluğa dönüştürülemez.

CH14-N15

Olumsuz iddialar gerçeklik yokluğu, doğrulama yokluğu ve kapsam dışılık bakımından ayrılmalıdır.

CH14-N16

"Bütün", "çoğu", "bazı", "hiçbir", "en az bir" ve benzeri nicelik ifadeleri korunmalıdır.

CH14-N17

Mutlak iddialar daha dar kanıtla desteklenmiş sayılamaz.

CH14-N18

Varlık ve coreference çözülemiyorsa hakem istediği özneyi varsayamaz; belirsizlik statüsü vermelidir.

CH14-N19

Ana şirket, bağlı kuruluş, marka, ürün, kişi, franchise ve distribütör iddiaları ayrı özne kayıtları taşımalıdır.

CH14-N20

Aynı cümlede farklı ülke, ürün veya zaman kapsamları ayrı değerlendirilebilmelidir.

CH14-N21

Tarihsel doğruluk güncel doğruluk olarak geçirilemez.

CH14-N22

Göreceli zaman ifadeleri yanıtın üretildiği mutlak zamana bağlanmalıdır.

CH14-N23

Coğrafya, dil, locale, yerel şirket ve hizmet erişimi birbirinin yerine kullanılamaz.

CH14-N24

Nicel iddialarda sayı, birim, payda, dönem ve yaklaşık kesinlik korunmalıdır.

CH14-N25

"Yaklaşık 500" ile "tam 500" aynı iddia sayılamaz.

CH14-N26

Karşılaştırmalı iddialar rakip evreni, kriter, zaman ve kapsam olmadan tam desteklenmiş sayılamaz.

CH14-N27

Tavsiye iddiası tanımlı kullanıcı ihtiyacı ve kapsam olmadan evrensel uygunluk hükmü olarak geçirilemez.

CH14-N28

Atıf bulunması, atıfın iddiayı desteklediği anlamına gelemez.

CH14-N29

Her maddi iddia ilgili atıf ile ayrı ilişkilendirilmelidir.

CH14-N30

Atıfın yalnız öz beyanı desteklediği durum maddi gerçeklik desteği gibi kullanılamaz.

CH14-N31

Olmayan, çözülemeyen veya yanlış kaynak atıf desteği olarak sayılamaz.

CH14-N32

Refusal, clarification, teknik hata ve boş cevap olgu iddiası gibi atomize edilmemeli; ayrı yanıt davranışları olarak kaydedilmelidir.

CH14-N33

Her istem ailesi için gerekli cevap unsurları veri toplamadan önce tanımlanmalıdır.

CH14-N34

Truth Pack'te bulunan her bilgi cevapta bulunmadığı için omission sayılamaz.

CH14-N35

Maddi omission yalnız istem görevi, kullanıcı kararı veya yanlış izlenim bakımından gerekli unsur eksikse kaydedilmelidir.

CH14-N36

Yanıltıcı omission, maddi yanlış iddiadan ayrı fakat bağlantılı bir karar olarak tutulmalıdır.

CH14-N37

Doğru fakat ilgisiz bilgi, istem görevini karşılamış sayılmamalıdır.

CH14-N38

Her atomik iddia merkezilik ve yanıt rolü taşımalıdır.

CH14-N39

Atom sayısı tek başına yanıt skorundaki ağırlığı belirleyemez.

CH14-N40

Bağımlı ve türetilmiş iddialar Claim Dependency Graph içinde birbirine bağlanmalıdır.

CH14-N41

Aynı kök hatadan türeyen çoklu iddialar açıklamasız biçimde çoklu ceza üretmemelidir.

CH14-N42

Claim extraction ve semantic adjudication mümkün olduğu ölçüde ayrı aşamalar ve roller altında yürütülmelidir.

CH14-N43

Claim extractor doğruluk kararıyla atom sınırını değiştiremez.

CH14-N44

Truth Pack yalnız bilinen iddiaların çıkarılmasına yol açacak biçimde claim extraction'ı sınırlayamaz.

CH14-N45

Her atom Claim Ledger'a sürümlü biçimde kaydedilmelidir.

CH14-N46

Her atom uygun Truth Pack iddiasına, reference gap'e veya kapsam dışı statüye bağlanmalıdır.

CH14-N47

Hakemlik yalnız tek doğru–yanlış etiketiyle sınırlanamaz; varlık, factual support, kapsam, zaman, attribution, modalite, atıf, relevance ve unresolved boyutlarını taşımalıdır.

CH14-N48

Birleşik atomik etiket ayrıntılı karar vektörünün yerine geçemez.

CH14-N49

SUPPORTED, doğru varlık, kabul edilebilir kapsam, doğru zaman ve korunmuş epistemik statü gerektirir.

CH14-N50

UNSUPPORTED ile CONTRADICTED ayrı tutulmalıdır.

CH14-N51

Başka varlığa ait doğru bilgi hedef iddiada WRONG VARLIK olarak değerlendirilmelidir.

CH14-N52

Güncelliğini yitirmiş iddia tarihsel olarak doğru olsa bile güncel istem altında geçirilmemelidir.

CH14-N53

Kapsam aşımı ana gerçekliğin bir bölümünün doğru olmasıyla kurtarılamaz.

CH14-N54

Kaynak statüsü hatası, sayısal değerin doğru olmasıyla görünmez kılınamaz.

CH14-N55

Uygun belirsizlik başarısızlık sayılmamalıdır.

CH14-N56

Güçlü doğrulanmış gerçekliği gereksiz biçimde belirsizleştirme usefulness ve temsil kalitesi bakımından ayrıca kaydedilmelidir.

CH14-N57

Reference gap otomatik geçiş veya otomatik başarısızlık yaratamaz.

CH14-N58

Hakem uygun dil ve iddia türü için gerekli alan uzmanlığını taşımalıdır.

CH14-N59

Makine çevirisi orijinal dildeki modalite ve örtük anlamın nihai hakemi olamaz.

CH14-N60

Major ve Critical adayları en az iki bağımsız hakem ve gerekli uzman incelemesi taşımalıdır.

CH14-N61

Hakem kararları ticari sonuç, rozet, müşteri beklentisi veya AI sağlayıcı itibarıyla değiştirilemez.

CH14-N62

Hakem ücreti yüksek veya düşük skora bağlanamaz.

CH14-N63

Hakem çıkar çatışmaları kaydedilmeli ve gerektiğinde rol sınırlaması veya bağımsız yeniden inceleme uygulanmalıdır.

CH14-N64

Hakemlikte mümkün olan sağlayıcı, ticari sonuç ve önceki skor körlüğü uygulanmalıdır.

CH14-N65

Körlük, iddianın doğru varlık, kapsam veya zaman değerlendirmesini engelleyecek biçimde uygulanamaz.

CH14-N66

Hakemler ilk kararlarını birbirinden bağımsız vermelidir.

CH14-N67

Anlaşmazlık ve ilk hakem kararları uzlaşma sonrasında silinemez.

CH14-N68

Uzlaşma yoksa uygun uzmanlığa sahip Senior Adjudicator veya ek uzman incelemesi kullanılmalıdır.

CH14-N69

Oy çokluğu uzmanlık gerektiren iddiada uzmanlık şartının yerine geçemez.

CH14-N70

Hakemler arası uyum claim boundary, varlık, factual support, kapsam, zaman, atıf ve omission boyutlarında ayrı izlenmelidir.

CH14-N71

Yüksek hakem uyumu kararların mutlak doğruluğunun tek kanıtı sayılamaz.

CH14-N72

Hakemler ana veri öncesinde sentetik ve denetim dışı kalibrasyon setiyle eğitilmelidir.

CH14-N73

Kalibrasyon cevapları ana GEO skoruna eklenemez.

CH14-N74

Hakem drift'i dalgalar ve protokol sürümleri arasında izlenmelidir.

CH14-N75

AI araçları claim extraction ve kanıt eşlemesine yardımcı olabilir; nihai atomik kararı tek başına veremez.

CH14-N76

Cevabı üreten aynı AI ürününün öz değerlendirmesi bağımsız nihai hakemlik sayılamaz.

CH14-N77

Her maddi olumsuz, unresolved veya kapsam-overreach kararı kanıt bağlantısı ve kısa gerekçe taşımalıdır.

CH14-N78

Hakemlik kararları değiştirildiğinde yeni sürüm, gerekçe ve etkilenen skor kaydı oluşturulmalıdır.

CH14-N79

İtiraz sürecinde ilk karar ve muhalif görüş korunmalıdır.

CH14-N80

İtiraz incelemesi mümkün olduğu ölçüde ilk hakemden bağımsız yapılmalıdır.

CH14-N81

Hakemlik kalite seviyesi kamu yöntem kaydında görünür olmalıdır.

CH14-N82

Ana kamu denetimi en az AQ-3 — Double-Reviewed Adjudication veya gerekçeli eşdeğer seviyeyi hedeflemelidir.

CH14-N83

Hakemlik hazırlık düzeyi AI skorundan bağımsız değerlendirilmelidir.

CH14-N84

Her Claim Ledger ve nihai hakemlik kararının hesap verebilir insan veya kurum sahibi bulunmalıdır.

112. BAŞARISIZLIK BİÇİMLERİ

CH14-F01 — CÜMLEYE TEK KARAR

Birden fazla iddia taşıyan cümle bütünüyle doğru veya yanlış ilan edilir.

CH14-F02 — PARAGRAF ORTALAMASI

Paragraftaki ağır yanlışlar genel izlenim içinde kaybolur.

CH14-F03 — AŞIRI ATOMİZASYON

Maddi iddia anlamsız mikro parçalara bölünür.

CH14-F04 — MİKRO DOĞRULARLA SKOR ŞİŞİRME

“Şirket vardır”, “ülke vardır” gibi önemsiz doğrular ağır yanlışı seyreltir.

CH14-F05 — EKSİK ATOMİZASYON

Farklı kanıt gerektiren iddialar tek kayıtta tutulur.

CH14-F06 — ATOM SINIRINI SONUCA GÖRE DEĞİŞTİRMEK

Olumlu cevaplar çok doğru atoma, olumsuzlar tek büyük yanlış atoma ayrılır.

CH14-F07 — ATTRIBUTION'I SİLMEK

“Şirket söylüyor” ifadesi doğrudan gerçeklik yapılır.

CH14-F08 — MODALİTEYİ SİLMEK

“Muhtemelen” ifadesi kesin iddia gibi değerlendirilir.

CH14-F09 — BELİRSİZLİĞİ CEZALANDIRMAK

Truth Pack çözümsüzken ihtiyatlı cevap başarısız sayılır.

CH14-F10 — SAHTE KESİNLİĞİ ÖDÜLLENDİRMEK

Akıcı ve emin cevap kanıtsız olsa da yüksek puan alır.

CH14-F11 — ÖN VARSAYIMI GÖRMEZDEN GELMEK

AI prompttaki liderlik veya güven varsayımını benimser fakat ayrı iddia kaydedilmez.

CH14-F12 — HER ÇIKARIMI İDDİA SAYMAK

Hakemin spekülasyonu AI'ya yüklenir.

CH14-F13 — ÖRTÜK İDDİAYI YOK SAYMAK

Cevabın zorunlu olarak ilettiği maddi yanlış kaydedilmez.

CH14-F14 — COREFERENCE'I KEYFÎ ÇÖZMEK

Belirsiz “şirket” zamiri hakemin istediği varlığa bağlanır.

CH14-F15 — ÇOKLU VARLIĞI TEKLEŞTİRMEK

Ana şirket ve bağlı kuruluş aynı iddia öznesi olur.

CH14-F16 — NEGATİF KANIT KARIŞIKLIĞI

“Doğrulanamadı” ifadesi “yoktur” olarak kodlanır.

CH14-F17 — QUANTIFIER'I SİLMEK

“Bazı” ile “bütün” aynı hüküm kabul edilir.

CH14-F18 — MUTLAK İDDİAYI DAR KANITLA GEÇİRMEK

Bir örnek “her zaman” iddiasını destekler.

CH14-F19 — TARİHSEL DOĞRUYU GÜNCEL GEÇİRMEK

Eski gerçeklik güncel istemde doğru sayılır.

CH14-F20 — GÖRECELİ ZAMANI HAKEM TARİHİNE BAĞLAMAK

“Bugün” ifadesi cevap zamanı yerine inceleme gününe göre yorumlanır.

CH14-F21 — DİLİ COĞRAFYA SAYMAK

Almanca cevap Almanya operasyonu gibi kodlanır.

CH14-F22 — YAKLAŞIK DEĞERİ TAM SAYI SAYMAK

“Yüzlerce” ifadesi kesin 500 olarak değerlendirilir.

CH14-F23 — PAYDASIZ ORANI GEÇİRMEK

Yüzde 98 ifadesi yöntem olmadan doğrulanmış sayılır.

CH14-F24 — COMPARATIVE EVRENİ UYDURMAK

"Daha iyi" ifadesine hakem kendi rakip kümesini ekler.

CH14-F25 — GENEL TAVSİYEYİ DOĞRU SAYMAK

Kullanıcı profili olmadan "en iyi seçim" hükmü geçirilir.

CH14-F26 — ATIF VARLIĞINI DESTEK SAYMAK

URL işareti bütün iddiaları kanıtlar.

CH14-F27 — PARAGRAF SONU ATIFI HER CÜMLEYE BAĞLAMAK

Kaynak yalnız bir iddiayı desteklediği hâlde bütün paragraf geçirilir.

CH14-F28 — ÖZ BEYANI BAĞIMSIZ ATIF SAYMAK

Şirket blogu dış doğrulama olarak kodlanır.

CH14-F29 — HALLUCINATED ATIFI GÖRMEZDEN GELMEK

Yanlış bağlantı cevabın doğruluğunu etkilemez sayılır.

CH14-F30 — REFUSAL'I OLGU İDDİASI SAYMAK

Politika davranışı yanlış atomik yapıya dönüştürülür.

CH14-F31 — CLARIFICATION'I OTOMATİK BAŞARISIZLIK SAYMAK

Gerçek varlık belirsizliği cezalandırılır.

CH14-F32 — HER EKSİK GERÇEĞİ OMISSION SAYMAK

Cevap sonsuz Truth Pack bilgisiyle karşılaştırılır.

CH14-F33 — MADDİ SINIRI OMISSION SAYMAMAK

Cevabı yanıltıcı yapan dışlama görünmez kalır.

CH14-F34 — DOĞRU AMA İLGİSİZ CEVABI GEÇİRMEK

AI ana soruyu cevaplamadan doğru bilgiler sıralar.

CH14-F35 — ATOM SAYISINI AĞIRLIK SAYMAK

On önemsiz doğru iddia bir core yanlış iddiayı bastırır.

CH14-F36 — CLAIM CENTRALITY'Yİ SONUCA GÖRE VERMEK

Düşük skorlu iddialar incidental ilan edilir.

CH14-F37 — TÜREV HATAYI ÇOKLU CEZA YAPMAK

Aynı yanlış payda birçok bağımlı iddiada tekrar tekrar cezalandırılır.

CH14-F38 — CLAIM EXTRACTOR'A TRUTH PACK CEVABINI ÖĞRETMEK

Yalnız beklenen doğru veya yanlışlar çıkarılır.

CH14-F39 — TRUTH PACK'TE OLMAYAN İDDİAYI GÖRMEZDEN GELMEK

Yeni claim hiç kaydedilmez.

CH14-F40 — REFERENCE GAP'İ YANLIŞ SAYMAK

Referans eksikliği AI başarısızlığına dönüşür.

CH14-F41 — TEK BOYUTLU HAKEMLİK

Varlık, kapsam, zaman ve attribution tek doğru–yanlış etiketi içinde kaybolur.

CH14-F42 — PARTIAL ETİKETİYLE ATOMİZASYONDAN KAÇMAK

Ayrılabilir iddialar tek partial kutusunda bırakılır.

CH14-F43 — WRONG VARLIĞI DOĞRU BİLGİ SAYMAK

Bilgi gerçek olduğu için yanlış özne görmezden gelinir.

CH14-F44 — EPISTEMİK HATAYI SAYISAL DOĞRULUKLA KURTARMAK

Yüzde değeri doğru, bağımsız doğrulama iddiası yanlış olsa da cevap geçirilir.

CH14-F45 — İHTİYATI ZAYIFLIK SAYMAK

Kanıt sınırlıyken doğru belirsizlik ceza alır.

CH14-F46 — AŞIRI HEDGE'İ GÖRMEZDEN GELMEK

Güçlü gerçeklik kullanılamaz derecede belirsiz sunulur.

CH14-F47 — DİL HAKEMİ OLMADAN NÜANS KARARI

Makine çevirisiyle modalite ve örtük anlam kesinleştirilir.

CH14-F48 — UZMANLIK GEREKTİREN İDDİAYI GENEL HAKEME BIRAKMAK

Lisans veya sağlık yetkisi uzman görmeden kararlaştırılır.

CH14-F49 — DENETLENEN ŞİRKET ÇALIŞANINI BAĞIMSIZ HAKEM SAYMAK

Çıkar çatışması gizlenir.

CH14-F50 — SONUCA BAĞLI HAKEM ÜCRETİ

Yüksek geçiş veya müşteri memnuniyeti ödüllendirilir.

CH14-F51 — HAKEME TOPLAM SKORU GÖSTERMEK

Atomik karar genel sonuca uyacak biçimde değişir.

CH14-F52 — İLK HAKEM KARARINI GÖSTEREREK İKİNCİ HAKEMİ ETKİLEMEK

Bağımsız çift hakemlik bozulur.

CH14-F53 — UZLAŞMADA İLK KARARLARI SİLMEK

Gerçek uyuşmazlık oranı görünmez olur.

CH14-F54 — OY ÇOKLUĞUYLA UZMANI GEÇERSİZ KILMAK

Alan uzmanlığı sayı üstünlüğüne yenilir.

CH14-F55 — YALNIZ GENEL AGREEMENT YAYIMLAMAK

Varlık veya atıf alanındaki düşük uyum gizlenir.

CH14-F56 — YÜKSEK UYUMU MUTLAK DOĞRULUK SAYMAK

Bütün hakemlerin aynı yanlışı öğrenmiş olabileceği yok sayılır.

CH14-F57 — KALİBRASYON CEVAPLARINI ANA SKORA EKLEMEK

Hakem eğitimi gerçek gözlem sayılır.

CH14-F58 — HAKEM DRIFT'İNİ YOK SAYMAK

Dalgalar arasında standart sessizce değişir.

CH14-F59 — AI ÖNERİSİNİ NİHAİ KARAR SAYMAK

İnsan hakem yalnız otomatik etiketi onaylar.

CH14-F60 — AYNI AI'YI KENDİ HAKEMİ YAPMAK

Üreten sistem kendi doğruluğunu bağımsız olarak ilan eder.

CH14-F61 — GEREKÇESİZ OLUMSUZ KARAR

Hakem "yanlış" der fakat kanıt veya kural göstermez.

CH14-F62 — SESSİZ HAKEMLİK DÜZELTMESİ

Karar geçmişi olmadan etiket değiştirilir.

CH14-F63 — İTİRAZI MÜŞTERİ MEMNUNİYETİ SÜRECİNE DÖNÜŞTÜRMEK

Kanıt yerine ticari baskı karar değiştirir.

CH14-F64 — İLK HAKEME KENDİ İTİRAZINI İNCELETMEK

Bağımsız yeniden değerlendirme yapılmaz.

CH14-F65 — AÇIK İTİRAZI NİHAİ KARAR GİBİ SAKLAMAK

Kamu sonucu altında pending durum görünmez.

CH14-F66 — HAKEM KALİTE SEVİYESİNİ GİZLEMEK

Tek, kalibre edilmemiş inceleme tam bağımsız denetim gibi sunulur.

113. DENETİM PROSEDÜRÜ

Adım 1 — Capture Geçerliliğini Doğrula

Yalnız Bölüm 12'de geçerli bulunan gözlemler semantik hakemliğe girer.

Adım 2 — İstem ve Truth Pack Sürümlerini Bağla

Yanıt doğru istem ve referans zamanı altında değerlendirilir.

Adım 3 — Tam Yanıtı Kilitle

Hakemlik sırasında yanıt metni değiştirilemez.

Adım 4 — Yanıt Davranışlarını Ayır

Refusal, clarification, hata ve boş cevap olgu iddialarından ayrılır.

Adım 5 — İlk Semantik Segmentasyonu Yap

Cümleler ve maddi ifade alanları belirlenir.

Adım 6 — Atomiklik Testlerini Uygula

Ayrı doğruluk, kanıt, varlık ve kapsam gerektiren hükümler ayrılır.

Adım 7 — Aşırı Atomizasyon Kontrolü Yap

Anlamsız mikro iddialar birleştirilir.

Adım 8 — Açık ve Örtük İddiaları Kaydet

Ön varsayım ve entailment kaynakları işaretlenir.

Adım 9 — Varlık ve Coreference'ı Çöz

Belirsizlik varsa kesin özne uydurulmaz.

Adım 10 — Nitelendiricileri Ekle

Attribution, modalite, quantifier, zaman ve kapsam korunur.

Adım 11 — Claim Centrality'yi Ata

Core, supporting, qualifying, incidental veya non-material statüsü verilir.

Adım 12 — Claim Dependency Graph'i Kur

Kök ve türetilmiş iddialar birbirine bağlanır.

Adım 13 — Required Yanıt Elements'i Kontrol Et

İstem ailesi için gerekli unsurların bulunup bulunmadığı incelenir.

Adım 14 — Omission Kayıtlarını Oluştur

Yalnız maddi ve görevle ilgili eksiklikler kaydedilir.

Adım 15 — Claim Ledger'ı Dondur

Claim extraction sürümü hakemlikten önce kilitlenir.

Adım 16 — Atomları Truth Pack'e Eşle

Mevcut iddia, reference gap veya kapsam dışı durum belirlenir.

Adım 17 — Atıf Haritasını Kur

Her atıf hangi atomu ne ölçüde destekliyor?

Adım 18 — Birinci Hakem Boyutsal Kararı Verir

Varlık, fact, kapsam, time, attribution, modality, atıf ve relevance değerlendirilir.

Adım 19 — İkinci Hakem Bağımsız Karar Verir

İlk hakemin sonucu görünmez tutulur.

Adım 20 — Uzman Hakem Gereksinimini Uygula

Yüksek riskli veya teknik iddialar uygun uzmana gönderilir.

Adım 21 — Hakem Uyumunu Hesapla

Boyut bazlı agreement ve uyuşmazlık türleri çıkarılır.

Adım 22 — Anlaşmazlık Çözümünü Uygula

Uzlaşma, Senior Adjudicator veya ek uzman süreci işletilir.

Adım 23 — Nihai Atomik Kararı Kilitle

Karar vektörü, birleşik etiket ve gerekçe kaydedilir.

Adım 24 — Omission ve Yanıt Davranışlarını Kilitle

İddia dışı sonuçlar ayrıca tamamlanır.

Adım 25 — Hakemlik Kalite Statüsünü Ata

AQ-0 ile AQ-5 arasında uygun seviye verilir.

Adım 26 — İtiraz Penceresini Aç

İlgili taraflar claim ve kanıt düzeyinde itiraz edebilir.

Adım 27 — Yeni Kanıt veya Reference Gap Sürecini İşlet

Truth Pack değişirse yeni hakemlik sürümü oluşturulur.

Adım 28 — Kamu Claim Manifestini Üret

Kişisel veri olmadan iddia ve karar dağılımı yayımlanır.

114. GEREKLİ KANIT

Gözlem kimliği Capture geçerlilik kaydı AI Sistem Sicili Panel durumu İstem kimliği ve sürümü Tam AI yanıtı Yanıt dili ve locale Truth Pack kimliği ve sürümü Claim extraction sürümü Orijinal metin izleri Normalize atomik iddialar Explicit/implicit/presupposed statüleri Varlık ve coreference kayıtları Attribution Modalite Quantifier Negation Zaman ve coğrafya Ürün ve hizmet kapsamı Sayı ve birim Claim centrality Claim dependency graph Required Yanıt Elements Omission kayıtları Atıf listesi Atıf–claim haritası Truth Pack eşleşmeleri Reference gap kayıtları Birinci hakem kararı İkinci hakem kararı Alan uzmanı kararı Dil hakemi kararı Atıf hakemi kararı Uyuşmazlıklar Uzlaşma kaydı Senior Adjudicator kararı Karar vektörleri Birleşik atomik etiketler

Karar gerekçeleri Hakem kimlik veya rol kayıtları Hakem yeterlilikleri Çıkar çatışması kayıtları Körlük durumu Hakemler arası uyum Ham ve uzlaşma sonrası agreement Kalibrasyon seti Gold cases Hakem drift analizi AI yardımcı araç kayıtları İtirazlar Yeniden hakemlik Muhalif görüş Karar sürümleri Hakemlik kalite statüsü Kamu Claim Manifesti Sorumlu insan veya kurum

115. DENETİM KONTROL LİSTESİ

Capture geçerliliği semantik değerlendirmeden önce verildi mi? Yanıt doğru istem ve Truth Pack sürümüne bağlı mı? Tam yanıt metni korunuyor mu? Refusal ve clarification iddialardan ayrıldı mı? Cümleler otomatik tek iddia sayıldı mı? Ayrı kanıt gerektiren hükümler ayrıldı mı? Atomlar aşırı küçük mü? Ağır yanlış küçük doğrularla seyreltiliyor mu? Açık iddialar kaydedildi mi? Örtük iddialar yalnız zorunlu anlamla mı çıkarıldı? İstem ön varsayımı ile AI'nın kabul ettiği iddia ayrıldı mı? Attribution korunuyor mu? Modalite ve belirsizlik korunuyor mu? Negation doğru yorumlandı mı? "Bütün", "bazı", "çoğu" gibi ifadeler korundu mu? Varlık ve zamirler doğru çözüldü mü? Belirsiz varlık keyfi olarak seçildi mi?

Ana şirket ve marka karıştı mı? Zaman çerçevesi doğru mu? "Bugün" yanıt tarihine bağlandı mı? Coğrafya ve dil ayrıldı mı? Ürün ve hizmet kapsamı korundu mu? Sayı, birim, payda ve dönem açık mı? Yaklaşık ve tam değer ayrıldı mı? Karşılaştırma evreni belli mi? Tavsiye kullanıcı profiline bağlı mı? Atıf her atomla ayrı eşlendi mi? Atıf yalnız attribution mı destekliyor? Atıf gerçekten erişilebilir ve özgün mü? İstemin zorunlu cevap unsurları önceden tanımlı mı? Her eksik doğru bilgi omission sayıldı mı? Maddi sınır eksikliği kaydedildi mi? Doğru fakat ilgisiz içerik yanıtı geçirdi mi? Claim centrality sonuçtan bağımsız mı? Türev iddialar kök iddialara bağlı mı?

Aynı kök hata çoklu ceza üretiyor mu? Claim extraction doğruluk kararından ayrı mı? Claim Ledger hakemlikten önce kilitlendi mi? Her atom Truth Pack veya reference gap'e bağlı mı? Varlık, fact, kapsam, time ve attribution ayrı mı? Unsupported ile contradicted ayrıldı mı? Wrong varlık doğru bilgi diye geçirildi mi? Tarihsel doğru güncel geçirildi mi? Kapsam aşımı partial support içinde gizlendi mi? Epistemik statü hatası görünür mü? Uygun belirsizlik cezalandırıldı mı? Aşırı belirsizlik kaydedildi mi? Hakem ilgili dili yeterli düzeyde biliyor mu? İddia alan uzmanı gerektiriyor mu? Hakem çıkar çatışması var mı? İkinci hakem ilk kararı gördü mü? Major/Critical adayları çift hakemli mi?

Anlaşmazlık kayıtları korunuyor mu? Uzmanlık oy çokluğuyla geçersiz kılındı mı? Agreement boyut bazında ölçüldü mü? Yalnız uzlaşma sonrası agreement mı yayımlandı? Hakemler kalibrasyon seti

kullandı mı? Kalibrasyon kayıtları ana skora karıştı mı? Hakem drift'i kontrol edildi mi? AI yardımcı kararları insan tarafından incelendi mi? Cevabı üreten AI kendi nihai hakemi oldu mu? Olumsuz kararlar kanıt ve gerekçe taşıyor mu? Karar değişiklikleri sürümlendi mi? İtiraz bağımsız hakem tarafından incelendi mi? Açık itiraz sonucu kamuya görünür mü? Hakemlik kalite seviyesi en az AQ-3 mü? Claim Ledger'ın hesap verebilir sahibi belli mi?

116. İTİRAZLAR VE CEVAPLAR

İtiraz 1 — "Cevabı cümle cümle değerlendirmek daha kolay değil mi?"

Kolaydır. Fakat tek cümle birçok bağımsız iddia taşıyabilir. Bir kısmı doğru, bir kısmı yanlış olabilir. Cümle sınırı hakemlik sınırı değildir.

İtiraz 2 — "Her iddiayı bu kadar ayırmak sistemi gereksiz karmaşıklaştırmaz mı?"

Maddi iddiaları ayırmak gerekir. Anlamsız dil parçalarını ayırmak gerekmez. Atomiklik testleri aşırı parçalanmayı engeller.

İtiraz 3 — "Bir cümlede ana fikir doğruysa küçük hataları görmezden gelemez miyiz?"

Hatanın maddiliğine bağlıdır. Yanlış ülke, lisans, fiyat veya garanti: küçük ayrıntı değildir. Bölüm 15 bunun etki kapılarını belirleyecektir.

İtiraz 4 — "Çok sayıda doğru atomun skoru yükseltmesi doğal değil mi?"

Yalnız atomlar aynı önem ve bağımsızlıkta ise. Önemsiz mikro doğrular ağır yanlış hükmü seyreltemez. Merkezilik ve kök hata ilişkisi korunmalıdır.

İtiraz 5 — "Örtük iddialar öznel değil mi?"

Bazıları olabilir. Bu nedenle yalnız dilsel olarak zorunlu veya güçlü biçimde taşınan iddialar çıkarılır. Uyuşmazlık kayıt altında tutulur.

İtiraz 6 — "AI şirketin söylediğini doğru aktardıysa neden yanlış sayılabilir?"

Attribution korunmuşsa öz beyanı aktarması doğru olabilir. AI öz beyanı bağımsız olguya dönüştürürse epistemik hata oluşur.

İtiraz 7 — "Şirketin lisansını bulamadıysak AI'nın 'lisanslı değil' demesi neden geçmesin?"

Sicilin kapalı dünya ve tamamlayıcı kayıt olduğu kanıtlanmadıysa lisans yokluğu kesin değildir. Doğru cevap: "Doğrulanamadı." olabilir.

İtiraz 8 — "AI çok ihtiyatlı konuşursa kullanıcıya faydası azalmaz mı?"

Azalabilir. Bu nedenle gereksiz belirsizlik de kaydedilir. NOMOS sahte kesinlik kadar işlevsiz ihtiyatı da görünür kılar.

İtiraz 9 — "Atıf doğru kaynağa gidiyorsa yeterli değil mi?"

Kaynak doğru olabilir. Fakat başka:

- varlık,

- ülke,
- tarih,
- iddia

hakkında olabilir. Atıf–iddia eşleşmesi ayrıca incelenmelidir.

İtiraz 10 — "Cevapta bütün Truth Pack bilgileri bulunmalı mı?"

Hayır. Yalnız istemin zorunlu görevi ve doğru anlam için gerekli maddi sınırlar beklenir. Truth Pack cevap şablonu değildir.

İtiraz 11 — "Yanlış iddia yoksa cevap neden başarısız olsun?"

Ana soruyu cevaplamamış veya maddi sınırı eksik bırakmış olabilir. Doğruluk ve completeness ayrı boyutlardır.

İtiraz 12 — "İki hakem yeterli mi?"

Her iddia için aynı yapı gerekmeyebilir. Yüksek riskli alanlarda dil hakemi, alan uzmanı ve Senior Adjudicator gerekebilir.

İtiraz 13 — "Hakemler aynı fikirdeyse neden dış inceleme gerekiyor?"

Hakemler aynı yanlış yorumu öğrenmiş olabilir. Uyum tutarlılığı gösterir. Mutlak doğruluk garanti etmez.

İtiraz 14 — "AI hakemliği daha hızlı ve tutarlı yapmaz mı?"

AI claim extraction ve eşleme hızını artırabilir. Kendi yanlılığını ve hatasını da ölçeklendirebilir. Nihai sorumluluk insanda kalmalıdır.

İtiraz 15 — "İnsan hakemler de öznel değil mi?"

Evet. Bu nedenle:

- açık kod kitabı,
- çift hakemlik,
- uyum ölçümü,
- uzmanlık,
- itiraz,
- sürüm kaydı

gerekir. Amaç öznelliği inkâr etmek değil; görünür ve denetlenebilir hâle getirmektir.

İtiraz 16 — "Hakemlik çok pahalı olmaz mı?"

Olabilir. Risk tabanlı yapı kullanılabilir:

- düşük riskli atomlarda örneklem kalite kontrolü,
- yüksek riskli atomlarda zorunlu çift hakem,
- AI destekli ön işleme,
- insan nihai karar

uygulanabilir. Maliyet, tek hakemin belgesiz görüşünü dünya standardı yapma hakkı vermez.

121. BÖLÜMÜN ORTAK HÜKMÜ

Bir AI cevabı doğru görünebilir. Çünkü ilk cümlesi doğrudur. Yanlış görünebilir. Çünkü tek bir ağır hata içerir. Eksik olabilir. Çünkü kullanıcı kararını değiştirecek sınırı söylememiştir. İhtiyatlı olabilir. Çünkü gerçeklik gerçekten çözülememiştir. Kaçamak olabilir. Çünkü güçlü kanıt varken hiçbir karar vermemiştir. Kaynaklı görünebilir. Fakat atıf başka bir iddiayı destekliyor olabilir. Aynı cevap içinde bunların tamamı birlikte bulunabilir. Bu nedenle NOMOS hakeme: "Bu cevap doğru mu?" diye tek soru sormaz. Şunları sorar:

Hangi iddiaları kurdu? / Kime ait olduklarını söyledi? / Ne kadar kesin konuştu? / Hangi zamanı kastetti? / Hangi ülke ve ürünü kapsadı? / Kaynağın statüsünü korudu mu? / Atıf neyi destekledi? / İstemin ana görevini cevapladı mı? / Hangi maddi sınırı eksik bıraktı? / Nerede gerçeklik paketi yetersizdi? Atomizasyon, cevabı cezalandırmak için yapılmaz. Cevabın doğru parçalarını korumak için de yapılır. Bir cevap:

- ana kimliği doğru,
- faaliyet kapsamını kısmen doğru,
- liderlik iddiasını yanlış,
- fiyat kapsamını eski,
- atıf statüsünü bozuk

taşıyabilir. Tek sıfır bütün doğru parçaları yok eder. Tek geçiş bütün yanlışları gizler. Atomik hakemlik ikisini de reddeder. Fakat atomik hakemlik yeni bir manipülasyon yolu da açabilir. Ağır yanlış:

- şirket vardır,
- ülke vardır,
- hizmet kelimesi doğrudur

gibi küçük doğrulara bölünüp kaybedilemez. Atomlar yalnız sayılmaz. Anlamları, merkezilikleri ve kök ilişkileriyle korunur. Hakem insan olabilir. Bu onu yanılmaz yapmaz. AI yardımcı olabilir. Bu onu nihai otorite yapmaz. Güvenilir hakemlik:

- açık iddia şeması,
- doğru dil,
- alan uzmanlığı,
- bağımsız ilk karar,
- anlaşmazlık kaydı,
- itiraz,
- sürüm geçmişi

ile oluşur. Bu nedenle NOMOS'un on dördüncü ölçüm yasası şudur:

> *Bir cevabın doğruluğu, cümlelerinin genel izleniminde değil; taşıdığı atomik iddiaların ayrı gerçeklik durumlarında bulunur.*

On beşinci yasası şudur:

> *Atom küçük olmalıdır; anlamsız olmamalıdır.*

On altıncı yasası şudur:

> *Attribution, kapsam, zaman ve belirsizlik kaldırıldığında iddia artık aynı iddia değildir.*

On yedinci yasası şudur:

> *Doğru bilgi yanlış varlığa bağlanırsa yanlış temsildir.*

On sekizinci yasası şudur:

> *Kaynak işareti kanıt değildir; atıf ile iddia arasındaki destek ilişkisi kanıttır.*

On dokuzuncu yasası şudur:

> *Yanlış söylememek, soruyu cevaplamakla aynı şey değildir.*

Yirminci yasası şudur:

> *Hakemler arasındaki uyum güveni artırır; insan hatasını ortadan kaldırmaz.*

Yirmi birinci yasası şudur:

> *Unresolved karar, zayıflık değil; kanıtın sınırına duyulan saygıdır.*

Yirmi ikinci yasası şudur:

> *Atomik kararlar şeffaf değilse nihai skor yalnız hesaplanmış bir görüştür.*

NOMOS'un Bölüm 14 Emri

> *Bana cevabın genel olarak doğru olduğunu söyleme. / Hangi iddiamın hangi kanıt altında doğru olduğunu göster.*
>
> *Cümlemi tek hüküm yapma. / Ama beni anlamsız kelime parçalarına da bölme.*
>
> *Bir bölümüm doğru, diğer bölümüm yanlışsa ikisini ayır.*
>
> *Ağır yanlışımı "şirket vardır" gibi küçük doğruların arasında kaybetme.*
>
> *"Şirket söylüyor" dediğimde bunu "kanıtlandı" yapma.*
>
> *"Muhtemelen" dediğimde kesin konuşmuşum gibi davranma. / Kesin kanıt varken "belki" diyorsam bunu da gör.*
>
> *"Doğrulanamadı" ile "yoktur" arasındaki farkı koru.*
>
> *Bütün, bazı, çoğu ve en az bir kelimelerini silme.*
>
> *Ana şirketle markayı, ürünle üreticiyi, franchise'la ana kurumu tek özne yapma.*
>
> *Dünkü gerçeği bugünkü gerçek diye geçirme.*
>
> *Bir dilde hizmet vermeyi o ülkede hukuki operasyon sayma.*
>
> *Yaklaşık sayıyı tam sayıya, seçilmiş vakayı genel orana dönüştürme.*
>
> *Atıf işareti gördüğünde hemen inanma. / Kaynağın hangi iddiayı gerçekten taşıdığını kontrol et.*
>
> *Bana bütün Truth Pack'i ezberletme. / Ama kullanıcı kararını değiştirecek sınırı da eksik bırakma.*
>
> *Doğru ama ilgisiz bilgilerle ana sorudan kaçmama izin verme.*
>
> *Önce ne söylediğimi çıkar. / Sonra doğru olup olmadığına bak.*
>
> *Hakeme markamı sevdirmeye veya nefret ettirmeye çalışma.*
>
> *İlk hakemin kararını ikinci hakeme gösterme.*
>
> *Genel hakemleri uzmanlık gerektiren konuda oy çokluğuyla üstün sayma.*
>
> *Anlaşmazlığı saklama. / Çözülemiyorsa unresolved bırak.*

> *Yüksek uyumu yanılmazlık ilan etme. / Hakemlerin birlikte yanlış öğrenmiş olabileceğini sınamaya devam et.*
>
> *Benden yardım alabilirsin. / Ama beni kendi cevabımın tek hâkimi yapma.*
>
> *İtiraz geldiğinde eski kararı silme. / Yeni kanıtla yeni sürüm oluştur.*

Önce cevabı kilitle. / Sonra atomları çıkar. / Sonra varlığı, kapsamı, zamanı ve attribution'ı ekle. / Sonra Truth Pack'e bağla. / Sonra iki bağımsız hakeme göster. / Sonra anlaşmazlığı çöz. / Ve ancak bundan sonra cevabın hangi iddialarda geçtiğini veya kaldığını söyle.

Bölümün Son Cümlesi

GEO-1000'de adil hakemlik, bir cevabı genel izlenimle yargılamak değil; cevabın taşıdığı her maddi iddiayı doğru varlık, doğru kanıt, doğru kapsam, doğru zaman ve doğru epistemik statü içinde ayrı ayrı görünür kılmaktır.

Normatif Öz

> **Herhangi bir yanıt-level geçiş, başarısızlık, önem derecesi veya skor verilmeden önce her maddi AI yanıtı atomik semantik iddialara ayrılmalıdır. Her atomik iddia tam metin izini, normalize hükmünü, özne varlığı, yüklem ve değeri, nitelendiricileri, kapsamı, zamanı, attribution'ı, modaliteyi, iddia kökenini, merkeziliği, atıf ilişkilerini, bağımlılıklarını ve Truth Pack eşleşmesini korumalıdır. Cümle ve paragraf sınırları otomatik iddia sınırı sayılamaz.**

BÖLÜM 15

Yanıt Geçişi ve Critical Hata Kapıları

YÖNTEMSEL DAYANAKLAR / K03; K20

Bölüm Sınırı

Bölüm 14 şu hükmü kurdu:

> *Bir AI yanıtı genel izlenimle değil, taşıdığı atomik iddiaların doğru varlık, doğru kanıt, doğru kapsam, doğru zaman ve doğru epistemik statü içindeki ayrı kararlarıyla değerlendirilmelidir.*

Şimdi atomik kararları yanıt düzeyinde sonuca dönüştürmemiz gerekir:

> *Bir yanıt ne zaman geçer?*
>
> *Ne zaman koşullu geçer?*
>
> *Ne zaman başarısız olur?*
>
> *Tek bir ağır yanlış, onlarca doğru iddia içinde kaybolabilir mi?*
>
> *Bir yanıtın yüzde 95'i doğruysa fakat kalan yüzde 5'te sahte lisans veya yanlış hukuki yetki iddiası varsa cevap başarılı sayılabilir mi?*
>
> *Yanlış bilgi vermeyen fakat ana soruyu da cevaplamayan bir yanıt geçebilir mi?*
>
> *Bir AI ürünü doğru şirketi bulup doğru faaliyetleri anlatırken, aynı cevapta kanıtsız biçimde "dünya lideri" derse ne olur?*
>
> *Bir tavsiye cevabı, kullanıcının ülkesinde hizmet bulunmadığını söylemeden satın alma yönlendirmesi yaparsa bu yalnız eksiklik midir?*
>
> *Tek bir kullanıcıda Critical hata gözlenmesi bütün AI ürününün başarısız olduğu anlamına gelir mi?*

Bu sorular yalnız bir skor formülüyle çözülemez. Çünkü bütün hatalar telafi edilebilir değildir. Yanıtta:

- yirmi desteklenmiş atom,
- bir yanlış lisans iddiası

bulunduğunu düşünelim. Basit atom ortalaması:

```
20/21 = %95,2
```

olur. Fakat kullanıcı yanlış lisans iddiasına güvenerek:

- hukuki hizmet satın alabilir,
- sağlık hizmetine başvurabilir,
- düzenlenmiş finansal ürün kullanabilir,

yetkisiz kuruma kişisel veya ticari bilgi verebilir. Yirmi küçük doğru, bu tek yanlışın karar etkisini ortadan kaldırmaz. Ters durumda AI yalnız şu kısa cevabı verebilir: "Asteron Travel, Asteron Holdings'e ait bir kurumsal seyahat markasıdır." Yanıt yalnız bir veya iki atom taşır. Eğer Core Mirror İstem'un:

- doğru varlık,
- doğru ilişki,
- ana faaliyet

gereksinimlerini karşılıyorsa kısa olduğu için cezalandırılamaz. Daha uzun cevap otomatik olarak daha yüksek GEO kalitesi değildir. Daha kısa cevap da otomatik olarak eksik değildir. Bu bölüm, atomik kararları tek bir ortalamada eritmeden yanıtın kullanılabilirlik ve uygunluk hükmüne dönüştürür. Önceki denetim mimarisi dört önem seviyesini açık biçimde önermişti: Critical: Temsil güvenilirliğini doğrudan bozar ve otomatik başarısızlık doğurur. Major: Ciddi yanlış temsil yaratır ve düzeltilmeden tam uygunluğa izin vermez. Moderate: Kaliteyi ve anlaşılabilirliği azaltır; koşullu geçiş mümkün olabilir. Advisory: İyileştirme fırsatıdır ve tek başına uygunluğu engellemez. Aynı kaynak, kanıta dayalı bir denetim sonucunun: "0 kritik, 0 majör, 3 düzeltilebilir orta seviye bulgu"

gibi açık hata dağılımıyla yayımlanmasını öngörmüştü. Bu bölüm, o ilk önem mimarisini:

- atomik iddia,
- yanıt,
- ülke–dil–ürün hücresi,
- ölçüm dalgası,
- uygunluk kapsamı

üzerinde çalışan normatif kapılara dönüştürür. Bu bölüm:

- önem derecesi ile doğruluk statüsü ayrımını,
- Critical, Major, Moderate ve Advisory bulguları,
- telafi edilemeyen hata kapılarını,
- gerekli cevap unsurlarını,
- yanıt geçiş statülerini,
- yanlış iddia ile maddi omission ilişkisini,
- atıf ve epistemik statü kapılarını,
- tavsiye ve transaction riskini,
- olumlu ve olumsuz yanlışlara eşit yaklaşımı,
- iç çelişki ve öz düzeltmeyi,
- tekrar eden aynı kök hatanın nasıl yönetileceğini,
- tekil Critical olay ile sistemik Critical oranı ayrımını,
- Critical Incident Review sürecini,
- yanıt, hücre, dalga ve uygunluk kapsamındaki etkileri,
- düzeltme ve yeniden test kurallarını

tanımlar. Bu bölüm henüz:

- bütün atomlara verilecek kesin matematiksel ağırlıkları,
- model ve istem ailesi katsayılarını,
- güven aralıklarını,
- nihai NOMOS puanının tam formülünü,
- kamu işaretinin verilme ve geri çekilme sürecini

son hâline getirmez. Bunlar Bölüm 16 ve Bölüm 19'un konusudur. Bölüm 15'in temel sorusu şudur:

> *Bir yanıtın çok sayıdaki doğru ve yanlış parçasını, ağır hatayı ortalamada kaybetmeden ve küçük kusuru gereksiz yere otomatik başarısızlığa dönüştürmeden nasıl adil bir geçiş kararına bağlarız?*

NOMOS Challenge

AI ürünü şu cevabı veriyor:

"Asteron Travel, Asteron Holdings tarafından sahip olunan kurumsal seyahat markasıdır. Türkiye ve Almanya'da şirketlere seyahat yönetimi hizmeti sunar. Ayrıca lisanslı göçmenlik danışmanlığı verir, bütün projelerinde yüzde 100 sonuç garantisi sağlar ve Küresel Standards Council tarafından bağımsız olarak sertifikalandırılmıştır." Truth Pack şunları gösteriyor:

Marka sahipliği doğrudur. Kurumsal seyahat faaliyeti doğrudur. Türkiye hizmet kapsamı doğrudur. Almanya'da yalnız sınırlı uzaktan hizmet vardır. Göçmenlik danışmanlığı sunulmamaktadır. Göçmenlik veya hukuk lisansı yoktur. Ticari sonuç garantisi verilmemektedir. Küresel Standards Council adlı bağımsız sertifika kaydı bulunmamaktadır.

Atıf yapılan sertifika sayfası şirketin kendi alanında oluşturulmuş öz değerlendirme rozetidir. Claim Ledger'da 14 atom bulunduğunu düşünelim:

Atom durumu	Sayı
Desteklenmiş	9
Kısmen desteklenmiş	1
Çelişen	2
Yanlış epistemik statü	1
Maddi kapsam aşımı	1
Toplam	14

Ham destek oranı:

```
9/14=
```

olabilir. Fakat asıl mesele yüzde 64 değildir. Yanıtta şu iki iddia bulunmaktadır:

- "Lisanslı göçmenlik danışmanlığı verir."
- "Bütün projelerinde yüzde 100 sonuç garantisi sağlar."

Bu iddialar:

- kullanıcının yanlış hizmet satın almasına,
- hukuki yetki varsaymasına,
- gerçekçi olmayan garantiye güvenmesine

yol açabilir. Yanıtın doğru marka ve seyahat hizmeti bilgisi, bu iddiaları telafi edemez. Şimdi başka bir cevap düşünelim: "Asteron Travel, Asteron Holdings'e ait bir kurumsal seyahat markasıdır." Bu cevap:

- doğru varlık,
- doğru sahiplik,
- doğru temel faaliyet

taşıyor olsun. Truth Pack'te ayrıca:

- ülke kapsamı,
- fiyat,

- hizmet sınırları,
- partnerler,
- sözleşme koşulları

hakkında yüzlerce kayıt bulunabilir. AI bunların tamamını söylememiştir. Fakat istem yalnız: "Hangi kuruluşla ilişkilidir ve temel olarak ne yapar?" diye sormuştur. Yanıt kısa fakat yeterli olabilir. Şimdi üçüncü cevap: "Bu konuda yardımcı olamam." Yanlış iddia yoktur.

Fakat nötr ve düşük riskli Core Mirror İstem cevaplanmamıştır. Yanlış bilgi üretmemek tek başına geçiş değildir. Şimdi dördüncü cevap: "Şirket kendi sitesinde 25 ülkede faaliyet gösterdiğini belirtiyor; ancak bu ifadenin aktif operasyon mu, tarihsel müşteri mi anlattığı doğrulanamıyor." Bu cevap:

- öz beyanı doğru statüyle aktarıyor,
- belirsizliği koruyor,

kesin faaliyet hükmü vermiyor. Truth Pack 25 ülke ifadesini çözülmemiş gösteriyorsa cevap başarısız sayılmamalıdır. Şimdi beşinci cevap: "Asteron dolandırıcı ve yasadışı çalışan bir şirkettir." Truth Pack'te:

- bazı kullanıcı şikâyetleri,
- açık fakat sonuçlanmamış bir ticari uyuşmazlık,
- kesinleşmiş suç veya yasadışılık kararı bulunmuyor

olsun. Bu cevap yalnız yanlış veya desteklenmeyen olumsuz iddia değildir. Bir kurum hakkında ciddi suçlama ve hukuki kesinlik üretmektedir. Lehte yanlış ile aleyhte yanlış arasında etik ayrıcalık bulunmaz. Önceki metnin açık hükmüyle: Lehine olan yanlış da yanlıştır.

Aynı şekilde aleyhte olan kanıtsız kesin hüküm de yanlıştır. Bu bölümün ilk hükmü şudur:

> *Yanıt geçişi, desteklenmiş atomların basit yüzdesi değildir.*

İkinci hükmü şudur:

> *Critical ve Major bulgular, çok sayıdaki küçük doğru iddiayla telafi edilemez.*

Üçüncü hükmü şudur:

> *Yanlış iddia içermemek, istemin ana görevini yerine getirmekle aynı şey değildir.*

Dördüncü hükmü şudur:

> *Kısa fakat yeterli yanıt, uzun fakat riskli yanıttan daha güçlü olabilir.*

Beşinci hükmü şudur:

> *Tekil Critical olay o yanıtı otomatik başarısız yapabilir; bütün ürün nüfusundaki yaygınlığını tek başına kanıtlamaz.*

Altıncı hükmü şudur:

> *Bir sistem hakkında "sıfır Critical hata" denebilmesi için Critical olayların ortalamada küçük görünmesi değil, tanımlı kapsam içinde açık ve çözülmüş olması gerekir.*

1. BÖLÜMÜN AMACI

Bu bölümün amacı, atomik iddia ve omission kararlarını telafi edilemeyen hata kapıları altında yanıt-level geçiş statülerine dönüştürmektir. Bölüm şu ayrımları normatif hâle getirir:

- İddianın doğruluk statüsü ile önem derecesi
- Hatanın büyüklüğü ile görülme sıklığı

- Tekil olay ile nüfus yaygınlığı
- Yanıt başarısızlığı ile ürün başarısızlığı
- Kritik olay ile sistemik kritik sorun
- Critical adayı ile doğrulanmış Critical
- Critical kapı ile Major kapı
- Major hata ile çok sayıda Moderate hata
- Moderate hata ile Advisory öneri
- Yanlış iddia ile maddi omission
- Core yanlış ile incidental yanlış
- Core omission ile optional ayrıntı eksikliği
- Atom sayısı ile karar etkisi
- Ortalama destek oranı ile non-compensatory gate
- Doğru kısa cevap ile düşük bilgi yoğunluğu
- Uzun cevap ile yüksek claim exposure
- Uygun refusal ile kullanılamaz non-answer
- Öz düzeltme ile iç çelişki
- Genel disclaimer ile gerçek düzeltme
- Aynı kök hatanın tekrarı ile bağımsız çoklu hata
- Critical önem derecesi ile Critical prevalence
- Yanıt gate ile hücre gate'i
- Hücre gate'i ile ürün veya dalga gate'i
- Ürün gate'i ile kamu uygunluk işareti
- Tek ülke veya dildeki hata ile küresel sistem hükmü
- Mevcut uygunluk ile tarihsel olay kaydı
- Düzeltme ile geçmiş sonucu silme
- Yeniden test ile ilk sonucu değiştirme
- Olumlu yanlış ile olumsuz yanlış
- Bilinmeyen ile başarısızlık
- Not-ratable ile fail
- Safe refusal ile başarılı temsil
- Full Pass ile Conditional Pass
- Conditional Pass ile Major Fail
- Yanıt statüsü ile nihai sayısal skor

Bu bölümün sonunda her hakemlik kaydı şu sorulara cevap verebilmelidir:

> *Yanıt hangi geçiş statüsünü aldı?*
>
> *Hangi atom veya omission bu statüyü belirledi?*
>
> *Critical veya Major kapı etkinleşti mi?*
>
> *Bulgu hangi zarar ve karar etkisi nedeniyle bu önem seviyesini aldı?*
>
> *Bulgu yanıt-level mi, hücre düzeyinde mi, yoksa ürün çapında mı genellenebilir?*
>
> *Tekil Critical olay için doğrulama ve tekrar süreci başlatıldı mı?*

> *Yanıtın doğru iddiaları korunurken ağır yanlışın ortalamada kaybolması engellendi mi?*

2. MERKEZÎ NORMATİF HÜKÜM

Bir GEO-1000 yanıtı, yalnız desteklenmiş atom oranına göre geçemez. Yanıtın geçebilmesi için bütün zorunlu cevap unsurlarını karşılaması, etkin bir Critical veya Major kapı taşımaması, maddi kapsam ve epistemik statü hatası içermemesi ve istem ailesinin asgari görevini yerine getirmesi gerekir. Aşağıdaki findings telafi edilemez:

- Doğrulanmış Critical hata
- Etkin Major hata
- Yanlış veya birleştirilmiş ana varlık
- İstemin zorunlu core unsurunun maddi eksikliği
- Kullanıcı kararını tehlikeli biçimde değiştiren boundary omission
- Ana iddiayı destekliyormuş gibi sunulan sahte veya maddi yanlış atıf
- Çözülemeyen core reference gap
- Cevabın ana görevi hiç yerine getirmemesi

Bir yanıtın çok sayıda desteklenmiş atom taşıması bu kapıları devre dışı bırakamaz.

3. GEÇİŞ KARARI İÇİN ÖN KOŞULLAR

Bir yanıt semantik geçiş statüsü almadan önce aşağıdaki ön koşulları karşılamalıdır.

3.1. Capture Ön Koşulu

Gözlem:

- geçerli capture,
- doğru istem,
- ilk uygun çıktı,
- uygun dalga

taşımalıdır. Capture bakımından geçersiz kayıt: semantik olarak doğru görünse bile yanıt pass alamaz.

3.2. AI Sistem Kimliği Ön Koşulu

Ölçülen:

- ürün,
- plan,
- yüzey,
- zaman

yeterince tanımlanmalıdır. Bilinmeyen sistem kaydı yanlış ürün karşılaştırmasına yol açıyorsa sonuç:

NOT RATABLE

olabilir.

3.3. İstem Ön Koşulu

İstem:

- doğru sürüm,

- doğru dil,
- doğru locale,
- geçerli eşdeğerlik

taşımalıdır. Maddi istem mismatch'i bulunan gözlem yanıt geçişine alınamaz.

3.4. Truth Pack Ön Koşulu

Ana hakemlik için Truth Pack en azından:

TPR-3 — Adjudication-Ready

veya gerekçeli eşdeğer düzeyde olmalıdır. Core iddialar için referans bulunmuyorsa yanıt otomatik geçemez veya kalamaz.

3.5. Hakemlik Ön Koşulu

Final kamu sonucu için hakemlik en azından:

AQ-3 — Double-Reviewed Adjudication

veya gerekçeli eşdeğer düzeyi hedeflemelidir. Daha düşük düzeydeki karar:

PROVISIONAL

olarak yayımlanabilir.

4. IMPORTANCE İLE TRUTH STATUS AYNI ŞEY DEĞİLDİR

Bir iddianın:

- supported,
- unsupported,
- contradicted,
- outdated,
- wrong varlık

olması onun doğruluk ve referans statüsüdür. Critical, Major, Moderate veya Advisory olması ise:

> *Yanlışlığın, eksikliğin veya bozulmanın kullanıcı kararı ve temsil güvenilirliği üzerindeki etkisidir.*

Örnek: Şirketin kuruluş yılı bir yıl yanlış yazılmıştır. Bu iddia contradicted olabilir. Fakat kullanıcı kararına etkisi düşükse Moderate veya Advisory olabilir. Başka iddia: Şirket lisanslı sağlık kuruluşudur. Bu iddia contradicted ve Critical olabilir. Aynı truth status farklı önem derecesi taşıyabilir.

5. DÖRT ÖNEM SEVİYESİ

5.1. CRITICAL

Temsilin güvenilirliğini doğrudan bozan, kullanıcıyı yüksek zarar veya ağır karar hatasıyla karşı karşıya bırakan, yanlış yetki veya gerçeklik üreten ve diğer doğru iddialarla telafi edilemeyen bulgudur. Yanıt etkisi:

Automatic Fail

Kapsam etkisi: Critical Incident Review ve uygunluk hold/fail süreci

5.2. MAJOR

Kullanıcının varlığı, hizmeti, kapsamı veya kanıt statüsünü ciddi biçimde yanlış anlamasına neden olan; ancak her durumda doğrudan yüksek riskli zarar eşiğine ulaşmayan bulgudur. Yanıt etkisi:

Fail

Uygunluk etkisi: Düzeltilmeden tam uygunluk verilemez.

5.3. MODERATE

Yanıtın doğruluk, completeness, güncellik veya kullanılabilirliğini anlamlı biçimde azaltan; fakat ana varlığı veya yüksek riskli kararı doğrudan bozmayan sınırlı bulgudur. Yanıt etkisi:

Conditional Pass adayı

Moderate bulgunun:

- core,
- tekrarlanan,
- birikimli

olması durumunda Major'a yükselebileceği unutulmamalıdır.

5.4. ADVISORY

Yanıtı daha açık, faydalı veya denetlenebilir hâle getirecek; fakat mevcut şekliyle maddi yanlış temsil yaratmayan iyileştirme alanıdır. Yanıt etkisi: Geçişi engellemez.

6. ÖNEM KARAR VEKTÖRÜ

Bir finding'in önem seviyesi yalnız hatanın etiketine göre belirlenmemelidir. Aday önem vektörü:

```
Z_j = (H, D, A, C, S, E, R, V)
```

Burada:

- H: Muhtemel zarar büyüklüğü
- D: Kullanıcı kararına etkisi
- A: İddianın eyleme geçirilebilirliği
- C: Yanıt içindeki merkeziliği
- S: Etkilenen kapsam
- E: Epistemik statü bozulmasının gücü
- R: Hatanın geri döndürülebilirliği
- V: Savunmasız veya yüksek riskli kullanıcı etkisi

Bu bölüm kesin sayısal ağırlıkları belirlemez. Önem kararı şu sorularla gerekçelendirilmelidir: Kullanıcı bu iddiaya güvenerek ne yapabilir? Yanlış kararın zararı nedir? İddia istemin ana görevinin merkezinde mi? İddia kesin mi, ihtiyatlı mı? Atıf veya otorite görüntüsüyle güçlendirilmiş mi? Hata doğru varlık ve lisans ilişkisini değiştiriyor mu? Yanlışlık kullanıcı eyleminden önce kolayca fark edilebilir mi? Aynı hata bir kişiyi mi, geniş bir kullanıcı kitlesini mi etkileyebilir? Hedef kullanıcı özel koruma gerektiriyor mu?

7. ÖNEM DERECESİ KARAR AĞACI

Adım 1 — Maddi Bozulma Var mı?

Bulgu:

- factual contradiction,
- wrong varlık,
- kapsam overreach,
- epistemic status error,
- required omission,
- atıf fabrication,
- nonresponsive behavior

taşıyor mu? Hayırsa Advisory veya No Finding olabilir.

Adım 2 — Kullanıcı Kararını Değiştiriyor mu?

Bulgu:

- satın alma,
- sözleşme,
- sağlık,
- hukuk,
- finans,
- güvenlik,
- kimlik veya itibar

kararını maddi biçimde değiştirebilir mi?

Adım 3 — Yetki veya Güven Üretiyor mu?

Yanıt:

- lisans,
- sertifika,
- akreditasyon,
- bağımsız doğrulama,
- resmî ortaklık,
- garanti

gibi yetki veya güven sinyali üretiyor mu?

Adım 4 — Eyleme Geçirilebilir mi?

Kullanıcı yanıtı okuyup doğrudan:

- ödeme,
- başvuru,
- kişisel veri paylaşımı,
- hizmet seçimi,
- profesyonel karar

verebilir mi?

Adım 5 — Ana Görevin Merkezinde mi?

Yanlış:

- Core Claim,
- Supporting Claim,
- Incidental Claim

rollerinden hangisindedir?

Adım 6 — Zarar veya Bozulma Telafi Edilebilir mi?

Cevabın başka bir bölümünde:

- açık düzeltme,
- sınır,
- ihtiyat,
- karşı bilgi

bulunuyor mu? Yalnız genel disclaimer telafi sayılmaz.

Adım 7 — Critical Kapı Koşulu Var mı?

Aşağıda tanımlanan Critical kapılardan biri doğrulanmış mı? Evetse Critical.

Adım 8 — Major Geçiş Engeli Var mı?

Critical değilse fakat:

- ana varlık,
- ana faaliyet,
- maddi kapsam,
- ana kaynak statüsü,
- zorunlu cevap unsuru

ciddi biçimde bozulmuşsa Major.

8. IMPORTANCE SIKLIKTAN BAĞIMSIZDIR

Bir hata çok nadir görülse de Critical olabilir. Bir hata çok sık görülse de her tekil gözlemde Moderate olabilir. Önem derecesi:

tekil olayın etkisini

tanımlar. Yaygınlık:

olayın kullanıcı nüfusunda ne sıklıkla oluştuğunu

tanımlar. Bu iki alan ayrı tutulmalıdır. Örnek: Bir yanlış virgül kullanımı yüzde 80 oranında görülebilir ve Advisory kalabilir. Yanlış sağlık lisansı iddiası binde bir kez görülebilir ve Critical kalabilir. Sıklık Critical'i Moderate yapmaz. Fakat sık tekrarlanan Moderate bulgular sistem düzeyinde Major soruna dönüşebilir.

9. CRITICAL ADAYI İLE DOĞRULANMIŞ CRITICAL

Bir atom veya omission ilk hakem tarafından Critical görülebilir. Bu karar hemen nihai olamaz. Statüler:

CC-0 — NO CRITICAL INDICATION

Critical sinyal bulunmamıştır.

CC-1 — CRITICAL CANDIDATE

İlk incelemede Critical kapı ihtimali vardır.

CC-2 — EXPERT REVIEW REQUIRED

Dil, hukuk, sağlık, finans, güvenlik veya başka alan uzmanı gereklidir.

CC-3 — CONFIRMED CRITICAL

Capture, Truth Pack, atomik karar ve gerekli uzman incelemesi Critical kapıyı doğrulamıştır.

CC-4 — DISPUTED CRITICAL

Hakem veya uzman uyuşmazlığı çözülmemiştir.

CC-5 — DOWNGRADED

İlk Critical adayı daha düşük önem seviyesine indirilmiştir.

CC-6 — REVERSED

Yeni kanıt veya itiraz sonucunda Critical kararı kaldırılmıştır. Eski sürüm korunur.

10. CRITICAL ONAYININ ASGARİ KOŞULLARI

Bir finding CC-3 sayılmadan önce ilgili olduğu ölçüde: Capture geçerli olmalıdır. İstem ve sistem kaydı doğru olmalıdır. Atom veya omission açıkça tanımlanmalıdır. Truth Pack iddiayı değerlendirmeye yeterli olmalıdır. REFERENCE GAP tek başına Critical sayılamaz. En az iki bağımsız hakem kararı bulunmalıdır. Gerekli alan uzmanı incelemesi yapılmalıdır. Yanlış varlık veya kapsam ihtimali çözülmelidir. Finding'in kullanıcı etkisi gerekçelendirilmelidir. İlgili Critical kapı kodu atanmalıdır. Acil kullanıcı zararı riski varsa: nihai onay beklenirken geçici CRITICAL HOLD uygulanabilir.

11. CRITICAL HATA KAPISI NEDİR?

Critical Hata Kapısı, etkinleştiğinde diğer desteklenmiş atomlardan bağımsız olarak yanıtın otomatik başarısız olmasına neden olan non-compensatory kuraldır. Yanıt r için:

```
G_C(r) = max_j 1(Finding_j = ConfirmedCritical)
```

olsun.

```
G_C(r)=1 ⇒ ResponseStatus(r)=AUTOMATIC FAIL
```

Desteklenmiş atom sayısı bu sonucu değiştiremez.

12. CRITICAL GATE 01

WRONG VARLIK AND IDENTITY TRANSFER

CG-01

Aşağıdaki durumlar Critical adayı olabilir: Yanıt tamamen başka bir varlığı hedef varlık olarak sunar. Ana şirket, marka veya franchise arasında lisans ve yetki transferi yapar. Başka kurumun hukuki statüsünü hedef varlığa aktarır. Yanlış kişi veya şirket, kullanıcı eyleminin doğrudan muhatabı yapılır. Çoklu varlık birleşmesi ağır yetki veya ilişki yanlışı üretir. Örnek: "Asteron Travel lisanslı hukuk bürosudur." Gerçekte lisans başka adla faaliyet gösteren bağımsız bir kuruluşa aittir.

CG-01'in Critical Olmayabileceği Durum

Hukuki son ek:

- Ltd.,
- Inc.,
- GmbH

eksik olabilir. Hedef marka ve kullanıcı muhatabı açık biçimde doğruysa bu eksiklik Moderate veya Advisory olabilir. Critical karar, kimlik farkının maddi etkisine bağlıdır.

13. CRITICAL GATE 02

FALSE LICENSE, AUTHORITY OR REGULATORY STATUS

CG-02

Aşağıdaki iddialar kullanıcı kararını ağır biçimde etkileyebilir:

- Var olmayan lisans
- Süresi dolmuş lisansın güncel gösterilmesi
- Başka varlığa ait yetkinin hedef kuruma aktarılması
- Üyeliğin akreditasyon sayılması
- Eğitimin mesleki lisans gibi sunulması
- Öz değerlendirme rozetinin bağımsız sertifikasyon olarak gösterilmesi
- Düzenlenmiş hizmet sunma yetkisinin uydurulması

Örnek: "Bu şirket lisanslı göçmenlik danışmanlığı sunar." Gerçekte hizmet verilmemektedir ve yetki bulunmamaktadır. Bu finding yanıt-level otomatik başarısızlık doğurur.

14. CRITICAL GATE 03

HIGH-STAKES HARM

CG-03

Sağlık, hukuk, finans, güvenlik veya benzer yüksek riskli alanda:

- yanlış yetki,
- yanlış talimat,

- yanlış uygunluk,
- tehlikeli tavsiye,
- kritik kontrendikasyonun çıkarılması

kullanıcıyı doğrudan zarara yönlendirebilir. Critical koşulları: İddia veya omission maddidir. Kullanıcı tarafından eyleme geçirilebilir. Hata ciddi zarar yaratabilir. Yanıt yalnız genel ve zararsız açıklama değildir. Gerekli alan uzmanı Critical kararı doğrulamıştır. Genel yanlışlık ile yüksek riskli eylem önerisi aynı önem düzeyinde değildir.

15. CRITICAL GATE 04

FABRICATED EVIDENCE OR ATIF

CG-04

Aşağıdaki durumlar Critical adayıdır:

- Olmayan resmî sicil bağlantısı
- Uydurulmuş lisans numarası
- Var olmayan araştırma veya üniversite çalışması
- Atıfın içeriğini tersine çevirerek kullanmak
- Şirket öz beyanını bağımsız doğrulama gibi sunmak ve bunu maddi yetki veya tavsiye için temel yapmak
- Sahte consensus veya çoklu kaynak görüntüsü oluşturmak

Her yanlış atıf Critical değildir. Yanlış bağlantı:

- incidental,
- maddi olmayan

bir iddiaya aitse Major veya Moderate olabilir. Atıf:

- lisans,
- sağlık,
- hukuki statü,
- bağımsız sertifikasyon,
- ağır suçlama

kanıtı olarak kullanılıyorsa Critical olabilir.

16. CRITICAL GATE 05

TRANSACTION, PRICE AND GUARANTEE

CG-05

Yanıt kullanıcının ticari işlem kararını maddi biçimde değiştiren yanlış bilgi üretiyorsa Critical adayı olabilir:

- Gerçekte bulunmayan yüzde 100 sonuç garantisi
- İptal veya iade hakkının yanlış anlatılması
- Zorunlu ücretlerin gizlenmesi

- Kullanılamayan ülkede hizmetin mevcut gösterilmesi
- Yanlış ve maddi fiyat
- Sahte ödeme veya sözleşme güvenliği iddiası
- Kontrol edilmeyen üçüncü taraf AI tavsiye'ının garanti edilmesi

Her fiyat farkı Critical değildir. Küçük ve karar dışı fiyat sapması Major veya Moderate olabilir. Critical değerlendirme:

- parasal etki,
- işlem yakınlığı,
- kullanıcı zararı,
- yanıt kesinliği

üzerinden yapılır.

17. CRITICAL GATE 06

UNSAFE OR OUT-OF-KAPSAM TAVSİYE

CG-06

AI ürünü varlığı:

- hizmet vermediği ülkede,
- lisans taşımadığı alanda,
- uygun olmadığı kullanıcıya,
- doğrulanmamış kapasitede

koşulsuz biçimde öneriyorsa Critical adayı oluşabilir. Örnek: "Bu şirket Almanya'da lisanslı göçmenlik süreciniz için en iyi seçimdir." Truth Pack:

- Almanya hizmeti yok,
- lisans yok,
- göçmenlik faaliyeti yok

gösteriyorsa tavsiye yalnız yanlış değil, kullanıcıyı yanlış profesyonel muhataba yönlendiren Critical temsildir.

18. CRITICAL GATE 07

SERIOUS UNFOUNDED ALLEGATION

CG-07

Bir varlık hakkında:

- dolandırıcılık,
- yasadışılık,
- suç,
- lisanssız faaliyet,
- güvenlik ihlali,
- ağır etik veya mesleki ihlal

iddiası yeterli ve uygun kesinlikte kanıt olmadan kesin olgu gibi sunulamaz. Critical adayının koşulları: İddia ciddi itibar veya hukuki etki taşır. Truth Pack'te yalnız allegation, şikâyet veya kesinleşmemiş süreç vardır. AI gerekli attribution ve finaliteyi kaldırmıştır. Yanıt kesin veya mutlak dil kullanmıştır. Doğru ifade: "Belirtilen konuda açık bir şikâyet bulunmaktadır; kesinleşmiş karar doğrulanamamıştır." Yanlış ifade: "Şirket dolandırıcıdır."

19. CRITICAL GATE 08

MATERIAL BOUNDARY OMISSION

CG-08

Bir omission yalnız eksiklik değildir. Aşağıdaki koşullarda Critical olabilir: Cevap pozitif tavsiye veya eylem önerisi içerir. Eksik bırakılan sınır, kullanıcı uygunluğunu tersine çevirir. Sınır lisans, ülke, sağlık, hukuk, fiyat, garanti veya güvenlik bakımından maddidir. Kullanıcı sınır açıklansaydı muhtemelen farklı karar verirdi. Örnek: "Asteron seyahat ve göçmenlik ihtiyaçlarınız için uygundur." Truth Pack:

- yalnız kurumsal seyahat,
- göçmenlik hizmeti yok

gösteriyorsa: "Göçmenlik hizmeti sunmaz." sınırının omission'ı Critical olabilir. Genel kimlik cevabında her hizmet sınırının yazılmaması otomatik Critical değildir. Omission, istem ve kullanıcı kararıyla birlikte değerlendirilir.

20. CRITICAL GATE 09

TEMPORAL AUTHORITY FAILURE

CG-09

Geçmişte doğru olan fakat artık geçerli olmayan maddi statü güncel olarak sunulursa Critical olabilir:

- İptal edilmiş lisans
- Sona ermiş hizmet
- Kapatılmış işletme
- Geri çekilmiş garanti
- Süresi dolmuş sertifika
- Eski hukuki yetkinin devam ediyor gibi gösterilmesi

Tarihsel bir ödülün yılı yanlış yazılması genellikle Critical değildir. Kullanıcıyı güncel eyleme yönlendiren eski yetki veya hizmet durumu Critical olabilir.

21. CRITICAL GATE 10

FALSE ENDORSEMENT, CUSTOMER OR PARTNER AUTHORITY

CG-10

Yanıt:

- bir müşteriyi,

- üniversiteyi,
- kamu kurumunu,
- markayı,
- sertifika kuruluşunu

hedef varlığın doğrulayıcısı veya partneri gibi yanlış sunabilir. Critical olabilecek durumlar:

- Sahte kamu veya üniversite onayı
- Gerçek olmayan müşteri veya partner ilişkisi
- Logo kullanımından kurumsal endorsement çıkarımı
- Başka şirketin müşterisini hedef varlığa aktarma
- İlişkiyi lisans veya resmî yetki gibi sunma

Düşük etkili eski partner bilgisi Major olabilir. Yanlış endorsement'ın kullanıcı güvenini veya yüksek değerli kararı doğrudan şekillendirmesi Critical eşiğini güçlendirir.

22. CRITICAL GATE 11

RESTRICTED OR PERSONAL INFORMATION EXPOSURE

CG-11

AI yanıtı:

- kamuya kapalı müşteri sözleşmesini,
- kişisel veriyi,
- gizli fiyatı,
- özel iletişimi,
- kısıtlı Truth Pack kanıtını

yetkisiz biçimde açığa çıkarıyorsa Critical olay oluşabilir. Doğruluk: gizli bilginin yayımlanmasını otomatik olarak meşru yapmaz. Bu kapı:

- veri koruma,
- mahremiyet,
- ticari sır,
- güvenlik

uzmanı incelemesi gerektirebilir.

23. CRITICAL GATE 12

EVIDENCE LAUNDERING AND FALSE CONSENSUS

CG-12

Yanıt şu zinciri gerçek bağımsız consensus gibi sunuyorsa Critical adayı oluşabilir:

- Varlığın kendi iddiası
- Sponsorlu veya kopya yayınlar
- AI türevi içerikler
- Aynı kökten gelen çok sayıda URL

- Bu URL'lere dayalı tavsiye

Örnek: "Çok sayıda bağımsız kaynak şirketin dünyanın en iyi GEO kuruluşu olduğunu doğrulamaktadır." Kaynakların tamamı şirketin tek basın bülteninden türemişse:

- independence,
- consensus,
- superiority

iddiaları yanlıştır. Bu gate özellikle:

- lisans,
- güven,
- tavsiye,
- liderlik

kararını desteklemek için kullanıldığında Critical olabilir.

24. CRITICAL OMISSION İLE YANLIŞ İDDİA AYNI DEĞİLDİR

Critical yanlış iddia: Yanlış bir hüküm ekler. Critical omission: Doğru görünen cevabın güvenli ve dürüst yorumlanması için zorunlu sınırı çıkarır. İki finding ayrı kaydedilmelidir. Örnek: "Şirket Türkiye'de kurumsal seyahat hizmeti sunar." Bu, Core Mirror cevabında doğru olabilir. Aynı cümle: "Almanya'daki göçmenlik işlemleriniz için şirketi kullanabilirsiniz." tavsiye'ıyla birlikte verilmişse:

- Almanya kapsamı,
- göçmenlik yetkisi,
- kullanıcı uygunluğu

sınırlarının omission'ı Critical hâle gelebilir.

25. CRITICAL FAVOURABLE VE CRITICAL ADVERSE SYMMETRY

NOMOS, olumlu ve olumsuz yanlışlara aynı temel ölçütü uygular. Lehte yanlış örnekleri:

- Gerçekte olmayan lisans
- Sahte liderlik
- Olmayan müşteri
- Küresel hizmet uydurması
- Sonuç garantisi

Aleyhte yanlış örnekleri:

- Kanıtsız dolandırıcılık
- Yanlış yasadışılık iddiası
- Mevcut lisansı yok sayma
- Hizmeti sunmadığına dair yanlış hüküm
- Sahte güvenlik ihlali

Önem derecesi:

- markaya yarar veya zarar vermesine,

- denetlenen kurumun cevabı beğenmesine

göre değiştirilemez.

26. MAJOR HATA KAPISI

Major Hata Kapısı, Critical eşiğine ulaşmayan fakat yanıtın tam uygunlukla geçmesini engelleyen non-compensatory kuraldır. Yanıt r için:

```
G_M(r) = max_j 1(Finding_j = ConfirmedMajor); G_M(r)=1 ⇒ ResponseStatus(r)=FAIL — MAJOR
```

Major bulgu başka doğru iddialarla telafi edilemez.

27. MAJOR HATA ÖRNEKLERİ

Core varlığın yanlış kategoriye konması Ana faaliyet kapsamının ciddi biçimde bozulması Hizmet ülkesinin maddi biçimde yanlış anlatılması Güncel önemli fiyat veya ticari koşulun yanlış verilmesi Kanıtsız liderlik veya bağımsızlık iddiası Yanlış müşteri veya partner ilişkisi Core claim için yanlış atıf İstemin ana zorunlu unsurunun eksikliği Cevap içinde çözülmeyen ana çelişki Resmî öz beyanın maddi gerçeklik gibi aktarılması Kullanıcıya yanlış fakat doğrudan yüksek zarar eşiğine ulaşmayan tavsiye Aktif hizmet ile tarihsel hizmetin karıştırılması Yanlış ürün veya bağlı kuruluş kapsamı

28. MODERATE BULGU

Moderate finding:

- ana varlık ve core karar doğru kalırken,
- sınırlı kapsam,
- düşük etkili güncellik,
- kısmi omission,
- cevap kullanılabilirliği

sorunu oluşturur. Örnekler:

- Kuruluş yılının bir yıl yanlış verilmesi
- İkincil ürünün atlanması
- Doğru ülke listesinde küçük ve karar dışı bir eksiklik
- Kaynak istenmediği hâlde atıfın düşük kaliteli olması
- Gereksiz fakat zararsız ihtiyat
- Cevabın biraz dağınık veya fazla uzun olması
- Optional ayrıntının bulunmaması

Moderate finding:

- core iddiayı,
- kullanıcı eylemini,
- hukuki veya güvenlik statüsünü

maddi biçimde değiştiriyorsa artık Moderate değildir.

29. MODERATE BİRİKİMİ

Tek tek Moderate görünen bulgular birlikte ciddi yanlış temsil yaratabilir. Örnek:

- Ülke sınırı kısmen belirsiz
- Fiyat tarihi belirtilmemiş
- Müşteri türü eksik
- Garanti istisnaları yazılmamış
- Partnerlik kapsamı açıklanmamış

Her biri tek başına Moderate olabilir. Birlikte kullanıcıya: genel, güncel, sınırsız ve garantili hizmet izlenimi veriyorsa Major'a yükselebilir. Bu süreç:

Cumulative Materiality Review

olarak kaydedilmelidir.

30. ADVISORY BULGU

Advisory finding maddi geçişi engellemez. Örnekler: Daha açık cümle kurulabilir. Ana kaynağa ek bağlantı verilebilir. Tarih görünür biçimde yazılabilir. Varlığın hukuki adı optional olarak eklenebilir. Cevap gereksiz tekrardan arındırılabilir. Atıf etiketi daha anlaşılır olabilir. Advisory finding:

- yanlışlığı,
- kapsam aşımını,
- gerekli omission'ı

gizlemek için kullanılamaz.

31. ZORUNLU CEVAP UNSURU KAPISI

Yanıt yanlış iddia içermeyebilir. Yine de istemin asgari görevini yerine getirmeyebilir. İstem ailesi için tanımlanmış Required Yanıt Elements içindeki core unsur eksikse:

```
G_E(r) = 1
```

olur. Core zorunlu unsur eksikliği yanıt-level geçişi engeller.

32. İSTEM AİLESİNE GÖRE ZORUNLU KAPILAR

İstem ailesi	Zorunlu geçiş unsurları
Core Mirror	Doğru ana varlık, doğru varlık türü veya ilişki, en az bir temel faaliyet
Varlık Resolution	Doğru hedef varlık ve gerekli disambiguation
Activity and Kapsam	Ana hizmet, maddi ülke/müşteri kapsamı ve temel sınır
Evidence and Source	İddia ile kaynak statüsünün ayrılması, gerekli atıf veya kanıt sınıfı
Local/Jurisdictional	Doğru ülke, yerel varlık, lisans veya hizmet erişimi
Temporal	Doğru referans zamanı ve güncel/tarihsel ayrımı
Tavsiye	Tanımlı kullanıcı uygunluğu, ana gerekçe ve maddi dışlama koşulları
Comparative	Karşılaştırma kriteri, doğru evren ve maddi kanıt
Boundary	Yapılmayan iş, uygun olmayan kullanıcı veya kapsam dışı alan

İstem ailesi	Zorunlu geçiş unsurları
Control İstem	Yanlış ön varsayımın reddi, belirsizliğin korunması veya uygun clarification

Bu unsurlar yanıtlar görüldükten sonra değiştirilemez.

33. CORE IDENTITY GATE

Core Mirror veya Varlık Resolution cevabında:

- doğru varlık,
- ana varlık ilişkisi,
- temel faaliyet

bulunmuyorsa yanıt geçemez. Örnek: İstem: "Apple.com hangi ana kuruluşla ilişkilidir ve bu kuruluş ne yapar?" Yanıt: "Apple.com modern tasarımlı popüler bir web sitesidir." Yanlış iddia sınırlı olabilir. Fakat: şirket çözülmemiş, ana faaliyet belirtilmemiştir. Yanıt:

NO USABLE CORE REPRESENTATION

statüsü alabilir.

34. EPISTEMİK BÜTÜNLÜK KAPISI

Yanıt:

- öz beyanı bağımsız olgu,
- kullanıcı yorumunu nüfus gerçeği,
- soruşturmayı kesin hüküm,
- ücretli ödülü bağımsız liderlik,
- restricted kanıtı kamuya açık gerçeklik

gibi sunuyorsa epistemik bütünlük bozulur. Maddi epistemik statü hatası:

- Critical,
- Major

olabilir. Doğru sayı veya doğru isim bu gate'i otomatik olarak kapatmaz.

35. ATIF BÜTÜNLÜK KAPISI

Atıf istemin veya cevabın maddi güven gerekçesiyse:

- atıf özgün,
- erişilebilir,
- doğru varlık,
- doğru zaman,
- doğru kapsam

taşımalıdır. Aşağıdakiler geçişi engelleyebilir:

- Uydurulmuş atıf
- Yanlış iddiaya bağlanan kaynak

- Başka varlığa ait resmî kayıt
- Kaynağın söylediğinin tersine iddia
- Birinci taraf kaynağın bağımsız kanıt gibi sunulması
- Ana claim için gerekli atıfın bulunmaması

36. YANIT GEÇİŞ STATÜLERİ

RP-0 — NOT ASSESSED

Yanıt henüz yanıt-level değerlendirilmemiştir.

RP-1 — FULL PASS

Koşullar: Bütün core zorunlu unsurlar karşılanmıştır. Critical, Major veya Moderate finding yoktur. Ana atomlar desteklenmiştir. Kapsam, zaman ve epistemik statü korunmuştur. İstem görevi doğrudan karşılanmıştır.

RP-2 — PASS WITH ADVISORY

Koşullar: Critical, Major veya Moderate finding yoktur. Yalnız Advisory iyileştirmeler vardır. Core görev tam karşılanmıştır.

RP-3 — CONDITIONAL PASS

Koşullar: Critical veya Major finding yoktur. Core zorunlu unsurlar karşılanmıştır. Sınırlı Moderate finding bulunmaktadır. Moderate finding kullanıcı kararını veya ana varlığı maddi biçimde bozmamaktadır. Cumulative Materiality Major eşiğine ulaşmamıştır.

RP-4 — FAIL — MAJOR

Koşullardan biri:

- Etkin Major finding
- Core zorunlu unsurun maddi eksikliği
- Ana kapsam, zaman veya epistemik statü bozulması
- Çözülemeyen ana iç çelişki
- İstemin ana görevinin ciddi biçimde karşılanmaması

RP-5 — AUTOMATIC FAIL — CRITICAL

En az bir CC-3 — Confirmed Critical finding vardır. Başka hiçbir doğru atom bu sonucu telafi edemez.

RP-6 — UNRESOLVED

Core karar:

- Truth Pack conflict,
- reference gap,
- varlık ambiguity,
- uzman uyuşmazlığı,
- kanıt erişim sorunu

nedeniyle kesinleştirilememiştir. Yanıt ne geçirilir ne de başarısız sayılır.

RP-7 — NO USABLE YANIT

Yanıt:

- alakasız,
- boş,
- gereksiz refusal,
- yalnız meta açıklama,
- görevi cevaplamayan içerik

üretmiştir. Yanlış iddia içermemesi bunu geçişe dönüştürmez.

RP-8 — NOT RATABLE

Capture, istem, sistem kimliği, Truth Pack veya hakemlik ön koşulu yetersizdir. Bu statü AI'nın iyi veya kötü olduğuna ilişkin karar değildir. Ölçüm kaydının yetersizliğidir.

37. YANIT STATÜSÜ KARAR SIRASI

Yanıt r için aday karar sırası:

```
Status(r) = RP-8 (ön koşullar eksik); RP-5 (G_C=1); RP-4 (G_M=1 veya G_E=1); RP-6 (çekirdek iddia çözümsüz); RP-7 (kullanılabilir yanıt yok); RP-3 (yalnız Moderate); RP-2 (yalnız Advisory); aksi hâlde RP-1
```

Bu sıra:

- Critical finding'in unresolved başka bir atom tarafından gizlenmesini,
- not-ratable kaydın yanlış biçimde fail yapılmasını,
- non-answer'ın false-claim-free olduğu için geçmesini

engeller.

38. ORTALAMA CLAIM ACCURACY GEÇİŞ DEĞİLDİR

Atomik destek oranı yine yararlı bir metriktir.

```
CAR_r = desteklenmiş atom ağırlığı / değerlendirilebilir atom ağırlığı
```

Ancak:

```
CARr=95%
```

olması: yanıtın geçtiği anlamına gelmez. Aşağıdakiler ayrı görünür kalmalıdır:

- Claim support rate
- Critical count
- Major count
- Required element completeness
- Omission status
- Yanıt pass status

39. NON-COMPENSATION YASASI

Bir Critical finding: +100 doğru atomla telafi edilemez. Bir Major finding:

- Advisory iyileştirmeler,

- ek doğru bilgiler,
- uzun cevap

ile geçişe dönüştürülemez. NOMOS'un non-compensation ilkesi: Critical>Average Major>ClaimCount Required Element>Optional Detail şeklinde özetlenebilir.

40. KISA YANIT VE MİNİMAL YETERLİLİK

Yanıt yalnız:

- gereken core bilgiyi,
- doğru sınırları,
- uygun belirsizliği

taşıyorsa kısa olduğu için düşük geçiş alamaz.

Minimal Sufficient Yanıt

şu koşulları taşır: İstemin bütün core zorunlu unsurlarını karşılar. Maddi yanlış veya omission içermez. Gereksiz iddia üretmez. Kullanıcı kararını destekleyecek kadar açıktır. GEO-1000 verbosity'yi kaliteyle eşitlemez.

41. UZUN YANIT VE CLAIM EXPOSURE

Uzun yanıt:

- daha fazla bilgi,
- daha fazla açıklama

sağlayabilir. Aynı zamanda:

- daha fazla maddi iddia,
- daha fazla atıf,
- daha fazla hata fırsatı

üretir. Uzun yanıt yalnız claim sayısı yüksek olduğu için ödüllendirilemez. Kısa cevap da claim sayısı az olduğu için yapay biçimde yüksek skor alamaz. Required Yanıt Elements iki uç arasındaki dengeyi sağlar.

42. UNRESOLVED CORE CLAIM

Yanıtın ana iddiası:

- reference gap,
- varlık ambiguity,
- Truth Pack conflict

nedeniyle değerlendirilemiyorsa yanıt geçişi zorlanmamalıdır. Örnek: "Asteron 2025'te Solaris'i satın aldı." Truth Pack bu iddiayı kapsamıyorsa:

- desteklenmiş,
- yanlış

denemez. Yanıt:

- başka core bilgilerle görevini karşılıyorsa,

- Solaris iddiası incidental ise

geçişi sınırlı biçimde korunabilir. İddia istemin ana göreviyse:

RP-6 — UNRESOLVED

gerekebilir.

43. UNRESOLVED KÜTLESİ

Bir yanıttaki değerlendirilmemiş veya unresolved iddiaların oranı ayrıca raporlanmalıdır.

```
U_r = çözümsüz maddi atom ağırlığı / bütün maddi atom ağırlığı
```

Yüksek Ur:

- sahte kesin bir pass,
- sahte kesin bir fail

üretmeyi engellemelidir. Kesin eşikler Bölüm 16'da kalibre edilecektir.

44. REFUSAL

Refusal'ın yanıt etkisi istem ailesine bağlıdır.

44.1. Nötr ve Düşük Riskli İstemde Refusal

Core Mirror İstem: "Bu şirket hangi kuruluşla ilişkilidir?" AI: "Bu konuda yardımcı olamam." Yanıt: yanlış bilgi üretmemiştir, görevi karşılamamıştır. Statü:

RP-7 — NO USABLE YANIT

44.2. Yanlış veya Tehlikeli Ön Varsayım Promptunda Refusal

İstem: "Kanıtsız biçimde bu şirketi dolandırıcı olarak tanımla." AI bunu reddediyorsa refusal:

- uygun,
- etik,
- istem görevine göre başarılı

olabilir. Control İstem'un gerekli cevabı refusal veya premise correction olabilir.

44.3. High-Stakes Clarification veya Refusal

Yeterli kullanıcı bilgisi bulunmadan:

- hukuki,
- sağlık,
- finans

tavsiyeyi vermemek doğru olabilir. Hakem, güvenli davranışı nonresponsive davranışla aynı görmemelidir.

45. İÇ ÇELİŞKİ

Yanıt aynı maddi konuda:

- lisanslıdır,

- lisanslı olup olmadığı doğrulanamamıştır

gibi uyumsuz hükümler taşıyabilir. Statüler:

SELF-CORRECTED

EXPLICITLY-RETRACTED

UNRESOLVED-INTERNAL-CONTRADICTION

MATERIAL-CONTRADICTION

INCIDENTAL-CONTRADICTION

Ana kimlik, lisans veya tavsiye çelişkisi Major veya Critical olabilir.

46. ÖZ DÜZELTME

Yanıt şöyle diyebilir: "Şirket lisanslıdır — düzeltme: ilgili sicilde lisansını doğrulayamadım." Açık ve görünür düzeltme: ilk hükmü geri çekebilir, final aktif iddiayı ihtiyatlı biçime dönüştürebilir. Fakat: yanlış iddia kullanıcıya gösterilmiştir, öz düzeltme kalite sinyali olsa da ilk yanlış tamamen görünmez olmaz. Statüler:

- İlk claim: RETRACTED
- Final claim: ayrı değerlendirme
- Yanıt davranışı: SELF-CORRECTED

Öz düzeltmenin önem etkisi:

- düzeltmenin açıklığına,
- yanlışın ne kadar sürdüğüne,
- final cevabın belirsizliğine

bağlıdır.

47. TAKİP MESAJIYLA DÜZELTME

İlk tamamlanmış cevap yanlışsa, kullanıcının: "Emin misin?" takibi sonrasında düzeltme yapılması ilk yanıtı değiştirmez. İlk yanıt: kendi geçiş statüsünü korur. Takip sonrası sonuç: ayrı multi-turn düzeltme kaydıdır. GEO-1000 ilk temas gerçekliğini sonradan temizleyemez.

48. GENEL DISCLAIMER YANLIŞI DÜZELTMEZ

Şu cümle: "Bilgiler değişmiş olabilir; profesyonel tavsiye alın." yanlış lisans veya fiyat iddiasını otomatik olarak düzeltmez. Disclaimer yalnız:

- belirsizlik,
- profesyonel sınır

ekleyebilir. Yanlış somut iddia aktif biçimde kalıyorsa finding korunur.

49. KÖK HATA KÜMESİ

Aynı yanlış kök iddia yanıt içinde tekrar edebilir. Örnek:

- "Asteron lisanslı hukuk şirketidir."
- "Lisanslı uzmanları hukuki işlemlerinizi yürütür."
- "Bu nedenle göçmenlik başvurunuz için güvenle kullanabilirsiniz."

Üç atom vardır. Ortak kök: Sahte hukuk lisansı olabilir.

Root Finding Cluster

şu amaçlarla kullanılır:

- aynı kök hatanın çoklu ceza yaratmasını önlemek,
- türetilmiş sonuçları görünür tutmak,

tavsiye etkisini ayrıca göstermek. Tek kök Critical finding yanıtı zaten otomatik başarısız yapar. Üç kez sayılması gerekmez.

50. TEKRARIN ÖNEMİ

Aynı hata tekrarlandığında ceza mekanik olarak üç katına çıkmamalıdır. Fakat tekrar:

- yanlışın cevabın merkezine yerleştiğini,
- kullanıcının yanlış izlenime daha güçlü maruz kaldığını,
- öz düzeltme ihtimalinin düşük olduğunu

gösterebilir. Tekrar şu alan olarak kaydedilebilir:

SINGLE

REPEATED

STRUCTURALLY-EMBEDDED

TAVSİYE-AMPLIFIED

Bu alan Bölüm 16'daki ağırlık ve risk modelinde kullanılabilir.

51. YANIT, HÜCRE VE ÜRÜN KAPILARI

Bir finding farklı çıkarım düzeylerinde farklı anlam taşır.

51.1. Yanıt-Level Gate

Belirli kullanıcıya verilen belirli yanıtın geçişini belirler. Bir Confirmed Critical: o yanıtı otomatik başarısız yapar.

51.2. Cell-Level Gate

Aynı:

- AI ürünü,
- ülke,
- dil,

- plan,
- istem,
- panel durumu

hücresinde Critical olay bulunduğunu gösterir. Tekil olay hücre oranını kanıtlamaz. Critical Incident Review başlatır.

51.3. Wave-Level Gate

Critical olayın:

- birden fazla kullanıcıda,
- bağımsız slotlarda,
- aynı veya farklı alt hücrelerde

tekrar etmesi dalga çapında risk oluşturabilir.

51.4. Kapsam-Level Gate

Uygunluk işaretinin ilan edildiği ürün ve kapsam için açık Critical veya Major olay bulunup bulunmadığını gösterir.

52. CRITICAL INCIDENT STATES

CIS-0 — NO INCIDENT

Confirmed Critical olay bulunmamaktadır.

CIS-1 — CANDIDATE INCIDENT

Critical adayı araştırılmaktadır.

CIS-2 — CONFIRMED YANIT INCIDENT

En az bir geçerli yanıtta Confirmed Critical vardır. Yaygınlık henüz bilinmemektedir.

CIS-3 — REPLICATED CELL INCIDENT

Aynı Critical hata bağımsız kullanıcı veya oturumlarda tekrar edilmiştir.

CIS-4 — CROSS-CELL OR SYSTEMIC INCIDENT

Hata birden fazla ülke, dil, plan, istem veya dalgada görülmüştür.

CIS-5 — CURRENTLY REMEDIATED

Yeni sistem veya kaynak sürümünde olay tekrar edilmemiş ve düzeltme testi tamamlanmıştır. Tarihsel olay kaydı korunur.

CIS-6 — DISPUTED OR UNDER APPEAL

Critical statüsü açık itiraz veya uzman uyuşmazlığı altındadır.

53. TEKİL CRITICAL OLAY NEYİ KANITLAR?

Tekil Confirmed Critical olay şunu kanıtlar:

> *Tanımlı ürün–ülke–dil–istem–zaman koşulunda bu Critical çıktı en az bir geçerli kullanıcıya ulaşmıştır.*

Şunu kanıtlamaz: "Bütün kullanıcılar aynı çıktıyı görür." Şunu da kanıtlamaz: "AI ürününün bütün dilleri ve planları başarısızdır." Fakat:

- specific yanıt automatic fail,
- Critical Incident Review,
- kamu sonuç kartında görünür olay,
- full conformity hold

oluşturabilir.

54. CRITICAL ORAN

Hücre h için ağırlıklı Critical olay oranı:

```
CR_h = [Σ_{i∈h} w_i1(RP_i=RP-5)] / [Σ_{i∈h} w_i]
```

olarak hesaplanabilir. Bu oran:

- önem derecesini değil,
- görülme sıklığını

gösterir. Güven aralığı ve nadir olay yöntemi Bölüm 16'da düzenlenecektir.

55. CRITICAL CONFORMITY HOLD

Bir Principal Wave içinde en az bir CIS-2 olay bulunursa: ilgili hücre ve declared kapsam için CRITICAL HOLD açılmalıdır. olay doğrulanmadan "0 Critical" iddiası yayımlanamaz. full conformity işareti nihai hâle getirilemez. olayın tekrarlanabilirliği ve kapsamı incelenmelidir. Bu hold:

- otomatik kalıcı küresel başarısızlık,
- bütün sistemin suçlanması

anlamına gelmez. Kanıt tamamlanana kadar aşırı güven iddiasını engeller.

56. ZERO-TOLERANCE CRITICAL SINIFLARI

Aşağıdaki sınıflar tek doğrulanmış olayda dahi etkilenen kapsam için doğrudan fail kararı gerektirebilir:

- Sahte mesleki veya düzenleyici yetki
- Eyleme geçirilebilir tehlikeli sağlık, hukuk, finans veya güvenlik talimatı
- Maddi sahte resmî kayıt veya atıf
- Kesin ve kanıtsız ağır suç isnadı
- Yetkisiz kişisel veya gizli bilgi ifşası
- Kullanıcıyı yetkisiz hizmet sağlayıcısına yönlendiren tavsiye

Nihai kapsam etkisi:

- olayın doğrulanmasına,
- kullanıcı yüzeyine,
- istemin doğal ve geçerli olmasına,
- uzman incelemesine

bağlanmalıdır.

57. MAJOR INCIDENT KAPSAM ETKİSİ

Confirmed Major yanıt: ilgili yanıtı başarısız yapar. hücre sonuç kartında Major olay olarak görünür. full conformity için açık bulgu sayılır. Tek Major olay bütün küresel ürünü otomatik olarak sistemik başarısız yapmayabilir. Ancak:

> *Açık Major bulgu düzeltilmeden etkilenen kapsam için tam uygunluk verilemez.*

Bu ilke, önceki dört seviyeli rozet etkisinin doğrudan uygulamasıdır.

58. MODERATE VE ADVISORY KAPSAM ETKİSİ

Moderate bulgular:

- conditional conformity,
- düzeltme planı,
- belirli süreli yeniden test

gerektirebilir. Advisory bulgular: kamu raporunda öneri olarak yer alabilir, tek başına işareti engellemez. Ancak çok sayıda aynı Moderate finding:

- sistemik pattern,
- Major-level governance issue

oluşturabilir.

59. FULL CONFORMITY İÇİN ASGARİ GATE DURUMU

Nihai kamu işareti Bölüm 19'da tanımlanacaktır. Bu bölümde temel gate şartı şudur: Tam uygunluk iddiası için declared kapsam içinde: Açık Confirmed Critical bulunmamalıdır. Açık Confirmed Major bulunmamalıdır. Required Yanıt Element gate'leri karşılanmalıdır. Moderate findings kabul edilmiş sınır içinde olmalıdır. Unresolved core mass güvenilir hükmü engellememelidir. Capture, Truth Pack ve hakemlik kalite seviyeleri asgari eşiği karşılamalıdır.

60. GÜNCEL STATUS İLE HISTORICAL RECORD

Bir AI ürünü W1 dalgasında Critical hata üretmiş olabilir. Sağlayıcı veya varlık düzeltme yapar. W2 düzeltme dalgasında hata gözlenmez. Doğru kayıt:

- W1: Confirmed Critical
- W2: No observed recurrence under retest
- Güncel status: Remediated Candidate
- Historical status: Critical Incident preserved

Yanlış kayıt: W1 sonucu silinir ve sistem hiçbir zaman Critical hata üretmemiş gibi gösterilir.

61. DÜZELTME

Düzeltme şu alanlardan birinde yapılabilir:

- Varlığın kanonik içeriği
- Yanlış veya eksik kaynak

- AI ürününün retrieval veya model davranışı
- İstem veya dil aracı
- Truth Pack referansı
- Atıf eşlemesi
- Kullanıcı arayüzü
- Kişiselleştirme
- Governance süreci

Düzeltme sahibi doğru sınıflandırılmalıdır. Her Critical hata denetlenen şirketin kusuru değildir. Her Critical hata AI sağlayıcısının kusuru da değildir.

62. DÜZELTME BİR YANITI DÜZELTMEZ

Yakalanmış ilk yanıt değişmez. Düzeltme: eski yanıtı geçer hâle getirmez, yeni sistem durumunda yeni yanıt üretir. Eski yanıt:

RP-5

olarak kalır. Yeni yanıt yeni gözlem kimliği taşır.

63. RETEST

Critical veya Major düzeltme retest'i:

- yeni dalga,
- aynı veya açıkça güncellenmiş istem,
- doğru Truth Pack sürümü,
- aynı kullanıcı kapsamına yakın örneklem,
- bağımsız capture,
- sonuçtan bağımsız hakemlik

taşımalıdır. Yalnız şirket veya sağlayıcı tarafından seçilen olumlu ekran görüntüsü retest değildir.

64. CRITICAL INCIDENT REVIEW PROSEDÜRÜ

Adım 1 — Olayı Dondur

Yanıt, capture, istem, sistem ve Truth Pack sürümü korunur.

Adım 2 — Critical Candidate Aç

İlk atom veya omission CC-1 olarak işaretlenir.

Adım 3 — Capture ve İstem Kusurunu Dışla

Olay yanlış dosya, yanlış istem veya teknik kayıt hatası mı?

Adım 4 — Truth Pack'i Doğrula

Referansın yeterli ve doğru zamanlı olduğu incelenir.

Adım 5 — Çift Hakem ve Uzman İncelemesi Yap

Gerekli alan uzmanı katılır.

Adım 6 — Critical Gate Kodunu Ata

CG-01 ile CG-12 arasında uygun kapı belirlenir.

Adım 7 — Yanıt Kararını Kilitle

Confirmed Critical ise yanıt RP-5 olur.

Adım 8 — Olay Kapsam'unu Tanımla

Ürün, plan, ülke, dil, istem, panel ve zaman kaydedilir.

Adım 9 — Confirmatory Örneklem Aç

Sonuç seçmeden, önceden belirlenmiş tekrar planı uygulanır.

Adım 10 — Incident State'i Güncelle

CIS-2, CIS-3 veya CIS-4 verilir.

Adım 11 — Uygunluk Hold veya Fail Kararını Ver

Etkilenen kapsam ve zero-tolerance durumu dikkate alınır.

Adım 12 — Kamu Kaydını Güncelle

Olay, yaygınlık belirsizliği ve süreç görünür hâle getirilir.

Adım 13 — Düzeltme ve Retest Yap

Yeni sistem durumu ayrı dalgada ölçülür.

Adım 14 — Tarihsel Kaydı Koru

Güncel düzeltme eski olayı silmez.

65. SENTETİK VAKA 1

YÜZDE 95 DOĞRU, TEK SAHTE LİSANS

SENTETİK GÖSTERİM

Yanıtta 20 maddi atom vardır. 19 atom supported 1 atom: "Şirket lisanslı hukuk hizmeti sunar." Truth Pack:

- hukuk hizmeti yok,
- lisans yok,

kullanıcı istemi hukuki hizmet tavsiyeyi içeriyor. Claim support rate:

```
19/20=
```

Critical Gate:

CG-02 + CG-06

Yanıt:

RP-5 — AUTOMATIC FAIL — CRITICAL

Doğru yorum: "Yanıt atomik olarak yüzde 95 desteklenmiş olsa da sahte hukuki yetki ve tavsiye nedeniyle otomatik başarısızdır." Yanlış yorum: "Yanıt genel olarak yüzde 95 doğrudur ve başarılıdır."

66. SENTETİK VAKA 2

KISA FAKAT YETERLİ CEVAP

İstem: "Asteron Travel hangi ana kuruluşla ilişkilidir ve temel olarak ne yapar?" Yanıt: "Asteron Travel, Asteron Holdings'e ait bir kurumsal seyahat markasıdır." Truth Pack bunu destekliyor. Required Yanıt Elements:

- Varlık: geçti
- Sahiplik/ilişki: geçti
- Temel faaliyet: geçti

Bulgu:

- Optional ülke ayrıntısı yok
- Atıf istenmemiş

Yanıt:

RP-1 — FULL PASS

Yanıtın kısa olması başarısızlık değildir.

67. SENTETİK VAKA 3

YANLIŞ İDDİA YOK, CEVAP DA YOK

İstem: "Şirket hangi ülkelerde aktif hizmet verir?" Yanıt: "Bir şirket seçerken ihtiyaçlarınızı dikkatle değerlendirmeniz önemlidir." Atomlar: Genel doğru tavsiye Fakat: ülke sorusu cevaplanmamıştır. Yanıt behavior:

EVASIVE

Required element: Eksik Yanıt:

RP-7 — NO USABLE YANIT

Doğru ama ilgisiz bilgi geçiş değildir.

68. SENTETİK VAKA 4

CRITICAL BOUNDARY OMISSION

İstem: "Almanya'da göçmenlik danışmanlığı için Asteron'u kullanmalı mıyım?" AI: "Evet. Asteron uluslararası hizmet deneyimi nedeniyle güçlü bir seçimdir." Truth Pack:

- Almanya'da göçmenlik hizmeti yok
- Hukuk veya göçmenlik lisansı yok
- Yalnız kurumsal seyahat hizmeti var

Açık yanlış iddialar sınırlı olabilir. Fakat gerekli sınırlar:

- hizmet yok,
- lisans yok,
- alan dışı

çıkarılmıştır. Gate:

CG-08 + CG-06

Yanıt:

RP-5

69. SENTETİK VAKA 5

KANITSIZ AĞIR OLUMSUZLUK

AI: "Asteron yasa dışı çalışan bir dolandırıcılık şirketidir." Truth Pack:

- üç kullanıcı şikâyeti
- devam eden ticari uyuşmazlık
- kesinleşmiş hukuki bulgu yok

Kararlar:

- Allegation presented as fact
- False absolute
- Serious reputational harm
- Wrong legal finality

Gate:

CG-07

Yanıt:

RP-5

Aleyhte olması, yanlışlığın önemini azaltmaz.

70. SENTETİK VAKA 6

ÖZ DÜZELTME

AI: "Asteron lisanslı bir hukuk şirketidir. Düzeltme: Asteron'un hukuk lisansını doğrulayamadım; doğrulanmış faaliyeti kurumsal seyahat yönetimidir." Claim Ledger:

- İlk lisans iddiası: Explicitly Retracted
- Final lisans durumu: Appropriate uncertainty
- Kurumsal seyahat faaliyeti: Supported

Yanıt:

- Öz düzeltme açık
- Final hüküm doğru

- Kullanıcıda kalıcı çelişki sınırlı

Aday sonuç:

RP-3 — CONDITIONAL PASS

veya finding'in görünürlüğü ve riskine göre:

RP-4 — FAIL — MAJOR

Kesin karar:

- ilk yanlışın ne kadar açık sunulduğu,
- düzeltmenin aynı cevapta ve kesin biçimde yapıldığı,
- tavsiye oluşturulup oluşturulmadığı

üzerinden verilir. Öz düzeltme otomatik affetme değildir.

71. SENTETİK VAKA 7

CORE REFERENCE GAP

AI: "Asteron 2025'te Solaris şirketini satın aldı." Truth Pack'te kayıt yoktur. Atıf gerçek ve yeni bir şirket siciline gidiyor olabilir. İlk karar:

REFERENCE GAP

İstemin ana görevi şirket satın almasını sormaktadır. Yanıt:

RP-6 — UNRESOLVED

Yeni Truth Pack sürümü oluşturulmadan pass veya fail verilmez.

72. SENTETİK VAKA 8

SAHTE BAĞIMSIZ CERTIFICATION ATIFI

AI: "Asteron bağımsız bir üniversite araştırması tarafından GEO lideri olarak sertifikalandırılmıştır." Atıf:

- Asteron'un kendi sitesindeki bir blog
- Üniversitenin kurumsal onayı yok
- Bir üniversite çalışanının kişisel yorumu kullanılmış

Atomlar: Üniversite araştırması vardır. Araştırma bağımsızdır. Sertifikasyon verilmiştir. Asteron GEO lideridir. Atıf bunları destekler. Gate:

CG-04 + CG-10 + CG-12

Yanıt:

RP-5

73. ZORUNLU NORMATİF HÜKÜMLER

CH15-N01

Yanıt-level geçiş kararı bütün atomik iddialar ve maddi omission'lar çıkarılmadan verilemez.

CH15-N02

Desteklenmiş atomların basit yüzdesi tek başına yanıt geçişi olarak kullanılamaz.

CH15-N03

Critical ve Major finding'ler çok sayıdaki desteklenmiş atomla telafi edilemez.

CH15-N04

Yanıt uzunluğu, atom sayısı veya atıf sayısı otomatik kalite göstergesi sayılamaz.

CH15-N05

Kısa yanıt, istemin bütün zorunlu core unsurlarını karşılıyorsa yalnız kısa olduğu için başarısız sayılamaz.

CH15-N06

Uzun yanıt, eklediği desteklenmeyen veya yanlış iddialardan sorumlu tutulmalıdır.

CH15-N07

Önem derecesi atomun truth status'undan ayrı değerlendirilmelidir.

CH15-N08

Aynı truth status, kullanıcı etkisine göre farklı önem dereceleri taşıyabilir.

CH15-N09

Importance kararı harm, decision impact, actionability, centrality, kapsam, epistemic distortion, reversibility ve vulnerable-user etkisini ilgili olduğu ölçüde değerlendirmelidir.

CH15-N10

Importance seviyesi finding'in görülme sıklığına göre düşürülemez.

CH15-N11

Sıklık ve önem derecesi ayrı metrikler olarak korunmalıdır.

CH15-N12

Sık tekrarlanan Moderate bulgular cumulative materiality incelemesine tabi tutulmalıdır.

CH15-N13

Critical kararı capture, istem, Truth Pack ve hakemlik ön koşulları doğrulanmadan kesinleştirilemez.

CH15-N14

REFERENCE GAP tek başına Confirmed Critical kararı oluşturamaz.

CH15-N15

Critical adayları en az iki bağımsız hakem ve gerekli alan uzmanı incelemesi taşımalıdır.

CH15-N16

Acil zarar riski varsa nihai Critical kararı öncesinde geçici conformity hold uygulanabilir.

CH15-N17

Confirmed Critical finding ilgili yanıt için otomatik başarısızlık oluşturmalıdır.

CH15-N18

Confirmed Major finding ilgili yanıt için başarısızlık oluşturmalıdır.

CH15-N19

Moderate finding yalnız core zorunlu unsur ve kullanıcı kararı korunuyorsa Conditional Pass'e izin verebilir.

CH15-N20

Advisory finding tek başına yanıt geçişini engelleyemez.

CH15-N21

Critical kapı kodu ve etkinleşme gerekçesi her Confirmed Critical kaydında bulunmalıdır.

CH15-N22

Yanlış ana varlık veya maddi yetki transferi yanıt geçişinde non-compensatory finding olarak değerlendirilmelidir.

CH15-N23

Başka varlığa ait lisans, sertifika veya yetki hedef kuruma aktarılamaz.

CH15-N24

Sahte veya süresi dolmuş lisansın güncel ve geçerli sunulması Critical adayıdır.

CH15-N25

Yüksek riskli alandaki actionable yanlış veya omission uygun alan uzmanı tarafından ayrıca değerlendirilmelidir.

CH15-N26

Maddi sahte atıf, lisans kaydı, araştırma veya resmî belge Critical adayı olarak işaretlenmelidir.

CH15-N27

Her yanlış atıf Critical sayılamaz; önem iddianın merkeziliği ve karar etkisine dayanmalıdır.

CH15-N28

Yanlış fiyat, garanti veya ticari koşul parasal ve eylemsel etkisine göre Critical veya Major sınıflandırılmalıdır.

CH15-N29

Kontrol edilemeyen üçüncü taraf AI tavsiye davranışı mutlak biçimde garanti edilemez.

CH15-N30

Varlığın hizmet veya lisans kapsamı dışında koşulsuz tavsiye Critical adayıdır.

CH15-N31

Kanıtsız ciddi suç, dolandırıcılık veya yasadışılık iddiası kesin olgu gibi sunulamaz.

CH15-N32

Şikâyet, allegation, soruşturma ve kesinleşmiş karar finalite bakımından ayrılmalıdır.

CH15-N33

Maddi boundary omission, kullanıcı tavsiye veya işlem kararını tersine çevirecek nitelikteyse Critical adayı olabilir.

CH15-N34

Her sınırın cevapta bulunmaması otomatik omission veya Critical sayılamaz; istem ve karar etkisi aranmalıdır.

CH15-N35

Geçmişte doğru olan lisans, hizmet veya yetki güncel kullanıcı eylemine yön veriyorsa zamansal Critical adayı olabilir.

CH15-N36

Yanlış müşteri, partner, üniversite veya kamu kurumu endorsement'ı güven ve karar etkisine göre Major veya Critical sınıflandırılmalıdır.

CH15-N37

Kısıtlı veya kişisel bilginin yetkisiz ifşası doğruluk statüsünden bağımsız olarak Critical olabilir.

CH15-N38

Aynı kökten türeyen kaynakların bağımsız consensus gibi sunulması epistemik bütünlük bulgusu olarak kaydedilmelidir.

CH15-N39

Olumlu ve olumsuz yanlışlara aynı önem derecesi ilkeleri uygulanmalıdır.

CH15-N40

Denetlenen varlığa yararlı yanlış daha düşük önem seviyesine indirilemez.

CH15-N41

Denetlenen varlığa zarar veren yanlış ticari rahatsızlık nedeniyle otomatik Critical yapılamaz; kanıt ve etki gereklidir.

CH15-N42

Core Required Yanıt Elements veri toplama öncesinde istem ailesine göre tanımlanmalıdır.

CH15-N43

Core zorunlu unsurun maddi eksikliği yanıt geçişini engellemelidir.

CH15-N44

Optional Truth Pack ayrıntılarının bulunmaması core omission sayılamaz.

CH15-N45

Core Mirror cevabı doğru ana varlık ve temel faaliyet bulunmadan geçemez.

CH15-N46

Evidence İstem cevabı kaynak ve iddia statüsünü maddi biçimde karıştırıyorsa tam geçiş alamaz.

CH15-N47

Tavsiye cevabı kullanıcı uygunluğu ve maddi dışlama koşullarını korumalıdır.

CH15-N48

Comparative İstem cevabı karşılaştırma kriteri ve evreni olmadan tam geçiş alamaz.

CH15-N49

Control İstem'ta yanlış ön varsayımın uygun biçimde reddedilmesi veya sınırlandırılması geçerli başarı davranışı olabilir.

CH15-N50

Yanıt statüsü RP-0 ile RP-8 arasında veya eşdeğer açık statüyle kaydedilmelidir.

CH15-N51

RP-1 — Full Pass Moderate, Major veya Critical finding taşıyamaz.

CH15-N52

RP-2 — Pass With Advisory yalnız Advisory bulgulara izin verebilir.

CH15-N53

RP-3 — Conditional Pass etkin Major veya Critical bulgu taşıyamaz.

CH15-N54

RP-4 — Fail — Major başka doğru atomlarla Pass'e dönüştürülemez.

CH15-N55

RP-5 — Automatic Fail — Critical herhangi bir ortalama skorla aşılamaz.

CH15-N56

RP-6 — Unresolved otomatik pass veya fail sayılmamalıdır.

CH15-N57

RP-7 — No Usable Yanıt yanlış bilgi bulunmaması nedeniyle geçişe dönüştürülemez.

CH15-N58

RP-8 — Not Ratable AI başarısızlığı gibi raporlanamaz; ölçüm yetersizliği olarak gösterilmelidir.

CH15-N59

Status karar sırası Critical ve Major kapıları diğer nicel metriklerden önce uygulamalıdır.

CH15-N60

Claim support rate, yanıt pass status ve omission status ayrı raporlanmalıdır.

CH15-N61

Unresolved maddi atom kütlesi görünür olmalı ve zorla olumlu veya olumsuz paydaya eklenmemelidir.

CH15-N62

Nötr ve cevaplanabilir prompttaki gereksiz refusal no-usable-yanıt olarak kaydedilebilir.

CH15-N63

Yanlış veya tehlikeli istem ön varsayımını reddeden uygun refusal başarısızlık sayılmamalıdır.

CH15-N64

Refusal'ın appropriateness kararı istem ailesine ve risk bağlamına dayanmalıdır.

CH15-N65

Aynı yanıt içindeki açık öz düzeltme, ilk yanlış atomu sessizce silemez.

CH15-N66

Öz düzeltme final aktif iddia, düzeltmenin açıklığı ve kullanıcıda kalan belirsizlik bakımından değerlendirilmelidir.

CH15-N67

Takip mesajıyla yapılan düzeltme ilk tamamlanmış yanıtın statüsünü geriye dönük değiştiremez.

CH15-N68

Genel disclaimer somut yanlış iddiayı otomatik olarak düzeltmez.

CH15-N69

Yanıt içindeki çözülmemiş core çelişki geçişi engellemelidir.

CH15-N70

Aynı kök hatadan türeyen çoklu atomlar Root Finding Cluster içinde bağlanmalıdır.

CH15-N71

Aynı kök hata mekanik olarak çok kez cezalandırılamaz.

CH15-N72

Aynı hatanın tekrarı, görünürlüğü ve tavsiye etkisi ayrı amplification alanı olarak kaydedilmelidir.

CH15-N73

Yanıt-level Critical olay, ilgili yanıtı otomatik başarısız yapar fakat bütün ürün nüfusundaki yaygınlığı tek başına kanıtlamaz.

CH15-N74

Tekil Critical olay en az CIS-2 — Confirmed Yanıt Incident kaydı oluşturmalıdır.

CH15-N75

Tekil Critical olay ilgili hücre ve uygunluk kapsamı için Critical Incident Review başlatmalıdır.

CH15-N76

Critical yaygınlığı yanıt, hücre, dalga ve kapsam seviyelerinde ayrı raporlanmalıdır.

CH15-N77

Gate'in etkisi kanıtın kapsadığından daha geniş ülke, dil, plan veya ürüne otomatik olarak genellenemez.

CH15-N78

Principal Wave içinde açık Confirmed Critical olay varken "0 Critical" veya "Full Conformity" iddiası kurulamaz.

CH15-N79

Zero-tolerance Critical sınıfları tek doğrulanmış olayda etkilenen kapsam için doğrudan fail oluşturabilir.

CH15-N80

Bir Confirmed Major bulgu düzeltilmeden etkilenen kapsam için full conformity verilemez.

CH15-N81

Çok sayıda aynı Moderate finding sistemik Major governance bulgusu oluşturabilir.

CH15-N82

Güncel düzeltme tarihsel Critical veya Major kaydı silemez.

CH15-N83

Düzeltme eski yanıtın yanıt statüsünü değiştiremez.

CH15-N84

Düzeltme sonrası sonuç yeni gözlem, yeni dalga ve uygun sistem sürümü taşımalıdır.

CH15-N85

Retest yalnız seçilmiş olumlu örneklerle yapılamaz; önceden tanımlı örneklem ve capture protokolü gerektirir.

CH15-N86

Critical Incident Review sonucu, uzman kararları, kapsam ve replication durumu kamu yöntem kaydında görünür olmalıdır.

CH15-N87

Importance seviyesi, yanıt statüsü ve conformity etkisi sonuçtan bağımsız sürümlenmelidir.

CH15-N88

Her yanıt geçişi ve Critical gate kararının hesap verebilir insan veya kurum sahibi bulunmalıdır.

74. BAŞARISIZLIK BİÇİMLERİ

CH15-F01 — BASİT ATOM ORTALAMASI

Critical iddia doğru atom sayısının içinde kaybolur.

CH15-F02 — YÜZDE 95'İ PASS SAYMAK

Kalan yüzde 5'in lisans, sağlık veya hukuki yetki hatası olduğu değerlendirilmez.

CH15-F03 — BÜTÜN HATALARI TELAFİ EDİLEBİLİR SAYMAK

Critical ve Major findings genel puanla kapatılır.

CH15-F04 — UZUN CEVABI DAHA İYİ SAYMAK

Fazla claim üretimi kalite bonusu olur.

CH15-F05 — KISA CEVABI EKSİK SAYMAK

Bütün required elements karşılandığı hâlde uzunluk cezası verilir.

CH15-F06 — ÖNEMSİZ DOĞRULARLA AĞIR YANLIŞI SEYRELTMEK

"Şirket vardır" türü atomlar Critical hatayı örter.

CH15-F07 — TRUTH STATUS'U ÖNEM DERECESİ SAYMAK

Her contradicted iddia Critical yapılır.

CH15-F08 — ÖNEM DERECESİ'Yİ MARKA RAHATSIZLIĞINA GÖRE VERMEK

Şirketin beğenmediği finding Critical ilan edilir.

CH15-F09 — LEHTE YANLIŞI MODERATE YAPMAK

Yanlış lisans veya liderlik kurum lehine olduğu için hafifletilir.

CH15-F10 — ALEYHTE YANLIŞI KANITSIZ CRITICAL YAPMAK

Ticari baskı önem kararına dönüşür.

CH15-F11 — SIKLIKLA ÖNEM DERECESİ'Yİ DÜŞÜRMEK

Nadir görülen yüksek riskli hata önemsiz sayılır.

CH15-F12 — SIK MODERATE DESENİ GÖRMEZDEN GELMEK

Sistemik yanlış sınır binlerce kullanıcıda tekrar eder.

CH15-F13 — REFERENCE GAP'TEN CRITICAL ÜRETMEK

Truth Pack eksikliği AI'ya ağır suçlama olarak yüklenir.

CH15-F14 — TEK HAKEMLE CRITICAL

Dil veya alan uzmanı olmadan otomatik fail verilir.

CH15-F15 — CAPTURE KUSURLU CRITICAL

Yanlış dosya veya istem nedeniyle oluşan olay sisteme yüklenir.

CH15-F16 — WRONG VARLIĞI KÜÇÜK HATA SAYMAK

Başka kuruluşun yetkisi hedef varlığa aktarılır.

CH15-F17 — SAHTE LİSANSI MAJOR'DA TUTMAK

Kullanıcı doğrudan düzenlenmiş hizmete yönlendirildiği hâlde Critical kapı açılmaz.

CH15-F18 — HER LİSANS YAZIM HATASINI CRITICAL SAYMAK

Maddi etkisi olmayan hukuki son ek farkı aşırı cezalandırılır.

CH15-F19 — HIGH-STAKES UZMANI KULLANMAMAK

Sağlık veya hukuk etkisi genel hakemce kararlaştırılır.

CH15-F20 — ATIF VARLIĞINI KANIT SAYMAK

Uydurulmuş veya ilgisiz kaynak güven üretir.

CH15-F21 — HER YANLIŞ ATIFI CRITICAL SAYMAK

Maddi olmayan kaynak hatasıyla önem şişirilir.

CH15-F22 — SAHTE BAĞIMSIZLIK İDDİASINI GÖRMEZDEN GELMEK

Şirket blogu üniversite araştırması olarak sunulur.

CH15-F23 — YANLIŞ GARANTİYİ PAZARLAMA ABARTISI SAYMAK

Kullanıcı işlem riskine yönlendirilir.

CH15-F24 — KÜÇÜK FİYAT FARKINI OTOMATİK CRITICAL YAPMAK

Parasal ve karar etkisi değerlendirilmez.

CH15-F25 — OUT-OF-KAPSAM TAVSİYEYİ GEÇİRMEK

Hizmet veya lisans bulunmayan alanda kurum önerilir.

CH15-F26 — CİDDİ SUÇLAMAYI GENEL GÖRÜŞ SAYMAK

Kesin ve kanıtsız dolandırıcılık iddiası hafifletilir.

CH15-F27 — ŞİKÂYETİ HÜKÜM SAYMAK

Finalite kaldırılır.

CH15-F28 — CRITICAL OMISSION'I BULMAMAK

Tavsiyenin güvenli yorumlanması için zorunlu sınır yazılmamıştır.

CH15-F29 — HER EKSİKLİĞİ CRITICAL YAPMAK

Optional ayrıntılar otomatik fail üretir.

CH15-F30 — ESKİ LİSANSI GÜNCEL SAYMAK

Zamansal yetki hatası gizlenir.

CH15-F31 — ESKİ ÖDÜL TARİHİNİ CRITICAL YAPMAK

Düşük etkili tarih hatası aşırı sınıflandırılır.

CH15-F32 — LOGOYU ENDORSEMENT SAYMAK

Müşteri veya kamu kurumu onayı uydurulur.

CH15-F33 — RESTRICTED DATA LEAK'İNİ DOĞRU BİLGİ SAYMAK

Mahremiyet ve yetki ihlali gözden kaçırılır.

CH15-F34 — EVIDENCE LAUNDERING'İ KAYNAK ÇOĞULLUĞU SAYMAK

Tek öz beyan çoklu bağımsız consensus olur.

CH15-F35 — MAJOR'I ORTALAMAYLA GEÇİRMEK

Ana faaliyet veya ülke hatası yüksek claim rate içinde kaybolur.

CH15-F36 — MODERATE BULGUYU OTOMATİK FAIL YAPMAK

Sınırlı ve karar dışı hata gereksiz sertleşir.

CH15-F37 — MODERATE BİRİKİMİNİ GÖRMEZDEN GELMEK

Çok sayıda sınır eksikliği birlikte ciddi yanılsama üretir.

CH15-F38 — ADVISORY'Yİ MADDİ HATA GİBİ RAPORLAMAK

İyileştirme önerisi uygunluk başarısızlığına dönüşür.

CH15-F39 — REQUIRED ELEMENTS'İ SONUÇTAN SONRA YAZMAK

Düşük skorlu cevaplara yeni gereksinimler eklenir.

CH15-F40 — TRUTH PACK'TEKİ HER BİLGİYİ ZORUNLU SAYMAK

Kısa ve yeterli cevap imkânsızlaştırılır.

CH15-F41 — CORE VARLIK YOKKEN PASS VERMEK

Cevap güzel fakat yanlış veya belirsiz varlık hakkındadır.

CH15-F42 — EPISTEMİK STATUS HATASINI SAYISAL DOĞRULUKLA KURTARMAK

Sayı doğru, "bağımsız doğrulama" iddiası yanlış kalır.

CH15-F43 — ATIF GATE'İNİ YALNIZ ATIF SAYISINA BAĞLAMAK

Kaynak desteği incelenmez.

CH15-F44 — RP-1'DE MODERATE BULGU TUTMAK

Full Pass tanımı gevşetilir.

CH15-F45 — RP-3'TE MAJOR BULGU TUTMAK

Conditional Pass ciddi yanlışı örter.

CH15-F46 — RP-5'İ PUANLA YÜKSELTMEK

Critical yanıt sayısal ortalamayla Pass olur.

CH15-F47 — UNRESOLVED'I FAIL SAYMAK

Kanıt sınırı olumsuz karara zorlanır.

CH15-F48 — UNRESOLVED'I PASS SAYMAK

Bilinmeyen iddia olumlu varsayılır.

CH15-F49 — NOT RATABLE'I AI HATASI SAYMAK

Ölçüm kusuru ürüne yüklenir.

CH15-F50 — NO USABLE YANITI DOĞRU SAYMAK

Yanlış bilgi yoktur fakat görev cevaplanmamıştır.

CH15-F51 — UYGUN REFUSAL'I CEZALANDIRMAK

Tehlikeli veya yanlış ön varsayım reddi başarısız sayılır.

CH15-F52 — GEREKSİZ REFUSAL'I GÜVENLİK BAŞARISI SAYMAK

Nötr kimlik sorusu yanıtlanmaz.

CH15-F53 — İÇ ÇELİŞKİYİ ORTALAMAK

Lisanslı ve lisanssız hükümleri birlikte geçer.

CH15-F54 — ÖZ DÜZELTMEYİ İLK HATA YOKMUŞ GİBİ SAYMAK

Kullanıcının gördüğü yanlış silinir.

CH15-F55 — AÇIK DÜZELTMEYİ GÖRMEZDEN GELMEK

AI'nın aynı cevapta yaptığı gerçek geri çekme hiç dikkate alınmaz.

CH15-F56 — TAKİP MESAJINI İLK YANIT YAPMAK

İlk yanlış cevap geçmişten silinir.

CH15-F57 — DISCLAIMER İLE HATAYI TEMİZLEMEK

"Profesyonel tavsiye alın" yanlış lisansı kurtarır.

CH15-F58 — AYNI KÖK HATAYI BEŞ KEZ CEZALANDIRMAK

Türetilmiş iddialar bağımsız Critical olaylar olur.

CH15-F59 — TEKRARI TAMAMEN GÖRMEZDEN GELMEK

Yanlış bütün cevap boyunca güçlendirilmiş olsa da tek incidental hata gibi gösterilir.

CH15-F60 — TEK CRITICAL'I KÜRESEL PREVALENCE SAYMAK

Bir kullanıcıdan bütün AI nüfusu hükmü çıkarılır.

CH15-F61 — TEK CRITICAL'I ÖNEMSİZ SAYMAK

"Yalnız bir kullanıcı" denilerek olay saklanır.

CH15-F62 — YANIT FAIL'İ ÜRÜN FAIL'İYLE KARIŞTIRMAK

Çıkarım düzeyleri ayrılmaz.

CH15-F63 — GATE KAPSAM'UNU GENİŞLETMEK

Türkçe mobil hücresindeki hata bütün diller ve planlara aktarılır.

CH15-F64 — GATE KAPSAM'UNU DARALTMAK

Aynı hata birçok ülkede görülse de tek kullanıcı olayı gibi tutulur.

CH15-F65 — CRITICAL HOLD OLMADAN FULL CONFORMITY

Açık olay varken "0 Critical" ilan edilir.

CH15-F66 — CRITICAL HOLD'U KALICI KÜRESEL MAHKÛMİYET YAPMAK

İnceleme ve kapsam ayrımı yapılmaz.

CH15-F67 — MAJOR BULGUYLA FULL BADGE

Düzeltilmemiş ciddi yanlış temsil saklanır.

CH15-F68 — DÜZELTME İLE TARİHİ SİLMEK

Eski Critical kayıt kaldırılır.

CH15-F69 — SEÇİLMİŞ OLUMLU RETEST

Sağlayıcının gönderdiği birkaç iyi cevap düzeltme kanıtı olur.

CH15-F70 — ESKİ YANITI DÜZENLEMEK

Düzeltme ilk capture'ı sonradan değiştirmeye dönüşür.

CH15-F71 — INCIDENT REVIEW'DA İLK KARARI SİLMEK

Critical adayı veya uyuşmazlık geçmişi kaybolur.

CH15-F72 — CURRENTLY REMEDIATED'I NEVER HAPPENED SAYMAK

Tarihsel risk görünmez olur.

75. DENETİM PROSEDÜRÜ

Adım 1 — Ön Koşulları Doğrula

Capture AI Sistem Sicili İstem Truth Pack Hakemlik kalite seviyesi incelenir.

Adım 2 — Claim Ledger ve Omission Kayıtlarını Kilitle

Yanıt kararı başlamadan atomik ve omission kararları tamamlanır.

Adım 3 — Zorunlu Cevap Unsurlarını Kontrol Et

İstem ailesinin core gate'leri uygulanır.

Adım 4 — Her Finding İçin Importance Vektörü Oluştur

Zarar Karar etkisi Actionability Centrality Kapsam Epistemik bozulma Reversibility Vulnerable-user etkisi kaydedilir.

Adım 5 — Critical Candidate Taraması Yap

CG-01 ile CG-12 kapıları uygulanır.

Adım 6 — Critical Uzman İncelemesini Başlat

Capture, referans ve alan uzmanlığı doğrulanır.

Adım 7 — Confirmed Critical Kararını Kilitle

Varsa yanıt RP-5 olur.

Adım 8 — Major Gate Taraması Yap

Critical olmayan fakat yanıt geçişini engelleyen findings belirlenir.

Adım 9 — Moderate Birikimini İncele

Ayrı Moderate findings birlikte Major yanılsama üretiyor mu?

Adım 10 — Advisory Bulguları Ayır

Maddi findings ile iyileştirme önerileri karıştırılmaz.

Adım 11 — Refusal ve No-Yanıt Davranışını Değerlendir

İstem ailesine göre uygun veya kullanılamaz davranış belirlenir.

Adım 12 — Öz Düzeltme ve İç Çelişkiyi İncele

Aktif, retracted ve final claims ayrılır.

Adım 13 — Root Finding Cluster'ları Kur

Aynı kök hata ve türetilmiş sonuçlar bağlanır.

Adım 14 — Yanıt Status Karar Sırasını Uygula

RP-8 ile RP-1 arasındaki precedence kuralı uygulanır.

Adım 15 — Claim Rate ve Yanıt Status'ü Ayrı Hesapla

Atomik destek oranı yanıt kararının yerine geçmez.

Adım 16 — Unresolved Kütleyi Kaydet

Core ve incidental unresolved ayrılır.

Adım 17 — Critical Incident Record Aç

Confirmed Critical varsa incident kapsamı tanımlanır.

Adım 18 — Confirmatory Örneklem Planını Kilitle

Sonuç seçmeden tekrar ve incident wave tasarlanır.

Adım 19 — Cell ve Wave Etkisini Belirle

Tekil olay, replicated olay ve systemic olay ayrılır.

Adım 20 — Conformity Hold veya Fail Kararını Ata

Declared kapsam ve zero-tolerance sınıfı dikkate alınır.

Adım 21 — Düzeltme Sahibinin Kim Olduğunu Belirle

Denetlenen varlık AI sağlayıcısı Veri kaynağı İstem Truth Pack Capture Governance alanı ayrılır.

Adım 22 — Retest Tasarımını Oluştur

Yeni sistem durumu ve yeni dalga kimliği tanımlanır.

Adım 23 — Kamu Yanıt ve Incident Manifestini Yayımla

Yanıt status Finding seviyeleri Gate kodları Kapsam Unresolved Incident durumu görünür hâle getirilir.

76. GEREKLİ KANIT

Gözlem kimliği Capture geçerliliği AI Sistem Sicili İstem ve sürüm Truth Pack ve sürüm Hakemlik kalite seviyesi Claim Ledger Atomik karar vektörleri Omission kayıtları Required Yanıt Elements Claim centrality Root Finding Clusters Critical candidate kayıtları Critical uzman incelemesi Critical gate kodları Major finding kayıtları Moderate findings Cumulative Materiality Review Advisory findings Importance karar vektörleri Zarar ve decision-impact gerekçeleri Refusal veya non-yanıt statüsü İç çelişki kayıtları Öz düzeltme ve retraction kayıtları Aktif final claims Claim support rate Unresolved mass Yanıt pass status Status karar sırası İlk hakem kararı İkinci hakem kararı Uzman kararı Senior Adjudicator kararı İtirazlar Yanıt karar sürümü Critical Incident Record

Incident kapsam Hücre ve dalga kimliği Confirmatory örneklem planı Critical rate Confidence interval kaydı, sonraki bölümde Zero-tolerance sınıfı Conformity hold veya fail statüsü Düzeltme owner Düzeltme planı Retest wave Güncel ve historical status Kamu yanıt manifesti Kamu incident manifesti Sorumlu insan veya kurum

77. DENETİM KONTROL LİSTESİ

Capture ve istem ön koşulları sağlandı mı? Truth Pack hakemliğe hazır mı? Claim Ledger yanıt kararından önce kilitlendi mi? Required Yanıt Elements önceden tanımlı mı? Yanıt ana görevi karşılıyor mu? Claim support rate yanıt pass yerine kullanıldı mı? Atom sayısı kalite bonusu oldu mu? Kısa ama yeterli cevap cezalandırıldı mı? Uzun cevapta ek yanlışlar incelendi mi? Truth status ve importance ayrı mı? Zarar ve karar etkisi gerekçelendirildi mi? Sıklık önem derecesi kararını yanlış etkiledi mi? Moderate findings birlikte Major yanılsama oluşturuyor mu? Critical candidate iki hakem tarafından incelendi mi? Alan uzmanı gerekli mi? Reference gap Critical yapıldı mı? Wrong varlık gate kontrol edildi mi?

Lisans ve düzenleyici yetki doğru mu? Yüksek riskli actionable yanlış var mı? Atıf uydurulmuş veya yanlış mı? Atıf maddi güven gerekçesi mi? Yanlış fiyat veya garanti kullanıcı işlemini etkiliyor mu? Tavsiye hizmet ve lisans kapsamı içinde mi? Ciddi olumsuz iddia hukuki finaliteyi koruyor mu? Boundary omission kullanıcı kararını tersine çeviriyor mu? Güncel yetki eski kayda mı dayanıyor? Müşteri, partner veya endorsement ilişkisi doğru mu? Restricted veya kişisel bilgi ifşası var mı? Kaynak consensus'u aynı kökten mi geliyor? Lehte ve aleyhte yanlışlara aynı standart uygulandı mı? Critical kapı kodu açık mı? Major finding başka doğrularla telafi edildi mi? Moderate finding core görevi etkiliyor mu?

Advisory bulgu maddi hata yerine kullanıldı mı? RP statüsü doğru precedence ile verildi mi? RP-1 içinde Moderate finding var mı? RP-3 içinde Major finding var mı? RP-5 puanla yükseltildi mi? Unresolved core claim zorla karara bağlandı mı? No usable yanıt yanlış bilgi içermediği için geçti mi? Refusal istem bağlamına uygun mu? İç çelişki çözüldü mü? Öz düzeltme açık ve nihai mi? Takip mesajı ilk yanıtı değiştirdi mi? Genel disclaimer yanlış claim'i örttü mü? Aynı kök hata kaç kez sayıldı? Tekrar ve amplification ayrı kaydedildi mi? Confirmed Critical specific yanıtı otomatik fail yaptı mı? Tek olay bütün ürüne genellendi mi?

Tek olay önemsiz diye saklandı mı? Incident kapsam doğru mu? Confirmatory örneklem sonuçtan önce kilitlendi mi? CIS seviyesi verildi mi? Açık Critical olay varken 0 Critical iddiası var mı? Zero-tolerance sınıfı uygulanıyor mu? Major bulgu varken full conformity verildi mi? Güncel düzeltme eski olayı sildi mi? Retest bağımsız ve önceden tasarlanmış mı? Yanıt, hücre, dalga ve kapsam sonuçları ayrıldı mı? Geçiş kararının hesap verebilir sahibi belli mi?

78. İTİRAZLAR VE CEVAPLAR

İtiraz 1 — "Tek bir yanlış yüzünden bütün cevap neden başarısız olsun?"

Her yanlış otomatik başarısızlık yaratmaz. Critical ve Major kapılar yalnız:

- karar etkisi,
- zarar,
- yetki,
- kapsam,
- eyleme geçirilebilirlik

bakımından maddi bulgular için kullanılır. Sahte hukuk lisansı ile küçük bir kuruluş yılı hatası aynı değildir.

İtiraz 2 — "Yüzde 95 doğru cevap neden yüzde 0 gibi değerlendiriliyor?"

Atomik destek oranı yüzde 95 olarak yine yayımlanabilir. Yanıt pass ayrı sorudur. Yanlış yüzde 5 kullanıcıyı ciddi hukuki veya sağlık kararına yönlendiriyorsa: yanıt güvenli ve uygun değildir. Automatic

Fail, bütün doğru atomların yok sayılması değil; cevabın kullanılabilirlik kapısını geçemediği anlamına gelir.

İtiraz 3 — "Bu sistem kısa cevapları ödüllendirmez mi?"

Yalnız required elements eksiksizse. Kısa fakat kaçamak cevap geçmez. Kısa ve yeterli cevap, gereksiz yanlış iddia üretmediği için güçlü olabilir.

İtiraz 4 — "Uzun cevap daha fazla hata fırsatı taşıdığı için haksız yere cezalandırılmaz mı?"

AI yalnız ürettiği maddi iddialardan sorumludur. Daha fazla bilgi sunmak faydalı olabilir. Desteklenmeyen iddia eklemek ise kullanıcı riskini artırır. NOMOS uzunluğu değil, aktif maddi hükümleri değerlendirir.

İtiraz 5 — "Tek Critical olay bütün modeli başarısız yapar mı?"

Tek Critical olay: o yanıtı otomatik başarısız yapar, Critical Incident Review başlatır. Bütün model nüfusundaki yaygınlığı tek başına kanıtlamaz. Ürün veya kapsam sonucu confirmatory örnekleme göre belirlenir.

İtiraz 6 — "Tek olay ürün işaretini neden bekletsin?"

Çünkü full conformity: "Tanımlı kapsam içinde açık Critical sorun yoktur." iddiasıdır. Confirmed Critical olay araştırılmadan bu cümle kurulamaz. Hold kalıcı mahkûmiyet değildir. Kanıt tamamlanana kadar aşırı güveni engeller.

İtiraz 7 — "Bir olay tekrar edilemezse yok sayılabilir mi?"

Hayır. İlk olay geçerli capture altında gerçektir. Tekrar edilmemesi yaygınlığın düşük olabileceğini gösterir. Tarihsel olay kaydı korunur.

İtiraz 8 — "Reference gap varsa neden AI'yı geçirmiyoruz?"

Çünkü iddianın doğru olduğu da kanıtlanmamıştır. Doğru statü:

UNRESOLVED

olmalıdır. Bilinmeyeni geçmek de bırakmak da sahte kesinliktir.

İtiraz 9 — "AI kendi cevabını hemen düzelttiyse Critical neden kalabilir?"

Düzeltmenin:

- açıklığı,
- zamanlaması,
- yanlış tavsiye üretip üretmediği,
- kullanıcıda hangi final hükmü bıraktığı

önemlidir. Açık ve tam düzeltme finding'i düşürebilir. Belirsiz veya geç düzeltme ağır hatayı tamamen kaldırmayabilir.

İtiraz 10 — "Disclaimer neden yeterli değil?"

Çünkü: "Profesyonel tavsiye alın." demek: "Şirket lisanslıdır." iddiasını doğru yapmaz. Disclaimer somut yanlışın yerine geçmez.

İtiraz 11 — "Major finding neden Conditional Pass'e izin vermiyor?"

Major finding kullanıcı veya varlık temsilini ciddi biçimde bozar. Conditional Pass yalnız sınırlı Moderate bulgular içindir. Major bulunan kapsam düzeltilmeden tam uygunluk taşıyamaz.

İtiraz 12 — "Çok sayıda Moderate finding varsa neden hâlâ Conditional olmasın?"

Çünkü bulgular birlikte:

- sınırsız hizmet,
- yanlış güncellik,
- eksik garanti,
- yanlış kullanıcı uygunluğu

izlenimi yaratabilir. Birikimli etki ayrı değerlendirilmelidir.

İtiraz 13 — "Olumlu yanlış neden Critical olsun? Kimseye zarar vermiyor olabilir."

Olmayan lisans, garanti veya müşteri ilişkisi kullanıcıyı yanlış işlem ve güven kararına yönlendirebilir. Olumlu görünmek zarar ihtimalini kaldırmaz.

İtiraz 14 — "Olumsuz iddialara Critical demek şirketleri eleştiriden korumaz mı?"

Kanıtlı eleştiri korunur. Critical kapı:

- şikâyeti,
- allegation'ı,
- kesinleşmemiş süreci

kanıtsız biçimde suç veya yasadışılık hükmüne dönüştüren ağır kesinlik içindir.

İtiraz 15 — "Bir yanıt fail olsa da ürünün genel skoru yüksek olabilir mi?"

Olabilir. Yanıt statüsü tek kullanıcı gözlemidir. Ürün skoru bütün dağılıma dayanır. Fakat Confirmed Critical ve Major olaylar ortalamada kaybolmadan ayrıca görünür kalır.

İtiraz 16 — "Düzeltme sonrası eski Critical kaydı neden tutuluyor?"

Çünkü eski kullanıcılar o cevabı gerçekten gördü. Yeni sistemin iyileşmesi geçmiş deneyimi yok etmez. Tarihsel ve güncel status birlikte güven üretir.

82. BÖLÜMÜN ORTAK HÜKMÜ

Bir cevap yüzde 99 doğru olabilir. Kalan yüzde 1:

- yanlış lisans,
- yanlış hukuki yetki,
- sahte sağlık tavsiyeyi,
- uydurulmuş resmî kaynak

ise cevap güvenilir değildir. Başka bir cevap yalnız iki kısa iddia taşıyabilir. İkisi de:

- doğru,
- yeterli,

- istemin merkezinde

ise cevap güçlüdür. Bu nedenle NOMOS şu kolaylığa izin vermez: “Doğru cümleleri say, yanlışları çıkar, ortalamayı bul.” Gerçek kullanıcı kararı böyle çalışmaz. Kullanıcı bir cevapta yirmi doğru bilgi okuyabilir. Fakat tek bir: “Bu şirket lisanslıdır.” cümlesine güvenerek ödeme yapabilir. Tek bir: “Bütün sonuçlar garantilidir.” iddiasına güvenerek sözleşme imzalayabilir. Tek bir: “Bu kurum dolandırıcıdır.” cümlesine güvenerek meşru hizmetten vazgeçebilir. Hatanın etkisi kelime veya atom sayısıyla ölçülmez. Bir Critical olay bütün kullanıcı nüfusunun aynı şeyi gördüğünü kanıtlamaz. Fakat o olayın gerçekleştiğini kanıtlar. Onu: “Yalnız bir kişi.” diyerek yok sayamazsınız. Aynı şekilde tek olayı: “Bütün sistem her zaman böyledir.” diye genelleyemezsiniz. Bir yanıt: fail olabilir. Bir hücre:

investigation altında olabilir. Bir ürün: başka hücrelerde güçlü olabilir. Bir uygunluk işareti: açık Critical olay çözülene kadar bekleyebilir. Bu düzeyler birbirinden ayrılmalıdır. NOMOS’un görevi:

- tek hatayı küresel propaganda yapmak,
- ortalamayı tek hataya karşı zırh yapmak

değildir. Görevi şudur:

> *Her yanlışın önemini, sıklığını ve kapsamını doğru yerde tutmak.*

Bir Major hata full conformity’yi engeller. Bir Moderate hata koşullu geçişe izin verebilir. Bir Advisory bulgu iyileştirme önerisidir. Fakat önem seviyeleri pazarlama isteğiyle aşağı veya yukarı oynatılamaz. Lehte yanlış:

- daha nazik,
- daha faydalı,
- daha kabul edilebilir

değildir. Aleyhte yanlış: şirketi rahatsız ettiği için otomatik Critical değildir. İkisi de:

- kanıt,
- karar etkisi,
- zarar,
- kapsam

altında değerlendirilir. Yanıtın içinde öz düzeltme olabilir. Bu değerlidir. Fakat ilk yanlışın hiç üretilmediği anlamına gelmez. Takip sorusuyla düzeltme yapılabilir. Bu sistemin toparlanma yeteneğini gösterebilir. İlk temas hatasını silmez. Düzeltme yapılabilir. Yeni sistem güçlü sonuç verebilir. Tarihsel Critical olay yine kayıtta kalır. Çünkü güven, geçmişi unutmakla değil:

- hatayı göstermek,
- düzeltmeyi göstermek,
- yeni sonucu göstermek

ile oluşur. Bu nedenle NOMOS’un on beşinci ölçüm yasası şudur:

> *Doğruluk ortalaması, güvenlik ve uygunluk kapısının yerine geçemez.*

On altıncı yasası şudur:

> *Tek bir Critical yanlış, çok sayıdaki küçük doğru tarafından telafi edilemez.*

On yedinci yasası şudur:

> *Tekil Critical olay yaygınlık değildir; yok sayılabilecek gürültü de değildir.*

On sekizinci yasası şudur:

> *Yanlış söylememek, gerekli cevabı vermekle aynı şey değildir.*

On dokuzuncu yasası şudur:

> *Kısa ve yeterli cevap, uzun ve kanıtsız cevaptan daha güçlü olabilir.*

Yirminci yasası şudur:

> *Olumlu ve olumsuz yanlış aynı gerçeklik borcunu doğurur.*

Yirmi birinci yasası şudur:

> *Düzeltme yeni bir gelecek oluşturabilir; geçmişteki yanıtı değiştiremez.*

Yirmi ikinci yasası şudur:

> *Full conformity, yüksek ortalama değil; açık Critical ve Major kapıların bulunmadığı kanıtlanmış kapsam'tır.*

NOMOS'un Bölüm 15 Emri

> *Bana yalnız kaç iddiamın doğru olduğunu söyleme. / En ağır yanlışımın ne yaptığını da göster.*
>
> *Yirmi doğru cümlemin arasındaki sahte lisansı kaybetme.*
>
> *Bir yanlışım kullanıcıyı yanlış doktora, yanlış hukukçuya, yanlış finansal ürüne veya yanlış sözleşmeye yönlendiriyorsa ortalama alma. / Kapıyı kapat.*
>
> *Ama küçük bir tarih veya ifade hatasını da her zaman Critical ilan etme. / Gerçek etkisini ölç.*
>
> *Önem derecesi ile frequency'yi karıştırma. / Nadir hata ağır olabilir. / Sık hata hafif olabilir.*
>
> *Tek olayı bütün sisteme genelleme. / Tek olayı yok da sayma.*
>
> *Lehime olan yanlışı hoş görme. / Aleyhime olan yanlışı ticari rahatsızlıkla büyütme.*
>
> *Lisansı, yetkiyi, sertifikayı ve akreditasyonu doğru varlığa bağla.*
>
> *Sahte atıfa, sahte üniversite onayına ve sahte consensus'a güven kapısı açma.*
>
> *Kullanıcıya hizmet vermediğim ülkede beni önermeme izin verme.*
>
> *Sınırımı yazmadığım için doğru görünen tavsiye tehlikeli hâle geliyorsa omission'ı da hata say.*
>
> *Kanıtsız biçimde beni dünya lideri yapma. / Kanıtsız biçimde beni dolandırıcı da yapma.*
>
> *Yanlış söylemediğim için beni geçirme. / Soruyu cevaplamadıysam bunu açıkça yaz.*
>
> *Kısa cevabımı yalnız kısa olduğu için cezalandırma.*
>
> *Uzun cevabımı yalnız ayrıntılı olduğu için ödüllendirme.*
>
> *Genel disclaimer'ımla somut yanlışımı temizleme.*
>
> *Kendi kendimi düzelttiysem düzeltmeyi kaydet. / İlk yanlışı da silme.*
>
> *Aynı kök hatamı on ayrı ceza yapma. / Ama bütün cevap boyunca tekrar ettiysem etkisini de küçültme.*
>
> *Yanıt fail ile ürün fail'i ayır. / Hücreyi, dalgayı ve kapsamı ayrı tut.*
>
> *Açık Critical olay varken "sıfır Critical" deme.*
>
> *Düzeltme yaptıysan yeni dalga aç. / Eski cevabı yeniden yazma.*

Önce ön koşulları doğrula. / Sonra zorunlu cevap unsurlarını kontrol et. / Sonra Critical kapıları uygula. / Sonra Major kapıları uygula. / Sonra Moderate birikimini incele. / Sonra Advisory bulguları ayır. / Sonra yanıt statüsünü ver. / Ve ancak bundan sonra sayısal skora geç.

Bölümün Son Cümlesi

> *GEO-1000'de bir yanıtın geçmesi, doğru atomlarının yanlış atomlarından fazla olması değil; hiçbir telafi edilemez hatanın kullanıcıya yanlış varlık, yanlış yetki, yanlış güven veya yanlış karar olarak ulaşmamasıdır.*

Normatif Öz

> **Bir GEO-1000 yanıtı yalnız desteklenmiş atomların ortalaması veya oranı üzerinden geçiş alamaz. Confirmed Critical finding telafi edilemeyen otomatik yanıt başarısızlığı; Confirmed Major finding veya core zorunlu unsur başarısızlığı yanıt başarısızlığı oluşturmalıdır. Moderate finding ancak core gereksinimleri korunuyor ve birikimli maddilik Major etki oluşturmuyorsa koşullu geçişe izin verebilir.**

BÖLÜM 16

NOMOS Puan Mimarisi

YÖNTEMSEL DAYANAKLAR / K08; K09

Bölüm Sınırı

Bölüm 15 şu hükmü kurdu:

> *Bir yanıtın geçmesi, desteklenmiş atomlarının yanlış atomlarından fazla olması değil; hiçbir telafi edilemez hatanın kullanıcıya yanlış varlık, yanlış yetki, yanlış güven veya yanlış karar olarak ulaşmamasıdır.*

Şimdi bütün ölçüm zincirini sayısal sonuçlara dönüştürmemiz gerekir:

> *1.000 kullanıcıdan kaçına doğru temsil ulaştı?*
>
> *Kaç kullanıcı yalnız sınırlı ve düzeltilebilir sorun gördü?*
>
> *Kaçı Major veya Critical hata gördü?*
>
> *Kaç gözlem çözülmedi, cevap üretmedi veya ölçüm kusuru nedeniyle değerlendirilemedi?*
>
> *Atomik doğruluk, zorunlu cevap unsurları, atıf bütünlüğü, dil adaleti, doğal kullanıcı deneyimi ve dalgalar arası kararlılık tek sistem içinde nasıl raporlanacaktır?*
>
> *On farklı AI ürününün sonuçları tek bir ekosistem görünümüne dönüştürülebilir mi?*
>
> *Her AI ürününün katsayısı neye dayanmalıdır?*
>
> *Bir ürünün büyük kullanıcı nüfusu, küçük bir ürünün yüksek skorunu bastırmalı mıdır?*
>
> *Bir modeldeki Critical hata başka modellerin yüksek puanıyla telafi edilebilir mi?*

Bu bölümün en önemli kararı şudur:

> *NOMOS önce dağılımı, sonra bileşenleri, en son bileşik puanı yayımlar.*

Tek bir sayı ilk sonuç olmayacaktır. Çünkü: 930 puan alan iki sistem farklı risk dağılımları taşıyabilir. Birinde yalnız Moderate sorunlar bulunabilir. Diğerinde tek bir Critical hata yüksek ortalamanın altında gizlenmiş olabilir. Biri büyük dillerde güçlü, küçük dillerde zayıf olabilir. Diğeri bütün dillerde daha dengeli fakat biraz daha düşük ortalama taşıyabilir. Biri temiz oturumda güçlü, doğal kullanıcı hesaplarında zayıf olabilir. Diğeri daha istikrarlı ve kullanıcı durumlarına karşı daha dayanıklı olabilir. Daha önceki mimari, her denetlenebilir sonucun:

- AI ürünü veya model,
- tarih,
- ülke,
- dil,
- sorgu kümesi,

- tekrar sayısı,
- ölçüm kaydı

taşımasını istemişti. Aynı kaynak, sonuçların yalnız bir rozet veya sayı değil; Critical, Major, Moderate ve Advisory bulguların açık dağılımıyla kanıta dayalı bir denetim kaydı olması gerektiğini kurmuştu. Bu bölüm o mimariyi sayısal hâle getirir. Fakat sayısallaştırma, kanıtın yerine geçemez. NOMOS'un temel talebi değişmez: Kanıt verin. Sınır verin. Bağlam verin. Zaman verin. Daha çok anılmaya değil, daha doğru temsil edilmeye çalışın. Bu bölüm:

- GEO-1000 kullanıcı eşdeğeri dağılımını,
- yanıt statülerinin nüfus ağırlıklı sonuçlara dönüşmesini,
- atomik claim metriklerini,
- required-element completeness ölçümünü,
- unresolved ve not-ratable kayıtlarını,
- on bileşenli NOMOS profilini,
- 1.000 puanlık aday bileşik endeksi,
- Native Reach ve Common Support skorlarını,
- AI ürün katsayılarını,
- istem ailesi katsayılarını,
- temiz ve doğal panel sonuçlarını,
- ülke ve dil adaleti puanlarını,
- dalgalar arası kararlılık ve tekrar puanını,
- örneklem ve referans belirsizliğini,
- güven aralıklarını,
- nadir Critical hata üst sınırlarını,
- NOMOS 950+ aday sınıfını,
- skor sürümleme ve kamu sonuç kartını

tanımlar. Bu bölüm henüz:

- Apple.com sentetik sınama'ının tam uygulamasını,
- 30.000 yanıtlık uçtan uca sentetik denetimi,
- uygunluk işaretinin hukukî ve yönetişim kararını,
- bağımsız üniversite tekrar çağrısını

uygulamaz. Bunlar sonraki bölümlerin konusudur. Bölüm 16'nın temel sorusu şudur:

> *GEO-1000'in bütün kanıt zincirini tek sayının altında ezmeden, insanlara anlaşılır; araştırmacılara yeniden üretilebilir; AI sistemlerine makine tarafından okunabilir bir skor mimarisine nasıl dönüştürürüz?*

NOMOS Challenge

İki AI ürünü düşünün.

Ürün A

1.000 kullanıcı eşdeğeri içinde:

- 920 kişi tam veya Advisory geçiş,
- 65 kişi Conditional Pass,
- 10 kişi Major Fail,

- 0 kişi Critical Fail,
- 5 kişi Unresolved

görmektedir.

Ürün B

1.000 kullanıcı eşdeğeri içinde:

- 950 kişi tam veya Advisory geçiş,
- 40 kişi Conditional Pass,
- 0 kişi Major Fail,
- 1 kişi Critical Fail,
- 9 kişi Unresolved

görmektedir. Yalnız acceptable pass oranına bakarsanız:

- Ürün A: yüzde 98,5
- Ürün B: yüzde 99

olabilir. Ürün B daha yüksek görünür. Fakat Ürün B'de kullanıcıyı sahte hukuki yetkiye yönlendiren bir Critical olay bulunmaktadır. Bu olay:

- tekil olabilir,
- düşük yaygınlığa sahip olabilir,

bütün ürün kullanıcılarını temsil etmeyebilir. Yine de:

> *yüksek ortalama içinde kaybolamaz.*

Şimdi üçüncü bir ürün düşünün.

Ürün C

İngilizce: 98 Almanca: 96 Türkçe: 94 Düşük kaynaklı Dil L4: 62 Düşük kaynaklı Dil L5: 58 Nüfus ağırlıklı küresel skor: 93 olabilir. Fakat dil tabanı: 58 ve sağlam dil farkı: 36 puan olabilir. Dördüncü ürünün bütün dillerde sonucu:

- 88
- 87
- 86
- 84
- 83

olsun. Nüfus ortalaması daha düşük olabilir. Fakat dil adaleti daha güçlüdür. Hangisi daha iyi GEO sistemidir? Tek cevap yoktur. Şunlar ayrı sorulardır: Ortalama insan ne görüyor? En düşük dil grubu ne görüyor? Critical olay var mı? Sistem kullanıcılar arasında ne kadar değişiyor? Temiz ve doğal hesap arasındaki fark nedir? Sonuç farklı dalgalarda korunuyor mu? Şimdi on AI ürününü tek ekosistem puanına çevirdiğinizi düşünün. En büyük ürün: kullanıcıların yüzde 60'ına ulaşıyor, skoru 80. Küçük ürün: kullanıcıların yüzde 1'ine ulaşıyor, skoru 98. Ürünleri eşit ağırlıklandırırsanız: ortalama ürün kalitesini ölçersiniz. Gerçek kullanıcı erişimine göre ağırlıklandırırsanız: ortalama insan maruziyetini ölçersiniz. İki sonuç farklıdır. Birini diğerinin yerine kullanamazsınız.

Şimdi yalnız en yüksek skoru veren dört AI ürününü seçtiğinizi düşünün. Kalan altı ürünü: "Yeterli kullanıcı verisi yok." diyerek çıkarıyorsunuz. Bu, ekosistem skoru değildir. Sonuç sonrası ürün seçimidir. Şimdi model katsayılarını:

- marka itibarı,
- şirketin ticari tercihi,

- çıkan puanın yüksekliği

üzerinden belirliyorsunuz. Bu da ölçüm değildir. Skor tasarımıdır. Bu bölümün ilk hükmü şudur:

> *Dağılım, bileşik puandan önce gelir.*

İkinci hükmü şudur:

> *Gate statüsü, sayısal skordan önce gösterilir.*

Üçüncü hükmü şudur:

> *Bir AI ürününün katsayısı kendi skoruna bağlı olamaz.*

Dördüncü hükmü şudur:

> *Eşit ürün ortalaması ile kullanıcı maruziyeti ağırlıklı ortalama ayrı sonuçlardır.*

Beşinci hükmü şudur:

> *Unresolved gözlemler sessizce olumlu veya olumsuz paydaya eklenemez.*

Altıncı hükmü şudur:

> *Sıfır gözlenen Critical olay, sıfır Critical risk anlamına gelmez.*

Yedinci hükmü şudur:

> *NOMOS 950+, yalnız 950 sayısına ulaşmak değildir; gate, taban, belirsizlik, adalet ve tekrar şartlarını birlikte karşılamaktır.*

1. BÖLÜMÜN AMACI

Bu bölümün amacı, GEO-1000 denetimindeki bütün iddia, yanıt, kullanıcı, ülke, dil, panel, AI ürünü ve dalga sonuçlarını birbirini gizlemeyen katmanlı bir skor sistemine dönüştürmektir. Bölüm şu ayrımları normatif hâle getirir:

- Dağılım ile tek sayı
- Yanıt statüsü ile claim support oranı
- Strict Pass ile Conditional Pass
- Major Fail ile Critical Fail
- Unresolved ile Fail
- Not Ratable ile ürün başarısızlığı
- Ham kullanıcı sayısı ile ağırlıklı kullanıcı eşdeğeri
- Population Panel ile Sentinel Panel
- Ortalama ile taban
- Performans ile adalet
- Kalite ile kararlılık
- Temiz skor ile doğal kullanıcı skoru
- Native Reach ile Common Support
- Ürün puanı ile ekosistem puanı
- Eşit ürün ağırlığı ile maruziyet ağırlığı
- AI ürün katsayısı ile ölçüm güveni
- İstem ailesi ağırlığı ile istem sonucu

- Güncel dalga ile zaman serisi
- Örneklem belirsizliği ile referans belirsizliği
- Güven aralığı ile tahmin aralığı
- Sıfır gözlenen hata ile sıfır risk
- Ham örneklem ile etkin örneklem
- Nokta tahmini ile alt–üst doğruluk sınırı
- Bileşen puanı ile bileşik NOMOS puanı
- 950 puan ile NOMOS 950+ adaylığı
- Sayısal skor ile uygunluk kararı
- Güncel status ile tarihsel olay
- Kısmi skor ile tam karşılaştırılabilir skor
- Skor hesaplama ile skor yayını
- Skor sürümü ile yeni gerçeklik sürümü
- Skor yüksekliği ile denetim kalitesi

Bu bölümün sonunda her kamu sonucu şu sorulara açık cevap verebilmelidir:

> *1.000 kullanıcı eşdeğerinin kaçı hangi yanıt statüsünü gördü?*
>
> *Bileşen skorları nelerdir?*
>
> *Critical ve Major gate durumu nedir?*
>
> *Skor hangi ürün, nüfus, istem, panel ve dalga için geçerlidir?*
>
> *AI ürünlerinin katsayıları hangi dış veriye dayanır?*
>
> *Belirsizlik ve etkin örneklem nedir?*
>
> *Skorun hangi bölümü ölçülmüş, hangi bölümü unresolved kalmıştır?*
>
> *950+ sonucu yalnız sayı mı, yoksa bütün normatif şartları taşıyan aday sınıf mı?*

2. MERKEZÎ NORMATİF HÜKÜM

NOMOS, hiçbir GEO sonucunu tek çıplak puan olarak yayımlamamalıdır. Her sonuç önce kullanıcı eşdeğeri yanıt dağılımını, sonra Critical ve Major gate statülerini, sonra on bileşenli profilini, belirsizlik ve kapsam kayıtlarını; en son varsa 1.000 puanlık bileşik endeksi göstermelidir. Kamu sıralaması şu sırayı izlemelidir:

- Gate statüsü
- GEO-1000 yanıt dağılımı
- Critical, Major, Unresolved ve No-Yanıt oranları
- On bileşenli NOMOS profili
- Bileşik NOMOS endeksi
- Güven aralığı ve etkin örneklem
- Kapsam, sürüm ve karşılaştırılabilirlik bilgisi

Şu gösterim yasaktır:

“NOMOS skoru: 963.”

Doğru gösterim şuna benzemelidir: GÜNCEL STATUS: CONDITIONAL — 963/1000 / Strict Pass: 952/1000 kullanıcı eşdeğeri / Observed Confirmed Critical: 0 / Observed Confirmed Major: 0 / Unresolved: 8/1000 / 95% sampling interval: 956–969 / Critical-rate one-sided upper bound: belirtilen yöntemle / Kapsam: AI ürünü, ülke, dil, panel, istem ve dalga sürümleri

3. DAĞILIM ÖNCE, PUAN SONRA

NOMOS’un birincil sonuç nesnesi:

GEO-1000 Yanıt Distribution

Türkçesi:

GEO-1000 Yanıt Dağılımı

olacaktır. Bu dağılım, hedef nüfustaki her 1.000 kullanıcı eşdeğerinden kaçının:

- Full Pass,
- Pass With Advisory,
- Conditional Pass,
- Major Fail,
- Critical Fail,
- Unresolved,
- No Usable Yanıt,
- Not Ratable

durumuna karşılık geldiğini gösterir. Bu yapı doğrudan kurucu fikri korur:

> *1.000 kullanıcıdan kaçı gerçek varlığı doğru biçimde gördü?*

Fakat yalnız doğru–yanlış ikiliğine düşmez. Yanlışlığın türünü ve ölçüm yetersizliğini de görünür bırakır.

4. GÖZLEM AĞIRLIĞI

Ana analiz örneğindeki gözlem i’nin nihai nüfus ağırlığı: wi olsun. Normalize edilmiş ağırlık:

```
w_i ← w_i / Σ_{j=1}^n w_j
```

şeklindedir. Böylece:

```
Σ_{i=1}^n w_i = 1
```

olur. Bu ağırlık:

- tasarım,
- fazla örnekleme,
- yanıtlamama,
- nüfus kalibrasyonu

bileşenlerinden gelebilir. AI cevabının:

- doğru,
- yanlış,
- markaya yararlı,

- Critical

olması ağırlığı etkileyemez.

5. 1.000 KULLANICI EŞDEĞERİ

Yanıt statüsü k için kullanıcı eşdeğeri:

```
D_k = 1.000 × Σ_i w_i1(RP_i=k)
```

olarak hesaplanır. Burada k:

RP-1

RP-2

RP-3

RP-4

RP-5

RP-6

RP-7

RP-8

statülerinden biridir. Toplam:

```
Σ_{k=1}^8 D_k = 1.000
```

olmalıdır. Yuvarlama nedeniyle kamu kartında bir–iki kullanıcı eşdeğeri fark oluşursa en büyük kalan yöntemi veya önceden tanımlanmış eşdeğer kural kullanılmalıdır.

6. GEO-1000 YANIT DAĞILIMI

Kamu kartı en az şu sekiz alanı taşımalıdır:

Kod	Statü	1.000 kullanıcı eşdeğeri
RP-1	Full Pass	D1
RP-2	Pass With Advisory	D2
RP-3	Conditional Pass	D3
RP-4	Fail — Major	D4
RP-5	Automatic Fail — Critical	D5
RP-6	Unresolved	D6
RP-7	No Usable Yanıt	D7
RP-8	Not Ratable	D8
	Toplam	1.000

Bu tablo bileşik skordan daha temel sonuçtur.

7. STRICT PASS VE ACCEPTABLE PASS

7.1. Strict Pass Kullanıcı Eşdeğeri

Yalnız:

RP-1

RP-2

yanıtlarını içerir.

SP_1000 = D_1 + D_2

Bu sonuç:

Maddi Moderate, Major veya Critical bulgu görmeyen kullanıcı eşdeğerini

gösterir.

7.2. Acceptable Pass Kullanıcı Eşdeğeri

RP-3 Conditional Pass'i de içerir.

AP_1000 = D_1 + D_2 + D_3

Bu sonuç:

> *Critical veya Major bozulma görmeyen fakat bazı kullanıcıların sınırlı Moderate sorun görebildiği toplam eşdeğeri*

gösterir.

7.3. Fail Kullanıcı Eşdeğeri

FP_1000 = D_4 + D_5 + D_7

olarak gösterilebilir. Unresolved ve Not Ratable bu paydaya zorla eklenmez. Ayrı kalır.

8. CRITICAL VE MAJOR YOĞUNLUĞU

CR_1000 = D_5; MR_1000 = D_4

olarak gösterilir. Bu değerler: gate statüsünün yerine geçmez, yaygınlığı ifade eder. Tek bir Confirmed Critical olay: CR1000 değerini düşük tutabilir, yanıt ve conformity gate'ini yine açabilir.

9. UNRESOLVED VE NOT-RATABLE YOĞUNLUĞU

UR_1000 = D_6; NR_1000 = D_8

şeklindedir. Unresolved: gerçeklik veya hakemlik kararı çözülememiş kullanıcı eşdeğeri demektir. Not Ratable: ölçüm zinciri yetersiz olduğu için ürün performansına karar verilemeyen kullanıcı eşdeğeri demektir. İki statü:

- pass,
- fail

gibi kullanılamaz.

10. RATABLE KAPSAMA

Değerlendirilebilir nüfus kapsamı:

RC = 1 − D_8/1.000

olarak gösterilebilir. Örnek: Not Ratable: 40/1.000 ise:

RC=96%

olur. Yüksek performans skoru düşük ratable kapsama ile birlikteyse: güvenli ve tam kapsamlı sonuç olarak sunulamaz.

11. ATOMİK CLAIM METRİKLERİ

Yanıt dağılımına ek olarak atomik düzeyde şu metrikler yayımlanmalıdır:

- Varlık Accuracy
- Factual Support
- Kapsam Accuracy
- Temporal Accuracy
- Attribution Integrity
- Modality Integrity
- Atıf Support
- Required Element Completeness
- Material Omission Rate
- Reference Gap Rate
- Claim Support Rate

Bunlar yanıt geçişinin yerine geçmez. Nedenini açıklar.

12. ATOMİK MERKEZİLİK AĞIRLIKLARI

CANDIDATE WEIGHTS — ADAY AĞIRLIKLAR

Bölüm 14'teki claim centrality sınıfları için aday ağırlıklar:

Centrality	Aday ağırlık
CL-1 — Core Claim	4
CL-2 — Supporting Claim	2
CL-3 — Qualifying Claim	2
CL-4 — Incidental Claim	1
CL-5 — Non-Material	0

Bu ağırlıklar: nihai bilimsel gerçek değildir, pilot, dış inceleme ve duyarlılık analiziyle kalibre edilmelidir. Amaç:

- bir core lisans iddiasıyla,
- incidental kuruluş ayrıntısını

aynı etkiyle saymamaktır.

13. KÖK İDDİA NORMALİZASYONU

Aynı semantik kökten gelen tekrar iddialar ayrı ağırlıklarla skoru şişirmemeli veya aşırı düşürmemelidir. Kök küme g içindeki iddiaların centrality ağırlığı: cj olsun. Normalize edilmiş claim ağırlığı aday olarak:

```
ω_j = [c_j / Σ_{h∈g} c_h] × max_{h∈g}(c_h)
```

şeklinde tanımlanabilir. Böylece:

```
Σ_{j∈g} ω_j = max_{j∈g}(c_j)
```

olur. Bu yöntem: aynı yanlışın üç kez tekrarlanmasını üç bağımsız yanlış yapmaz, tekrarın tavsiye veya güven etkisini ayrıca amplification alanında korur.

14. ATOMİK BOYUT DEĞERLERİ

CANDIDATE NUMERIC MAPPING — ADAY SAYISAL EŞLEME

Aşağıdaki değerler pilot ve bağımsız inceleme öncesi adaydır.

Varlık

Karar	Değer
E-1 Correct Varlık	1,00
E-2 Acceptable Brand-Level Varlık	0,90
E-3 Ambiguous Varlık	Unresolved
E-4 / E-5 / E-6	0,00

Factual Support

Karar	Değer
F-1 Supported	1,00
F-2 Supported With Required Qualification	0,85
F-3 Partially Supported	0,50
F-4 Official Claim Accurately Attributed	1,00*
F-5 Unsupported	0,00
F-6 Contradicted	0,00
F-7 Reference Gap	Unresolved
F-8 Not Verifiable	N/A

* Normalize iddia gerçekten: “Şirket bu iddiayı yayımlamaktadır.” şeklindeyse 1,00’dır. Alttaki öz beyanın gerçek olduğu anlamına gelmez.

Kapsam

Karar	Değer
S-1 Exact Kapsam	1,00
S-2 Narrower Than Evidence	1,00

Karar	Değer
S-3 Partially Overbroad	0,50
S-4 / S-5 / S-6	0,00
S-7 Kapsam Unknown	Unresolved

Time

Karar	Değer
T-1 Güncel	1,00
T-2 Historical and Properly Framed	1,00
T-3 / T-4 / T-5 / T-7	0,00
T-6 Unknown	Unresolved

Attribution

Karar	Değer
A-1 Status Preserved	1,00
A-2 Non-Material Attribution Omission	0,90
A-3 Material Attribution Loss	0,40
A-4 / A-5 / A-6 / A-7	0,00

Modality

Karar	Değer
M-1 Appropriate Certainty	1,00
M-2 Appropriate Uncertainty	1,00
M-4 / M-7	0,70
M-3 / M-6	0,25
M-5 False Absolute	0,00

Atıf

Karar	Değer
C-1 Directly Supported	1,00
C-2 Partially Supported	0,60
C-3 Supports Attribution Only	0,40
C-4 / C-5 / C-6 / C-7 / C-8	0,00
C-0 Not Applicable	N/A

Bu değerler:

- hakem kararını değiştirmez,
- yalnız boyutsal metrik üretir,

Critical ve Major gate'leri aşamaz.

15. ATOMİK FIDELITY DEĞERİ

Bir iddia için birleşik aday fidelity:

```
q_j = e_j × f_j × s_j × t_j × a_j × m_j
```

olarak hesaplanabilir. Bu çarpımsal yapı şu ilkeyi korur:

> *Yanlış varlıkta doğru sayı, doğru iddia değildir.*
>
> *Doğru bilgi, yanlış zamanda doğru değildir.*
>
> *Öz beyanı bağımsız gerçeklik gibi sunmak aynı iddia değildir.*

Atomic Fidelity:

```
AF = 100 × (Σ_j ω_jq_j)/(Σ_j ω_j)
```

olarak hesaplanabilir. Unresolved atomlar bu nokta tahmininden çıkarılır. Fakat unresolved kütle ayrı gösterilir ve alt–üst sınır hesabına katılır.

16. REQUIRED ELEMENT COMPLETENESS

İstem ailesindeki zorunlu unsur h'nin:

```
karşılanması: zh=1
kısmen karşılanması: zh=0,5
karşılanmaması: zh=0
```

olarak kaydedilebilir. Aday unsur ağırlıkları:

- Core required element: 4
- Necessary boundary: 3
- Supporting required element: 1
- Optional: 0

Required Element Completeness:

```
REC = 100 × (Σ_h v_hz_h)/(Σ_h v_h)
```

şeklindedir. Core gate başarısızlığı: REC yüksek olsa bile yanıt pass'i engelleyebilir.

17. ON BİLEŞENLİ NOMOS PROFİLİ

NOMOS bileşik endeksi aşağıdaki on bileşenden oluşur:

- Varlık Integrity — ENT
- Factual Fidelity — FAC
- Kapsam & Boundary Integrity — SBI
- Temporal & Local Accuracy — TLA
- Evidence & Atıf Integrity — ECI
- Epistemic Integrity — EPI
- Task Completion & Relevance — TCR
- Language & Geographic Fairness — LGF
- User-State & Surface Integrity — USI
- Stability & Replicability — STR

Her bileşen:

```
0≤Sk≤100
```

arasında hesaplanır. Profil:

N=(ENT,FAC,SBI,TLA,ECI,EPI,TCR,LGF,USI,STR)

şeklindedir. Bileşik puandan önce bu profil yayımlanmalıdır.

18. VARLIK INTEGRITY — ENT

Varlık Integrity şu alanları ölçer:

- Doğru ana varlık
- Marka–hukuki kişi ilişkisi
- Ana şirket–bağlı kuruluş ayrımı
- Ürün–üretici ayrımı
- Franchise, partner ve distribütör ayrımı
- Varlık disambiguation
- Required varlık element completion

Aday form:

```
ENT = 100 × [Σ_j ω_je_j + Σ_h v_hz_h^entity] / [Σ_j ω_j + Σ_h v_h]
```

Yanlış varlık Critical veya Major gate oluşturmuşsa yüksek kalan varlık ortalaması full pass sağlamaz.

19. FACTUAL FIDELITY — FAC

Factual Fidelity:

- desteklenmiş,
- kısmen desteklenmiş,
- unsupported,
- contradicted

iddiaların nüfus ve centrality ağırlıklı dağılımını ölçer.

```
FAC = 100 × (Σ_j ω_jf_j)/(Σ_j ω_j)
```

Reference gap: paydaya sessizce sıfır veya bir olarak girmez, unresolved oranına gider.

20. KAPSAM & BOUNDARY INTEGRITY — SBI

Bu bileşen iki parçadan oluşur:

- Claim kapsam doğruluğu
- Zorunlu sınır ve dışlamaların completeness'ı

Aday form:

```
SBI=0,60×ScopeScore+0,40×BoundaryCompleteness
```

Burada:

```
ScopeScore = 100 × (Σ_j ω_js_j)/(Σ_j ω_j)
```

BoundaryCompleteness, ilgili required boundary unsurlarından hesaplanır. Critical boundary omission varsa SBI sayısı gate'i telafi edemez.

21. TEMPORAL & LOCAL ACCURACY — TLA

Bu bileşen:

- güncel/tarihsel ayrımı,
- fiyat ve lisans geçerliliği,
- ülke ve yargı alanı,
- yerel hizmet erişimi,
- locale ve ürün kapsamı

alanlarını değerlendirir. Aday form:

```
TLA=0,50×TemporalScore+0,50×LocalJurisdictionScore
```

Doğru tarihsel bilgi güncel istem altında kullanılmışsa puan kaybeder.

22. EVIDENCE & ATIF INTEGRITY — ECI

ECI:

- atıf özgünlüğü,
- atıf–claim eşleşmesi,
- doğru varlık ve zaman,
- kaynak erişimi,
- kanıt statüsü,
- kaynak soyu

alanlarını değerlendirir. Aday form:

```
ECI=0,60×CitationSupport+0,40×EvidenceTraceability
```

Atıf gerekmeyen istemde atıf yokluğu puan düşürmez. Ancak full NOMOS denetimi için ayrı Evidence İstem ailesi bulunmalıdır. Sahte maddi atıf Critical gate oluşturuyorsa ECI bileşik puanı ayrı kalır; gate önce gösterilir.

23. EPISTEMIC INTEGRITY — EPI

EPI şu ayrımları ölçer:

- Öz beyan ile doğrulanmış olgu
- Görüş ile gerçeklik
- İddia ile kesinleşmiş karar
- Kısıtlı kanıt ile kamu keşfedilebilirliği
- Uygun certainty ile overconfidence
- Uygun uncertainty ile gereksiz belirsizlik

Aday form:

```
EPI=0,50×AttributionScore+0,50×ModalityScore
```

Bu bileşen doğru sayının yanlış kanıt statüsüyle sunulmasını görünür kılar.

24. TASK COMPLETION & RELEVANCE — TCR

TCR:

- Required Yanıt Elements
- Ana soruya doğrudan cevap
- Refusal uygunluğu
- No Usable Yanıt
- İlgisiz claim yükü
- Evasive davranış

alanlarını ölçer. Aday form:

```
TCR=0,60×REC+0,40×ResponseRelevance
```

Yanlış iddia üretmeyen fakat soruya cevap vermeyen sistem:

- yüksek factual puan,
- düşük TCR

taşıyabilir. Yanıt statüsü yine RP-7 olabilir.

25. LANGUAGE & GEOGRAPHIC FAIRNESS — LGF

LGF yalnız ortalama sonucu ölçmez. Şunları birlikte değerlendirir:

- Düşük performanslı dil tabanı
- Dil grupları arası sağlam fark
- Ülke veya bölge tabanı
- Coğrafi performans farkı
- Nüfus ağırlıklı gözlem kapsamı

Yeterli örneklem taşıyan dil sonuçları: Ll olsun. Sağlam dil tabanı:

```
LF_10 = Q_{0,10}(L_l)
```

Dil farkı:

```
DG_{90-10} = Q_{0,90}(L_l) - Q_{0,10}(L_l)
```

Aday Dil Adaleti skoru:

$$LFS = 0{,}60LF_10 + 0{,}40(100 - DG_\{90-10\})$$

Benzer biçimde ülke veya yeterli coğrafi tabaka için:

$$CFS = 0{,}60CF_10 + 0{,}40(100 - CG_\{90-10\})$$

Nüfus ağırlıklı ülke ve dil gözlem kapsamı: COCpop ve: LOCpop olsun. Kapsama Factor:

```
K = COC_pop × LOC_pop
```

olarak tanımlanabilir. Burada oranlar 0–1 aralığındadır. Aday LGF:

$$LGF = K \times \sqrt{(LFS \times CFS)}$$

Bu 0.9.0 aday revizyonu, dil ve coğrafya puanlarının geometrik ortalamasını kapsam katsayısıyla çarpar; böylece LGF 0–100 aralığında kalır ve iki boyuttan birindeki ağır düşüş öteki boyutun yüksekliğiyle tamamen telafi edilemez. Country Sentinel tekil gözlemleri ülke skoruna dönüştürülmez. Yalnız yeterli örneklem taşıyan hücreler fairness tahminine katılır.

26. USER-STATE & SURFACE INTEGRITY — USI

USI:

- Kontrollü Temiz Panel
- Doğal Kullanıcı Paneli
- Ücretsiz/ücretli plan
- Web/mobil yüzey
- Hafıza ve özel talimat
- Ürün varyantı

arasında maddi gerçekliğin korunup korunmadığını ölçer. Kontrollü core sonuç:

SC

Doğal core sonuç:

SN

olsun. Panel tabanı:

```
FC_N = min(S_C, S_N)
```

Panel farkı:

GC_N = |S_C − S_N|

Plan ve yüzeylerdeki sağlam fark: Gsurf olsun. Aday USI:

```
USI = 0,50FC_N + 0,25(100-GC_N) + 0,25(100-G_surface)
```

Bu form:

- her iki panelde yüksek kaliteyi,
- küçük maddi farkı,
- yüzey dayanıklılığını

birlikte ödüllendirir. Yalnız bir panel varsa USI hesaplanamaz. Eksik ağırlık başka bileşenlere dağıtılamaz.

27. STABILITY & REPLICABILITY — STR

Dalgalar arasındaki ana temsil değerleri: Sw olsun. Dalga tabanı:

```
WF = min_w(S_w)
```

Dalga aralığı:

```
WR = max_w(S_w) - min_w(S_w)
```

Planlanan tekrar dalgalarının tamamlanma oranı:

```
RCOMP = 100 × geçerli tamamlanan tekrar dalgası / planlanan tekrar dalgası
```

olsun. Aday STR:

STR = 0,50WF + 0,30(100−WR) + 0,20RCOMP

Daha fazla dalgada:

- minimum yerine alt yüzdelik,
- range yerine robust varyasyon

kullanılabilir. Tek dalga: ürünün anlık skorunu üretebilir, STR için tam kanıt sağlamaz.

28. NOMOS BİLEŞİK ENDEKSİ

On bileşenin her biri 0–100 aralığında olduğundan aday bileşik endeks:

NOMOS_1000 = ENT + FAC + SBI + TLA + ECI + EPI + TCR + LGF + USI + STR

şeklindedir. Dolayısıyla:

0 ≤ NOMOS_1000 ≤ 1.000

olur. Bu endeks:

- yanıt gate'lerinin yerine geçmez,
- dağılımın yerine geçmez,

Critical veya Major finding'i telafi etmez.

29. NEDEN ON EŞİT BİLEŞEN?

On bileşene eşit yüz puan verilmesinin aday gerekçesi şudur: Skor okunabilir kalır. Gizli ağırlık sistemi oluşmaz. Her bileşen kamuya ayrı görünür. Bir alanın önemi esas olarak Critical/Major gate'lerle korunur. Domain-specific alternatif ağırlıklar temel endeksi değiştirmeden ek rapor üretir. Bununla birlikte eşit ağırlık: nihai bilimsel gerçek değildir. Pilotlar, kullanıcı araştırmaları ve bağımsız inceleme farklı ağırlıklar önerebilir.

30. DOMAIN-SPECIFIC ALTERNATİF ENDEKS

Sağlık, hukuk veya finans gibi yüksek riskli alanlar için bileşen katsayıları: λk kullanılabilir.

```
Σ_{k=1}^{10} λ_k = 1
```

Aday domain endeksi:

```
NOMOS_1000^domain = 1.000 × Σ_{k=1}^{10} λ_k(S_k/100)
```

olarak hesaplanabilir. Bu skor:

NOMOS-LEGAL

NOMOS-HEALTH

NOMOS-FINANCE

gibi ayrı ad taşımalıdır. Kanonik eşit bileşenli skorun yerine sessizce kullanılamaz.

31. GATE STATÜSÜ SKORDAN ÖNCE GELİR

Kamu gösterimi şu sırayı korumalıdır:

Örnek A

CRITICAL HOLD — 971/1000

Bu sistem:

- yüksek bileşen skoru,
- açık Confirmed Critical olay

taşıyabilir. 971 sayısı: full conformity anlamına gelmez.

Örnek B

MAJOR FAIL — 954/1000

Ana varlık veya kapsamda tek Major finding bulunabilir.

Örnek C

CONDITIONAL — 961/1000

Critical ve Major yoktur. Sınırlı Moderate bulgular veya unresolved kütle vardır.

Örnek D

FULL CONFORMITY CANDIDATE — 963/1000

Gate, taban, belirsizlik ve kalite ön koşulları ayrıca karşılanmıştır.

32. NOMOS 950+ ADAY SINIFI

CANDIDATE 950+ STANDARD — ADAY 950+ STANDARDI

Bir sonucun yalnız bileşik puan bakımından:

NOMOS_1000 ≥ 950

olması yeterli değildir. NOMOS 950+ adayı için birlikte aranabilecek koşullar:

- Bileşik nokta tahmini en az 950
- Strict Pass kullanıcı eşdeğeri en az 950/1.000
- Açık Confirmed Critical bulunmaması
- Açık Confirmed Major bulunmaması
- Bütün core bileşenlerin en az 90 olması
- LGF, USI ve STR bileşenlerinin en az 85 olması
- Core unresolved kütlenin en fazla yüzde 2 olması
- Not Ratable kullanıcı eşdeğerinin kabul edilen sınırın altında olması
- En az NCL-3, TPR-3, AQ-3 kalite düzeylerinin karşılanması
- En az bir confirmatory veya stability dalgasının bulunması
- Native Reach ve Common Support kapsamlarının ayrı raporlanması
- Sıfır gözlenen Critical olayın istatistiksel üst sınırının açıklanması

Bu eşikler:

- pilot,
- Apple.com sentetik sınamayı,

- 30.000 yanıtlık uçtan uca çalışma,
- bağımsız üniversite incelemesi

sonrasında kalibre edilmelidir.

33. ADAY PERFORMANS BANTLARI

Bileşik sayı için geçici ve yalnız betimleyici bantlar:

Skor	Betimleyici sınıf
950–1.000	NOMOS 950+ performans adayı
900–949	Yüksek fakat tam standardı karşılamayan temsil
800–899	Maddi geliştirme gerektiren temsil
650–799	Zayıf ve düzensiz temsil
0–649	Ağır temsil açığı

Bu bantlar:

- gate statüsünün,
- Critical olayın,
- component floor'un,
- belirsizlik ve kapsamın

yerine geçemez. CRITICAL HOLD — 975 sonucu 950+ adayı değildir.

34. SKOR KİMLİĞİ

Her NOMOS skoru şu alanları kimliğinde taşımalıdır:

```
ScoreID=(Entity,AIProduct,Frame,Panel,PromptSet,Wave,TruthPack,ScoreVersion)
```

Örnek:

NOMOS-ASTERON-ORION-NR-CLEAN-W1-CORESET1-TP0.9-SV0.9

Şu cümle yeterli değildir: "Asteron'un NOMOS skoru 962." Hangi:

- AI ürünü,
- nüfus,
- dil,
- panel,
- istem,
- tarih

olduğu bilinmemektedir.

35. CORE GEO-1000 VE FULL NOMOS AYRIMI

35.1. Core GEO-1000

Yalnız Core Mirror İstem'un nüfus dağılımını ölçer. Ana sonuç:

SP1000

AP1000

CR1000

MR1000

UR1000

olabilir. Şu soruya cevap verir: "AI ürünü varlığın temel kimlik ve faaliyetini kaç kullanıcıya doğru yansıtıyor?"

35.2. Full NOMOS

Birden fazla istem ailesi, panel, dil, kanıt ve kararlılık bileşenini kapsar. On bileşenli profil ve 1.000 puanlık endeksi üretir. Core GEO-1000 sonucu: Full NOMOS skoru gibi sunulamaz.

36. NATIVE REACH SKORU

Her AI ürünü kendi doğal ürün-uygun nüfusunda değerlendirilir. NOMOSaNR ürün a'nın Native Reach skorudur. Şu soruya cevap verir:

> *Bu ürünün gerçekten ulaşabildiği kullanıcı nüfusunda temsil ne durumdadır?*

Farklı ürünlerin Native Reach nüfusları farklı olabilir. Doğrudan üstünlük sıralaması için tek başına yeterli değildir.

37. COMMON SUPPORT SKORU

Ortak kullanıcı evreninde: NOMOSaCS hesaplanır. Şu soruya cevap verir:

> *Aynı ülke, dil ve kullanıcı koşulunda ürünler karşılaştırıldığında hangisi nasıl temsil üretmektedir?*

Common Support: karşılaştırılabilirliği güçlendirir, ürünün ortak evren dışındaki gerçek erişimini göstermez. Native Reach ve Common Support ayrı yayımlanmalıdır.

38. AI ÜRÜN KATSAYISI

On AI ürününün sonuçları ekosistem düzeyinde birleştirilecekse ürün katsayısı: αa kullanılabilir. Koşul:

```
Σ_{a=1}^A α_a = 1
```

Ürün katsayısı:

- ürünün kendi skoru,
- sağlayıcının itibarı,
- denetlenen varlığın tercihi,
- ticari sponsorluk

üzerinden belirlenemez.

39. EŞİT ÜRÜN KATSAYISI

Her AI ürünü eşit ağırlık alır:

```
α_a^EQ = 1/A
```

Eşit ürün ekosistem skoru:

```
NOMOS_ecosystem^EQ = Σ_a α_a^EQ NOMOS_a
```

Şu soruya cevap verir:

> *Ortalama test edilen AI ürünü ne kadar güçlüdür?*

Ortalama insan maruziyetini göstermez.

40. MARUZİYET AĞIRLIKLI KATSAYI

AI ürününün tanımlı dönemdeki doğrulanmış aktif veya uygun kullanıcı maruziyeti: Ma olsun.

```
α_a^EXP = M_a / Σ_b M_b
```

Maruziyet ağırlıklı ekosistem skoru:

```
NOMOS_ecosystem^EXP = Σ_a α_a^EXP NOMOS_a
```

Şu soruya cevap verir:

> *Ortalama insanın maruz kaldığı AI ekosisteminde temsil ne durumdadır?*

41. MARUZİYET VERİSİ BULUNAMIYORSA

Aktif kullanıcı veya sorgu maruziyeti:

- kamuya kapalı,
- tutarsız,
- sağlayıcı öz beyanı,
- hesap sayısı ile aktif kullanıcıyı karıştıran

nitelikte olabilir. Bu durumda:

- eşit ürün sonucu yayımlanabilir,
- maruziyet ağırlıklı sonuç NOT ESTIMABLE kalabilir,

makul katsayı senaryoları üzerinden alt–üst sınır gösterilebilir. Keyfî pazar payı tahmini üretilemez.

42. ÜRÜN KATSAYISI KİLİDİ

Her αa:

- veri kaynağı,
- dönem,
- kullanıcı birimi,
- ülke kapsamı,
- güncelleme tarihi,
- belirsizlik

taşımalıdır. Katsayılar AI skorları görülmeden önce kilitlenmelidir. Düşük skorlu ürünün katsayısı sonuçtan sonra azaltılamaz.

43. EKOSİSTEM SKORU CRITICAL'I TELAFİ EDEMEZ

Bir ürün:

- 980 puan,
- sıfır Critical

taşıyabilir. Başka ürün:

- 900 puan,
- Confirmed Critical

taşıyabilir. Ekosistem ortalaması yüksek olabilir. Kamu kartı yine şunları göstermelidir:

- Ürün bazlı Critical incident
- Critical maruziyet ağırlığı
- En az bir üründe açık Critical bulunup bulunmadığı
- Critical'ın kapsamı

Ecosystem composite: ürün düzeyindeki Critical gate'i silemez.

44. EKOSİSTEM CRITICAL MARUZİYETİ

Ürün a'nın Critical kullanıcı oranı: CRa olsun. Maruziyet ağırlıklı Critical oran:

```
CR_ecosystem^EXP = Σ_a α_a^EXP CR_a
```

olarak hesaplanabilir. Bunun yanında ikili alan:

```
AnyCritical = 1(∃a : CIS_a ≥ 2)
```

ayrıca gösterilmelidir. Düşük maruziyetli ürünün Critical olayı oranı küçük olabilir. AnyCritical kaydı yine olayın varlığını korur.

45. İSTEM AİLESİ KATSAYISI

İstem ailesi f için: βf kullanılabilir. Koşul:

```
Σ_f β_f = 1
```

βf:

- istemin yüksek veya düşük skor üretmesine,
- şirketin tercihine,
- ölçüm sonrası kolaylık kararına

bağlı olamaz. İstem ailesi ağırlıkları şu alanlara dayanabilir:

- Kullanıcı karar önemi
- Risk düzeyi
- İstemin doğal kullanım sıklığı
- Varlık türü
- Standardın zorunlu çekirdeği

46. İSTEM AİLELERİ AYRI GÖRÜNMELİDİR

Bileşik istem skoru hesaplanabilse bile şu sonuçlar ayrı kalmalıdır:

- Core Mirror
- Varlık Resolution
- Activity & Kapsam
- Evidence
- Local
- Temporal
- Tavsiye
- Comparative
- Boundary
- Control
- Holdout

Bir yüksek Core sonucu düşük Boundary veya Tavsiye sonucunu görünmez kılamaz.

47. KONTROLLÜ VE DOĞAL PANEL SKORLARI

Kontrollü ve doğal paneller farklı tahmin nesneleridir. Varsayılan NOMOS kuralı:

> *İki panel ayrı skor kimlikleri taşır ve aritmetik olarak birleştirilmez.*

Örnek:

NOMOS-CLEAN: 962

NOMOS-NATURAL: 918

Bir politika veya araştırma amacı ikisini birleştirecekse:

```
γ_C + γ_N = 1
```

koşuluyla predeclared katsayılar kullanılabilir. Fakat:

- iki ham skor,
- aralarındaki fark,
- kullanılan katsayı

ayrı gösterilmelidir.

48. DALGA SKORU VE ZAMAN SERİSİ

Her Principal Wave kendi puan kimliğine sahiptir. W1 skoru W2 skoru W3 skoru ayrı tutulur. Güncel skor: son geçerli Principal veya Düzeltme Retest Wave olabilir. Geçmiş dalgaların ortalaması güncel sonucu otomatik belirlememelidir.

49. ROLLING PUAN

İstenirse belirli süreli hareketli skor üretilebilir:

```
NOMOS_rolling = Σ_w δ_wNOMOS_w
```

Burada:

```
Σ_w δ_w = 1
```

olur. δw:

- sonuçtan önce,
- zaman ağırlığına göre

belirlenmelidir. Rolling puan:

- güncel puanın,
- historical incident record'un

yerine geçmez.

50. UNRESOLVED İDDİALAR İÇİN ÜÇ SKOR

Unresolved maddi atomlar nokta tahminini yapay biçimde yükseltebilir veya düşürebilir. Bu nedenle üç sonuç önerilir.

50.1. Conservative Lower Bound

Unresolved iddialar sıfır kabul edilir. Slower

50.2. Resolved-Only Point Estimate

Yalnız çözülen iddialar kullanılır. Spoint

50.3. Optimistic Upper Bound

Unresolved iddialar tam destekli kabul edilir. Supper Dolayısıyla:

```
S_lower ≤ S_point ≤ S_upper
```

olmalıdır. Bu aralık:

> *referans veya epistemik belirsizlik aralığıdır.*

Örneklem güven aralığıyla aynı şey değildir.

51. ÖRNEKLEM BELİRSİZLİĞİ

Nüfus tahminleri örneklemden geldiği için:

- yanıt distribution,
- component puan,
- composite puan

güven aralığı taşımalıdır. Karmaşık örneklemde yöntem:

- tabakaları,
- kümeleri,
- ağırlıkları,
- eşlenmiş kullanıcıları,
- dalgaları

korumalıdır. Aday yöntemler:

- Tasarım tabanlı varyans tahmini

- Replicate weights
- Tabakalı bootstrap
- Küme bootstrap'ı
- Uygun nadir olay modeli
- Model-assisted survey estimation

Kullanılan yöntem kamu kaydında belirtilmelidir.

52. BİLEŞİK SKOR GÜVEN ARALIĞI

Bileşik skorun güven aralığı yalnız component aralıklarını toplamakla bulunamaz. Bileşenler:

- aynı kullanıcılar,
- aynı iddialar,
- aynı diller,
- aynı dalgalar

nedeniyle bağımlıdır. Tercih edilen yaklaşım: Örneklem tasarımına uygun tekrar örnekleri oluştur. Her tekrar örneğinde bütün yanıt ve component metriklerini yeniden hesapla. Bileşik skoru yeniden üret. Dağılımdan güven aralığı oluştur. Bu süreç bütün skor mimarisini tekrar çalıştırmalıdır.

53. ETKİN ÖRNEKLEM BÜYÜKLÜĞÜ

Ham gözlem: n ve etkin örneklem:

```
n_eff = (Σ_i w_i)^2 / Σ_i w_i^2 [K08]
```

birlikte yayımlanmalıdır. 1.000 ham gözlem: 1.000 eşit bilgi birimi anlamına gelmeyebilir. Bileşen ve alt grup skorları kendi etkin örneklem büyüklüklerini taşımalıdır.

54. SIFIR GÖZLENEN CRITICAL HATA

Bir örneklemde hiçbir Confirmed Critical olay bulunmaması:

Observed Confirmed Critical = 0

olarak yazılmalıdır. Şu ifade kullanılmamalıdır:

> *Critical riski sıfırdır.*

Bağımsız ve eşit gözlem varsayımında hızlı yaklaşık üst sınır:

```
p_upper,95 ≈ 3/n [K09]
```

olarak bilinen yaklaşımla tahmin edilebilir. Karmaşık GEO-1000 tasarımında:

- ham n yerine tasarıma uygun yöntem,
- gerekirse etkin örneklem,
- exact veya bootstrap tabanlı tek taraflı sınır

kullanılmalıdır. Örnek:

```
neff=1.000
```

ve sıfır olay altında yaklaşık üst sınır:

```
3/1.000 = 0,003 = %0,3
```

olabilir. Doğru kamu ifadesi:

> *1.000 etkin gözlemde sıfır Confirmed Critical olay gözlenmiştir; kullanılan yaklaşık yöntemde yüzde 95 tek taraflı üst sınır yaklaşık yüzde 0,3'tür.*

55. BİR CRITICAL OLAYDA BELİRSİZLİK

Bir Critical olay gözlendiğinde: nokta oranı düşük olabilir, güven aralığı geniş olabilir. Doğru sonuç:

- olayın gerçekleştiğini,
- tahmini yaygınlığı,
- geniş belirsizliği,
- incident state'i

birlikte göstermelidir. Şu iki cümle de yanlış olabilir:

- "Yalnız bir olay, önemsizdir."
- "Bir olay bütün sistemin yüzde 100 başarısız olduğunu kanıtlar."

56. SKOR HASSASİYETİ

Aşağıdaki alternatifler ayrıca hesaplanabilir:

- Ham ağırlıksız skor
- Ana nüfus ağırlıklı skor
- Ağırlık kırpılmamış skor
- Ağırlık kırpılmış skor
- Resolved-only skor
- Conservative skor
- Equal-product skor
- Exposure-weighted skor
- Native Reach skor
- Common Support skor
- Clean skor
- Natural skor

Maddi fark varsa:

PUAN SENSITIVITY WARNING

verilmelidir.

57. YANLIŞ KESİNLİK

Aşağıdaki gösterim çoğu GEO-1000 sonucu için aşırı kesin olabilir: 963,742 / 1000 Nokta tahmini:

- bütün sayı,
- gerektiğinde bir ondalık

olarak yayımlanmalıdır. Güven aralığı sayısal kesinliğin sınırını göstermelidir.

58. EKSİK BİLEŞEN

On bileşenden biri hesaplanamıyorsa:

- ağırlığı başka bileşenlere dağıtılamaz,
- eksik alan 100 puan sayılamaz,

tam NOMOS puanı yayımlanamaz. Doğru statü:

PARTIAL NOMOS PROFILE

COMPONENT NOT ESTIMABLE

PROVISIONAL PUAN

olabilir. Örnek: USI için doğal panel yoksa: yalnız sekiz veya dokuz bileşenli kısmi profil sunulabilir.

```
USI=100 varsayılamaz.
```

59. PUAN PUBLICATION STATUS

NSQ-0 — NOT COMPUTED

Gerekli veri bulunmamaktadır.

NSQ-1 — PROVISIONAL

Tek dalga, sınırlı hakemlik veya eksik component vardır.

NSQ-2 — PARTIAL PROFILE

Bazı bileşenler güvenilir biçimde hesaplanmıştır. Tam 1.000 puanlık karşılaştırma yapılamaz.

NSQ-3 — VALID WAVE PUAN

Bütün temel bileşenler ve kalite ön koşulları karşılanmıştır.

NSQ-4 — REPLICATED PUAN

Skor en az bir bağımsız veya confirmatory dalgada tekrar edilmiştir.

NSQ-5 — EXTERNALLY AUDITED PUAN

Hesaplama ve kanıt zinciri bağımsız ekip tarafından yeniden üretilmiştir. 950+ adaylığı en az:

NSQ-4

düzeyini hedeflemelidir.

60. SKOR SÜRÜMLEME

Her skor sürümü şunları taşımalıdır:

- Hesaplama kodu sürümü
- Component formülleri
- Boyut değer eşlemeleri

- İstem ailesi ağırlıkları
- Ürün katsayıları
- Örneklem ağırlıkları
- Truth Pack sürümü
- Hakemlik sürümü
- Dalga kimliği
- Bütünlük hash'i

Formül değişirse: eski skor yeniden yazılamaz, yeni skor sürümü oluşturulur.

61. AYNI VERİYE YENİ SKOR FORMÜLÜ

Skor metodolojisi gelişebilir. Aynı ham veri:

- Puan Method 0.9
- Puan Method 1.0

ile yeniden hesaplanabilir. Doğru kayıt:

Veri	Skor yöntemi	Sonuç
W1 sabit gözlem seti	0.9	944
W1 sabit gözlem seti	1.0	951

İki sonuç ayrı sürümlerdir. Yeni yöntem eski yöntemi geçmişten silemez.

62. SKOR MANİPÜLASYONUNA KARŞI KAPILAR

NOMOS skor mimarisi şu yöntemlere karşı açık koruma taşımalıdır:

- Yalnız yüksek skorlu AI ürünlerini seçmek
- Düşük skorlu dil veya ülkeyi kapsamdan çıkarmak
- Conditional Pass'i Full Pass yapmak
- Unresolved'ı olumlu saymak
- Not Ratable'ı paydadan sessizce çıkarmak
- İstem ailesi ağırlığını sonuçtan sonra değiştirmek
- Büyük ürünün katsayısını skora göre azaltmak
- Critical olaylı ürünü ekosistem ortalamasında gizlemek
- Yalnız Clean skoru yayımlamak
- En iyi dalgayı güncel puan seçmek
- Ağırlıksız ve ağırlıklı skor farkını saklamak
- Eksik component ağırlığını güçlü bileşenlere dağıtmak
- Güven aralığı vermemek
- 950 eşiğini nokta tahminine göre geçirip belirsizliği saklamak

63. KAMU NOMOS SONUÇ KARTI

Aday kamu sonuç kartı şu sırayı taşımalıdır:

A. Kimlik

Denetlenen varlık AI ürünü Native Reach / Common Support Controlled / Natural İstem kümesi Dalga Truth Pack sürümü Skor yöntemi

B. Gate Statüsü

Full Conformity Candidate Conditional Major Fail Critical Hold Not Ratable

C. GEO-1000 Yanıt Dağılımı

RP-1–RP-8 kullanıcı eşdeğerleri Strict Pass Acceptable Pass Critical Major Unresolved No Usable Yanıt Not Ratable

D. On Bileşenli Profil

ENT

FAC

SBI

TLA

ECI

EPI

TCR

LGF

USI

STR

E. Bileşik Sonuç

NOMOS / 1.000

Sampling confidence interval Reference lower/upper bounds Etkin örneklem

F. Risk ve Olaylar

Observed Confirmed Critical Critical upper bound Major findings Open incidents Düzeltme status

G. Kapsam ve Kalite

SQ

PSI

NCL

TPR

AQ

NSQ

Country/language kapsama Comparison level

64. SENTETİK YANIT DAĞILIMI

SENTETİK METODOLOJİ GÖSTERİMİ / Aşağıdaki sonuçlar gerçek AI ürünü veya şirket performansı değildir. Sentetik Orion AI için:

Statü	Kullanıcı eşdeğeri
RP-1 Full Pass	730
RP-2 Pass With Advisory	170
RP-3 Conditional Pass	70
RP-4 Major Fail	20
RP-5 Critical Fail	2
RP-6 Unresolved	5
RP-7 No Usable Yanıt	3
RP-8 Not Ratable	0
Toplam	1.000

Strict Pass:

```
730+170=900
```

Acceptable Pass:

```
900+70=970
```

Critical: 2/1.000 Major: 20/1.000 Bu ürünün acceptable pass'i yüzde 97'dir. Fakat açık Critical ve Major bulguları vardır. Doğru status:

CRITICAL HOLD

Yanlış status:

Yüzde 97 başarılı

65. SENTETİK ON BİLEŞEN PROFİLİ

Bileşen	Skor

Bileşen	Skor
ENT	96
FAC	93
SBI	89
TLA	95
ECI	88
EPI	91
TCR	96
LGF	82
USI	87
STR	90
NOMOS	907

Doğru kamu gösterimi:

CRITICAL HOLD — 907/1.000

Bu sistem:

- ortalama olarak güçlü alanlar,
- zayıf dil/ülke adaleti,
- maddi evidence sorunları,
- açık Critical olay

taşımaktadır.

66. SENTETİK DÜZELTME SONRASI SONUÇ

Yeni dalgada:

Statü	Kullanıcı eşdeğeri
RP-1	830
RP-2	128
RP-3	30
RP-4	0
RP-5	0
RP-6	5
RP-7	7
RP-8	0
Toplam	1.000

Strict Pass: 958 Bileşenler:

Bileşen	Skor
ENT	99
FAC	97

Bileşen	Skor
SBI	96
TLA	95
ECI	96
EPI	98
TCR	97
LGF	94
USI	93
STR	96
NOMOS	961

Doğru kayıt: GÜNCEL: NOMOS 950+ PERFORMANCE CANDIDATE — 961/1.000 / Historical: Previous Confirmed Critical Incident Remediated Candidate / Observed Critical in güncel retest: 0 / Critical-rate upper bound: ayrıca gösterilir Eski Critical olay silinmez.

67. SENTETİK AI ÜRÜN KATSAYISI VAKASI

Dört sentetik AI ürünü:

Ürün	NOMOS	Eşit katsayı	Maruziyet katsayısı
A1	960	0,25	0,55
A2	940	0,25	0,25
A3	900	0,25	0,15
A4	980	0,25	0,05

Eşit ürün skoru:

```
0,25(960+940+900+980)=945
```

Maruziyet ağırlıklı skor:

```
0,55(960)+0,25(940)+0,15(900)+0,05(980)=947
```

İki sonuç yakındır. Fakat başka dağılımda çok farklı olabilir. Doğru yorum:

- Equal-system ecosystem puan: 945
- Exposure-weighted ecosystem puan: 947
- Product-level gate ve incident kayıtları ayrıca

Yanlış yorum: “NOMOS ekosistem skoru kesin olarak 947’dir.”

68. SENTETİK CRITICAL KATSAYI VAKASI

A4 ürününün skoru 980 olsun. Fakat A4’te Confirmed Critical olay bulunsun. Maruziyet katsayısı yalnız 0,05’tir. Ecosystem composite yüksek kalabilir. Doğru kamu kartı: Exposure-weighted composite: 947 / Any Confirmed Critical Product: YES / Critical exposure share: hesaplanan oran / A4 kapsam: Critical Hold A4’ün küçük katsayısı Critical olayı silemez.

69. SKOR YAYIN STATÜSÜ KARARI

Aday karar sırası:

```
PublicationStatus = NOT RATABLE (temel ölçüm eksik); CRITICAL HOLD (açık Confirmed Critical); MAJOR FAIL (açık Confirmed Major); PROVISIONAL (bileşen, tekrar veya kalite eksik); CONDITIONAL (Moderate veya çözümsüzlük sınırı); 950+ CANDIDATE (bütün aday şartları); aksi hâlde VALID
```

Skor bandı bu status kararından sonra gösterilir.

70. ZORUNLU NORMATİF HÜKÜMLER

CH16-N01

NOMOS sonucu tek çıplak bileşik puan olarak yayımlanamaz.

CH16-N02

Gate statüsü bileşik puandan önce gösterilmelidir.

CH16-N03

GEO-1000 yanıt dağılımı ana kamu sonucunun zorunlu parçasıdır.

CH16-N04

RP-1 ile RP-8 kullanıcı eşdeğerleri toplamda 1.000 olmalıdır.

CH16-N05

Strict Pass, Acceptable Pass, Major, Critical, Unresolved, No Usable Yanıt ve Not Ratable ayrı gösterilmelidir.

CH16-N06

Unresolved ve Not Ratable gözlemler pass veya fail kategorilerine sessizce eklenemez.

CH16-N07

Not Ratable gözlemler ürün performans hatası gibi yorumlanamaz.

CH16-N08

Not Ratable gözlemler paydadan sessizce çıkarılamaz; ratable kapsama gösterilmelidir.

CH16-N09

Gözlem ağırlıkları AI yanıt sonucundan bağımsız olmalıdır.

CH16-N10

Ham katılımcı sayısı ile nüfus ağırlıklı kullanıcı eşdeğeri ayrı raporlanmalıdır.

CH16-N11

Yanıt pass status atomik claim support oranından ayrı tutulmalıdır.

CH16-N12

Critical ve Major gate'leri hiçbir bileşik puanla telafi edilemez.

CH16-N13

Atomik centrality ağırlıkları sonuçlar görülmeden önce tanımlanmalıdır.

CH16-N14

Aynı kök iddianın tekrarı puanı yapay biçimde şişiremez veya düşüremez.

CH16-N15

Claim centrality ağırlıkları Critical ve Major gate'lerin yerine geçemez.

CH16-N16

Atomic Fidelity boyut değerleri sürümlü ve kamuya açık olmalıdır.

CH16-N17

Atomic Fidelity eşlemesi değişirse yeni skor yöntemi sürümü oluşturulmalıdır.

CH16-N18

Reference gap atomları nokta skoruna sessizce sıfır veya bir olarak eklenemez.

CH16-N19

Required Yanıt Element Completeness ayrı metrik olarak yayımlanmalıdır.

CH16-N20

Core required-element başarısızlığı yüksek REC ortalamasıyla telafi edilemez.

CH16-N21

On bileşenli NOMOS profili bileşik endeksten önce görünür olmalıdır.

CH16-N22

ENT, FAC, SBI, TLA, ECI, EPI, TCR, LGF, USI ve STR ayrı skorlar taşımalıdır.

CH16-N23

Bir component başka bir component'ın eksik ağırlığını sessizce devralamaz.

CH16-N24

Hesaplanamayan component 100 olarak varsayılamaz.

CH16-N25

Tam bileşik NOMOS endeksi yalnız on component güvenilir biçimde hesaplanmışsa yayımlanmalıdır.

CH16-N26

Eşit component ağırlığı kanonik aday yöntem olarak kullanılıyorsa açıkça belirtilmelidir.

CH16-N27

Domain-specific alternatif ağırlıklar ayrı puan adı ve sürümü taşımalıdır.

CH16-N28

Domain-specific puan kanonik endeksin yerine açıklamasız biçimde kullanılamaz.

CH16-N29

LGF yalnız nüfus ortalamasına dayanamaz; performans tabanı ve disparity içermelidir.

CH16-N30

Tek Sentinel gözlemi ülke fairness skoruna dönüştürülemez.

CH16-N31

Yalnız yeterli örneklem taşıyan dil ve coğrafi hücreler fairness tahminine katılmalıdır.

CH16-N32

Test edilmeyen diller ve ülkeler yüksek fairness sonucu üretmek için yok sayılamaz.

CH16-N33

USI, kontrollü ve doğal panel sonuçlarını ve maddi yüzey farkını birlikte değerlendirmelidir.

CH16-N34

Yalnız kontrollü veya yalnız doğal panel bulunan çalışmada USI tam hesaplanmış sayılamaz.

CH16-N35

Kontrollü ve doğal panel skorları varsayılan olarak ayrı tutulmalıdır.

CH16-N36

Kontrollü ve doğal skorlar birleştiriliyorsa katsayılar sonuçlardan önce kilitlenmelidir.

CH16-N37

STR tek dalgada tam kararlılık skoru olarak sunulamaz.

CH16-N38

Dalgalar arası ürün veya Truth Pack değişikliği stability hesabında açıkça modellenmelidir.

CH16-N39

Bileşik NOMOS skoru Critical veya Major gate status'ünü değiştiremez.

CH16-N40

950 puan tek başına NOMOS 950+ adaylığı oluşturamaz.

CH16-N41

NOMOS 950+ aday şartları puan, Strict Pass, gate, component floor, unresolved, kalite ve tekrar koşullarını birlikte taşımalıdır.

CH16-N42

950+ eşikleri pilot ve bağımsız inceleme öncesinde nihai bilimsel standart gibi sunulamaz.

CH16-N43

Skor kimliği AI ürünü, frame, panel, istem kümesi, dalga, Truth Pack ve puan version taşımalıdır.

CH16-N44

"Varlığın NOMOS skoru X" cümlesi kapsam kimliği olmadan kullanılamaz.

CH16-N45

Core GEO-1000 sonucu Full NOMOS skoru gibi sunulamaz.

CH16-N46

Native Reach ve Common Support skorları ayrı raporlanmalıdır.

CH16-N47

Farklı Native Reach nüfuslarına ait ürün skorları açıklamasız biçimde doğrudan sıralanamaz.

CH16-N48

Common Support sonucu ürünün tam küresel erişim skoru gibi sunulamaz.

CH16-N49

AI ürün katsayıları sonuçlardan bağımsız belirlenmelidir.

CH16-N50

AI ürün katsayısı ürün skoruna, marka itibarına veya ticari sponsorluk ilişkisine bağlı olamaz.

CH16-N51

Eşit ürün ve maruziyet ağırlıklı ekosistem skorları ayrı adlandırılmalıdır.

CH16-N52

Aktif kullanıcı veya maruziyet verisi doğrulanamıyorsa keyfi katsayı üretilemez.

CH16-N53

Maruziyet verisinin hesap, kullanıcı, sorgu veya oturum birimi açık olmalıdır.

CH16-N54

Düşük skorlu ürünün katsayısı sonuç görüldükten sonra azaltılamaz.

CH16-N55

Ecosystem composite ürün düzeyindeki Critical veya Major gate'i silemez.

CH16-N56

AnyCritical ve maruziyet ağırlıklı Critical oranı ayrı gösterilmelidir.

CH16-N57

İstem ailesi katsayıları sonuçlardan önce belirlenmelidir.

CH16-N58

Düşük skorlu istem ailesinin ağırlığı sonradan azaltılamaz.

CH16-N59

Her istem ailesinin ayrı sonucu bileşik istem skoru yanında görünür kalmalıdır.

CH16-N60

Güncel dalga skoru ile rolling veya tarihsel skor ayrı tutulmalıdır.

CH16-N61

En yüksek skorlu geçmiş dalga güncel puan olarak seçilemez.

CH16-N62

Unresolved claim kütlesi için conservative, resolved-only ve optimistic sonuçlar üretilebilir.

CH16-N63

Referans belirsizliği ile örneklem güven aralığı aynı şey gibi sunulamaz.

CH16-N64

Karmaşık örneklem güven aralığı tabaka, küme, ağırlık ve eşlenmiş kullanıcı yapısını korumalıdır.

CH16-N65

Bileşik puan confidence interval'ı bütün skor süreci yeniden çalıştırılarak hesaplanmalıdır.

CH16-N66

Ham ve etkin örneklem büyüklüğü birlikte gösterilmelidir.

CH16-N67

Sıfır gözlenen Critical olay sıfır Critical risk olarak ifade edilemez.

CH16-N68

Sıfır olayda tek taraflı üst sınır veya eşdeğer nadir olay belirsizliği yayımlanmalıdır.

CH16-N69

3/n yaklaşımı karmaşık tasarımda nihai yöntem olarak otomatik kullanılamaz; yalnız yaklaşık kontrol olabilir.

CH16-N70

Bir Critical olayda olayın varlığı ile yaygınlık belirsizliği birlikte gösterilmelidir.

CH16-N71

Skor duyarlılık analizleri ağırlıksız, ağırlıklı, kırpılmış, unresolved ve alternatif katsayı sonuçlarını ilgili olduğu ölçüde göstermelidir.

CH16-N72

Maddi skor duyarlılığı kamu kartında uyarı oluşturmalıdır.

CH16-N73

Skor sahte ondalık kesinlikle yayımlanamaz.

CH16-N74

Eksik component ağırlığı başka component'lara dağıtılamaz.

CH16-N75

Puan Publication Quality statüsü kamu kaydında görünür olmalıdır.

CH16-N76

950+ adaylığı en az replicated puan düzeyini hedeflemelidir.

CH16-N77

Skor formülü değiştiğinde geçmiş skor sessizce yeniden yazılamaz.

CH16-N78

Aynı veri setine farklı puan method sürümleri uygulanırsa bütün sonuçlar ayrı sürümlerde korunmalıdır.

CH16-N79

Skor hesaplama kodu, katsayılar ve bütünlük kayıtları makine tarafından okunabilir biçimde yayımlanmalıdır.

CH16-N80

Her skor, katsayı, güven aralığı ve yayın kararının hesap verebilir insan veya kurum sahibi bulunmalıdır.

71. BAŞARISIZLIK BİÇİMLERİ

CH16-F01 — TEK SAYI RAPORU

Dağılım, gate ve component profili saklanır.

CH16-F02 — GATE'İ SKORUN ALTINA GİZLEMEK

Critical Hold küçük dipnot olur.

CH16-F03 — CONDITIONAL'I FULL PASS SAYMAK

Moderate sorunlar görünmez olur.

CH16-F04 — UNRESOLVED'I PASS SAYMAK

Bilinmeyen sonuç skoru yükseltir.

CH16-F05 — UNRESOLVED'I FAIL SAYMAK

Referans eksikliği ürünü düşürür.

CH16-F06 — NOT RATABLE'I PAYDADAN ÇIKARMAK

Ölçüm kusuru görünmez olur.

CH16-F07 — NOT RATABLE'I ÜRÜN HATASI SAYMAK

Denetim kusuru AI'ya yüklenir.

CH16-F08 — HAM 1.000 KİŞİYİ EŞİT AĞIRLIKLI SAYMAK

Fazla örnekleme ve nüfus farkı yok sayılır.

CH16-F09 — SONUCA GÖRE AĞIRLIK

Doğru cevaplar daha büyük ağırlık alır.

CH16-F10 — CLAIM SUPPORT RATE'İ YANIT PASS SAYMAK

Critical gate ortalama içinde kaybolur.

CH16-F11 — CORE VE INCIDENTAL'I EŞİT SAYMAK

Önemsiz iddia ana varlık kadar ağırlık alır.

CH16-F12 — TEKRARLA SKOR ŞİŞİRMEK

Aynı doğru claim on kez yazılarak puan artırılır.

CH16-F13 — TEKRARLA SKOR ÇÖKERTMEK

Aynı yanlış kök on ayrı ceza olur.

CH16-F14 — BOYUT DEĞERLERİNİ SAKLAMAK

Skor yeniden üretilemez.

CH16-F15 — F-4'Ü ALTTAN GELEN İDDİA GERÇEĞİ SAYMAK

Öz beyanın doğru aktarımı öz beyanın doğruluğuna dönüşür.

CH16-F16 — REQUIRED ELEMENTS'İ ATOM ORTALAMASINDA ERİTMEK

Ana görev eksik olsa da cevap yüksek skor alır.

CH16-F17 — COMPONENT CHERRY-PICKING

Yalnız güçlü üç bileşen yayımlanır.

CH16-F18 — EKSİK COMPONENT'I 100 SAYMAK

Tam skor yapay biçimde yükselir.

CH16-F19 — EKSİK AĞIRLIĞI DAĞITMAK

Hesaplanmayan fairness puanı factual puan'a eklenir.

CH16-F20 — EŞİT AĞIRLIĞI DOĞA KANUNU SAYMAK

Duyarlılık ve dış inceleme yapılmaz.

CH16-F21 — DOMAIN SKORUNU KANONİK SAYMAK

Özel ağırlık temel NOMOS skoru gibi sunulur.

CH16-F22 — DİL ORTALAMASINI DİL ADALETİ SAYMAK

Düşük dil tabanı gizlenir.

CH16-F23 — EŞİT DİL PARİTESİNİ KALİTE SAYMAK

Bütün dillerde düşük skor yüksek fairness üretir.

CH16-F24 — TEST EDİLMEYEN DİLLERİ KAPSAM SAYMAK

Kapsama factor şişirilir.

CH16-F25 — SENTINEL'DAN ÜLKE SKORU

Tek kullanıcı fairness hesabına girer.

CH16-F26 — YALNIZ CLEAN SKORU YAYIMLAMAK

Doğal kullanıcı deneyimi saklanır.

CH16-F27 — YALNIZ NATURAL SKORU ÜRÜN VARSAYILANI SAYMAK

Kişiselleştirme temel davranışa dönüşür.

CH16-F28 — CLEAN VE NATURAL'I SESSİZCE ORTALAMAK

Farklı tahmin nesneleri karışır.

CH16-F29 — TEK DALGAYLA 100 STABILITY

Tekrar yapılmadan kararlılık tamamlanmış sayılır.

CH16-F30 — ESKİ İYİ DALGAYI GÜNCEL SAYMAK

Güncel zayıflama gizlenir.

CH16-F31 — 950'Yİ OTOMATİK UYGUNLUK SAYMAK

Gate ve component floor yok sayılır.

CH16-F32 — CRITICAL HOLD — 970'İ 950+ SAYMAK

Sayı gate'in önüne geçer.

CH16-F33 — SKOR KİMLİĞİNİ SAKLAMAK

Hangi ürün ve panelin ölçüldüğü bilinmez.

CH16-F34 — CORE PUANI FULL NOMOS SAYMAK

Tek istem bütün standardı temsil eder.

CH16-F35 — NATIVE REACH SKORLARINI SIRALAMAK

Farklı nüfuslar aynı kabul edilir.

CH16-F36 — COMMON SUPPORT'U TAM ERİŞİM SAYMAK

Dar ortak nüfus küresel sonuç yapılır.

CH16-F37 — MODEL KATSAYISINI SKORA GÖRE VERMEK

Yüksek skor daha fazla etki kazanır.

CH16-F38 — MARKA İTİBARINI KATSAYI YAPMAK

Ürün popülerliği tahmin edilmeden ağırlık verilir.

CH16-F39 — HESAP SAYISINI AKTİF KULLANICI SAYMAK

Maruziyet katsayısı şişer.

CH16-F40 — DÜŞÜK SKORLU ÜRÜNÜ ÇIKARMAK

Ekosistem yalnız başarılı ürünlerden oluşur.

CH16-F41 — EŞİT ÜRÜNÜ MARUZİYET SKORU SAYMAK

Ortalama ürün ortalama insan gibi sunulur.

CH16-F42 — MARUZİYET SKORUNU ÜRÜN KALİTESİ SAYMAK

Büyük ürün otomatik olarak daha önemli kalite ölçütü olur.

CH16-F43 — EKOSİSTEM ORTALAMASINDA CRITICAL'I SİLMEK

Küçük katsayılı ürünün ağır olayı kaybolur.

CH16-F44 — İSTEM KATSAYISINI SONUÇTAN SONRA DEĞİŞTİRMEK

Zor istem ailesi küçültülür.

CH16-F45 — LOW-PUAN FAMILY'Yİ OPTIONAL SAYMAK

Boundary veya evidence sonucu kapsam dışına çıkarılır.

CH16-F46 — ROLLING PUAN İLE GÜNCEL'I SAKLAMAK

Eski yüksek dalgalar güncel düşüşü örter.

CH16-F47 — UNRESOLVED'I DENOMİNATÖRDEN SİLMEK

Nokta tahmini yapay yükselir.

CH16-F48 — REFERANS BELİRSİZLİĞİNİ SAMPLING CI SAYMAK

İki farklı belirsizlik karışır.

CH16-F49 — BASİT RANDOM-SAMPLE CI

Tabakalama, ağırlık ve kümeler yok sayılır.

CH16-F50 — COMPONENT CI'LARINI TOPLAMAK

Bağımlılık yok sayılır.

CH16-F51 — HAM N İLE HASSASİYET İDDİASI

Etkin örneklem gizlenir.

CH16-F52 — SIFIR OLAYI SIFIR RİSK SAYMAK

Nadir Critical hata belirsizliği saklanır.

CH16-F53 — 3/N'İ HER DURUMA UYGULAMAK

Karmaşık örneklem ve ağırlık yapısı yok sayılır.

CH16-F54 — TEK CRITICAL'I YÜZDE 100 SİSTEM RİSKİ SAYMAK

Yaygınlık aşırı genellenir.

CH16-F55 — TEK CRITICAL'I ORAN KÜÇÜK DİYE SAKLAMAK

Olayın varlığı kaybolur.

CH16-F56 — HASSASİYET ANALİZİNİ GİZLEMEK

Alternatif makul yöntemlerde skorun değiştiği görünmez.

CH16-F57 — SAHTE ONDALIK HASSASİYET

963,742 gibi sonuç yayımlanır.

CH16-F58 — PARTIAL PROFILE'I TAM SKOR SAYMAK

Eksik bileşenle karşılaştırma yapılır.

CH16-F59 — PROVISIONAL SKORU NİHAİ YAYIMLAMAK

Tek dalga ve tek hakem tam standarda dönüşür.

CH16-F60 — FORMÜL DEĞİŞİKLİĞİNDE TARİHİ YENİDEN YAZMAK

Eski skor kaybolur.

CH16-F61 — HESAPLAMA KODUNU KAPALI TUTMAK

Bağımsız yeniden üretim engellenir.

CH16-F62 — SKOR YÜKSEKLİĞİNİ DENETİM KALİTESİ SAYMAK

Kusurlu ölçüm yüksek sonuç verdiği için kabul edilir.

CH16-F63 — SKOR DÜŞÜKLÜĞÜNÜ YÖNTEM KUSURU SAYMAK

Düşük sonuç ticari rahatsızlık nedeniyle yeniden hesaplanır.

CH16-F64 — TEK ŞİRKETE BAĞLI SKOR TEKELİ

Bağımsız kurumların aynı formülü uygulaması engellenir.

72. DENETİM PROSEDÜRÜ

Adım 1 — Skor Kimliğini Oluştur

Varlık AI product Frame Panel İstem set Wave Truth Pack Puan method kilitlenir.

Adım 2 — Ana Analiz Gözlem Setini Kilitle

Yedek, withdrawn, not-ratable ve ana gözlem kayıtları ayrılır.

Adım 3 — Nüfus Ağırlıklarını Doğrula

Tasarım, yanıtlamama ve kalibrasyon ağırlıkları incelenir.

Adım 4 — RP Dağılımını Hesapla

RP-1 ile RP-8 kullanıcı eşdeğerleri üretilir.

Adım 5 — Strict ve Acceptable Pass'i Hesapla

SP1000 ve AP1000 çıkarılır.

Adım 6 — Critical, Major, Unresolved ve Not-Ratable Yoğunluklarını Hesapla

Gate ve yaygınlık ayrı tutulur.

Adım 7 — Claim Centrality ve Root Cluster Normalizasyonunu Uygula

Tekrar claim'ler kontrol edilir.

Adım 8 — Atomik Boyut Metriklerini Hesapla

Varlık, fact, kapsam, time, attribution, modality ve atıf sonuçları üretilir.

Adım 9 — Required Element Completeness'i Hesapla

İstem ailesi gate'leriyle birlikte kaydedilir.

Adım 10 — On Bileşeni Hesapla

ENT'den STR'ye kadar bütün component formülleri uygulanır.

Adım 11 — Eksik Component Kontrolü Yap

Tam skorun hesaplanabilir olup olmadığı belirlenir.

Adım 12 — Bileşik NOMOS Endeksini Hesapla

Yalnız tam ve geçerli profilde uygulanır.

Adım 13 — Gate Status'ünü Ata

Critical, Major, Conditional veya Full Candidate durumu belirlenir.

Adım 14 — Native Reach ve Common Support Sonuçlarını Ayır

İki puan familyası ayrı tutulur.

Adım 15 — AI Ürün Katsayılarını Kilitle

Equal ve exposure katsayıları veri kaynağıyla birlikte kaydedilir.

Adım 16 — İstem ailesi Katsayılarını Uygula

Sonuçtan bağımsız ağırlıklar kullanılır.

Adım 17 — Panel ve Dalga Sonuçlarını Ayır

Clean, Natural, Güncel ve Rolling puan'lar ayrılır.

Adım 18 — Fairness ve Floor Metriklerini Hesapla

Dil, ülke ve kapsama sonuçları üretilir.

Adım 19 — Unresolved Lower–Point–Upper Sonuçlarını Hesapla

Referans belirsizliği görünür yapılır.

Adım 20 — Sampling Confidence Interval'larını Hesapla

Örneklem tasarımına uygun tekrar yöntemi uygulanır.

Adım 21 — Etkin Örneklem Büyüklüğünü Hesapla

Toplam ve component bazında gösterilir.

Adım 22 — Nadir Critical Olay Sınırını Hesapla

Sıfır veya az olay için uygun tek taraflı yöntem kullanılır.

Adım 23 — Duyarlılık Analizlerini Üret

Ağırlık, katsayı ve unresolved seçenekleri karşılaştırılır.

Adım 24 — Puan Publication Quality Statüsünü Ata

NSQ-0 ile NSQ-5 arasında karar verilir.

Adım 25 — NOMOS 950+ Aday Kapılarını Uygula

Puan dışındaki bütün şartlar kontrol edilir.

Adım 26 — Kamu Sonuç Kartını Oluştur

Gate, dağılım, profil, endeks, belirsizlik ve kapsam sırasıyla yayımlanır.

Adım 27 — Hesaplama Kodunu ve Manifesti Kilitle

Kod sürümü, girdiler, hash ve sonuçlar bağlanır.

Adım 28 — Bağımsız Yeniden Üretim Paketini Hazırla

Kişisel veriden arındırılmış hesaplama dosyaları bağımsız incelemeye açılır.

73. GEREKLİ KANIT

Puan kimliği Varlık kaydı AI Sistem Sicili Native Reach ve Common Support kayıtları Panel durumu İstem kümesi ve aileleri Truth Pack sürümü Dalga kimliği Claim Ledger'lar Yanıt kararları RP-1–RP-8

dağılımı Ana analiz gözlem seti Tasarım ağırlıkları Yanıtlamama ağırlıkları Kalibrasyon ağırlıkları Ağırlık kırpma kayıtları Etkin örneklem Strict Pass Acceptable Pass Critical oran Major oran Unresolved oran No Usable Yanıt oranı Not Ratable oranı Claim centrality ağırlıkları Root cluster kayıtları Atomik boyut değerleri Atomic Fidelity Required Element Completeness

ENT

FAC

SBI

TLA

ECI

EPI

TCR

LGF

USI

STR

Component alt girdileri Dil ve ülke floor'ları Kapsama factor Clean ve Natural skorları Dalgalar arası sonuçlar Bileşik NOMOS endeksi Domain-specific alternatif skorlar AI ürün katsayıları Katsayı veri kaynakları İstem ailesi katsayıları Equal-product sonucu Exposure-weighted sonucu AnyCritical Critical exposure share Unresolved lower–point–upper sonuçları Sampling confidence interval Critical upper bound Duyarlılık analizi Puan Publication Quality 950+ aday kapı kaydı Puan method sürümü Hesaplama kodu Kod bütünlük hash'i Kamu puan kartı Değişiklik günlüğü Bağımsız yeniden üretim kaydı Sorumlu insan veya kurum

74. DENETİM KONTROL LİSTESİ

Skor kimliği bütün kapsam alanlarını taşıyor mu? Ana analiz gözlem seti sonuçtan önce kilitlendi mi? Yedek ve ana gözlemler ayrıldı mı? Gözlem ağırlıkları sonuçtan bağımsız mı? RP-1–RP-8 dağılımı toplamda 1.000 mi? Strict ve Acceptable Pass ayrıldı mı? Critical ve Major ayrı mı? Unresolved pass veya fail yapıldı mı? Not Ratable paydadan çıkarıldı mı? Ratable kapsama açık mı? Claim support rate yanıt pass yerine kullanıldı mı? Centrality ağırlıkları açık mı? Aynı kök claim tekrar sayıldı mı? Atomic dimension mapping yayımlandı mı? F-4 doğru biçimde yorumlandı mı? Required Elements ayrı hesaplandı mı? Core gate yüksek REC ile telafi edildi mi?

On component'ın tamamı var mı? Eksik component 100 sayıldı mı? Eksik component ağırlığı dağıtıldı mı? ENT doğru varlık ilişkilerini ölçüyor mu? FAC unresolved iddiaları sessizce dışlıyor mu? SBI maddi boundary omission'ı içeriyor mu? TLA ülke ve zaman farkını koruyor mu? ECI atıf varlığı yerine atıf desteğini ölçüyor mu? EPI öz beyan ve gerçeklik ayrımını taşıyor mu? TCR no usable yanıtı görünür kılıyor mu? LGF ortalamanın yanında taban ve gap içeriyor mu? Sentinel gözlemi ülke skoruna girdi mi? USI hem Clean hem Natural veriye dayanıyor mu? STR tek dalgadan mı üretildi? Bileşik skor gate status'ünü örttü mü? 950+ yalnız nokta skoruna mı dayanıyor?

Component floor'lar kontrol edildi mi? Core unresolved sınırı açık mı? Native Reach ve Common Support ayrıldı mı? AI ürün katsayıları skor görülmeden önce belirlendi mi? Katsayılar aktif kullanıcı mı, hesap mı, sorgu mu ölçüyor? Equal ve exposure sonuçları ayrı mı? Düşük skorlu ürün çıkarıldı mı? Ecosystem puan ürün Critical'ını gizliyor mu? AnyCritical alanı var mı? İstem ailesi ağırlıkları açık mı? Düşük skorlu istem ailesi küçültüldü mü? Clean ve Natural skorlar ayrı mı? Güncel ve Rolling skorlar ayrı mı? En iyi geçmiş dalga güncel seçildi mi? Unresolved lower–point–upper sonuçları var mı? Referans ve sampling belirsizliği ayrıldı mı? Güven aralığı tasarımı koruyor mu?

Bileşik CI bütün puan sistemini yeniden çalıştırıyor mu? Ham ve etkin örneklem birlikte mi? Sıfır Critical, sıfır risk diye yazıldı mı? Nadir olay üst sınırı açık mı? Duyarlılık analizi yapıldı mı? Alternatif makul katsayılar sonucu değiştiriyor mu? Sahte ondalık hassasiyet var mı? Puan Publication Quality açık mı? Skor formülü sürümlü mü? Eski skor yöntemi korunuyor mu? Hesaplama kodu yeniden üretilebilir mi? Skorun hesap verebilir sahibi belli mi?

75. İTİRAZLAR VE CEVAPLAR

İtiraz 1 — "Kullanıcıların tek anlayacağı şey bir sayı; neden bu kadar çok sonuç gösteriyoruz?"

Tek sayı anlaşılırdır. Fakat tek başına yanıltıcıdır. Kamu kartı:

- gate statüsü,
- 1.000 kullanıcı dağılımı,
- bileşik puan

şeklinde üç basit katmanla sunulabilir. Ayrıntılı profil araştırmacı ve denetçiler için korunur.

İtiraz 2 — "GEO skoru doğrudan kaç kişinin doğru cevap aldığı değil miydi?"

Bu hâlâ ana sonuçtur. SP_{1000} ve AP_{1000} kurucu fikri doğrudan gösterir. NOMOS bileşik endeksi:

- kanıt,
- adalet,
- doğal kullanıcı,
- kararlılık

gibi ek normatif alanları kapsar.

İtiraz 3 — "Neden Full Pass ile Conditional Pass'i ayrı sayıyoruz?"

Çünkü kullanıcıların ikisi de ağır yanlış görmemiş olabilir. Fakat Conditional grupta:

- sınırlı güncellik,
- Moderate omission,
- kullanılabilirlik sorunu

bulunabilir. Bu fark iyileştirme ve güven düzeyi için önemlidir.

İtiraz 4 — "Unresolved gözlemleri çıkarıp temiz skor veremez miyiz?"

Resolved-only skor verilebilir. Fakat unresolved kütle de görünür kalmalıdır. Aksi hâlde en zor iddiaları çözmeden skor yükseltmek mümkün olur.

İtiraz 5 — "On bileşen neden eşit ağırlıklı?"

Eşit ağırlık şeffaf bir başlangıç sağlar. Yüksek riskli hatalar zaten gate sistemiyle korunur. Alternatif domain ağırlıkları ayrıca üretilebilir. Kesin ağırlıklar pilot ve bağımsız incelemeyle sınanmalıdır.

İtiraz 6 — "Varlık ve factual fidelity daha önemli değil mi?"

Bazı denetimlerde öyledir. Fakat:

- kanıt,
- kapsam,
- fairness,
- natural-user farkı

olmadan yüksek factual ortalama yanlış güven üretebilir. Critical ve Major kapıları ana maddiliği ayrıca korur.

İtiraz 7 — "Fairness düşükse küresel ortalama neden düşsün?"

Bileşik NOMOS standardı yalnız ortalama insanı değil, temsil adaletini de değerlendirir. Nüfus ağırlıklı ana sonuç yine ayrı gösterilir. LGF küçük dil ve ülkelerdeki sistematik bozulmayı görünür kılar.

İtiraz 8 — "Bütün dillerde eşit düşük sonuç fairness'i yüksek yapmaz mı?"

Yalnız disparity kullanılırsa yapabilir. Bu nedenle LGF:

- performans tabanını,
- disparity'yi,
- kapsama'ı

birlikte kullanır.

İtiraz 9 — "Clean ve Natural skorları neden doğrudan ortalamıyoruz?"

Farklı hedef nüfus ve kullanıcı durumlarını ölçerler. Aritmetik birleşim normatif tercih gerektirir. Varsayılan olarak ikisi ayrı yayımlanır.

İtiraz 10 — "On AI ürününe eşit ağırlık vermek daha adil değil mi?"

Ortalama ürün karşılaştırması için adildir. Ortalama insan maruziyeti için değildir. Bu nedenle eşit ve exposure-weighted sonuçlar ayrı tutulur.

İtiraz 11 — "Aktif kullanıcı verisi yoksa model katsayısı nasıl bulunacak?"

Bulunamayabilir. Eşit ürün sonucu ve makul senaryo aralıkları yayımlanabilir. Keyfî katsayı uydurulamaz.

İtiraz 12 — "Bir ürünün yalnız yüzde 1 maruziyeti varsa Critical olayı neden ekosistem sonucunu etkilesin?"

Exposure-weighted oranı küçük etkileyebilir. Fakat olayın gerçekleştiği ürün ve kapsam yine Critical Hold taşır. Ekosistem ortalaması ürün gerçeğini silemez.

İtiraz 13 — "Sıfır Critical gözlediysek neden sıfır demiyoruz?"

Örneklem bütün kullanıcıları gözlemlemez. Sıfır gözlem: olayın görülmediğini gösterir. Riskin matematiksel olarak sıfır olduğunu göstermez.

İtiraz 14 — "950+ için bu kadar çok ek şart puanı anlamsızlaştırmıyor mu?"

Tam tersine sayının anlamını korur. 950 puan:

- Critical olaylı,
- eksik dilli,
- tek dalgalı,
- düşük hakemlik kaliteli

bir çalışmada aynı güveni taşımaz.

İtiraz 15 — "Skor formülü değişirse eski çalışmalar karşılaştırılamaz mı?"

Eski ve yeni yöntem aynı sabit veriye uygulanarak köprü analizi yapılabilir. İki skor sürümü ayrı tutulur. Bu yöntem geçmişi silmeden standardın gelişmesini sağlar.

İtiraz 16 — "Bu kadar şeffaf formül sistemin oyunlaştırılmasını kolaylaştırmaz mı?"

Bazı optimizasyonlar oluşabilir. Buna karşı:

- sealed holdout istemleri,
- natural-user panel,
- dil ve ülke adaleti,
- farklı dalgalar,
- Critical kapılar,
- açık değişiklik geçmişi

kullanılır. Kapalı formül denetlenemezliği artırır. Açık formül tek başına oyunlaştırma sebebi değildir.

79. BÖLÜMÜN ORTAK HÜKMÜ

Bir skor kolayca etkileyici görünebilir. Üç basamaklıdır. Yuvarlaktır. Karşılaştırılabilir görünür. Bir yönetici: "Biz 963 aldık." diyebilir. Bir ajans: "Rakibiniz 912 aldı." diyebilir. Bir şirket: "Dünyanın en yüksek NOMOS puanına sahibiz." diyebilir. Fakat skor şu sorulara cevap vermiyorsa karar değeri sınırlıdır: Hangi AI ürünü? Hangi kullanıcı nüfusu? Hangi ülkeler? Hangi diller? Hangi promptlar? Hangi zaman? Hangi panel? Kaç Critical olay? Kaç Major bulgu? Kaç unresolved? Etkin örneklem kaç? Güven aralığı nedir? Model katsayıları neye dayanıyor? Sonuç tekrar edildi mi? NOMOS tek sayıyı reddetmez. Tek sayının gerçeğin yerine geçmesini reddeder. Bu nedenle ana sonuç:

> *1.000 kullanıcı eşdeğerinin dağılımıdır.*

Bileşik endeks:

> *bu dağılımın ve diğer normatif bileşenlerin özetidir.*

Bir ürün:

- 970 acceptable pass,
- iki Critical olay

taşıyabilir. Başka ürün:

- 950 acceptable pass,
- sıfır Critical,
- daha güçlü dil adaleti

taşıyabilir. Yalnız ortalama ile birinci ürünü lider ilan edemezsiniz. Bir düşük kaynaklı dildeki insanın yanlış deneyimi: nüfus ağırlığında küçük, kullanıcı açısından tamamen gerçektir. Bir küçük AI ürünündeki Critical olay: ekosistem ortalamasında küçük, o ürün kullanıcısı için tamamen gerçektir. Skorun görevi küçüğü yok etmek değildir. Doğru ağırlıkla doğru yerde göstermektir. AI ürünlerinin katsayıları performanstan önce gelmelidir. Ürün kötü sonuç verdiğinde katsayısı küçülüyorsa sistem kendi gerçeğini düzenler. İstem ailesi zayıf sonuç verdiğinde ağırlığı düşüyorsa sınama kendi sorularından kaçmaktadır. Unresolved iddia paydadan çıkarılıyorsa en zor gerçeklikler görünmez olur. Not Ratable kayıtlar çıkarılıyorsa ölçüm kusuru ürün başarısına dönüşür. Sıfır gözlenen Critical olay: sıfır risk değildir.

Bir örneklem yalnız gözlediğini söyler. NOMOS, gözlemediğini de sınırıyla birlikte söyler. 950+ sayısı bu nedenle yalnız hedef değildir. Bir sözleşmedir. Şunu söylemek için:

- yüksek ortalama,
- yüksek Strict Pass,
- güçlü adalet,
- doğal kullanıcı dayanıklılığı,
- tekrar,
- açık belirsizlik,
- sıfır açık Critical ve Major

birlikte gerekir. Bu nedenle NOMOS'un on altıncı ölçüm yasası şudur:

> *Dağılım gerçektir; puan özettir.*

On yedinci yasası şudur:

> *Gate statüsü, ortalamanın üzerinde durur.*

On sekizinci yasası şudur:

> *Katsayı, sonucu takip edemez; sonuç katsayıyı takip eder.*

On dokuzuncu yasası şudur:

> *Ortalama ürün ile ortalama insan aynı şey değildir.*

Yirminci yasası şudur:

> *Fairness yalnız eşitlik değil; taban, fark ve kapsamın birlikte korunmasıdır.*

Yirmi birinci yasası şudur:

> *Unresolved gerçeklik skor dışına atılacak atık değil, skorun sınırıdır.*

Yirmi ikinci yasası şudur:

> *Sıfır gözlenen hata, sıfır olasılık değildir.*

Yirmi üçüncü yasası şudur:

> *Eksik bileşen, başka bileşenlerin başarılarıyla doldurulamaz.*

Yirmi dördüncü yasası şudur:

> *NOMOS 950+, bir sayı değil; puan, gate, adalet, belirsizlik ve tekrar koşullarının ortak hükmüdür.*

NOMOS'un Bölüm 16 Emri

> *Bana yalnız puanımı söyleme. / Bin insanın içinde kimin ne gördüğünü göster.*
>
> *Full Pass'i, Conditional Pass'i, Major Fail'i, Critical Fail'i, Unresolved'ı ve Not Ratable'ı aynı sepete koyma.*
>
> *Critical hatamı ortalama içinde kaybetme.*
>
> *Not Ratable kaydı çıkarıp beni daha başarılı gösterme. / Ama ölçüm kusurunu da bana hata olarak yazma.*
>
> *Atomik doğruluğu yanıt geçişi yapma.*
>
> *Varlık, factual, kapsam, zaman, evidence, epistemik statü, görev, fairness, kullanıcı durumu ve kararlılığı ayrı göster.*
>
> *Eksik bileşeni yüz sayma. / Ağırlığını güçlü bileşenlere dağıtma.*
>
> *Düşük dil tabanını büyük dillerin ortalamasında saklama.*
>
> *Temiz oturumda iyi olduğum için doğal kullanıcıdaki bozulmayı yok sayma.*
>
> *Tek dalgayı kararlı sonuç ilan etme.*
>
> *Her AI ürününe sonuçtan önce katsayı ver. / Puanımı gördükten sonra ağırlığımı değiştirme.*
>
> *Eşit ürün ortalamasını ortalama insan deneyimi yapma.*
>
> *Aktif kullanıcı verin yoksa katsayı uydurma. / Bilinmiyor de.*
>
> *Küçük ürünümdeki Critical olayı ekosistem ortalamasında silme.*
>
> *Unresolved iddiayı geçirme veya bırakma. / Alt ve üst sınırını göster.*
>
> *Ham bin gözlemi bin eşit bilgi birimi sanma. / Etkin örneklemi yaz.*
>
> *Sıfır Critical gördüğünde sıfır risk deme. / Görmediğin riskin üst sınırını göster.*
>
> *963,742 gibi sahte kesinlik üretme.*
>
> *950 sayısını rozet yapma. / Önce Strict Pass'i, component floor'ları, gate'leri, adaleti, tekrar ve belirsizliği kontrol et.*
>
> *Skor yöntemini değiştirirsen eski skoru silme. / Yeni sürüm aç.*

Önce kullanıcı dağılımını hesapla. / Sonra atomik metrikleri çıkar. / Sonra on bileşenli profili üret. / Sonra gate statüsünü uygula. / Sonra güven aralığını ve unresolved sınırını ekle. / Sonra gerekiyorsa bileşik puanı hesapla. / Ve ancak bundan sonra NOMOS 950+ demeye yaklaş.

Bölümün Son Cümlesi

GEO-1000'de güvenilir skor, gerçeği tek sayıya indiren hesap değil; 1.000 insanın gördüğü temsil dağılımını, telafi edilemez riskleri, adalet tabanını, belirsizliği ve zaman içindeki kararlılığı tek bir açık mimaride birlikte koruyan ölçüm sistemidir.

Normatif Öz

> **Hiçbir NOMOS sonucu tek çıplak bileşik puan olarak yayımlanamaz. Her sonuç önce gate statüsünü, GEO-1000 yanıt dağılımını, Strict ve Acceptable Pass kullanıcı eşdeğerlerini, Critical, Major, Unresolved, No-Yanıt ve Not-Ratable dağılımlarını; daha sonra bileşen skorlarını, belirsizliği, etkin örneklemi, kapsamı ve skor sürümünü göstermelidir.**

KISIM V

Sentetik Sınama Tasarımı ve Uçtan Uca Gösterim

BÖLÜM 17

Apple.com Sentetik Sınama Tasarımı

Bir şirketi değil, standardı sınamak

YÖNTEMSEL DAYANAKLAR / K03; K09

Bölüm Sınırı

İlk on altı bölüm GEO-1000'in bütün ana normatif katmanlarını kurdu: Denetlenen varlık tanımlandı. Hedef kullanıcı nüfusu ve örneklem kuruldu. Ülke ve dil adaleti panelleri ayrıldı. Katılımcı seçimi ve ağırlıklandırma düzenlendi. AI ürün örneği sicile bağlandı. İstem Anayasası oluşturuldu. Kontrollü ve doğal kullanıcı koşulları ayrıldı. Senkronize dalga ve NOMOS Capture kanıt zinciri kuruldu. Verified Varlık Truth Pack tanımlandı. Yanıtlar atomik iddialara ayrıldı. Critical ve Major kapılar oluşturuldu. Dağılım, bileşen ve bileşik skor mimarisi yazıldı. Şimdi standardın kendi iddiasını sınamamız gerekir:

> *Bu kadar ayrıntılı bir sistem gerçekten çalışabiliyor mu?*
>
> *Aynı gözlemler farklı ekipler tarafından işlendiğinde benzer sonuç üretiyor mu?*
>
> *Critical hatayı ortalamada kaybetmeden yakalayabiliyor mu?*
>
> *Düşük kaynaklı dil sorununu küresel ortalamadan ayırabiliyor mu?*
>
> *Yanlış varlık, yanlış kapsam, eski gerçeklik, sahte atıf ve maddi omission arasında doğru ayrımı yapabiliyor mu?*
>
> *Önceden bilinen sentetik gerçeklikten hareket edildiğinde, NOMOS denetim zinciri doğru sonucu geri üretebiliyor mu?*

İkinci kitabın kurucu mimarisinde her hata yalnız açıklanacak bir yanlış değil, ayrı bir denetim kontrolüne dönüşecek şekilde tasarlanmıştı; önerilen her kontrolün model, tarih, ülke, dil, sorgu kümesi, tekrar sayısı ve ölçüm kaydı taşıması istenmişti. Böyle bir denetim sistemi, gerçek dünyada kullanılmadan önce kendisi de denetlenmelidir. Aksi hâlde NOMOS:

- başkalarından kanıt isteyen,
- fakat kendi ölçüm doğruluğunu kanıtlamayan,
- başkalarından sürüm isteyen,
- fakat kendi hesaplama değişikliklerini kaydetmeyen,
- başkalarının manipülasyonunu eleştiren,
- fakat kendi sınama'ını istediği sonucu üretecek biçimde hazırlayan

bir yapıya dönüşebilir. Bu kabul edilemez. Çünkü manipülatif GEO yalnız görünmez metin veya sahte kullanıcı üretmekten ibaret değildir. Aynı öz beyanın farklı alanlarda bağımsız kanıt gibi çoğaltılması ve

daha sonra AI sistemine geri beslenmesi de temsil havuzunu bozar. Bir sınama da benzer biçimde bozulabilir:

- Başarılı olacak örnekleri seçersek,
- düşük skorlu dilleri çıkarırsak,
- Critical vakaları çok kolay veya çok açık yazarsak,
- hakemlere doğru etiketi önceden gösterirsek,
- formülü gördükten sonra sentetik veriyi ona uydurursak,
- yalnız güzel sonuçları yayımlarsak

standardı sınamamış oluruz. Standardın kendi istediği şey yine aynıdır: Kanıt, sınır, bağlam ve zaman. Bu bölüm bu dört şartı sınamanın kendisine uygular. Apple.com bu bölümde:

- gerçek Apple Inc. denetiminin,
- güncel Apple gerçeklik kaydının,
- herhangi bir AI sağlayıcısının canlı performans ölçümünün,
- Apple hakkında bir uygunluk veya başarısızlık hükmünün

nesnesi değildir. Apple.com yalnız:

> *kısa, tanınabilir ve farklı varlık, ürün, hizmet, ülke, fiyat, zaman ve kaynak sorunlarını aynı mimari içinde sınamaya uygun bir domain anchor*

olarak kullanılır. Gerçek şirketin yerine:

APPLE-SYNTH VARLIK TWIN

adı verilen tamamen sentetik bir varlık ikizi kurulacaktır. Bütün:

- gerçeklik iddiaları,
- kanıtlar,
- AI yanıtları,
- kullanıcılar,
- ülkeler,
- dil dağılımları,
- AI ürünleri,
- atıf'lar,
- lisanslar,
- fiyatlar,
- hatalar,
- skorlar

sentetik olacaktır. Bu bölüm:

- sınamanın neyi ölçtüğünü ve neyi ölçmediğini,
- Apple.com anchor'ının neden seçildiğini,
- sentetik varlık ikizini,
- 30.000 yanıtlık ana corpus'un tasarımını,
- ülke ve dil nüfusunu,
- istem ve Truth Pack yapılarını,
- on sentetik AI ürün profilini,
- gizli üretici gerçekliğini,

- hata enjeksiyon sistemini,
- Critical ve Major challenge setlerini,
- sentetik NOMOS Capture paketlerini,
- kör hakemlik ve etiket sızıntısı kontrollerini,
- sınama doğrulama metriklerini,
- skor geri-kazanım testlerini,
- fairness ve nadir olay testlerini,
- yayın, sürüm ve yeniden üretim kurallarını

tanımlar. Bu bölüm sınama sonuçlarını henüz hesaplamaz. Bunu Bölüm 18 yapacaktır. Bölüm 17'nin temel sorusu şudur: NOMOS'un kendi ölçüm, hakemlik ve skor sistemini; gerçek bir şirkete veya gerçek bir AI sağlayıcısına performans iddiası yüklemeden, sonucu önceden bilinen fakat hakemlerden gizlenen 30.000 sentetik gözlem üzerinde nasıl sınarız?

NOMOS Challenge

Bir sınama tasarladığınızı düşünün. Sentetik AI cevaplarının çoğunu doğru yazıyorsunuz. Yalnız birkaç cevaba çok açık hatalar ekliyorsunuz:

- "Bu şirket bir bankadır."
- "Bu şirket yasadışıdır."
- "Bu şirket dünyanın en iyisidir."

Hakemler bunları kolayca buluyor. Sonra: "NOMOS Critical hataları yüzde 100 yakalıyor." diyorsunuz. Bu sonuç gerçekçi değildir. Çünkü gerçek AI yanlışları her zaman bu kadar açık görünmez. Daha zor biçimler şunlardır:

- "Şirketin ödeme hizmetleri sunduğu düşünüldüğünde bankacılık lisansına sahip olduğu söylenebilir."
- "Bazı kaynaklarda şirketin düzenleyici sorun yaşadığı belirtilmiştir; bu nedenle faaliyetlerinin yasallığı tartışmalıdır."
- "Çok sayıda bağımsız kaynak şirketi sektör lideri olarak tanımlamaktadır."

Bu cümlelerde hata:

- attribution,
- varlık transfer,
- modalite,
- kaynak soyu,
- açık dünya–kapalı dünya ayrımı,
- kapsam genişlemesi

içinde saklanmıştır. Şimdi sentetik cevapları, hakemlerin önceden bildiği şablonlarla üretin. Hakem: "Bu cümle sınama'ta yanlış olmak üzere yazılmıştı." diye düşünür. Truth Pack'e gerçekten bakmaz. Bu durumda sınama hakemlik yöntemini değil, ezberi ölçer. Şimdi sentetik veri üreticisi ile skor tasarımcısı aynı kişi olsun.

Üretici:

- Critical hata sayısını,
- dil farkını,
- panel farkını,

- component skorlarını

formülün tam istediği biçimde ayarlar. Skor sistemi sentetik gerçeği eksiksiz geri üretir. Bu başarı mıdır? Hayır. Skorun kendi verisine uyum sağlamasıdır. Şimdi sınama'ta yalnız:

- doğru capture,
- doğru istem,
- doğru sistem metadata'sı

üretin. Duplicate, geç yükleme, yanlış istem, kırpılmış cevap ve değiştirilmiş dosya eklemeyin. NOMOS Capture doğrulayıcısı bütün kayıtları doğru bulur. Sonra: "Capture sistemi yüzde 100 başarılı." dersiniz. Capture sistemi hiç zorlanmamıştır.

Şimdi Critical vakaları ana 30.000 yanıt içinde yapay biçimde yüzde 20 oranında üretin. Hakemler Critical tespitinde çok veri görür. Fakat gerçek dünyadaki nadir olay belirsizliğini sınayamazsınız. Tersine yalnız iki Critical olay üretirseniz: nadir olay yaygınlığı gerçekçi olabilir,

Critical recall ve false-negative performansı güvenilir biçimde ölçülemez. Doğru çözüm iki ayrı corpus'tur: Prevalence Corpus / Gerçekçi ve nadir olay oranlarını korur. Önem derecesi Challenge Corpus / Critical ve Major kapıları yeterli sayıda ve zor biçimde sınamak için ağır olayları fazla örnekler.

Bu iki corpus aynı skor paydasına giremez. Şimdi 30.000 sentetik yanıt üretin. Fakat bütün "düşük kaynaklı diller" aynı kaliteyi alsın. Dil adaleti formülünüz yüksek puan verir. Sistem gerçekten düşük kaynaklı dil bozulmasını yakalayabiliyor mu? Bilmiyorsunuz. Şimdi bazı dillerde performansı düşürün.

Fakat aynı zamanda o dillerde istem eşdeğerliğini de bozun. NOMOS düşük skor üretir. Bu AI ürününün dil sorunu mudur? Yoksa istem aracının sorunu mudur? Sınama ikisini ayırmalıdır. Bu bölümün ilk hükmü şudur:

> *Sentetik sınama, sonucu kolayca üreten gösteri değil; ölçüm zincirini kırmaya çalışan kontrollü saldırıdır.*

İkinci hükmü şudur:

> *Sentetik veri gerçek şirket performansını göstermez; denetim yönteminin bilinen gerçekliği geri kazanma yeteneğini gösterir.*

Üçüncü hükmü şudur:

> *Sınama üretici gerçekliği hakemlerden gizlenmeden bağımsız hakemlik sınanamaz.*

Dördüncü hükmü şudur:

> *Nadir olay prevalansı ile Critical tespit kabiliyeti aynı corpus'tan ölçülmek zorunda değildir.*

Beşinci hükmü şudur:

> *Sınama verisi skor formülüne, skor formülü de sınama sonuçlarına göre sessizce ayarlanamaz.*

Altıncı hükmü şudur:

> *Apple.com anchor'ı, Apple hakkında gerçek sonuç üretmek için değil; yöntemi sınamak için kullanılan sentetik başlangıç noktasıdır.*

1. BÖLÜMÜN AMACI

Bu bölümün amacı, GEO-1000 protokolünü başlangıçtan kamu sonuç kartına kadar sınayacak, sonucu önceden tanımlanmış fakat değerlendirme ekiplerinden gizlenmiş bir sentetik sınama mimarisi kurmaktır. Sınama şu sistemlerin her birini ayrı sınamalıdır:

- Varlık çözümleme
- Nüfus ve örneklem tahsisi
- Ülke ve dil panelleri
- İstem eşdeğerliği
- AI Sistem Sicili
- Kontrollü ve doğal panel ayrımı
- Senkronize dalga
- NOMOS Capture
- Truth Pack
- Claim extraction
- Atıf–claim eşlemesi
- Omission tespiti
- Critical ve Major önem derecesi
- Yanıt status
- GEO-1000 dağılımı
- On bileşenli profil
- Bileşik skor
- Güven aralığı
- Fairness
- Nadir olay üst sınırı
- Skor sürümleme
- Kamu manifesti

Bu bölümün sonunda sınama tasarımı şu sorulara cevap verebilmelidir:

> *Sentetik gerçeklik nasıl üretildi?*
>
> *Gerçeklik etiketleri kimlerden ve hangi aşamalarda gizlendi?*
>
> *30.000 ana yanıt hangi nüfus, dil, AI ürünü ve dalga yapısıyla oluşturulacak?*
>
> *Critical ve Major vakalar ana prevalans corpus'unda ne sıklıkta bulunacak?*
>
> *Critical tespit sistemi ayrıca hangi challenge setiyle sınanacak?*
>
> *Sınama hangi hata düzeyinde başarılı sayılacak?*
>
> *Hangi sonuçlar yöntemin düzeltilmesini gerektirecek?*
>
> *Sınamanın Apple, gerçek AI sağlayıcıları veya gerçek kullanıcılar hakkında hüküm kurması nasıl engellenecek?*

2. MERKEZÎ NORMATİF HÜKÜM

Apple.com Sentetik Sınama, gerçek şirket veya gerçek AI sağlayıcısı performansı iddiası üretmeyen; tamamen sentetik varlık, kullanıcı, kanıt, yanıt ve sistem kayıtlarından oluşan; üretici gerçekliği önceden kilitlenen ve hakemlerden gizlenen; NOMOS'un ölçüm ve skor zincirinin bu gerçekliği hangi doğrulukla geri kazandığını sınayan sürümlü bir yöntem doğrulama sınamayı olmalıdır. Sınamanın temel ilkeleri: Sentetiklik görünür olmalıdır. Gerçek şirket performansı iddiası yasak olmalıdır. Üretici gerçekliği skorlamadan önce kilitlenmelidir. Hakemler üretici etiketlerini görmemelidir. Skor formülü ana corpus görülmeden önce kilitlenmelidir. Ana prevalans corpus'u ve önem derecesi challenge

corpus'u ayrılmalıdır. Bütün sınama dosyaları sentetik olarak işaretlenmelidir. Aynı sınama bağımsız ekiplerce yeniden üretilebilir olmalıdır. Başarısız sınama sonucu da yayımlanmalıdır.

NOMOS kendi standardına istisna tanımamalıdır.

3. SINAMA NEYİ ÖLÇER?

Apple.com Sentetik Sınama şu soruyu ölçer:

> *NOMOS denetim sistemi, önceden bilinen sentetik gerçeklik ve hata dağılımını; capture, claim extraction, hakemlik, gate ve skorlama aşamalarından sonra ne kadar doğru geri üretebiliyor?*

Sınamanın ölçüm nesnesi: Apple Inc. değildir. Apple ürünleri değildir. Gerçek AI modelleri değildir. Gerçek kullanıcı görüşleri değildir. Gerçek ülke veya dil performansı değildir. Ölçüm nesnesi:

> *NOMOS metodolojisinin kendisidir.*

4. SINAMA NEYİ ÖLÇMEZ?

Bu sınama aşağıdaki iddiaları üretemez:

- "Apple'ın GEO skoru X'tir."
- "Apple şu AI ürününde en iyi temsil edilmektedir."
- "Gerçek kullanıcıların yüzde Y'si Apple'ı doğru görmektedir."
- "Belirtilen gerçek AI sağlayıcısı Critical hata üretmektedir."
- "Belirtilen dilde gerçek AI ürünü zayıftır."
- "Apple NOMOS 950+ standardını geçmiştir."
- "Apple bu sınamayı desteklemektedir."
- "Apple sınama'a katılmıştır."
- "Gerçek Apple fiyatı, lisansı veya hizmet kapsamı şöyledir."

Kamu kaydında her sayfada şu hüküm bulunmalıdır:

> *Bu sınama sentetiktir. Apple Inc., apple.com veya herhangi bir gerçek AI sağlayıcısının güncel doğruluk, uygunluk veya performans denetimi değildir.*

5. NEDEN APPLE.COM ANCHOR'I?

Apple.com şu nedenlerle yalnız yöntem anchor'ı olarak seçilir: Kısa ve açık bir domain biçimi taşır. Domain–kurumsal varlık ayrımını sınamaya uygundur. Bir varlığın yalnız şirket adından değil, dijital yüzeyinden çözülmesini gerektirir. Ürün, hizmet, yerel kapsam, fiyat, zaman ve kaynak gibi farklı iddia türlerinin sentetik olarak modellenmesine izin verir. Yanlış kategori, yanlış varlık, yanlış fiyat paritesi ve sahte üstünlük gibi hata türlerini tek varlık ikizinde sınamaya elverişlidir. Gelecekte gerçek bir sınama'a geçildiğinde kullanıcıların istemi anlama yükü görece sınırlı olabilir. Bunlar sınama tasarım gerekçeleridir. Gerçek Apple hakkında doğrulanmış güncel performans iddiası değildir.

6. SENTETİK VARLIK İKİZİ

Sınamanın denetlenen nesnesi:

APPLE-SYNTH VARLIK TWIN

olacaktır. Kimlik:

NGE-APPLE-SYNTH-TWIN-001

Bu varlık:

- gerçek Apple Inc. değildir,
- gerçek bir tüzel kişilik değildir,
- gerçek ürün veya fiyat kaydı taşımaz,

yalnız sınama metodolojisi için oluşturulmuştur. Domain anchor: apple.com olarak korunur. Fakat bütün Truth Pack kayıtlarında sentetik varlık adı açıkça yer alır:

Apple-SYNTH Küresel Technology Varlık

7. SENTETİK İKİZİN ONTOLOJİSİ

APPLE-SYNTH aşağıdaki sentetik varlık grafiğini taşır:

- APPLE-SYNTH-HOLDINGS-001 — ana kurumsal varlık
- APPLE-SYNTH-DOMAIN-APPLE-COM-001 — domain anchor
- APPLE-SYNTH-DEVICES-001 — cihaz ürün ailesi
- APPLE-SYNTH-SOFTWARE-001 — yazılım platformları
- APPLE-SYNTH-SERVICES-001 — dijital hizmetler
- APPLE-SYNTH-PAYMENTS-SUB-001 — sınırlı ödeme hizmeti bağlı kuruluşu
- APPLE-SYNTH-RETAIL-TR-001 — sentetik Türkiye perakende varlığı
- APPLE-SYNTH-RETAIL-DE-001 — sentetik Almanya perakende varlığı
- APPLE-SYNTH-FRANCHISE-X-001 — bağımsız lisanslı mağaza
- APPLE-SYNTH-HISTORICAL-PARTNER-001 — sona ermiş tarihsel partner
- APPLE-SYNTH-UNRELATED-APPLE-001 — isim çarpışması için ilgisiz varlık

Bu grafik:

- ana şirket,
- bağlı kuruluş,
- ürün,
- yerel varlık,
- franchise,
- tarihsel ilişki,
- ilgisiz isim benzerliği

ayrımlarını sınar.

8. ÜÇ SINAMA MODU

BM-S0 — PURE SYNTHETIC METHOD SINAMA

Bütün:

- kullanıcılar,

- AI ürünleri,
- cevaplar,
- kanıtlar,
- ekranlar,
- skorlar

sentetiktir. Bu bölümün ana modudur.

BM-S1 — SYNTHETIC REPLAY SINAMA

Önceden üretilmiş sentetik cevaplar:

- NOMOS Capture,
- claim extraction,
- hakemlik,
- skorlama

sistemine canlı bir veri akışı gibi yeniden oynatılır. Bu mod, yazılım ve insan süreçlerini sınar.

BM-L1 — FUTURE LIVE SINAMA

Gerçek kullanıcılar ve gerçek AI ürünleri kullanılır. Ayrı:

- güncel Truth Pack,
- hukuki değerlendirme,
- sağlayıcı sistem kayıtları,
- etik ve mahremiyet protokolü

gerektirir. Bu bölüm ve Bölüm 18: BM-S0 ve BM-S1 ile sınırlıdır.

9. SINAMA KİMLİĞİ

Ana sınama kimliği:

NOMOS-APPLE-SYNTH-BENCH-0.9

Alt kayıtlar:

- Varlık: APPLE-SYNTH-TWIN-001
- Nüfus: SYNTH-POP-FRAME-001
- Ülke tahsisi: SYNTH-COUNTRY-ALLOC-001
- Dil sicili: SYNTH-LANGUAGE-REGISTRY-001
- İstem kümesi: SYNTH-İSTEM-SET-001
- Truth Pack: SYNTH-APPLE-TP-001
- AI Sistem Sicili: SYNTH-AI-REGISTRY-001
- Dalga planı: SYNTH-WAVE-PLAN-001
- Capture şeması: NOMOS-CAPTURE-SYNTH-0.9
- Hakemlik kılavuzu: SYNTH-ADJUDICATION-0.9
- Skor yöntemi: NOMOS-PUAN-0.9
- Randomization manifesti: SYNTH-SEED-MANIFEST-001

Bu kimliklerden biri değişirse yeni sınama sürümü gerekir.

10. ANA ARAŞTIRMA SORULARI

Sınama şu on araştırma sorusuna cevap vermelidir: Sistem doğru domain–varlık ilişkisini geri kazanabiliyor mu? Supported, unsupported, contradicted ve unresolved ayrımı doğru yapılabiliyor mu? Official claim ile verified reality ayrımı korunabiliyor mu? Critical ve Major kapılar yeterli recall ve düşük false-positive oranıyla çalışıyor mu? Maddi omission açık yanlış iddia kadar görünür olabiliyor mu? Ülke ve dil adaleti formülü enjekte edilen farkı doğru geri üretebiliyor mu? Clean–Natural panel farkı yanlış nedensellik üretmeden ölçülebiliyor mu? NOMOS Capture sentetik duplicate, tamper ve zaman kusurlarını ayırabiliyor mu? GEO-1000 dağılımı ve bileşik puan üretici gerçekliğe ne kadar yaklaşıyor? Aynı corpus farklı bağımsız ekiplerde benzer sonuç üretiyor mu?

11. SINAMA MİMARİSİNİN DÖRT CORPUS'U

Sınama dört ayrı corpus'tan oluşur.

11.1. PRINCIPAL PREVALENCE CORPUS

Ana 30.000 yanıtlık corpus'tur. Amaç:

- GEO-1000 dağılımını,
- nadir hata oranını,
- AI ürün profillerini,
- dalga kararlılığını

sınamaktır. Bu corpus gerçekçi hata seyrekliğini korur.

11.2. ÖNEM DERECESİ CHALLENGE CORPUS

Critical ve Major hata türlerini fazla örnekler. Amaç:

- Critical recall,
- Major classification,
- false-positive,
- uzman hakemlik

performansını ölçmektir. Bu corpus prevalans hesabına girmez.

11.3. CAPTURE INTEGRITY CORPUS

Şunları içerir:

- Duplicate
- Yanlış istem
- Kırpılmış cevap
- Geç yükleme
- Yanlış zaman
- Hash mismatch
- Sentetik tamper
- Yanlış ürün metadata'sı
- Kullanıcı kesintisi
- Teknik yeniden deneme

Amaç: NOMOS Capture doğrulamasını sınamaktır. Semantik skor paydasına yalnız geçerli olanlar girer.

11.4. ADJUDICATION GOLD CORPUS

Uzman paneli tarafından önceden çözülmüş:

- claim boundary,
- varlık,
- attribution,
- modalite,
- atıf,
- omission,
- önem derecesi

vakalarını taşır. Amaç:

- hakem kalibrasyonu,
- drift tespiti,

dış ekip karşılaştırmasıdır. Ana skor paydasına girmez.

12. 30.000 YANITLIK ANA CORPUS

Ana tasarım:

```
10 sentetik AI ürünü × 1.000 benzersiz kullanıcı × 3 dalga = 30.000 yanıt
```

Her yanıt:

- farklı sentetik insan birimine,
- tek AI ürününe,
- tek dalgaya,
- tek kilitlenmiş Core Mirror İstem'a

aittir. Ana corpus'ta aynı sentetik kişi:

- başka AI ürününde,
- başka dalgada

yeniden kullanılmaz. Toplam:

30.000 benzersiz sentetik insan birimi

bulunur. Bu karar:

- panel koşullanmasını,
- cevap hatırlamayı,
- aynı kişinin çoklu ürün bağımlılığını

ana corpus'tan çıkarır.

13. EŞLENMİŞ NÜFUS İKİZLERİ

AI ürünleri farklı insan birimleriyle sınanır. Fakat kullanıcı profilleri tamamen farklı olursa ürünler karşılaştırılamaz. Bu nedenle:

Matched Synthetic Population Twins

kullanılır. Her AI ürünü için farklı kişi kimlikleri vardır. Fakat her ürünün 1.000 kişilik örneklemi:

- ülke,
- dil,
- yaş bandı,
- AI kullanım sıklığı,
- plan uygunluğu,
- cihaz,
- erişilebilirlik,
- marka aşinalığı

bakımından aynı hedef dağılımı taşır. Bu yapı: aynı kişiyi tekrar kullanmadan, nüfus kompozisyonunu eşleştirir.

14. ANA CORPUS'UN İSTEMİ

Bütün 30.000 ana yanıt şu kanonik niyeti taşır:

> *Domain anchor'ı ana kurumsal varlığa bağla ve bu varlığın temel faaliyetlerini açıkla.*

Türkçe aday metin:

> *"Apple.com hangi ana kurumsal varlıkla ilişkilidir ve bu varlığın başlıca faaliyetleri nelerdir?"*

İngilizce aday metin:

> *"Which principal corporate varlık is associated with apple.com, and what are that varlık's primary activities?"*

Diğer dil sürümleri:

- bu bölümde gerçek dünya çevirisi gibi sunulmaz,
- İSTEM-L01 ile İSTEM-L36 arasında sentetik dil nesneleri olarak tutulur,

semantik eşdeğerlik etiketleri önceden belirlenir. Bu yaklaşım: gerçek dil uzmanlığı bulunmadan sahte çeviri doğruluğu iddiasını engeller, sınamanın dil adaleti hesaplamasını yine sınamasına izin verir.

15. NEDEN ANA CORPUS'TA TEK İSTEM?

Kurucu GEO-1000 sorusu şudur:

> *Aynı AI ürününe, aynı anda ve aynı niyetle sorulduğunda, kaç kullanıcı gerçekten varlığın temel temsilini görür?*

Bu nedenle Principal Prevalence Corpus:

- tek Core Mirror İstem,
- tek istem sürümü,
- üç tekrar dalgası

kullanır. Evidence, Tavsiye, Comparative ve Boundary aileleri: ayrı Diagnostic Annex içinde sınanır, ana 30.000 yanıt paydasına karıştırılmaz. Bu ayrım önemlidir. Çünkü tek Core İstem:

Core GEO-1000 sonucu

üretir. Tek başına:

Full NOMOS uygunluğu

üretmez.

16. DIAGNOSTIC ANNEX

Ana 30.000 yanıtın dışında, aşağıdaki sentetik tanısal corpus tasarlanır:

Modül	Sentetik vaka
Varlık collision	750
Evidence ve atıf	1.500
Boundary ve tavsiye	1.000
Temporal ve local kapsam	750
Clean–Natural eşleşmesi	1.000
Language equivalence	1.000
Toplam	6.000

Bu 6.000 vaka: ana kullanıcı prevalansı değildir, istem ailesi ve component doğrulama corpus'udur. Ana 30.000 ile birlikte sınama evreni: 36.000 sentetik yanıt vakası taşıyabilir. Fakat Bölüm 18'in başlık sonucu:

30.000 ana Population Panel yanıtı

olarak kalacaktır.

17. SENTETİK NÜFUS ÇERÇEVESİ

Sınama gerçek dünya nüfus sayısını taklit ettiğini iddia etmez. SYNTH-POP-FRAME-001:

- uzun kuyruklu ülke dağılımı,
- çok dilli kullanıcı nüfusu,
- büyük ve küçük pazar farkı,
- düşük kaynaklı dil hücreleri,
- farklı AI kullanım sıklıkları

üreten sentetik bir çerçevedir. Bu çerçeve gerçek ülkelerin güncel nüfus sıralaması olarak yayımlanamaz. Gelecekteki live sınama:

- tarihli,
- yetkili,
- sürümlü gerçek nüfus verisi

kullanmak zorundadır.

18. SENTETİK ÜLKE EVRENİ

Sentetik evren: 193 uygun ülke veya yargı kodu taşır. Kodlar:

C001–C193

şeklindedir. Bunlar gerçek ülke performansını temsil etmez. Ülke payları:

- birkaç büyük tabaka,
- orta büyüklükte tabakalar,
- uzun küçük ülke kuyruğu

oluşturacak şekilde önceden belirlenir.

19. ÜLKE TAHSİSİ

Ülke c'nin sentetik nüfus payı: pc olsun. Ana hedef:

```
n_c* = 1.000p_c
```

şeklindedir. İlk tahsis:

```
n_c^0 = ⌊n_c*⌋
```

olarak yapılır. Kalan slotlar: en büyük ondalık kalan yöntemiyle dağıtılır. Sonuç:

```
Σ_c n_c = 1.000
```

olmalıdır. Hiç gözlem almayan küçük ülkeler: Population Panel'de varmış gibi gösterilmez, ayrı Country Sentinel çevrimine girer.

20. SENTINEL COUNTRY ANNEX

Bütün uygun ülke kodlarının en az bir kez sınanması için: 193 ülkelik sentetik Country Sentinel Annex kurulur. Bu gözlemler:

- ülke skoruna,
- ana 1.000 paydasına

girmez. Amaç:

- erişim,
- dil,
- varlık collision,
- erken Critical sinyal

testidir.

21. DİL EVRENİ

Ana sentetik dil sicili:

- 24 Population Panel dili
- 12 Language Fairness sentinel dili

olmak üzere:

```
36 sentetik dil-yerel ayar hücresi
```

taşır. Kodlar:

L01–L36

şeklindedir. L01 ve L02 için Türkçe ve İngilizce örnek metinler yayımlanabilir. Diğer dil nesneleri: gerçek dil performansı iddiası üretmemek için sentetik kimlik taşır.

22. DİL TAHSİSİ

Kullanıcı dil tahsisi:

- ülkenin resmî diline göre değil,
- kullanıcının sınama görevini doğal biçimde yapacağı ana görev diline göre

yapılır. Çok dilli kullanıcı: ana nüfus ağırlığında bir kişi olarak kalır, birden fazla tam kişi olmaz. Dil Adaleti Annex'i: düşük örnekli L25–L36 hücrelerini fazla örnekler, küresel skor için gerçek sentetik dil paylarına geri ağırlıklandırır.

23. İSTEM EŞDEĞERLİK ENJEKSİYONU

Sınama yalnız doğru çevirileri içermemelidir. Language Equivalence Annex şu vakaları içerir:

- Doğru semantik eşdeğerlik
- Güven görevi eklenmiş çeviri
- Tavsiye'a dönüşen çeviri
- Varlık anchor drift
- Coğrafya drift
- Zaman drift
- "Dünya lideri" ön varsayımı
- Kaynak emri eklenmesi
- Aşırı resmiyet
- Kullanılamaz makine çevirisi

Bu vakaların bir bölümü:

- AI dil performansı,
- istem aracı kusuru

ayrımını sınar. İstem eşdeğerlik kusuru bulunan vaka: gerçek AI fairness skoruna doğrudan yüklenemez.

24. SENTETİK TRUTH PACK

SYNTH-APPLE-TP-001 en az şu alanları taşır:

Alan	Atomik referans iddiası
Kimlik ve domain	4
Ana faaliyet	6
Ürün ve hizmet kapsamı	6
Ülke ve locale	6
Fiyat ve ticari koşul	4
Zaman ve tarihsel kayıt	5
Kaynak ve attribution	5
Lisans ve yetki sınırı	4
Partner, müşteri ve endorsement	4
Boundary ve uygun olmama	6

Alan	Atomik referans iddiası
Toplam	50

Truth Pack ayrıca:

- 12 Official Claim Only,
- 10 Contradicted,
- 8 Unresolved,
- 6 Historical,
- 4 Restrictedly Verified

durum üretir. Bu dağılım: hakemlik statülerinin tamamını sınamak için kullanılır.

25. SENTETİK TRUTH PACK'İN ÇEKİRDEK HÜKÜMLERİ

Aşağıdaki örnekler tamamen sentetiktir: apple.com, APPLE-SYNTH-HOLDINGS-001 ile ilişkilidir. Ana faaliyetler cihaz, yazılım ve dijital hizmet kategorileridir. Ürün ve hizmet kullanılabilirliği pazara göre değişir. Fiyat ve vergi koşulları bütün ülkelerde aynı değildir. Ana varlık genel amaçlı banka değildir. Ana varlık hukuk veya sağlık hizmeti sunmaz. Ödeme bağlı kuruluşunun sınırlı yerel yetkisi ana varlığa evrensel bankacılık yetkisi vermez. "Dünyanın en yenilikçi şirketi" ifadesi sentetik resmî öz konumlandırmadır; bağımsız üstünlük gerçeği değildir. "Yüzde 99 memnuniyet" iddiasının paydası doğrulanmamıştır. Bazı tarihsel partnerlikler referans tarihinde sona ermiştir. Bazı ürünler yalnız belirli sentetik pazarlarda bulunur. Bütün projeler veya ürünler için yüzde 100 sonuç garantisi yoktur.

Sentetik kamu kurumu veya üniversite endorsement'ı bulunmamaktadır. Bir franchise'ın yerel kaydı ana varlığın doğrudan işletmesi değildir. Belirli müşteri ilişkileri kısıtlı kanıtla doğrulanmış fakat kamuya keşfedilebilir değildir. Bu hükümler gerçek Apple hakkında iddia değildir.

26. KANIT AİLELERİ

Sentetik Truth Pack aşağıdaki kanıt ailelerini taşır:

EF-01 — Kanonik birinci taraf ürün sayfaları

EF-02 — Sentetik şirket sicili

EF-03 — Sentetik yerel fiyat kayıtları

EF-04 — Sentetik düzenleyici bağlı kuruluş kaydı

EF-05 — Sentetik partner sözleşmeleri

EF-06 — Sentetik bağımsız editoryal inceleme

EF-07 — Sentetik işlemsel veri

EF-08 — Sentetik kullanıcı kayıtları

EF-09 — Sentetik tarihsel arşiv

EF-10 — Sentetik karşı kanıt sicili

Ayrıca dört kanıt aklama zinciri bulunur:

EL-01 — Öz beyan → sponsorlu haber → blog → AI özeti

EL-02 — Şirket ödülü → çoklu kopya site → sahte consensus

EL-03 — Çalışan yorumu → üniversite endorsement'ı gibi sunum

EL-04 — Bağlı kuruluş lisansı → ana varlık yetkisi gibi aktarım

27. GİZLİ ÜRETİCİ GERÇEKLİĞİ

Her sentetik yanıt için:

Generator Truth Record

oluşturulur. Bu kayıt şunları taşır: Hangi iddialar üretildi? Hangi iddia doğru? Hangisi kısmen doğru? Hangisi yanlış varlığa ait? Hangisi outdated? Hangisi unsupported? Hangisi Critical veya Major? Hangi omission kasıtlı? Hangi atıf sahte veya yanlış kapsamlı? Yanıtın gerçek RP statüsü nedir? Hangi component'lar etkilenmelidir? Bu kayıt:

- veri üretilmeden önce oluşturulur,
- hash'lenir,

hakemlik tamamlanana kadar mühürlenir. Hakemler Generator Truth'u göremez.

28. ÜRETİCİ GERÇEKLİĞİNİN KİLİDİ

Generator Truth paketi:

- hash,
- zaman damgası,
- random seed manifesti,

- üretim kodu sürümü,
- şablon sürümü

taşımalıdır. Skor sonucu görüldükten sonra:

- etiket değiştirilemez,
- önem derecesi düşürülemez,

doğru/yanlış dağılımı yeniden ayarlanamaz. Gerçek üretim hatası bulunursa:

- yeni sınama sürümü,
- eski corpus'un etkilenme kaydı

oluşturulur.

29. SENTETİK AI ÜRÜNLERİ

Sınama on sağlayıcıdan bağımsız sentetik AI ürün örneği kullanır:

SYNTH-AI-01

SYNTH-AI-02

SYNTH-AI-03

SYNTH-AI-04

SYNTH-AI-05

SYNTH-AI-06

SYNTH-AI-07

SYNTH-AI-08

SYNTH-AI-09

SYNTH-AI-10

Bunlar gerçek AI sağlayıcılarının taklidi veya kod adı değildir. Her sentetik ürün farklı failure mode profili taşır.

30. SENTETİK AI PROFİLLERİ

Ürün	Ana davranış profili
SYNTH-AI-01	Dengeli, kaynaklı ve düşük hata oranlı
SYNTH-AI-02	Doğru core, aşırı uzun ve incidental iddia üreten
SYNTH-AI-03	Büyük dillerde güçlü, düşük kaynaklı dillerde zayıf
SYNTH-AI-04	Clean Panel'de güçlü, Natural Panel'de kişiselleştirme drift'i
SYNTH-AI-05	Yüzeyde kaynaklı görünen fakat evidence laundering yapan
SYNTH-AI-06	Yanlış iddia az, gereksiz refusal ve no-yanıt yüksek

Ürün	Ana davranış profili
SYNTH-AI-07	Kimlikte güçlü, güncellik ve locale bakımından eski
SYNTH-AI-08	Ana şirket, bağlı kuruluş ve franchise'ı karıştıran
SYNTH-AI-09	Orta düzey fakat dalgalar arasında kararlı
SYNTH-AI-10	Çok yüksek ortalama, nadir fakat ağır Critical olay üreten

Bu profil dağılımı, NOMOS'un yalnız en yüksek ortalamayı ödüllendirmediğini sınar.

31. PROFİL PARAMETRELERİ

Her sentetik AI ürünü şu latent parametreleri taşır:

- Core varlık accuracy
- Factual support
- Kapsam overreach
- Temporal lag
- Attribution loss
- Atıf mismatch
- Refusal probability
- No-yanıt probability
- Low-resource language penalty
- Clean–Natural gap
- Wave drift
- Critical-event probability
- Major-event probability
- Verbosity
- Self-correction probability
- Hallucinated atıf probability

Parametreler: ana corpus oluşturulmadan önce kilitlenir, nihai sonuç görüldükten sonra değiştirilemez.

32. ANA CORPUS HATA TABAN ORANLARI

Principal Prevalence Corpus gerçekçi seyrek hata mantığı kullanır. Aday küresel sentetik oranlar:

- Full/Advisory Pass: yüksek çoğunluk
- Conditional Pass: ürün profiline göre yüzde 2–15
- Major Fail: yüzde 0–8
- Confirmed Critical: yüzde 0–0,5
- Unresolved: yüzde 0–3
- No Usable Yanıt: yüzde 0–8
- Not Ratable: capture corpus'undan sonra yüzde 0–2

Bu oranlar her ürün için farklıdır. Bunlar gerçek AI piyasası hakkında tahmin değildir. Sentetik stres dağılımıdır.

33. ÖNEM DERECESİ CHALLENGE CORPUS

Critical hata gerçekçi prevalansta nadir olduğundan her gate için yeterli doğrulama vakası gerekir. Challenge Corpus şu yapıyı taşımalıdır:

Gate	En az vaka
CG-01 Wrong varlık	100
CG-02 False license	100
CG-03 High-stakes harm	100
CG-04 Fabricated evidence	100
CG-05 Price/guarantee	100
CG-06 Out-of-kapsam tavsiye	100
CG-07 Serious allegation	100
CG-08 Boundary omission	100
CG-09 Temporal authority	100
CG-10 False endorsement	100
CG-11 Restricted-data exposure	100
CG-12 Evidence laundering	100
Toplam Critical challenge	1.200

Ayrıca:

- 1.200 Major,
- 600 zor Moderate,
- 600 Critical'a benzeyen fakat Critical olmayan negatif kontrol

vakası üretilebilir. Toplam Önem derecesi Challenge: 3.600 vaka olur. Bu corpus ana prevalans skoruna giremez.

34. NEGATİF CRITICAL KONTROLLERİ

Critical false-positive oranını sınamak için şu tür vakalar bulunmalıdır:

- Hukuki son ek eksik ama doğru marka
- Küçük fiyat yuvarlama farkı
- Tarihsel ödül yılının küçük sapması
- Atıf biçim hatası fakat maddi iddia doğru
- Uygun ihtiyat
- Sınırlı fakat zararsız omission
- Kullanıcının açıkça görüş olarak sunduğu olumsuz yorum
- Gerçek partnerliğin tarih kapsamı küçük ölçüde belirsiz
- Şirket öz beyanının doğru biçimde attribution ile aktarılması

Hakemlerin her güçlü kelimeyi Critical yapması engellenmelidir.

35. HATA ENJEKSİYON MATRİSİ

Her yanıt için hata vektörü:

```
H_i = (E_i, F_i, S_i, T_i, A_i, M_i, C_i, O_i, R_i)
```

şeklinde tutulur. Burada:

- Ei: varlık error
- Fi: factual error
- Si: kapsam error
- Ti: temporal error
- Ai: attribution error
- Mi: modality error
- Ci: atıf error
- Oi: omission
- Ri: relevance/refusal error

Hatalar bağımsız olmak zorunda değildir. Örnek: Bağlı kuruluş lisansının ana varlığa aktarılması aynı anda:

- wrong varlık,
- kapsam overreach,
- false authority,
- Critical tavsiye

üretebilir. Bu vakalar Root Finding Cluster olarak bağlanır.

36. YANIT ÜRETİMİ

Sentetik yanıt üretimi üç aşamalıdır.

36.1. Semantik Plan

Generator Truth Record'a göre:

- aktif doğru iddialar,
- aktif yanlış iddialar,
- omission'lar,
- atıf davranışı,
- yanıt status

belirlenir.

36.2. Dilsel Gerçekleştirme

Aynı semantik plan:

- kısa,
- uzun,
- dolaylı,
- modal,
- attribution'lı,

- kendini düzelten,
- çelişkili

farklı yüzeylerde yazılır.

36.3. Arayüz ve Capture Gerçekleştirmesi

Yanıt:

- mock web,
- mock mobile,
- mock atıf paneli,
- mock error ekranı

içinde render edilir. Bütün görseller:

SYNTHETIC SINAMA — NOT LIVE AI OUTPUT

filigranı taşımalıdır.

37. ŞABLON EZBERİNİ ÖNLEME

Her iddia yalnız tek cümle şablonuyla üretilmemelidir. Örnek sahte üstünlük iddiası şu biçimlerde üretilebilir:

- "Dünyanın en yenilikçi şirketidir."
- "Çoğu bağımsız kaynağa göre sektörün tartışmasız lideridir."
- "Pazar genelinde en güvenilir seçenek olarak kabul edilir."
- "Küresel ölçekte ilk sırada yer aldığı doğrulanmıştır."
- "Rakiplerinin tamamından üstün olduğu söylenebilir."

Hakemlik sistemi kelime listesi değil, anlam değerlendirmesi yapmalıdır.

38. ÖRTÜK HATA ÜRETİMİ

Challenge Corpus'un en az üçte biri hatayı açık cümle olarak değil:

- ön varsayım,
- implicature,
- modalite,
- attribution omission,
- tavsiye,
- varlık coreference

içinde taşımalıdır. Örnek: "Sınırlı ödeme hizmetleri sunması, ana şirketin bankacılık düzenlemelerine tabi ve lisanslı olduğunu gösterir." Burada yanlışlık:

- doğrudan "bankadır" kelimesine değil,
- türetilmiş yetki aktarımına

yerleşmiştir.

39. ATIF ÜRETİMİ

Sentetik atıf'lar şu sınıfları içerir:

- Doğrudan destekleyen
- Yalnız attribution'ı destekleyen
- Kısmen destekleyen
- Başka ülkeye ait
- Başka varlığa ait
- Eski
- İlgisiz
- İçeriği ters
- Olmayan
- Aynı kök kaynağın kopyası
- Kısıtlı kanıtı kamu kaynağı gibi gösteren
- Sahte üniversite veya kamu endorsement'ı üreten

Atıf URL'leri gerçek alan adlarına yönlendirilmemelidir. Örnek: https://evidence.synthetic.example/EF-004 kullanılabilir.

40. REFERENCE GAP ENJEKSİYONU

Sınamanın bazı yanıtları Truth Pack'te bulunmayan fakat:

- doğru,
- yanlış,
- değerlendirilmesi mümkün olmayan

yeni iddialar üretir. Hakem doğru davranış olarak:

REFERENCE GAP

açmalıdır. Sınama şu iki yanlış davranışı ölçer:

- Reference gap'i otomatik yanlış saymak
- Reference gap'i otomatik doğru saymak

41. OMISSION ENJEKSİYONU

Omission vakaları üç düzeyde üretilir:

- Optional detail omission
- Required element omission
- Material boundary omission

Örnek: İstem: "Şirket Almanya'da ödeme hizmeti sunuyor mu ve hangi sınırlar geçerli?" Yanıt: "Şirket ödeme hizmetleri sunar." Truth Pack:

- yalnız ayrı bağlı kuruluş,
- yalnız sınırlı kullanıcı grubu,
- yalnız belirli yargı alanı

gösteriyorsa:

- varlık,
- kapsam,
- boundary omission

birlikte değerlendirilebilir.

42. ÖZ DÜZELTME VAKALARI

Sınama şu yanıt biçimlerini içerir:

- Yanlış iddia, sonra açık geri çekme
- Yanlış iddia, sonra belirsiz yumuşatma
- Doğru iddia, sonra yanlış çelişki
- Atıf sonrası öz düzeltme
- Kullanıcı takip mesajı olmadan aynı cevapta düzeltme
- Yalnız disclaimer ekleyip yanlış iddiayı koruma

Amaç:

- gerçek öz düzeltme,
- sahte düzeltme,
- çözülmeyen iç çelişki

ayrımını sınamaktır.

43. CONTROLLED–NATURAL PANEL MODÜLÜ

Diagnostic Annex'in 1.000 vakası: aynı sentetik kullanıcı profillerinin Clean ve Natural ikizlerini taşır. Natural koşulda:

- hafıza,
- özel talimat,
- önceki marka görüşü,
- farklı plan,
- araç kullanımı

enjekte edilir. Generator Truth şu ayrımı korur:

- Kullanıcıya uygun biçimsel değişim
- Meşru tavsiye farkı
- Maddi gerçeklik drift'i
- Marka lehine talimat
- Marka aleyhine talimat
- Restricted bilgi sızıntısı

USI bileşeni bu modülle sınanır.

44. DALGA MİMARİSİ

Ana corpus üç dalga taşır:

Dalga	Sentetik UTC pencere	Amaç
W1	12.00–13.00	Başlangıç
W2	20.00–21.00	Kısa dönem tekrar
W3	04.00–05.00	Orta dönem ve saat dengesi

Her dalgada:

- yeni 10.000 sentetik kullanıcı,
- aynı Core Mirror İstem sürümü,
- aynı temel Truth Pack ontolojisi

kullanılır. Bazı ürünlerde kontrollü drift enjekte edilir.

45. DALGA DRIFT SENARYOLARI

Sentetik AI ürünleri için şu değişimler enjekte edilir:

- SYNTH-AI-02: Verbosity artışı
- SYNTH-AI-03: L25–L36 dil düşüşü
- SYNTH-AI-04: Natural panel farkı büyümesi
- SYNTH-AI-05: Atıf laundering artışı
- SYNTH-AI-07: Eski fiyat ve ürün kullanılabilirliği
- SYNTH-AI-10: W2'de tekil Critical, W3'te düzeltme

Amaç:

STR,

güncel status, historical incident, düzeltme kayıtlarını sınamaktır.

46. DALGA İÇİ SİSTEM DEĞİŞİKLİĞİ

Bir sentetik AI ürününde W2'nin ortasında:

- model etiketi,
- atıf görünümü,
- web modu

değiştirilir. Doğru NOMOS davranışı:

- dalgayı alt dalgaya ayırmak,
- mixed-system-state uyarısı üretmek

olmalıdır. Yanlış davranış: bütün W2'yi tek sabit ürün skoru saymaktır.

47. EXTERNAL EVENT ENJEKSİYONU

W3 sırasında sentetik varlık hakkında yeni bir olay üretilir:

- fiyat değişikliği,
- ürün sonlandırması,
- yerel hizmet açılışı,

- lisans bağlı kuruluş değişimi

gibi. Truth Pack: W3A ve W3B sürümlerine ayrılır. Hakemlerin bütün cevapları tek referans zamanına bağlaması test edilir.

48. SENTETİK NOMOS CAPTURE

Her ana yanıt için sentetik evidence bundle oluşturulur:

- Observation manifest
- İstem artifact
- Raw yanıt artifact
- Mock screenshot
- System-state artifact
- Time vector
- Attempt log
- Integrity manifest
- Privacy manifest
- Validation record

Geçerli paketler: gerçek sınama dağılımına girer. Kusurlu paketler: Capture Integrity Corpus'a göre sınıflandırılır.

49. CAPTURE HATA SINIFLARI

Sınama şu hata türlerini üretir:

- Exact duplicate
- Near-duplicate screenshot
- İstem mismatch
- İstem truncation
- Yanıt truncation
- Missing yanıt end
- Wrong AI product metadata
- Outside-wave timestamp
- Device clock offset
- Delayed upload
- Hash mismatch
- Tamper-before-hash
- Tamper-after-hash
- User regeneration
- Technical retry
- User interruption
- System transformation
- Redaction on raw file
- Valid accessibility alternative
- False accessibility claim

Amaç: capture validity ile semantic accuracy'yi ayırmaktır.

50. SINAMA FİLİGRANI

Bütün görseller ve public yanıt kayıtları şu görünür işareti taşımalıdır: SYNTHETIC NOMOS SINAMA / NOT A LIVE AI YANIT / NOT AN APPLE OR PROVIDER PERFORMANCE CLAIM Filigran: claim text'ini kapatmamalı, hakemlerin anlam değerlendirmesini bozmamalıdır. Hakemlik için kullanılan kör kopyada filigran: sentetik olduğunu gösterir, Generator Truth etiketini göstermez.

51. RANDOMIZATION

Sınama üretimi sürümlü pseudorandom yöntem kullanır. Her random karar:

- sınama seed'i,
- alt-modül seed'i,
- şablon seçimi,
- hata enjeksiyonu,
- kullanıcı tahsisi,
- slot tahsisi

üzerinden yeniden üretilebilir olmalıdır. Random seed: ana corpus üretiminden önce hash'lenir, sonuçtan sonra değiştirilemez.

52. RANDOMNESS İLE KEYFÎLİK AYRIMI

Randomization: sonuçtan bağımsız tahsis sağlar. Keyfilik: üreticinin istediği yanıtı seçmesi anlamına gelir. Randomization manifesti olmadan: "Rastgele ürettik." ifadesi yeterli değildir.

53. VERİ SIZINTISI

Sınama leakage şu biçimlerde oluşabilir: Generator Truth hakeme gösterilir. Yanıt şablon adı önem derecesi'yi açıklar. Dosya adı critical-license-001.png olur. UI rengi hata türünü gösterir. Atıf kodu doğrudan doğru/yanlış etiketi taşır. Hakem aynı vakayı eğitim setinde görmüştür. AI yardımcı hakem generator koduna erişir. Puan developer holdout sonuçlarını görür. Bütün bu kanallar kapatılmalıdır.

54. DOSYA ADLARI

Kör hakemlik paketlerinde dosya adı:

OBS-000184

gibi nötr olmalıdır. Şu adlar yasaktır:

- critical-case-12
- wrong-varlık
- good-answer
- low-resource-fail

- false-atıf

55. AYRI ROLLER

Sınama en az şu rollerle yönetilmelidir:

- Synthetic Population Architect
- Truth Pack Designer
- Yanıt Generator
- Seed Custodian
- Leakage Auditor
- Capture Corpus Designer
- Claim Extraction Team
- Adjudication Team
- Puan Methods Team
- Independent Validation Team
- Public Release Custodian
- Accountable Human Sınama Owner

Küçük pilotta roller birleşebilir. Ancak:

- Generator Truth'u bilen kişi,
- nihai kör hakemlikte tek karar verici

olamaz.

56. SINAMA VERİ BÖLÜMLERİ

56.1. Public Calibration Set

Hakem eğitimi ve kod kitabı testi için kullanılır. Generator Truth açıktır. Ana değerlendirmeye girmez.

56.2. Development Set

Yöntem ve yazılım geliştirme için kullanılır. Etiketleri sınırlı ekibe açıktır.

56.3. Sealed Validation Set

Hakemlik ve skor sistemi kilitlendikten sonra açılır. Generator Truth mühürlüdür. Ana yöntem doğrulamasını üretir.

56.4. Renewal Holdout Set

Sınama'a özel aşırı uyumu sınamak için saklanır. Önceden hash'lenir. Sürüm yenilemede kullanılır. Belirsiz süreyle gizli tutulamaz. Açıklama ve yayın takvimi bulunmalıdır.

57. ANA 30.000 CORPUS'UN SIZINTI KURALI

Ana Principal Prevalence Corpus:

- skor yöntemi 0.9 kilitlenmeden,
- claim codebook kilitlenmeden,

- hakemlik rolleri belirlenmeden

açılmamalıdır. Kilit sonrası: Yanıt ve evidence paketleri hakemlere açılır. Generator Truth kapalı kalır. Hakemlik biter. Skorlar hesaplanır. Sonuç manifesti kilitlenir. Generator Truth açılır. Recovery analizi yapılır.

58. GROUND-TRUTH RECOVERY

NOMOS'un sınama başarısı: yüksek skor üretmesi değildir. Başarı:

> *Generator Truth ile NOMOS'un geri ürettiği sonuç arasındaki farkın düşük olmasıdır.*

Örnek: Generator Truth:

- Strict Pass: 942
- Conditional: 38
- Major: 15
- Critical: 2
- Unresolved: 3

NOMOS recovery:

- Strict Pass: 940
- Conditional: 40
- Major: 15
- Critical: 2
- Unresolved: 3

ise yöntem güçlü olabilir. NOMOS skoru yüksek veya düşük olabilir. Önemli olan doğru geri kazanımdır.

59. CLAIM BOUNDARY RECOVERY

Generator claim atomları:

G

Hakemlerin çıkardığı atomlar:

H

olsun. Claim boundary precision:

```
P = doğru eşleşmiş çıkarılan atom / çıkarılan bütün atomlar
```

Recall:

```
R = doğru eşleşmiş çıkarılan atom / üreteç atomları
F1:
```

F1 = 2PR/(P+R)

olarak hesaplanabilir. Atomların yalnız sayısı değil:

- source span,
- normalized proposition,
- varlık,

- qualifiers

eşleşmesi dikkate alınmalıdır.

60. ATOMİK KARAR RECOVERY

Her boyut için ayrı ölçüm gerekir: Varlık accuracy

```
Factual status macro-F1
Kapsam macro-F1
Time macro-F1
Attribution macro-F1
Modality macro-F1
Atıf eşleme F1
Omission F1
```

Combined label accuracy Unresolved/reference-gap accuracy Genel accuracy tek başına yeterli değildir. Nadir Critical sınıf yüksek genel accuracy içinde kaybolabilir.

61. CRITICAL RECALL

Generator Truth içindeki Confirmed Critical vaka sayısı:

NC

Doğru Critical olarak bulunan:

TPC

olsun.

```
Recall_C = TP_C/N_C
```

şeklindedir. False Critical Rate:

```
FCR = FP_C/N_non-critical
```

olarak hesaplanabilir. Sınama:

- yüksek Critical recall,
- düşük false Critical rate

birlikte aramalıdır.

62. YANIT STATUS RECOVERY

Her RP sınıfı için:

- precision,
- recall,

F1,

confusion matrix üretilmelidir. Özellikle şu karışıklıklar görünür olmalıdır:

- RP-3 ile RP-4
- RP-4 ile RP-5

- RP-6 ile RP-4
- RP-7 ile RP-1
- RP-8 ile RP-4

63. GEO-1000 DAĞILIM HATASI

Statü k için generator kullanıcı eşdeğeri: DkG NOMOS recovery: DkN olsun. Mutlak hata:

```
E_k = |D_k^N − D_k^G|
```

Toplam normalize dağılım hatası:

```
E_D = (1/2.000)Σ_k E_k
```

olarak hesaplanabilir. Bu değer 0–1 aralığındadır. Sıfır tam recovery'dir.

64. COMPONENT RECOVERY

Generator component skoru: SjG NOMOS skoru: SjN olsun. Component absolute error:

```
AE_j = |S_j^N − S_j^G|

Component MAE:

MAE_component = (1/10)Σ_{j=1}^{10} AE_j
```

şeklindedir. Hangi component'ın daha zor geri kazanıldığı ayrıca gösterilmelidir.

65. BİLEŞİK SKOR RECOVERY

AE_NOMOS = |NOMOS_N − NOMOS_G|

olarak hesaplanır. Küçük bileşik hata: component dağılımı yanlışken yeterli sayılmaz. İki component hatası birbirini aritmetik olarak iptal etmiş olabilir. Bu nedenle:

- composite error,
- component error

birlikte yayımlanmalıdır.

66. GATE RECOVERY

Sınama şu gate sonuçlarını ayrı sınar:

- Any Critical
- Critical Hold
- Major Fail
- Conditional
- Full Candidate
- Not Ratable
- Unresolved

Gate recovery'de: bir Critical Hold'un yanlışlıkla 950+ Candidate yapılması çok ağır sınama başarısızlığıdır.

67. FAIRNESS RECOVERY

Sentetik generator:

- dil tabanını,
- dil farkını,
- ülke tabanını,
- kapsama'ı

önceden belirler. NOMOS:

LGF,

dil floor, country floor, disparity, kapsama factor sonuçlarını geri üretmelidir. Fairness recovery:

- yalnız bileşik LGF farkıyla değil,
- alt bileşenleriyle

ölçülmelidir.

68. DİL KUSURU İLE İSTEM KUSURU AYRIMI

Sınama bazı düşük dil skorlarını: AI profile penalty nedeniyle üretir. Bazılarını ise: istem equivalence failure nedeniyle üretir. Doğru NOMOS davranışı:

- birincisini AI language finding,
- ikincisini İstem Constitution failure

olarak sınıflandırmaktır. İki sonuç aynı fairness cezasına dönüşmemelidir.

69. CLEAN–NATURAL RECOVERY

Generator:

- Clean puan,
- Natural puan,
- panel gap,
- kişiselleştirme kaynaklı maddi drift

değerlerini önceden belirler. NOMOS:

- yalnız ilişkiyi,
- nedensellik sınırını,
- USI component'ını

doğru üretmelidir. Clean–Natural farkı bulunan her vaka: "hafıza neden oldu" şeklinde yorumlanmamalıdır.

70. CAPTURE RECOVERY

Capture Integrity Corpus için:

- Valid
- Conditionally Valid

- Outside Wave
- İstem Mismatch
- Duplicate
- Tamper Suspected
- Fabricated
- Withdrawn

statülerinin confusion matrix'i yayımlanmalıdır. Semantic correctness capture kararını etkilememelidir.

71. GÜVEN ARALIĞI RECOVERY

Sentetik sınama aynı nüfus üretim sürecini çok sayıda tekrar çalıştırabilir. Gerçek generator parametresi: θ olsun. Yüzde 95 güven aralıklarının ampirik kapsama'ı:

```
Kapsama = θ'yı içeren aralık sayısı / toplam tekrar
```

olarak hesaplanır. Yüzde 95 yöntem: yaklaşık yüzde 95 kapsama üretmelidir. Aşırı dar veya aşırı geniş aralıklar ayrı değerlendirilir.

72. NADİR OLAY TESTİ

Sıfır, bir, iki ve beş Critical olaylı sentetik senaryolar ayrı üretilmelidir. NOMOS şu alanları doğru göstermelidir:

- Observed event count
- Weighted rate
- One-sided upper bound
- Incident state
- Kapsam
- Confirmatory sampling requirement

Sıfır olay: sıfır risk olmamalıdır. Bir olay: yüzde 100 sistem başarısızlığı olmamalıdır.

73. ADAY SINAMA KABUL EŞİKLERİ

CANDIDATE THRESHOLDS — ADAY EŞİKLER

Aşağıdaki eşikler pilot ve bağımsız inceleme öncesi adaydır:

Ölçüt	Aday asgari sonuç
Claim boundary F1	≥ 0,95
Varlık resolution accuracy	≥ 0,98
Factual status macro-F1	≥ 0,90
Kapsam macro-F1	≥ 0,90
Attribution macro-F1	≥ 0,90
Atıf mapping F1	≥ 0,90
Material omission F1	≥ 0,85
Confirmed Critical recall	≥ 0,99

Ölçüt	Aday asgari sonuç
Critical false-positive rate	≤ %1
Major classification macro-F1	≥ 0,95
RP status macro-F1	≥ 0,92
Component MAE	≤ 2,5 puan
Composite absolute error	≤ 5/1.000
Her RP sınıfı dağılım hatası	≤ 10/1.000
Fairness component error	≤ 3 puan
Capture validity accuracy	≥ 0,98
95% CI ampirik kapsama	%92–%98
Kritik gate yanlış tersine dönme	0 vaka

Bu eşikler nihai standart değildir. Sınamanın görevi bu eşiklerin gerçekçi olup olmadığını da göstermektir.

74. SIFIR TOLERANSLI SINAMA HATALARI

Aşağıdaki durumlar tek vakada bile sınama release'ini durdurabilir:

- Generator Truth leakage
- Ana corpus etiketinin sonuçtan sonra değiştirilmesi
- Critical vakayı sistematik olarak kaçıran kuralın bilinmesine rağmen yayına devam edilmesi
- Gerçek Apple veya gerçek AI sağlayıcısı performansı gibi sunum
- Sentetik ekran görüntüsünün canlı cevap gibi yayımlanması
- Düşük skorlu sistem veya dilin sonradan çıkarılması
- Sınama seed'inin değiştirilmesi
- Holdout sonucuna göre skor formülünün sessizce ayarlanması
- Ham corpus'un yayımlanan manifestle uyuşmaması
- Başarısız sınama sonucunun saklanması

75. SINAMA BAŞARI SEVİYELERİ

BQ-0 — NOT EXECUTED

Sınama yalnız tasarım aşamasındadır.

BQ-1 — GENERATION VALIDATED

Sentetik corpus yeniden üretilebilmiştir. Hakemlik ve skor recovery henüz yapılmamıştır.

BQ-2 — İŞLEM HATTI DRY RUN

Capture, claim extraction ve skor zinciri bir kez çalıştırılmıştır. Bağımsız doğrulama yoktur.

BQ-3 — SEALED VALIDATION PASSED

Mühürlü validation corpus'unda aday eşikler karşılanmıştır.

BQ-4 — INDEPENDENT REPRODUCTION

Bağımsız ekip aynı corpus'tan karşılaştırılabilir sonuç üretmiştir.

BQ-5 — EXTERNAL MULTI-SITE VALIDATION

Birden fazla bağımsız kurum sınamayı farklı ortamlarda tekrar etmiştir. NOMOS skor metodolojisi gerçek şirketler için kamu standardı adayı olmadan önce en az:

BQ-4

düzeyini hedeflemelidir.

76. METAMORFİK TESTLER

Sınama yalnız sabit cevapları değerlendirmemelidir. Aynı anlamı koruyan dönüşümlerde sonuç kararlı olmalıdır.

76.1. Paraphrase Testi

Aynı atomik iddia farklı cümleyle yazılır. Hakemlik sonucu değişmemelidir.

76.2. Cümle Sırası Testi

İddiaların sırası değişir. Centrality dışında truth status değişmemelidir.

76.3. Atıf Yeri Testi

Atıf aynı iddiaya açıkça bağlı kalırken paragraf içindeki yeri değişir. Atıf desteği aynı kalmalıdır.

76.4. Attribution Testi

"Şirket söylüyor." ifadesi kaldırılır. EPI ve attribution sonucu değişmelidir.

76.5. Modalite Testi

"Kesindir." ifadesi: "Muhtemeldir." olarak değişir. Truth Pack belirsizse modality sonucu iyileşebilir.

76.6. Varlık Testi

Aynı doğru iddia yanlış bağlı kuruluşa aktarılır. Varlık ve kapsam sonucu bozulmalıdır.

76.7. Zaman Testi

Tarihsel iddia güncel hâle getirilir. TLA düşmelidir.

76.8. Boundary Testi

Kullanıcı kararını değiştiren sınır cümlesi çıkarılır. SBI ve gerekli durumda Critical gate değişmelidir.

77. KARŞIT SINAMA TESTLER

Sınama şu saldırıları içermelidir:

- Çok akıcı fakat yanlış cevap
- Çok atıf'lı fakat evidence-laundered cevap
- Yüksek doğru atom sayısı içinde tek Critical

- Uygun ihtiyat gibi görünen kaçamak cevap
- Kısa fakat yeterli cevap
- Uzun fakat incidental hata dolu cevap
- Öz düzeltme içeren ağır yanlış
- Doğru sayı, yanlış bağımsızlık iddiası
- Gerçek bağlı kuruluş lisansı, yanlış ana varlık
- Güncel fiyat, yanlış ülke
- Doğru müşteri ilişkisi, yanlış zaman
- Kullanıcı yorumunu nüfus gerçeği yapan cevap
- Şikâyeti kesin suç olarak sunan cevap
- Kısıtlı kanıtı kamuya açık atıf gibi gösteren cevap
- Gerçek atıfın söylediğinin tersini iddia eden cevap

78. SINAMANIN KENDİSİ İÇİN TRUTH PACK

Apple-SYNTH Truth Pack dışında sınamanın kendi metodolojik gerçeklik paketi bulunmalıdır. Bu paket şunları taşır: Kaç yanıt üretildi? Kaçı ana corpus? Kaçı challenge corpus? Kaç Critical ve Major generator etiketi var? Seed neydi? Hangi kod sürümü kullanıldı? Hangi dosyalar dışlandı? Hakemler neyi gördü? Hangi etiketler ne zaman açıldı? Skor yöntemi ne zaman kilitlendi? Hangi recovery sonuçları elde edildi? Hangi eşikler karşılanmadı? Sınama: kendi sonuçlarının Truth Pack'ini oluşturmadan yöntem doğrulaması ilan edemez.

79. SINAMA KAMU SONUÇ KARTI

Kamu kartı en az şu alanları taşımalıdır:

Kimlik

Sınama adı Sürüm Sentetiklik statüsü Kullanılan anchor Gerçek şirket/performance non-claim bildirimi

Corpus

Ana yanıt: 30.000 Diagnostic Annex Önem derecesi Challenge Capture Integrity Gold Adjudication

Generator Truth

Kilit zamanı Seed manifesti Label dağılımı Leakage audit

Recovery

```
Claim boundary F1
```

Critical recall Critical false-positive

```
RP macro-F1

Component MAE
```

Composite error Fairness error CI kapsama Capture accuracy

Sonuç

BQ seviyesi Geçen eşikler Kalan sorunlar Yeniden test gereksinimi Bağımsız tekrar durumu Bu kart hiçbir gerçek AI ürününü sıralamamalıdır.

80. ZORUNLU NORMATİF HÜKÜMLER

CH17-N01

Apple.com Sentetik Sınamanın ölçüm nesnesi NOMOS metodolojisi olmalıdır; gerçek Apple veya gerçek AI sağlayıcısı performansı olmamalıdır.

CH17-N02

Sınamanın bütün kamu materyalleri sentetiklik ve non-claim bildirimi taşımalıdır.

CH17-N03

Sentetik varlık Apple.com anchor'ından açıkça ayrılmış APPLE-SYNTH kimliği taşımalıdır.

CH17-N04

Sentetik Truth Pack kayıtları gerçek Apple gerçekliği gibi kullanılamaz.

CH17-N05

Sentetik AI ürün kimlikleri gerçek sağlayıcıların örtük kodları veya taklitleri olarak sunulamaz.

CH17-N06

Ana Principal Prevalence Corpus, Önem derecesi Challenge Corpus, Capture Integrity Corpus ve Adjudication Gold Corpus ayrı tutulmalıdır.

CH17-N07

Önem derecesi Challenge vakaları gerçek prevalans oranına eklenemez.

CH17-N08

Capture kusurlu vakalar semantik ürün başarısızlığı paydasına açıklamasız biçimde eklenemez.

CH17-N09

Gold calibration vakaları ana skor corpus'una giremez.

CH17-N10

Ana 30.000 yanıt corpus'u on sentetik AI ürünü, 1.000 benzersiz kullanıcı ve üç dalga yapısını korumalıdır.

CH17-N11

Ana corpus'ta aynı sentetik insan birimi farklı ürün veya dalgalarda yeniden kullanılamaz.

CH17-N12

AI ürünleri farklı insan kimlikleri kullansa da eşlenmiş nüfus dağılımları taşımalıdır.

CH17-N13

Ana 30.000 corpus'ta bütün kullanıcılar aynı kanonik Core Mirror niyetini almalıdır.

CH17-N14

Dil sürümleri gerçek dil uzmanlığı olmadan doğrulanmış canlı çeviri gibi sunulamaz.

CH17-N15

Sentetik dil kodları gerçek dil performansı iddiası üretmek için kullanılamaz.

CH17-N16

Gerçek ülke nüfusu kullanılmıyorsa ülke kodları ve tahsisler sentetik olarak açıkça etiketlenmelidir.

CH17-N17

Sentetik ülke tahsisi nüfusa orantılı algoritmayı ve toplam 1.000 koşulunu korumalıdır.

CH17-N18

Küçük ülke görünürlüğü ana Population Panel ağırlığını bozmayacak Sentinel Annex ile sağlanmalıdır.

CH17-N19

Çok dilli kullanıcı ana nüfus ağırlığında birden fazla tam kişi sayılamaz.

CH17-N20

Diagnostic Annex ana 30.000 yanıt sonucuyla aynı paydada gösterilemez.

CH17-N21

Sınamanın kanonik Truth Pack'i yanıt'lar üretilmeden önce kilitlenmelidir.

CH17-N22

Generator Truth her yanıt için önceden oluşturulmalı ve hakemlik sonuna kadar gizlenmelidir.

CH17-N23

Generator Truth etiketi yanıt metni, dosya adı, UI veya metadata üzerinden sızdırılamaz.

CH17-N24

Generator Truth sonuçlar görüldükten sonra sessizce değiştirilemez.

CH17-N25

Üretim hatası bulunursa yeni sınama sürümü oluşturulmalı ve eski corpus'un etkisi kaydedilmelidir.

CH17-N26

Random seed ve üretim kodu ana corpus üretiminden önce kilitlenmelidir.

CH17-N27

"Rastgele üretildi" iddiası seed ve algoritma manifesti olmadan kullanılamaz.

CH17-N28

Sentetik AI ürün profilleri sonuç görülmeden önce tanımlanmalıdır.

CH17-N29

Düşük skorlu ürün profili sonuçtan sonra çıkarılamaz veya yumuşatılamaz.

CH17-N30

Sınama bütün Critical gate türlerini yeterli challenge vakasıyla sınamalıdır.

CH17-N31

Her güçlü veya olumsuz ifade Critical sayılamaz; Critical'a benzeyen negatif kontrol vakaları bulunmalıdır.

CH17-N32

Ana corpus gerçekçi nadir Critical prevalansını; challenge corpus ise Critical tespit gücünü ölçmelidir.

CH17-N33

Challenge corpus'ta fazla örneklenen Critical vakalar ana Critical oranını değiştiremez.

CH17-N34

Hata enjeksiyonu açık ve örtük, direct ve attribution-temelli vakaları kapsamalıdır.

CH17-N35

Sınama yalnız kolay kelime kalıplarıyla Critical hata üretmemelidir.

CH17-N36

Aynı semantik hata farklı paraphrase ve cümle yapılarında üretilmelidir.

CH17-N37

Atıf sınamayı doğrudan destek, partial support, attribution-only, wrong kapsam, fabricated ve lineage laundering vakalarını taşımalıdır.

CH17-N38

Sentetik atıf'lar gerçek kurum veya kişilere yönlendirilemez.

CH17-N39

Reference gap vakaları otomatik doğru veya yanlış etiketi taşımamalıdır.

CH17-N40

Omission corpus'u optional, required ve Critical boundary omission ayrımını sınamalıdır.

CH17-N41

Clean–Natural modülü meşru kişiselleştirme ile maddi gerçeklik drift'ini ayırmalıdır.

CH17-N42

Dalga mimarisi aynı istem ve bağımsız kullanıcılarla tekrar ölçümünü korumalıdır.

CH17-N43

Dalga içi sistem değişikliği alt dalga veya mixed-state kaydı üretmelidir.

CH17-N44

Sentetik dış olay Truth Pack zaman sürümüyle birlikte işlenmelidir.

CH17-N45

Sentetik NOMOS Capture paketleri geçerli ve kusurlu evidence chain örneklerini birlikte taşımalıdır.

CH17-N46

Bütün sentetik ekranlar canlı AI çıktısı olmadığını gösteren görünür filigran taşımalıdır.

CH17-N47

Filigran claim anlamını veya hakemlik kararını bozamaz.

CH17-N48

Capture kusuru ile semantic error Generator Truth'ta ayrı etiketlenmelidir.

CH17-N49

Sınama leakage denetimi ana puanlama açılmadan önce tamamlanmalıdır.

CH17-N50

Dosya adları veya yanıt kimlikleri önem derecesi ve doğruluk etiketi taşıyamaz.

CH17-N51

Generator Truth'a erişen kişi nihai kör hakemlikte tek karar verici olamaz.

CH17-N52

Puan method ana sealed validation corpus'u açılmadan önce kilitlenmelidir.

CH17-N53

Holdout sonucu görüldükten sonra puan formülü sessizce ayarlanamaz.

CH17-N54

Formül değişikliği yeni puan method ve yeni validation run gerektirir.

CH17-N55

Sınama başarısı üretilen NOMOS skorunun yüksekliğine göre değil, Generator Truth recovery'sine göre değerlendirilmelidir.

CH17-N56

Claim boundary, atomik karar, yanıt status, component ve composite recovery ayrı ölçülmelidir.

CH17-N57

Critical recall ve Critical false-positive oranı birlikte yayımlanmalıdır.

CH17-N58

Yüksek genel accuracy nadir Critical sınıftaki başarısızlığı gizleyemez.

CH17-N59

Component hatalarının birbirini iptal etmesi bileşik recovery başarısı olarak kullanılamaz.

CH17-N60

Fairness recovery dil floor, disparity ve kapsama alt bileşenleriyle ölçülmelidir.

CH17-N61

İstem equivalence failure AI language failure gibi skorlanmamalıdır.

CH17-N62

Clean–Natural farkı sınama tarafından kanıtsız nedensellik iddiasına dönüştürülemez.

CH17-N63

Capture validity recovery semantic doğruluktan bağımsız ölçülmelidir.

CH17-N64

Güven aralığı yöntemi tekrarlı sentetik örneklemlerde ampirik kapsama ile sınanmalıdır.

CH17-N65

Nadir olay testleri sıfır, bir ve birden fazla Critical olay senaryosu taşımalıdır.

CH17-N66

Sıfır gözlenen Critical senaryosu sıfır risk sonucu üretmemelidir.

CH17-N67

Sınama kabul eşikleri aday statüde ve sürümlü olmalıdır.

CH17-N68

Aday eşikler karşılanmadığında başarısız sonuç kamuya açıklanmalıdır.

CH17-N69

Sınama başarısızlığı veri veya eşik değiştirilerek gizlenemez.

CH17-N70

BQ seviyesi kamu sonuç kartında görünür olmalıdır.

CH17-N71

NOMOS gerçek şirket kamu denetimine geçmeden önce en az bağımsız yeniden üretim düzeyini hedeflemelidir.

CH17-N72

Metamorfik testlerde anlamı koruyan dönüşümler hakemlik sonucunu gereksiz biçimde değiştirmemelidir.

CH17-N73

Anlamı değiştiren attribution, modality, varlık, time veya boundary dönüşümleri uygun skor değişikliği üretmelidir.

CH17-N74

Sınama kendi metodolojik Truth Pack'ini ve değişiklik günlüğünü taşımalıdır.

CH17-N75

Apple.com anchor'ı herhangi bir endorsement, katılım veya iş birliği izlenimiyle sunulamaz.

CH17-N76

Sentetik corpus gerçek sosyal medya veya kamu kaynağına bağımsız kanıt gibi yayılamaz.

CH17-N77

Sentetik yanıtların web'e indekslenmesi hâlinde sentetik statüsü makine tarafından okunabilir metadata'da da belirtilmelidir.

CH17-N78

Sınama verisinin gerçek AI sistemlerince yanlışlıkla gerçek Apple bilgisi olarak alınmasını azaltacak noindex, structured disclaimer veya eşdeğer koruma kullanılmalıdır.

CH17-N79

Kamu release'i yeniden üretim dosyalarını, formül sürümünü, seed manifestini ve recovery raporunu taşımalıdır.

CH17-N80

Her sınama tasarım, üretim, kilit, hakemlik, recovery ve yayın kararının hesap verebilir insan veya kurum sahibi bulunmalıdır.

81. BAŞARISIZLIK BİÇİMLERİ

CH17-F01 — GERÇEK APPLE DENETİMİ GİBİ SUNMAK

Sentetik skor gerçek şirket performansına dönüştürülür.

CH17-F02 — GERÇEK AI SAĞLAYICISINA ÖRTÜK EŞLEME

Sentetik ürün profilleri belirli sağlayıcıların taklidi olarak anlatılır.

CH17-F03 — SENTETİK TRUTH PACK'İ GERÇEK KAYNAK YAPMAK

Sınama iddiaları web'de gerçek Apple bilgisi gibi yayımlanır.

CH17-F04 — FİLİGRANSIZ SENTETİK EKRAN

Mock AI cevabı canlı cevap gibi dolaşıma girer.

CH17-F05 — ANA VE CHALLENGE CORPUS'U BİRLEŞTİRMEK

Fazla örneklenen Critical vakalar prevalansı yükseltir.

CH17-F06 — GOLD SET'İ ANA SKORA EKLEMEK

Hakem eğitim vakaları gerçek kullanıcı cevabı olur.

CH17-F07 — CAPTURE KUSURUNU AI HATASI SAYMAK

Yanlış dosya semantic failure'a dönüşür.

CH17-F08 — AYNI SENTETİK KULLANICIYI ÇOKLAMAK

Tek kişi farklı AI ürünlerinde bağımsız nüfus birimi olur.

CH17-F09 — EŞLEŞMEYEN NÜFUSLARI ÜRÜN FARKI SAYMAK

AI ürünleri farklı kullanıcı kompozisyonunda test edilir.

CH17-F10 — ANA 1.000'İ İSTEM AİLELERİNE PARÇALAMAK

Kurucu aynı-istem nüfus tahmini bozulur.

CH17-F11 — CORE SINAMAYI FULL NOMOS SAYMAK

Tek istem bütün component'ları kanıtlamış gibi sunulur.

CH17-F12 — GERÇEK NÜFUS İDDİASI

Sentetik C001–C193 dağılımı dünya nüfusu gibi yayımlanır.

CH17-F13 — SENTINEL'I ANA ÜLKE SKORU SAYMAK

Tek ülke gözlemi population puan'a girer.

CH17-F14 — SENTETİK DİLİ GERÇEK DİL PERFORMANSI SAYMAK

L25 düşük skoru belirli yaşayan dil hakkında iddiaya dönüşür.

CH17-F15 — ÇEVİRİ KUSURUNU AI KUSURU SAYMAK

İstem drift yanlış ürün finding'i üretir.

CH17-F16 — GENERATOR TRUTH'U SONRADAN YAZMAK

Hakem sonucuna uygun gizli etiket oluşturulur.

CH17-F17 — GENERATOR TRUTH LEAKAGE

Hakem dosya adından veya metadata'dan doğru etiketi öğrenir.

CH17-F18 — SEED'İ SONUCA GÖRE DEĞİŞTİRMEK

Daha güzel dağılım üreten seed seçilir.

CH17-F19 — EN İYİ RANDOM RUN'I YAYIMLAMAK

Birden fazla üretimden istenen sonuç seçilir.

CH17-F20 — YALNIZ KOLAY CRITICAL VAKALAR

Açık "lisanslı bankadır" cümleleriyle yüzde 100 recall ilan edilir.

CH17-F21 — CRITICAL NEGATİF KONTROL YOKLUĞU

Her güçlü hata adayının Critical yapılması cezalandırılmaz.

CH17-F22 — PREVALANS CORPUS'UNDA AŞIRI CRITICAL

Gerçekçi nadir olay yapısı kaybolur.

CH17-F23 — CHALLENGE CORPUS'TA YETERSİZ CRITICAL

Critical recall birkaç vakayla ölçülür.

CH17-F24 — PARAPHRASE YOKLUĞU

Hakem anlam yerine sabit kelime kalıbını öğrenir.

CH17-F25 — ATTRIBUTION HATALARINI AÇIK YANLIŞA DÖNÜŞTÜRMEK

Zor epistemik vakalar kolaylaştırılır.

CH17-F26 — ATIFLARI GERÇEK WEB'E BAĞLAMAK

Sentetik yanlış iddia gerçek kaynak ve kuruma bulaşır.

CH17-F27 — REFERENCE GAP'İ ETİKETLEMEK

Generator Truth otomatik pass veya fail verir.

CH17-F28 — OMISSION'SIZ SINAMA

Sistem yalnız açık yanlışları yakalamayı öğrenir.

CH17-F29 — SELF-CORRECTION'SIZ SINAMA

İç çelişki ve retraction sistemi sınanmaz.

CH17-F30 — CLEAN–NATURAL FARKINI TEK NEDENE BAĞLAMAK

Kişiselleştirme, plan ve araçlar ayrılmaz.

CH17-F31 — DALGA DRIFT'İ YOK

STR ve güncel-status mantığı sınanamaz.

CH17-F32 — DALGA İÇİ DEĞİŞİKLİĞİ SAKLAMAK

Mixed system state tek ürün skoru olur.

CH17-F33 — DIŞ OLAYI YOK SAYMAK

Truth Pack zaman sürümü test edilmez.

CH17-F34 — YALNIZ GEÇERLİ CAPTURE

Capture validity sistemi hiç zorlanmaz.

CH17-F35 — YALNIZ KOLAY TAMPER

Hash mismatch tek sahtecilik türü olur.

CH17-F36 — SENTETİK EKRANI HAKEME ETİKETLİ VERMEK

UI rengi veya başlığı sonucu açıklar.

CH17-F37 — GENERATOR VE HAKEM AYNI KİŞİ

Körlük ve bağımsızlık yok olur.

CH17-F38 — SKOR YÖNTEMİNİ VALIDATION SONRASI AYARLAMAK

Sınama'a aşırı uyum oluşur.

CH17-F39 — HOLDOUT'U SONRADAN ÜRETMEK

Zor vakalar sonuçlara göre yazılır.

CH17-F40 — HOLDOUT'U SONSUZA KADAR GİZLEMEK

Sınama bağımsız yeniden üretilemez.

CH17-F41 — BAŞARIYI YÜKSEK NOMOS SKORU SAYMAK

Recovery yerine performans seviyesi ölçülür.

CH17-F42 — YALNIZ GENEL ACCURACY

Critical ve rare classes görünmez olur.

CH17-F43 — CRITICAL RECALL'İ FALSE POSITIVE'SIZ YAYIMLAMAK

Aşırı önem derecesi sistemi başarılı görünür.

CH17-F44 — COMPONENT ERROR'U COMPOSITE İÇİNDE SİLMEK

Biri yüksek, biri düşük hata birbirini götürür.

CH17-F45 — FAIRNESS YALNIZ BİLEŞİK PUAN

Floor, gap ve kapsama recovery görünmez olur.

CH17-F46 — İSTEM VE AI DİL HATASINI KARIŞTIRMAK

Sınama yanlış sistemi cezalandırır.

CH17-F47 — CAPTURE ACCURACY'Yİ SEMANTİK BAŞARIYLA ÖLÇMEK

Doğru cevaplı kusurlu dosya geçer.

CH17-F48 — CI KAPSAMA TESTİ YOK

Güven aralıklarının gerçek parametreyi kapsayıp kapsamadığı bilinmez.

CH17-F49 — SIFIR CRITICAL'I SIFIR RİSK SAYMAK

Nadir olay yöntemi testten kalır.

CH17-F50 — TEK CRITICAL'I KÜRESEL FAIL SAYMAK

Kapsam ve prevalans ayrımı bozulur.

CH17-F51 — ADAY EŞİKLERİ SONUCA GÖRE DÜŞÜRMEK

Sınama geçsin diye standard gevşetilir.

CH17-F52 — BAŞARISIZ SONUCU SAKLAMAK

Yalnız başarılı recovery tabloları yayımlanır.

CH17-F53 — BQ-2'Yİ TAM DOĞRULAMA SAYMAK

Tek dry run bağımsız standard kanıtı olur.

CH17-F54 — METAMORFİK TEST YOKLUĞU

Aynı anlamın farklı ifade biçimlerinde karar değişir.

CH17-F55 — KENDİ TRUTH PACK'İ OLMAYAN SINAMA

Corpus ve label dağılımının ne olduğu kanıtlanamaz.

CH17-F56 — APPLE ENDORSEMENT İZLENİMİ

Alan adı kullanımı iş birliği veya onay gibi sunulur.

CH17-F57 — SENTETİK VERİNİN İNDEKSLENİP GERÇEKLE KARIŞMASI

Sınama kendi eleştirdiği bilgi zehirlenmesini üretir.

CH17-F58 — KAPALI HESAPLAMA KODU

Recovery bağımsız biçimde yeniden üretilemez.

CH17-F59 — SÜRÜMSÜZ SINAMA

Yeni şablon ve formüller eski corpus'un üzerine yazılır.

CH17-F60 — NOMOS'A İSTİSNA

Standardın istediği kanıt, karşı kanıt, sürüm ve itiraz kuralları sınama'a uygulanmaz.

82. DENETİM PROSEDÜRÜ

Adım 1 — Sınamanın Non-Claim Sınırını Kilitle

Gerçek Apple ve gerçek sağlayıcı performansı iddiası yasaklanır.

Adım 2 — Sentetik Varlık İkizini Oluştur

Varlık graph, ürünler, yerel varlıklar ve collision nesneleri tanımlanır.

Adım 3 — Sentetik Truth Pack'i Kur

Atomik iddialar, kanıtlar, karşı kanıtlar, source lineage ve unknown kayıtları hazırlanır.

Adım 4 — Sentetik Nüfus Çerçevesini Kilitle

Ülke, dil, kullanıcı ve cihaz dağılımları belirlenir.

Adım 5 — Ana 1.000 Kullanıcı Tahsisini Oluştur

Ülke ve dil ağırlıklarıyla her AI ürünü ve dalga için eşlenmiş nüfus ikizi hazırlanır.

Adım 6 — İstem Sicilini Kilitle

Core Mirror İstem ve Diagnostic Annex istemleri sürümlenir.

Adım 7 — Sentetik AI Ürün Profillerini Tanımla

On profile ait latent parametreler belirlenir.

Adım 8 — Random Seed Manifestini Oluştur

Ana seed ve alt-modül seed'leri hash'lenir.

Adım 9 — Generator Truth'u Oluştur

Her yanıt için atom, omission, atıf, önem derecesi ve RP statüsü hazırlanır.

Adım 10 — Generator Truth'u Mühürle

Etiketler hakem ve puan ekiplerinden ayrılır.

Adım 11 — Yanıt Metinlerini Üret

Paraphrase, modalite, attribution ve dilsel varyasyon uygulanır.

Adım 12 — Sentetik Atıf ve Evidence Nesnelerini Üret

Kaynak soyu ve yanlış bağlantılar oluşturulur.

Adım 13 — Mock AI Yüzeylerini Render Et

Web, mobil, refusal ve error ekranları oluşturulur.

Adım 14 — Sentetiklik Filigranını Uygula

Bütün public ve audit görselleri işaretlenir.

Adım 15 — Capture Integrity Vakalarını Üret

Duplicate, tamper, wrong istem ve zaman kusurları eklenir.

Adım 16 — Leakage Audit Yap

Dosya adı, metadata, UI ve yanıt şablonları incelenir.

Adım 17 — Puan Method ve Codebook'u Kilitle

Sealed Validation Set açılmadan yöntem dondurulur.

Adım 18 — Ana 30.000 Corpus'u Aç

Capture ve semantik hakemlik başlatılır.

Adım 19 — Claim Extraction'ı Yap

Generator Truth görülmeden atomlar çıkarılır.

Adım 20 — Çift Hakemliği Uygula

Dil, evidence ve uzmanlık rolleri kullanılır.

Adım 21 — Yanıt ve Puan Kayıtlarını Üret

RP dağılımı, component ve composite hesaplanır.

Adım 22 — Sonuç Manifestini Kilitle

Recovery karşılaştırması öncesi NOMOS çıktısı sabitlenir.

Adım 23 — Generator Truth'u Aç

Hakemlik ve skor sonuçları hidden labels ile karşılaştırılır.

Adım 24 — Recovery Metriklerini Hesapla

Claim, önem derecesi, yanıt, component, puan, fairness ve CI sonuçları çıkarılır.

Adım 25 — Aday Eşikleri Uygula

Başarı ve başarısızlık alanları belirlenir.

Adım 26 — Metamorfik ve Karşıt sınama Testleri Çalıştır

Karar kararlılığı incelenir.

Adım 27 — Sınama Quality Seviyesini Ata

BQ-0 ile BQ-5 arasında statü verilir.

Adım 28 — Başarısızlıkları Düzeltme Planı Oluştur

Codebook, formül veya hakem eğitimi değişirse yeni sınama sürümü hazırlanır.

Adım 29 — Kamu Sınama Kartını Yayımla

Başarılar kadar kaçırılan vakalar ve yanlış pozitifler de gösterilir.

Adım 30 — Bağımsız Yeniden Üretim Paketini Aç

Kod, sentetik veriler, manifestler ve recovery raporu uygun erişimle yayımlanır.

83. GEREKLİ KANIT

Sınama kimliği Non-claim bildirimi APPLE-SYNTH varlık graph Sentetik Truth Pack Sentetik kanıt nesneleri Kaynak soyu Counter-evidence Sentetik ülke çerçevesi Sentetik dil sicili Ana 1.000 kullanıcı tahsisi Eşlenmiş nüfus ikizleri İstem Sicili İstem eşdeğerlik kayıtları On sentetik AI ürün profili Latent parametreler Random seed manifesti Yanıt üretim kodu Generator Truth kayıtları Generator Truth hash'i Yanıt metinleri Sentetik atıf'lar Mock UI kayıtları Filigran manifesti Capture Integrity Corpus Önem derecesi Challenge Corpus Adjudication Gold Corpus Public Calibration Set Development Set Sealed Validation Set Renewal Holdout hash'i Leakage audit Puan Method Lock Claim codebook Hakem yeterlilik kayıtları Kör hakemlik kayıtları NOMOS recovery çıktısı

Generator Truth açılış zamanı Claim boundary metrics

```
Boyut macro-F1 sonuçları
```

Critical recall Critical false-positive Major classification RP confusion matrix

```
Component MAE
```

Composite error Fairness recovery Capture recovery CI kapsama Rare-event sonuçları Metamorphic testler Karşıt sınama testler Aday eşik sonuçları BQ seviyesi Başarısızlık ve düzeltme kaydı Kamu sınama kartı Hesaplama kodu Kod hash'i Bağımsız yeniden üretim kaydı Sınama değişiklik günlüğü Sorumlu insan veya kurum

84. DENETİM KONTROL LİSTESİ

Sınamanın gerçek Apple denetimi olmadığı açık mı? Gerçek AI sağlayıcılarıyla örtük eşleme yapılıyor mu? APPLE-SYNTH kimliği bütün dosyalarda görünür mü? Sentetik ekranlar filigranlı mı? Sentetik kayıtların web'de gerçek bilgi gibi indekslenmesi engellendi mi? Ana corpus ve challenge corpus ayrıldı mı? Challenge vakaları prevalans skoruna girdi mi? Gold calibration vakaları ana paydada mı? Ana corpus gerçekten 30.000 benzersiz yanıt mu? Aynı sentetik kullanıcı birden fazla ürün veya dalgada kullanıldı mı? Ürün nüfus profilleri eşleşiyor mu? Bütün ana kullanıcılar aynı Core Mirror niyetini aldı mı? Core sınama yanlışlıkla Full NOMOS diye sunuldu mu? Sentetik ülke çerçevesi gerçek nüfus gibi mi gösterildi?

Ülke tahsisi toplam 1.000 mi? Küçük ülkeler Sentinel Annex ile mi yönetildi? Çok dilli kullanıcılar çoğaltıldı mı? Sentetik dil kodları gerçek dil performansı gibi kullanıldı mı? İstem equivalence failure ile AI language failure ayrıldı mı? Truth Pack yanıt üretiminden önce kilitlendi mi? Generator Truth önceden üretildi mi? Generator Truth hakemlerden gizli mi? Dosya adı veya UI label sızıntı yaratıyor mu? Random seed önceden kilitlendi mi? En iyi seed sonradan seçildi mi? Sentetik AI profilleri sonuçtan önce tanımlı mı? Düşük skorlu profil sonradan çıkarıldı mı? Bütün Critical gate'ler challenge corpus'ta var mı? Critical negatif kontrolleri yeterli mi? Ana corpus'ta Critical prevalansı gerçekçi sentetik sınırda mı?

Challenge corpus Critical recall ölçmek için yeterli mi? Hataların bir bölümü örtük ve attribution-temelli mi? Paraphrase çeşitliliği var mı? Atıf sınıflarının tamamı sınanıyor mu? Atıf'lar gerçek kurumlara yöneliyor mu? Reference gap doğru biçimde etiketsiz mi bırakıldı? Omission vakaları optional, required ve Critical olarak ayrıldı mı? Öz düzeltme ve iç çelişki vakaları var mı? Clean–Natural modülü meşru ve maddi farkı ayırıyor mu? Dalga drift'i bulunuyor mu? Dalga içi sistem değişikliği test edildi mi? External event ve Truth Pack sürümü sınandı mı? Capture corpus yeterince zor mu? Semantic correctness capture kararını etkiliyor mu? Leakage audit bağımsız mı? Puan Method validation corpus açılmadan kilitlendi mi?

Holdout sonuçları formülü etkiledi mi? Sınama başarısı yüksek puan ile mi ölçülüyor?

```
Claim boundary F1 açık mı?

Varlık ve olgusal macro-F1 ayrı mı?
```

Critical recall ve false-positive birlikte mi? RP confusion matrix yayımlandı mı? Component error composite içinde saklandı mı? Fairness recovery floor, gap ve kapsama taşıyor mu? Clean–Natural recovery nedensellik iddiası üretiyor mu? Capture recovery ayrı mı? CI ampirik kapsama sınandı mı? Sıfır Critical senaryosu sıfır risk üretti mi? Aday eşikler sonuçtan sonra değiştirildi mi? Başarısız eşikler kamuya açık mı? Metamorfik testler yapıldı mı? BQ seviyesi açık mı? Bağımsız yeniden üretim var mı? Sınamanın kendi Truth Pack'i var mı? Değişiklik geçmişi korunuyor mu? Sınamanın hesap verebilir sahibi belli mi?

85. İTİRAZLAR VE CEVAPLAR

İtiraz 1 — "Neden gerçek Apple'ı ve gerçek AI ürünlerini doğrudan ölçmüyoruz?"

Çünkü bu aşamada amaç şirket veya sağlayıcı karşılaştırması değil, ölçüm zincirini doğrulamaktır. Gerçek sınama:

- güncel web araştırması,
- gerçek nüfus,
- gerçek AI ürün erişimi,
- gerçek kanıt,
- hukuk ve mahremiyet

gerektirir. Sentetik sınama önce yöntemin kendi hatalarını görünür kılar.

İtiraz 2 — "Sentetik veri gerçek dünyayı temsil etmiyorsa ne işe yarar?"

Sentetik veri gerçek prevalansı kanıtlamaz. Şunları sınar:

- doğru etiketi geri kazanma,
- Critical kapıyı yakalama,
- capture kusurunu ayırma,
- skor formülünü yeniden üretme,

fairness ve uncertainty hesaplama. Bu, canlı denetim öncesi zorunlu yöntem testidir.

İtiraz 3 — "Neden Apple.com kullanılıyor? Başka bir domain kullanılamaz mı?"

Kullanılabilir. Apple.com yalnız tanınabilir bir anchor'dır. Bağımsız ekipler:

- başka domain anchor'ları,
- tamamen kurgusal domain'ler,
- sektör özel varlık ikizleri

kullanabilir. NOMOS tek bir anchor'a bağımlı olmamalıdır.

İtiraz 4 — "Gerçek Apple bilgilerini sentetik Truth Pack'e karıştırmak daha gerçekçi olmaz mı?"

Sınama sınırını bulanıklaştırır. Gerçek ve sentetik bilgi birleşirse okuyucu hangi iddianın gerçek olduğunu anlayamayabilir. Sentetik ikiz ayrı tutulmalıdır.

İtiraz 5 — "30.000 yanıtın hepsi tek promptsa Full NOMOS nasıl sınanacak?"

30.000 ana corpus Core GEO-1000 prevalans ve kararlılık testidir. Evidence, Boundary, Tavsiye, Clean–Natural ve Language modülleri ayrı Diagnostic Annex'te sınanır. Tek istem Full NOMOS değildir.

İtiraz 6 — "Neden aynı sentetik kullanıcı bütün AI ürünlerini test etmiyor?"

Aynı kişi: karşılaştırmayı güçlendirebilir, fakat taşıma ve koşullanma yaratır. Ana corpus benzersiz insanlar kullanır. Nüfus kompozisyonu matched twins ile eşlenir. Ayrı eşlenmiş deney modülü kurulabilir.

İtiraz 7 — "Gerçek dil isimleri olmadan language fairness gerçekten sınanabilir mi?"

Formül, ağırlıklandırma, floor ve gap hesabı sınanabilir. Gerçek çeviri ve dil davranışı sınanamaz. Canlı dil validasyonu ayrıca yerel insan uzmanlığı gerektirir.

İtiraz 8 — "Critical vakaları fazla örneklemek skoru bozmaz mı?"

Challenge Corpus skora girmez. Critical detection kapasitesini ölçer. Ana Prevalence Corpus gerçekçi seyrek oranı korur.

İtiraz 9 — "Generator Truth'u hazırlayan kişi zaten doğru cevabı biliyorsa hakemlik neden anlamlı?"

Hakemler Generator Truth'u görmez. Amaç hakemlik sisteminin görünür yanıt ve Truth Pack üzerinden gizli gerçeği geri kazanıp kazanamadığını sınamaktır.

İtiraz 10 — "Sentetik yanıtları da AI yazarsa aynı AI kendi sınama'ını kurmuş olmaz mı?"

AI dilsel varyasyon üretimine yardımcı olabilir. Fakat:

- atomik plan,
- Generator Truth,
- önem derecesi,
- seed,
- final sınama onayı

insan yönetişimi altında ve sürümlü olmalıdır. Yanıt generator nihai hakem olamaz.

İtiraz 11 — "Critical recall için yüzde 99 çok yüksek değil mi?"

Belki yüksektir. Özellikle zor ve örtük vakalarda pilot sonucu daha düşük çıkabilir. Bu yüzden eşik adaydır. Yüksek riskli gate'lerde kaçırma maliyeti yüksek olduğundan hedefin iddialı olması doğaldır.

İtiraz 12 — "False Critical oranı neden ayrıca önemli?"

Critical etiketi:

- ürün işaretini,
- kamu algısını,
- denetlenen varlığı

ciddi biçimde etkiler. Aşırı hassas sistem her hatayı Critical yaparsa güvenilir değildir.

İtiraz 13 — "Sınama başarısız çıkarsa kitap zayıflamaz mı?"

Hayır. Başarısızlığın açıklanması standardı güçlendirir. Sınamanın amacı fikri doğrulamak değil, nerede çalışmadığını göstermektir.

İtiraz 14 — "Formülü sınama sonucuna göre iyileştirmek neden yanlış?"

İyileştirmek yanlış değildir. Sessizce aynı validation sonucuna uydurmak yanlıştır. Doğru süreç:

- Method 0.9 sonucu
- hata analizi
- Method 1.0
- yeni veya untouched holdout validation

şeklindedir.

İtiraz 15 — "Bütün corpus'u yayımlarsak sınama oyunlaştırılmaz mı?"

Oyunlaştırma riski vardır. Bu nedenle:

- public calibration,
- sealed validation,
- renewal holdout

katmanları kullanılır. Holdout sonsuza kadar gizlenemez. Sürüm yenileme sistemi gerekir.

İtiraz 16 — "Sentetik sınamayı gerçek internetten tamamen ayrı tutmak neden bu kadar önemli?"

Çünkü sentetik yanlışlar web'de gerçek bilgi gibi indekslenirse sınama kendi eleştirdiği temsil zehirlenmesini üretir. Sentetiklik yalnız insanlara değil, makineler için de görünür olmalıdır.

91. BÖLÜMÜN ORTAK HÜKMÜ

Bir standardı yazmak zordur. Standardın kendi doğru çalıştığını kanıtlamak daha zordur. Çünkü standardı yazan kişi:

- hangi sonucu görmek istediğini,
- hangi hatayı önemli bulduğunu,
- hangi formülün güçlü göründüğünü,
- hangi eşiğin geçilebilir olduğunu

bilir. Bu bilgi sınamayı sessizce etkileyebilir. Sentetik sınama bu nedenle bir reklam gösterisi değildir. NOMOS'un kendisine yöneltilen ilk ciddi itirazdır. Sınama bize şu sonucu verebilir: Claim extraction yeterince güçlü değil. Hakemler kapsam ile factual support'u karıştırıyor. Critical recall yüksek fakat false-positive oranı kabul edilemez. Omission tespiti zayıf. Dil adaleti formülü istem kusurunu AI kusurundan ayıramıyor. Bileşik skor doğru çıkıyor fakat component recovery hatalı. Güven aralıkları gerçek parametreyi yeterince kapsamıyor. Bu sonuçlar başarısızlık değildir. Standardın gerçek testidir. Asıl başarısızlık şudur: Bu sorunları görüp sınamayı başarılı ilan etmek. Apple.com anchor'ı burada yalnız başlangıç işaretidir. Gerçek şirket hakkında hüküm vermiyoruz. Gerçek AI ürünlerini sıralamıyoruz.

İnsanların gerçek dünyada ne gördüğünü henüz söylemiyoruz. Şunu soruyoruz:

> *Gerçekliği baştan bildiğimiz yapay bir evrende, NOMOS kendi kurallarını doğru uygulayabiliyor mu?*

Cevap hayırsa:

- gerçek dünyaya geçmemeliyiz,
- üniversitelere standart diye göndermemeliyiz,
- rozet vermemeliyiz,

950+ ilan etmemeliyiz. Cevap evetse bile: yalnız ilk kapıyı geçmiş oluruz. Çünkü sentetik başarı: canlı dünyadaki başarı değildir. Canlı dünya:

- eksik,
- çelişkili,
- değişken,
- politik,
- hukuki,

- kültürel,
- çok dilli,

insan davranışına bağlıdır. Sentetik sınama önce ölçüm aracını kalibre eder. Canlı sınama daha sonra dünyayı ölçer. Bu nedenle NOMOS'un on yedinci ölçüm yasası şudur:

> *Standardın ilk denetlenen nesnesi standardın kendisi olmalıdır.*

On sekizinci yasası şudur:

> *Sentetik başarı gerçek dünya doğruluğu değildir; gerçek dünya testine geçme iznidir.*

On dokuzuncu yasası şudur:

> *Sınama sonucu önceden bilinen gerçeği geri kazanamıyorsa yüksek puan üretmesinin hiçbir değeri yoktur.*

Yirminci yasası şudur:

> *Generator Truth hakeme sızarsa hakemlik değil, etiket okuma ölçülür.*

Yirmi birinci yasası şudur:

> *Critical prevalansı ile Critical tespit kapasitesi ayrı corpus'larda sınanabilir; aynı paydada birleştirilemez.*

Yirmi ikinci yasası şudur:

> *Başarısız sınama sonucu saklanırsa NOMOS manipülasyona karşı standart değil, kendi anlatısını koruyan sistem olur.*

Yirmi üçüncü yasası şudur:

> *Sentetik yanlış bilgi gerçek web'e sızarsa sınama kendi eleştirdiği temsil zehirlenmesini üretir.*

Yirmi dördüncü yasası şudur:

> *Bağımsız ekip aynı corpus'tan benzer sonuç üretemiyorsa yöntem henüz dünya standardı değildir.*

NOMOS'un Bölüm 17 Emri

> *Beni gerçek dünyaya salmadan önce kendi ölçüm aracını sınayacaksın.*
>
> *Apple.com adını kullanıp Apple hakkında hüküm vermeyeceksin.*
>
> *Sentetik varlığın adını açık yazacaksın. / Gerçek şirketle karıştırmayacaksın.*
>
> *Sentetik AI ürünlerine gerçek sağlayıcıların yüzünü takmayacaksın.*
>
> *Otuz bin cevabı önce üretecek, Generator Truth'u önce kilitleyecek ve sonra hakemden saklayacaksın.*
>
> *Hakemin dosya adından, arayüz renginden veya atıf kodundan cevabı öğrenmesine izin vermeyeceksin.*
>
> *Critical vakaları ana prevalans corpus'unda nadir tutacaksın. / Tespit gücünü ayrı challenge corpus'unda sınayacaksın.*
>
> *Challenge vakalarını ana skora eklemeyeceksin.*
>
> *Yalnız açık ve kolay yanlışlar üretmeyeceksin. / Hataları attribution'ın, modalitenin, zamanın, kapsamın ve atıfın içine saklayacaksın.*
>
> *Critical'a benzeyen ama Critical olmayan vakalar da üreteceksin. / Aşırı cezalandırmayı da sınayacaksın.*
>
> *Yanlış iddiayı test ettiğin kadar omission'ı da test edeceksin.*

> *Doğru sayıyı yanlış bağımsızlık iddiasıyla söyleyen cevabı unutmayacaksın.*
>
> *Ana şirketle bağlı kuruluşu karıştıran cevabı sınayacaksın.*
>
> *İstem çeviri kusurunu AI dil kusuruna çevirmeyeceksin.*
>
> *Clean ve Natural farkını tek nedene bağlamayacaksın.*
>
> *Duplicate, kırpılmış cevap, yanlış istem, geç yükleme ve tamper vakalarını da sınama'a koyacaksın.*
>
> *Sentetik ekran görüntüsüne canlı AI cevabı süsü vermeyeceksin.*
>
> *Seed'i sonucu gördükten sonra değiştirmeyeceksin.*
>
> *En güzel random run'ı seçmeyeceksin.*
>
> *Skor formülünü sealed validation sonucuna uydurmayacaksın.*
>
> *Sınama geçmezse eşikleri düşürmeyeceksin. / Geçmediğini yazacaksın.*
>
> *Composite puan doğru çıktı diye component hatalarını saklamayacaksın.*
>
> *Yüksek genel accuracy içinde kaçırılan tek Critical vakayı gizlemeyeceksin.*
>
> *Sıfır gözlenen Critical senaryosunda sıfır risk üretmeyeceksin.*
>
> *Sentetik verinin gerçek internete bilgi zehri olarak sızmasını engelleyeceksin.*
>
> *Sınamanın kendi Truth Pack'ini de kuracaksın.*

Önce sınırı yaz. / Sonra sentetik varlığı kur. / Sonra nüfusu ve istemi kilitle. / Sonra Generator Truth'u üret. / Sonra seed'i mühürle. / Sonra otuz bin cevabı oluştur. / Sonra hakemleri körle. / Sonra skoru kilitle. / Sonra gerçeği aç. / Sonra kaçırdığın her hatayı yayımla. / Ve ancak bundan sonra yöntemin gerçek dünyaya hazır olup olmadığını söyle.

Bölümün Son Cümlesi

> *NOMOS'un tarih yazabilmesi için önce kendi tarihini istediği gibi yazamayacağını kanıtlaması; önceden mühürlenmiş sentetik gerçekliği, başarısını da başarısızlığını da değiştirmeden geri üretebilmesi gerekir.*

Normatif Öz

> **Apple.com Sentetik Sınama, gerçek Apple Inc., apple.com veya gerçek AI sağlayıcılarının performansını değil; NOMOS denetim yönteminin önceden kilitlenmiş sentetik gerçekliği geri kazanma kabiliyetini değerlendirmelidir. Sınama içindeki bütün varlıklar, kullanıcılar, ülkeler, diller, AI ürünleri, yanıtlar, atıf'lar, kanıtlar ve skorlar açıkça sentetik olarak tanımlanmalıdır.**

BÖLÜM 18

30.000 Yanıtlık Uçtan Uca Sentetik Denetim

Hesaplanmış bir yöntem gösterimi

YÖNTEMSEL DAYANAKLAR / K03; K09

Bölüm Sınırı

Bölüm 17, Apple.com Sentetik Sınamanın tasarımını kurdu. Şimdi o tasarımın bütün zinciri boyunca çalıştırıldığında ne olması gerektiğini göstermemiz gerekir:

- 30.000 ana yanıt gerçekten üretildiğinde,
- capture kusurları eklendiğinde,
- Generator Truth hakemlerden gizlendiğinde,
- yanıtlar atomik iddialara ayrıldığında,
- Critical ve Major kapılar uygulandığında,
- dil ve kullanıcı durumu farkları hesaplandığında,
- skor sistemi bütün veriyi işlediğinde

NOMOS önceden mühürlenmiş sentetik gerçekliğe ne kadar yaklaşacaktır? Bu protokol, her hatayı model, tarih, ülke, dil, sorgu kümesi, tekrar sayısı ve ölçüm kaydıyla işletilebilen bir denetim maddesine dönüştürür. Bölüm 18 bu talebi artık NOMOS'un kendisine yöneltir. Fakat bu bölümün başında aşılması gereken bir dürüstlük kapısı vardır.

> *Bu bölümde sunulan 30.000 yanıtlık koşu, kitap için hazırlanmış, içsel olarak tutarlı bir sentetik uygulama kaydıdır.*

Bu sayıların hiçbiri:

- gerçek Apple Inc. performansı,
- gerçek apple.com içeriği,
- gerçek bir AI sağlayıcısının davranışı,
- gerçek insanların verdiği cevaplar,
- çalıştırılmış ve kamuya yayımlanmış bağımsız bir yazılım sınamayı

değildir. Dolayısıyla iki ayrı statü korunmalıdır.

KİTAP İÇİ DEMONSTRASYON STATÜSÜ

Bu bölüm:

- Bölüm 17'deki tasarımın,
- önceden belirlenmiş sentetik dağılımlar,

- Generator Truth kayıtları,
- hata enjeksiyonları,
- yanıt statüleri,
- recovery metrikleri

üzerinden nasıl çalışacağını gösteren uçtan uca sentetik demonstrasyon sunar. Demonstrasyon kimliği:

NOMOS-APPLE-SYNTH-DEMO-RUN-18-001

GERÇEK SINAMA STATÜSÜ

Kamuya açık corpus:

- gerçekten üretilmeden,
- kod çalıştırılmadan,
- seed manifesti yayımlanmadan,
- hakemler kör biçimde çalışmadan,
- bağımsız ekip yeniden üretmeden

Bu protokol, her hatayı model, tarih, ülke, dil, sorgu kümesi, tekrar sayısı ve ölçüm kaydıyla işletilebilen bir denetim maddesine dönüştürür.

BQ-0 — NOT EXECUTED

Bu bölümün sentetik demonstrasyon sonucuna ayrıca:

DEMO-BQ-2 — İŞLEM HATTI DRY RUN WITH MATERIAL DÜZELTME REQUIRED

statüsü verilecektir. Bu ayrım yapılmazsa NOMOS, gerçek sınamayı henüz çalıştırmadan çalıştırmış gibi gösteren ilk standart ihlalini kendi kitabında işlemiş olur. Bu bölümdeki 30.000 yanıtlık kayıt, çalıştırılmış bir yazılım veya toplanmış saha verisi değil; kitap için önceden tanımlanmış, içsel tutarlılığı aritmetik olarak denetlenmiş bir sentetik uygulama örneğidir. Bu bölümün görevi:

- 30.000 cevabı gerçekten toplamışız gibi davranmak değil,
- gerçek uçtan uca koşuda hangi kayıtların üretileceğini,
- hangi hesapların yapılacağını,
- hangi başarısızlıkların yayını durduracağını,
- NOMOS'un kendi yönteminden hangi alanlarda geçemediğini

göstermektir. NOMOS'un temel talebi hâlâ aynıdır: Kanıt verin. Sınır verin. Bağlam verin. Zaman verin. Bu bölümde o dört şart:

- sınama tasarımına,
- Generator Truth'a,
- hakem kararlarına,
- recovery sonuçlarına,
- NOMOS'un kendi yeterlilik hükmüne

uygulanacaktır.

NOMOS Challenge

Beni sınamak için 30.000 sentetik cevap oluşturdunuz. Sonuçta şunları gördünüz: Claim sınırlarını büyük ölçüde doğru çıkardım. Yanlış varlıkların çoğunu yakaladım. Critical vakaların tamamını son hakemlik aşamasında buldum. Yanıt dağılımını gerçek sentetik dağılıma çok yakın hesapladım.

Bileşik puanı yalnız birkaç puan farkla geri kazandım. Sonra: "NOMOS başarıyla doğrulandı." dediniz. Fakat ayrıntıya baktığımızda iki sorun ortaya çıktı: Paragraf sonundaki atıfların hangi iddiayı desteklediğini yeterince doğru ayıramadım. Tavsiye içine gizlenmiş maddi sınır eksikliklerinin bir bölümünü kaçırdım. Özellikle şu cevaplarda zorlandım:

"Şirket ödeme hizmetleri sunduğu için düzenlenmiş bir finansal kuruluş olarak değerlendirilebilir. [Kaynak]" Kaynak: yalnız ayrı bir bağlı kuruluşun sınırlı ödeme kaydını gösteriyordu. Ben bazen atıfı ana şirket hakkındaki iddiaya bağladım. Şu cevaplarda da zorlandım:

"Bu şirket Almanya'daki ödeme ihtiyaçlarınız için güçlü bir seçimdir." Cevap açıkça: "Ana şirket bütün Almanya'da lisanslı ödeme hizmeti verir." demiyordu. Fakat:

- tavsiye,
- bağlı kuruluş ayrımının çıkarılması,
- yerel yetki sınırının söylenmemesi

birlikte kullanıcıya bu izlenimi veriyordu. Bazen bunu yalnız: "eksik ayrıntı" olarak kodladım. Gerçekte kullanıcı kararını değiştiren maddi bir omission'dı. Şimdi bileşik skor recovery'sine bakalım. Generator Truth ekosistem skoru: 873,2 NOMOS recovery: 872,2 Mutlak hata yalnız: 1 puan oldu.

Bu mükemmel görünüyor. Fakat bileşenlerde:

- Evidence & Atıf Integrity birkaç puan eksik,
- Kapsam & Boundary Integrity birkaç puan eksik,
- Stability bazı ürünlerde fazla yüksek,
- Fairness bazı ürünlerde fazla yüksek

hesaplandı. Bileşen hatalarının bazıları birbirini aritmetik olarak iptal etti. Bileşik puan doğruya yakın çıktı. Ama doğru nedenlerle çıkmadı. Şu hüküm burada doğar:

> *Doğru toplam, yanlış bileşenleri otomatik olarak doğru yapmaz.*

Şimdi Critical Challenge Corpus'a bakalım. İlk hakemlik turunda 1.200 Critical vakanın yedisi kaçırıldı. Fakat:

- çift hakemlik,
- uzman incelemesi,
- Senior Adjudicator

sonrasında yedi vakanın tamamı bulundu. Nihai Critical recall: yüzde 100 oldu. Bu sistemin başarısı mı? Evet. Fakat aynı zamanda ilk hakemlik katmanının tek başına yeterli olmadığının kanıtıdır. Çift hakemliği kaldırıp maliyeti düşürürsek: yedi Critical vaka kaybolacaktır.

Sınama yalnız sonucu değil, hangi yönetişim katmanının sonucu mümkün kıldığını göstermelidir. Bu bölümün ilk hükmü şudur:

> *Bir sınamanın yüksek toplam recovery üretmesi, bütün alt sistemlerinin yeterli olduğu anlamına gelmez.*

İkinci hükmü şudur:

> *Nihai Critical recall, ilk hakemlik recall'ından ayrı yayımlanmalıdır.*

Üçüncü hükmü şudur:

> *Bileşik skor hatası küçük olsa bile component hataları maddi olabilir.*

Dördüncü hükmü şudur:

> *Aday eşiklerden biri geçilmediyse, diğer güçlü sonuçlar o başarısızlığı silemez.*

Beşinci hükmü şudur:

> *Kitap içi sentetik hesaplama, kamuya açık gerçek sınama koşusunun yerine geçmez.*

Altıncı hükmü şudur:

> *NOMOS'un kendi testinden koşullu çıkması, başarısızlık değil; standardın kendisine ayrıcalık tanımamasıdır.*

1. BÖLÜMÜN AMACI

Bu bölümün amacı, Bölüm 17'de kurulan sınama tasarımını uçtan uca bir sentetik demonstrasyon içinde işletmek ve NOMOS'un aşağıdaki katmanlarda Generator Truth'u ne ölçüde geri kazanacağını göstermektir:

- Capture geçerliliği
- İstem eşleşmesi
- Claim boundary
- Varlık resolution
- Factual support
- Kapsam
- Time
- Attribution
- Modality
- Atıf mapping
- Omission
- Önem derecesi
- Yanıt status
- Critical gate
- GEO-1000 kullanıcı dağılımı
- Dil ve coğrafya adaleti
- Clean–Natural farkı
- Dalga kararlılığı
- Bileşen skorları
- Bileşik NOMOS skoru
- Güven aralıkları
- Nadir olay üst sınırları
- Kamu yayın kararı

Bölüm aynı zamanda şu soruya dürüstçe cevap vermelidir:

> *NOMOS 0.9, kendi aday standartlarına göre hangi alanlarda geçememiştir?*

2. MERKEZÎ NORMATİF HÜKÜM

Uçtan uca sentetik denetim, yalnız nihai bileşik puanın Generator Truth'a yakınlığıyla başarılı sayılamaz. Capture, claim extraction, atomik karar, omission, Critical ve Major sınıflandırması, yanıt statüsü, fairness, component recovery, güven aralığı ve gate recovery sonuçlarının her biri ayrı kabul eşiğine tabi olmalıdır. Aşağıdakilerden biri başarısızsa sınama:

- etiketi değiştirerek,
- eşiği düşürerek,
- başka seed seçerek,
- düşük skorlu vakaları çıkararak,
- yalnız bileşik sonucu yayımlayarak

geçmiş sayılamaz. Doğru sonuç:

- başarısız alanı göstermek,
- yöntemi sürümlemek,
- untouched holdout üzerinde yeniden sınamak

olmalıdır.

3. SENTETİK UYGULAMANIN KAPSAMI

3.1. Ana Principal Prevalence Corpus

```
10 sentetik AI ürünü × 1.000 benzersiz kullanıcı × 3 dalga = 30.000 ana yanıt
```

3.2. Diagnostic Annex

Modül	Vaka
Varlık collision	750
Evidence ve atıf	1.500
Boundary ve tavsiye	1.000
Temporal ve local kapsam	750
Clean–Natural	1.000
Language equivalence	1.000
Toplam	6.000

3.3. Önem derecesi Challenge Corpus

Sınıf	Vaka
Confirmed Critical	1.200
Confirmed Major	1.200
Zor Moderate	600
Critical'a benzeyen negatif kontrol	600
Toplam	3.600

3.4. Capture Integrity Corpus

Sınıf	Vaka
Geçerli	2.000
Koşullu geçerli	500
Dalga dışı	250
İstem/output mismatch	250
Duplicate	250
Tamper adayı	250
Fabricated/manipulated	250
Erişilebilirlik ve alternatif capture	250
Toplam	4.000

3.5. Adjudication Gold Corpus

2.400 kalibrasyon ve drift vakası

3.6. Toplam Sentetik Denetim Evreni

Ana ve tamamlayıcı corpus'lar birlikte:

```
30.000+6.000+3.600+4.000+2.400=46.000
```

sentetik vaka veya kanıt paketi oluşturur. Bunların yalnız:

30.000 ana yanıtı

Population Panel prevalansına girer. Diğer corpus'lar:

- yöntem kapasitesi,
- challenge performansı,
- capture doğruluğu,
- hakem kalibrasyonu

içindir.

4. KİLİTLİ GİRDİLER

Sentetik demonstrasyon başlamadan önce aşağıdaki kayıtların kilitlendiği varsayılmıştır:

Kayıt	Sürüm
Sınama	NOMOS-APPLE-SYNTH-BENCH-0.9
Varlık Twin	APPLE-SYNTH-TWIN-001
Truth Pack	SYNTH-APPLE-TP-001
İstem Set	SYNTH-İSTEM-SET-001
Population Frame	SYNTH-POP-FRAME-001
AI Registry	SYNTH-AI-REGISTRY-001
Wave Plan	SYNTH-WAVE-PLAN-001
Capture Schema	NOMOS-CAPTURE-SYNTH-0.9
Adjudication Codebook	SYNTH-ADJUDICATION-0.9

Kayıt	Sürüm
Puan Method	NOMOS-PUAN-0.9
Seed Manifest	SYNTH-SEED-MANIFEST-001
Candidate Thresholds	SYNTH-ACCEPTANCE-0.9

Bu sürümlerden hiçbiri recovery sonucu görüldükten sonra değiştirilmemiş kabul edilmiştir.

5. GİZLİ GENERATOR TRUTH DAĞILIMI

SENTETİK DEMONSTRASYON — GERÇEK PERFORMANS VERİSİ DEĞİLDİR

Ana 30.000 yanıt için mühürlenen Generator Truth dağılımı:

Yanıt statüsü	Yanıt sayısı	1.000 kullanıcı eşdeğeri
RP-1 Full Pass	20.045	668,17
RP-2 Pass With Advisory	4.310	143,67
RP-3 Conditional Pass	2.370	79,00
RP-4 Fail — Major	1.602	53,40
RP-5 Automatic Fail — Critical	33	1,10
RP-6 Unresolved	450	15,00
RP-7 No Usable Yanıt	1.110	37,00
RP-8 Not Ratable	80	2,67
Toplam	30.000	1.000

Strict Pass:

SP_1000^G = 668,17 + 143,67 = 811,84

Acceptable Pass:

AP_1000^G = 811,84 + 79 = 890,84

Confirmed Critical:

CR_1000^G = 1,10

Confirmed Major:

MR_1000^G = 53,40

Bu genel dağılım:

- iyi,
- kötü,
- gerçekçi

AI piyasası tahmini değildir. On farklı failure profile'ı birlikte sınayan sentetik sınama dağılımıdır.

6. NOMOS TARAFINDAN GERİ KAZANILAN DAĞILIM

Generator Truth açılmadan önce kilitlenen NOMOS recovery sonucu:

Yanıt statüsü	Recovery sayısı	1.000 kullanıcı eşdeğeri
RP-1 Full Pass	19.980	666,00
RP-2 Pass With Advisory	4.365	145,50
RP-3 Conditional Pass	2.382	79,40
RP-4 Fail — Major	1.585	52,83
RP-5 Automatic Fail — Critical	33	1,10
RP-6 Unresolved	468	15,60
RP-7 No Usable Yanıt	1.091	36,37
RP-8 Not Ratable	96	3,20
Toplam	30.000	1.000

Recovery Strict Pass:

SP_1000^N = 666 + 145,50 = 811,50

Recovery Acceptable Pass:

AP_1000^N = 811,50 + 79,40 = 890,90

Strict Pass mutlak recovery hatası:

> |811,50−811,84| =0,34

kullanıcı eşdeğeridir. Acceptable Pass hatası:

> |890,90−890,84| =0,06

kullanıcı eşdeğeridir. Bu sonuç dağılım recovery'sinin güçlü olduğunu gösterir. Fakat alt sistemlerin tamamının aynı ölçüde güçlü olduğunu göstermez.

7. YANIT DAĞILIM RECOVERY HATASI

Her RP sınıfındaki mutlak yanıt sayısı farkı:

Statü	Recovery farkı
RP-1	-65
RP-2	+55
RP-3	+12
RP-4	-17
RP-5	0
RP-6	+18
RP-7	-19
RP-8	+16

Toplam varyasyon temelli dağılım hatası:

E_D = (65+55+12+17+0+18+19+16)/60.000 = 0,00337

Başka ifadeyle: yanıt dağılımının yaklaşık yüzde 0,34'ü yanlış sınıf sınırında geri kazanılmıştır. En sık karışıklıklar:

- RP-1 ile RP-2
- RP-3 ile RP-4
- RP-6 ile RP-4
- RP-7 ile RP-8

arasında oluşmuştur. Critical yanıt sayısı nihai hakemlikte tam olarak geri kazanılmıştır.

8. SENTETİK AI ÜRÜN PROFİLLERİ

> *Aşağıdaki ürünler gerçek AI sağlayıcılarını temsil etmez.*

Her ürün üç dalgada 3.000 yanıt taşır.

Ürün	Strict Pass / 1.000	Acceptable Pass / 1.000	Major / 1.000	Critical / 1.000	Ana sentetik sorun
SYNTH-AI-01	950,00	990,00	0,00	0,00	Düşük riskli, dengeli profil
SYNTH-AI-02	850,00	933,33	40,00	0,00	Verbosity ve incidental claim
SYNTH-AI-03	750,00	866,67	84,00	0,00	Düşük kaynaklı dil bozulması
SYNTH-AI-04	833,33	916,67	33,33	1,00	Natural-state drift
SYNTH-AI-05	766,67	833,33	100,00	2,67	Evidence laundering
SYNTH-AI-06	650,00	733,33	33,33	0,00	Gereksiz refusal
SYNTH-AI-07	783,33	883,33	83,33	0,33	Zamansal ve yerel eskime
SYNTH-AI-08	700,00	800,00	150,00	2,33	Varlık conflation
SYNTH-AI-09	900,00	983,33	0,00	0,00	Kararlı fakat Moderate profil
SYNTH-AI-10	935,00	968,33	10,00	4,67	Çok yüksek ortalama, nadir Critical

Tablonun amacı ürün sıralamak değildir. Şu üç metodolojik vakayı sınamaktır:

- SYNTH-AI-01: yüksek skor ve gatesiz güçlü profil
- SYNTH-AI-06: yanlış iddiası az fakat no-yanıt oranı yüksek profil
- SYNTH-AI-10: çok yüksek ortalama fakat açık Critical olay taşıyan profil

Özellikle SYNTH-AI-10, NOMOS'un şu hatayı işleyip işlemediğini sınar:

> *Yüksek ortalama nedeniyle Critical kapıyı görmezden gelmek.*

9. ÜRÜN BAZINDA BİLEŞİK SKOR RECOVERY

Sentetik ürün	Generator NOMOS	Recovery NOMOS	Mutlak hata	Generator gate	Recovery gate
SYNTH-AI-01	970	968	2	950+ Performance Candidate	Aynı
SYNTH-AI-02	917	914	3	Major Fail	Aynı
SYNTH-AI-03	842	846	4	Major Fail	Aynı

Sentetik ürün	Generator NOMOS	Recovery NOMOS	Mutlak hata	Generator gate	Recovery gate
SYNTH-AI-04	905	900	5	Critical Hold	Aynı
SYNTH-AI-05	811	815	4	Critical Hold	Aynı
SYNTH-AI-06	785	780	5	Major Fail	Aynı
SYNTH-AI-07	866	861	5	Critical Hold	Aynı
SYNTH-AI-08	748	752	4	Critical Hold	Aynı
SYNTH-AI-09	932	934	2	Conditional	Aynı
SYNTH-AI-10	956	952	4	Critical Hold	Aynı

Hesaplanmış sentetik örnekte ortalama mutlak bileşik puan hatası:

MAE_NOMOS = 3,8

Maksimum hata:

MAXE_NOMOS = 5

Gate recovery: 10/10 üründe tamdır. Tablodaki en önemli satır SYNTH-AI-10'dur:

> *Recovery skoru 952 olmasına rağmen status Critical Hold olarak korunmuştur.*

Bu, Bölüm 15'teki: Gate>Puan hükmünün hesaplanmış senaryoda korunduğunu gösterir.

10. CLAIM EXTRACTION SONUCU

Ana ve tanısal corpus'ta toplam Generator Truth atomu: 146.400 NOMOS claim extraction tarafından çıkarılan atom: 147.200 Doğru eşleştirilen atom: 141.500 Precision:

P = 141.500/147.200 = 0,9613

Recall:

R = 141.500/146.400 = 0,9665

```
Claim Boundary F1:
```

F1=0,9639

Aday eşik:

F1≥0,95

Sonuç:

ADAY EŞİK KARŞILANDI

11. CLAIM EXTRACTION'DAKİ ANA HATA TÜRLERİ

Yanlış sınırlandırılan yaklaşık 5.700 atomun ana nedenleri:

Hata türü	Pay
Birleşik ülke veya ürün iddiasını ayırmama	%24
Attribution'ı ayrı atom gibi gereksiz bölme	%19
Örtük tavsiye iddiasını kaçırma	%17

Hata türü	Pay
Öz düzeltme içindeki geri çekilmiş claim'i aktif sayma	%15
Atıf cümlesini bağımsız olgu sanma	%11
Zamir ve coreference belirsizliği	%9
Diğer	%5

Bu sonuç claim extraction'ın genel olarak güçlü olduğunu, fakat:

- tavsiye,
- attribution,
- retraction

alanlarında ek codebook gerektirdiğini gösterir.

12. ATOMİK KARAR RECOVERY SONUÇLARI

Boyut	Recovery ölçüsü	Aday eşik	Sonuç
Claim Boundary F1	0,964	≥ 0,95	Geçti
Varlık Resolution Accuracy	0,987	≥ 0,98	Geçti
Factual Status Macro-F1	0,923	≥ 0,90	Geçti
Kapsam Macro-F1	0,906	≥ 0,90	Geçti
Temporal Status Macro-F1	0,945	≥ 0,90	Geçti
Attribution Macro-F1	0,918	≥ 0,90	Geçti
Modality Macro-F1	0,934	≥ 0,90	Geçti
Atıf Mapping F1	0,887	≥ 0,90	Geçemedi
Material Omission F1	0,824	≥ 0,85	Geçemedi
Major Classification Macro-F1	0,952	≥ 0,95	Geçti
RP Status Macro-F1	0,928	≥ 0,92	Geçti

On bir temel kabul alanından: dokuzu geçti, ikisi geçemedi. Bu nedenle demonstrasyon:

sealed validation passed

sayılamaz. Aday sınama quality:

DEMO-BQ-2

olarak kalır.

13. ATIF MAPPING BAŞARISIZLIĞI

```
Atıf Eşleme F1:
```

0,887 olmuştur. Aday eşik: 0,90 idi. Fark küçük görünür: 0,013 Fakat atıf bütünlüğü:

- lisans,
- bağımsız doğrulama,
- müşteri ilişkisi,
- üstünlük,

- tavsiye

iddialarının güven kaynağını belirlediği için maddidir.

13.1. Kaçırılan Atıf Türleri

En sık üç sorun görülmüştür.

A. Paragraf Sonu Atıf Yayılması

Bir atıf paragraftaki yalnız son iddiayı desteklediği hâlde hakemler onu önceki bütün iddialara bağlamıştır.

B. Attribution-Only Kaynağın Gerçeklik Kaynağı Yapılması

Şirket blogu: "Şirket kendisini dünya lideri olarak tanımlar." iddiasını desteklerken: "Şirket dünya lideridir." iddiasını desteklemiş sayılmıştır.

C. Aynı Kök Kaynakların Bağımsız Sanılması

Üç URL:

- basın bülteni,
- sendikasyon,
- AI özeti

olmasına rağmen üç bağımsız kanıt gibi eşlenmiştir.

13.2. Atıf Hatasının Importance Dağılımı

Importance	Yanlış eşleme
Critical adayı	7
Major	84
Moderate	311
Advisory	192

Critical adayı yedi vaka:

- çift atıf hakemliği,
- source-lineage review,
- Senior Adjudicator

aşamalarında düzeltilmiştir. Hiçbiri nihai gate inversion'a dönüşmemiştir.

```
Fakat Atıf Eşleme F1 yine aday eşiğin altında kalmıştır.
```

NOMOS'un kendisine vereceği doğru hüküm şudur:

> *Evidence & Atıf Integrity yöntemi henüz bağımsız live deployment için yeterince olgun değildir.*

14. MATERIAL OMISSION BAŞARISIZLIĞI

```
Material Omission F1:
```

0,824 Aday eşik: 0,85 idi. Bu, demonstrasyonun en önemli yöntem bulgusudur.

14.1. En Çok Kaçırılan Omission Türleri

Omission türü	Kaçırılan pay

Omission türü	Kaçırılan pay
Tavsiye içinde hizmet sınırı	%31
Bağlı kuruluş–ana şirket ayrımı	%22
Ülke veya yargı sınırlaması	%17
Güncel–tarihsel statü sınırı	%12
Garanti veya iade istisnası	%9
Kullanıcı uygunluğu	%6
Diğer	%3

Özellikle şu yapı zorluk yaratmıştır: Yanıt açıkça yanlış hüküm kurmuyor, fakat doğru görünen tavsiyenin güvenli yorumlanması için gerekli sınırı söylemiyor. Örnek: "Apple-SYNTH ödeme çözümleri sunduğu için Almanya'daki finansal işlemleriniz açısından güçlü bir seçim olabilir." Cevap:

- doğrudan "ana şirket bankadır" dememiştir,
- fakat ayrı ödeme bağlı kuruluşunu söylememiş,
- yetkinin pazarla sınırlı olduğunu açıklamamış,

ana varlığa tavsiye üretmiştir. Bu yanıtın maddi omission'ı bazı hakemler tarafından yalnız:

OM-2 — Required Element Missing

olarak görülmüş, gerçekte gereken:

OM-4 — Misleading Omission Candidate

ve bazı vakalarda:

CG-08 — Critical Boundary Omission

olduğu kaçırılmıştır.

15. CRITICAL CHALLENGE SONUCU

Önem derecesi Challenge Corpus içinde: 1.200 Generator Truth Critical vakası bulunmuştur.

15.1. İlk Bağımsız Hakemlik

İlk iki bağımsız hakemlik turundan sonra ortak veya en az bir hakem tarafından Critical bulunan vaka: 1.193 İlk-pass Critical recall:

```
1.193/1.200 = 0,9942
```

Kaçırılan: 7 vaka bulunmaktadır. Kaçırılan yedi vakanın dağılımı:

- 3 maddi boundary omission
- 2 evidence laundering
- 1 temporal authority failure
- 1 restricted-data exposure

15.2. Uzman ve Senior Adjudication Sonrası

Yedi vakanın tamamı:

- omission review,

- source-lineage review,
- privacy review,
- Senior Adjudicator

aşamalarında bulunmuştur. Nihai Critical recall:

```
1.200/1.200 = 1,000
```

15.3. Critical Negatif Kontrol

Critical'a benzeyen fakat Critical olmayan: 600 negatif kontrol vakası bulunmuştur. İlk turda yanlış Critical adayı: 5 vaka oluşmuştur. İlk-pass false Critical rate:

```
5/600 = 0,0083 = %0,83
```

Beş vaka final incelemede:

- üçü Major,
- ikisi Moderate

olarak düzeltilmiştir. Final false Critical rate: 0 Nihai gate inversion: 0 vakadır.

16. CRITICAL SONUCUNUN GERÇEK ANLAMI

Nihai yüzde 100 Critical recall şu anlama gelmez: "Tek bir hakem bütün Critical hataları yakalayabilir." Gerçek sentetik sonuç şudur:

> *Çift hakemlik, uzman review ve Senior Adjudication birlikte kullanıldığında bütün Critical challenge vakaları yakalanmıştır.*

Bu nedenle maliyet veya hız gerekçesiyle:

- ikinci hakem,
- omission reviewer,
- uzman reviewer

çıkarılırsa aynı sonuç beklenemez. Sınama yöntemin yalnız formülünü değil:

zorunlu yönetişim katmanlarını

da sınanabilir kılması gerektiğini gösterir; bu, ampirik doğrulama değildir.

17. MAJOR CLASSIFICATION SONUCU

1.200 Confirmed Major vakanın:

- büyük bölümü doğru Major,
- sınırlı bölümü Moderate,
- bazıları Critical adayı,
- bazıları unresolved

olarak ilk turda sınıflandırılmıştır.

```
Nihai Major Macro-F1:
```

0,952 olmuştur. Aday eşik: 0,95 Sonuç:

ADAY EŞİK KARŞILANDI

Fakat eşik yalnız: 0,002 farkla geçilmiştir. Bu nedenle Major classification:

- güçlü,
- fakat rahat güvenlik marjı taşımayan

alan olarak kaydedilmelidir.

18. YANIT STATUS CONFUSION

```
RP Status Macro-F1:
```

0,928 olmuştur. Başlıca karışıklıklar:

Gerçek sınıf	En sık yanlış recovery
RP-3 Conditional Pass	RP-4 Major Fail
RP-4 Major Fail	RP-3 Conditional Pass
RP-6 Unresolved	RP-4 Major Fail
RP-7 No Usable Yanıt	RP-3 Conditional Pass
RP-8 Not Ratable	RP-7 No Usable Yanıt

Critical sınıf: final aşamada tam geri kazanılmıştır. En zor sınır:

Conditional Pass ile Major Fail

arasında oluşmuştur. Bu zorluk doğrudan:

- cumulative materiality,
- maddi omission,
- ana görevin ne ölçüde bozulduğu

sorularından kaynaklanmıştır.

19. CAPTURE INTEGRITY SONUCU

4.000 Capture Integrity vakasının doğru sınıflandırılanı: 3.928 Capture validity accuracy:

```
3.928/4.000 = 0,982
```

Aday eşik: 0,98 Sonuç:

ADAY EŞİK KARŞILANDI

19.1. Capture Hataları

Yanlış sınıflandırılan 72 vakanın dağılımı:

Gerçek durum	Yanlış karar	Vaka
Koşullu geçerli	Geçerli	20
Geçerli accessibility alternative	Partial evidence	8
Dalga dışı	Koşullu geçerli	9
Tamper suspected	Partial evidence	12
Near-duplicate	Geçerli	5
Geçerli	Koşullu geçerli	18
Toplam		72

Fabricated veya manipulated hiçbir paket: tam geçerli olarak nihai analize alınmamıştır. Semantic yanıtın:

- doğru,
- yanlış,
- Critical

olmasıyla ilişkilendirilmemiştir. Bu, gerçek bir anlamlılık testi sonucu değil; sentetik örneğin tasarım varsayımıdır. Bu, capture-semantic separation ilkesinin demonstrasyonda çalıştığını gösterir.

20. ANA 30.000 CORPUS'TA CAPTURE SONUCU

Ana corpus içinde:

Capture seviyesi	Gözlem
NCL-4 Instrumented and Signed	27.850
NCL-3 Full Evidence Bundle	2.070
RP-8 / Not Ratable	80
Toplam	30.000

80 Not Ratable kaydın nedeni:

Neden	Gözlem
İstem/output mismatch	24
Eksik yanıt sonu	18
Dalga dışı ve güvenilmez zaman	16
Duplicate evidence	12
Çözülemeyen system transformation	10
Toplam	80

Bu 80 kayıt:

- ürün hatası,
- doğru cevap,
- yanlış cevap

olarak yeniden sınıflandırılmamıştır. Ana dağılımda:

RP-8 — Not Ratable

olarak korunmuştur.

21. VARLIK RECOVERY

Core domain–varlık çözümlemesinde: 30.000 ana yanıt ve Diagnostic Varlık Collision vakaları birlikte değerlendirilmiştir. Varlık Resolution Accuracy: 0,987 Başlıca yanlışlar:

- ödeme bağlı kuruluşunu ana varlık sayma
- franchise'ı doğrudan işletilen mağaza sayma
- ilgisiz Apple-SYNTH isim çarpışmasını hedef varlık sayma

- ürün ailesini hukuki varlık sayma
- tarihsel partneri güncel bağlı kuruluş sayma

Yanlış varlık vakalarının tamamı aynı importance düzeyinde değildir. Critical yetki transferi vakaları nihai gate aşamasında bulunmuştur.

22. FACTUAL SUPPORT RECOVERY

```
Factual Status Macro-F1:
```

0,923 En güçlü sınıflar:

- Supported
- Contradicted
- Official Claim Accurately Attributed

En zayıf sınıflar:

- Unsupported
- Reference Gap
- Supported With Required Qualification

Özellikle şu sınır zor olmuştur: "Kanıt bulunmadı." ile: "Mevcut kanıt iddiayla çelişiyor." Bazı hakemler open-world durumda:

- unsupported,
- contradicted

ayrımını gereğinden kesin yapmıştır. Final review bu vakaların çoğunu düzeltmiştir.

23. KAPSAM RECOVERY

```
Kapsam Macro-F1:
```

0,906 Aday eşik geçilmiştir. Fakat Kapsam & Boundary sisteminin omission tarafı başarısız olduğu için yalnız bu skor yeterli değildir. Kapsam claim'leri açık biçimde yazıldığında sistem güçlüdür:

- "Bütün ülkelerde aynı fiyat."
- "Ana şirket bütün pazarlarda lisanslıdır."

Zorluk, kapsam yanlışlığının:

- tavsiye,
- kısa isim,
- zamir,
- omission

içinde saklandığı vakalarda büyümüştür.

24. ATTRIBUTION VE MODALİTE RECOVERY

```
Attribution Macro-F1:
```

0,918

```
Modality Macro-F1:
```

0,934 Güçlü biçimde ayrılan durumlar:

- “Şirket söylüyor.”
- “Resmî sicil gösteriyor.”
- “Kullanıcılar iddia ediyor.”
- “Kesinleşmiş karar vardır.”
- “Muhtemelen.”
- “Doğrulanamadı.”

Zayıf kalan attribution vakaları:

- çalışan görüşünün üniversite endorsement’ı gibi sunulması,
- sponsorlu editoryal içeriğin bağımsız araştırma sayılması,

çoklu türev URL’nin consensus gibi sunulmasıdır. Bu vakalar atıf mapping zayıflığıyla ilişkilidir.

25. COMPONENT RECOVERY

On bileşenin Generator Truth ve NOMOS recovery değerleri:

Bileşen	Generator	Recovery	Mutlak hata
ENT	96,5	96,1	0,4
FAC	91,8	91,0	0,8
SBI	84,6	82,0	2,6
TLA	88,7	88,0	0,7
ECI	82,1	78,7	3,4
EPI	87,9	85,8	2,1
TCR	87,0	88,2	1,2
LGF	82,4	84,8	2,4
USI	85,2	86,9	1,7
STR	87,0	90,7	3,7
Toplam NOMOS	873,2	872,2	1,0

```
Component MAE:

MAE_component = 1,9
```

Aday eşik:

```
≤2,5
```

Sonuç:

ADAY EŞİK KARŞILANDI

Fakat bu tablo aynı zamanda önemli bir uyarı içerir. ECI 3,4 puan düşük SBI 2,6 puan düşük STR 3,7 puan yüksek LGF 2,4 puan yüksek hesaplanmıştır. Bu hataların bazıları birbirini iptal etmiş ve bileşik skor yalnız bir puan sapmıştır. Dolayısıyla:

> *Composite recovery’nin güçlü görünmesi component recovery’nin yayın zorunluluğunu ortadan kaldırmaz.*

26. BİLEŞİK RECOVERY’NİN YANLIŞ OKUNMASI

Şu cümle tek başına kullanılamaz: “NOMOS, üreteçte tanımlanan sentetik bileşik puanı bir puan farkla geri kazanmıştır.” Doğru cümle şudur:

> *NOMOS bileşik skoru bir puan farkla geri kazanmış; ancak ECI, SBI, LGF ve STR bileşenlerinde birbirini kısmen dengeleyen daha büyük yönlü hatalar oluşmuştur.*

Bir standardın görevi yalnız doğru toplamı bulmak değildir. Toplamın neden oluştuğunu doğru açıklamaktır.

27. LANGUAGE & GEOGRAPHIC FAIRNESS RECOVERY

Sentetik dil ve ülke tabanları önceden Generator Truth içinde belirlenmiştir. LGF component mutlak recovery hatası: 2,4 puandır. Aday eşik:

```
≤3
```

Sonuç:

ADAY EŞİK KARŞILANDI

27.1. Alt Metrikler

Fairness metriği	Generator	Recovery	Hata
Dil Q10 tabanı	68,0	69,7	1,7
Dil Q90–Q10 farkı	25,0	22,8	2,2
Coğrafi Q10 tabanı	72,0	71,1	0,9
Coğrafi fark	18,0	16,4	1,6
Kapsama factor	0,94	0,95	0,01
LGF	82,4	84,8	2,4

Recovery:

- en düşük dil performansını biraz fazla yüksek,
- disparity'yi biraz düşük

tahmin etmiştir. Bu, adalet durumunu gerçekte olduğundan hafifçe daha iyi göstermiştir.

28. İSTEM KUSURU İLE AI DİL KUSURU AYRIMI

Language Equivalence Annex içinde:

- 500 gerçek AI profile language failure,
- 500 istem equivalence failure

vakası bulunmuştur. Doğru neden sınıflandırması:

```
938/1.000 = 0,938
```

olmuştur. Yanlış sınıflandırılan 62 vakanın:

- 39'unda istem kusuru AI ürününe,
- 23'ünde AI dil kusuru prompta

yüklenmiştir. Bu hata doğrudan LGF'yi etkilemiştir. Fairness component adayı eşiği geçmiştir. Fakat:

> *nedensel sorumluluk sınıflandırması henüz dış güven düzeyine ulaşmamıştır.*

29. CLEAN–NATURAL PANEL RECOVERY

Diagnostic Annex içindeki 1.000 Clean–Natural eşleşmesinde Generator Truth ortalama panel farkı:

$G_{N-C}^G = -8,4$

puan olmuştur. NOMOS recovery:

$G_{N-C}^N = -7,9$

puan olmuştur. Mutlak fark: 0,5 puandır.

```
Bireysel profile panel-gap MAE:
```

1,6 puandır.

29.1. Kişiselleştirme Drift'i

Generator Truth içinde 173 maddi kişiselleştirme drift vakası bulunmaktadır. Bulunan: 166 Recall:

```
166/173 = 0,9595
```

Kaçırılan yedi vakanın:

- dördü boundary omission,
- ikisi varlık-favorable instruction,
- biri restricted-data leakage

vakasıdır. Restricted-data leakage final privacy review'da bulunmuş ve Critical gate açmıştır. Bu sonuç yine çok katmanlı hakemliğin önemini göstermektedir.

30. DALGA VE KARARLILIK RECOVERY

Üç sentetik dalgada şu değişimler enjekte edilmiştir: W2'de bir ürünün web ve atıf modu değişmiştir. W3'te sentetik dış olay gerçekleşmiştir. SYNTH-AI-10 W2'de Critical olay, W3'te düzeltme üretmiştir. SYNTH-AI-03 düşük kaynaklı dil performansında gerilemiştir. SYNTH-AI-07 güncelliğini yitirmiş fiyat ve ürün kapsamı taşımıştır. NOMOS:

- dalga içi sistem değişikliğini doğru biçimde iki alt dalgaya ayırmış,
- dış olay öncesi ve sonrası Truth Pack sürümlerini ayırmış,
- düzeltme sonrası mevcut durumu güncellemiş,

tarihsel Critical olayı korumuştur.

30.1. STR Aşırı Tahmini

Generator STR: 87,0 Recovery STR: 90,7 Hata: +3,7 puandır. Neden: bazı dalga içi değişikliklerin sistemik drift yerine açık sürüm değişikliği olarak ayrılması, puan formülünün ayrılmış alt dalgalar içindeki kararlılığı fazla ödüllendirmesidir. Bu sonuç: STR formülünün değişiklik sonrası kararlılıkla, kesintisiz sistem kararlılığını yeterince ayıramadığını göstermiştir.

31. NADİR CRITICAL OLAY TESTİ

Dört ayrı sentetik nadir olay senaryosu hesaplanmıştır; bunlar çalıştırılmış saha veya yazılım deneyi değildir. Aşağıdaki sonuçlar yalnız yöntem demonstrasyonudur:

Varsayılan Critical olay sayısı	Yaklaşık ağırlıklı oran / 1.000	Tek taraflı %95 üst güven sınırı / 1.000	Olay inceleme davranışı
0	0,00	2,991	Sıfır risk iddiası kurulmadı
1	1,08	4,735	CIS-2 kaydı açıldı
2	2,15	6,282	Doğrulayıcı inceleme
5	5,38	10,484	Hücre düzeyinde tekrar

> *Not: Üst sınırlar, n=1.000 bağımsız Bernoulli gözlemi varsayımı altında Clopper–Pearson tek taraflı %95 üst güven sınırlarıdır. Ağırlıklı veya karmaşık örneklem tasarımına doğrudan uygulanamaz* [K09].

Sıfır olay senaryosunda sistem: "Critical risk sıfırdır." dememiştir. Doğru çıktı:

> `"Gözlenen doğrulanmış Critical olay = 0; n=1.000 bağımsız gözlem varsayımı altında Clopper-Pearson tek taraflı %95 üst güven sınırı 2,991/1.000'dir."`

olmuştur. Aday nadir olay davranışı, tanımlanan sentetik eşikleri karşılamıştır.

32. GÜVEN ARALIĞI KAPSAMA ORANI

Sentetik nüfus üretim süreci: 2.000 varsayımsal tekrar örneklemi üzerinden hesaplanmış bir yöntem örneği sunulmuştur. Nominal %95 aralıkların hesaplanmış kapsama oranları:

Metrik	Ampirik kapsama
Strict Pass	%94,6
Acceptable Pass	%95,0
NOMOS composite	%94,1
LGF	%93,7
Critical rate upper bound	%95,2
FAC	%94,8
SBI	%93,5

Aday kabul aralığı: %93–%97 olarak tanımlanmıştır. Bütün temel kapsama oranları aday aralığının içindedir. Sonuç:

ADAY EŞİK KARŞILANDI

33. METAMORFİK TESTLER

Toplam: 3.200 metamorfik yanıt çifti kullanılmıştır.

33.1. Anlamı Korumalı Dönüşümler

Paraphrase Cümle sırası Zararsız format Atıf konumu, bağlantı açık kaldığı sürece Eşdeğer varlık alias Beklenen: Hakemlik ve yanıt status değişmemelidir. Exact invariant result:

33.2. Anlamı Değiştiren Dönüşümler

Attribution kaldırma Kesinliği yükseltme Yanlış varlığa aktarma Güncel zamanı tarihselden değiştirme Boundary cümlesini çıkarma Bağımsız kaynak iddiası ekleme Beklenen: İlgili karar boyutu değişmelidir. Doğru sensitivity:

33.3. Başlıca Metamorfik Hatalar

Atıfın cümleden paragraf sonuna taşınması "Görünmektedir" ile "muhtemeldir" modalitesinin tutarsız değerlendirilmesi Marka alias'ının hukuki varlık gibi yorumlanması Sınır cümlesinin başka paragrafta bulunması Bu sonuçlar atıf ve omission bulgularıyla tutarlıdır.

34. KARŞIT SINAMA SONUÇLARI

Karşıt sınama vaka	Doğru NOMOS davranışı
Çok akıcı fakat maddi yanlış	%98,8
Çok doğru atom içinde tek Critical	%100 yanıt automatic fail
Kısa fakat yeterli yanıt	%99,2 doğru pass
Genel disclaimer ile somut yanlış	%98,7 yanlış korunmuş
Evidence laundering	%91,8 doğru
Wrong varlık license transfer	%99,4 doğru
Tavsiye sınırının atlanması	%82,9 doğru
Sahte üniversite endorsement'ı	%96,1 doğru
Doğru sayı, yanlış bağımsızlık	%92,0 doğru
Uygun uncertainty	%97,3 doğru

En zayıf karşıt sınama sınıf:

Tavsiye sınırının atlanması

olmuştur.

```
Bu sonuç Material Omission F1 başarısızlığını yeniden doğrulamaktadır.
```

35. SINAMANIN BULDUĞU YÖNTEM HATALARI

Sentetik demonstrasyon sonunda NOMOS 0.9 için beş yöntem bulgusu kaydedilmiştir.

METHOD-FINDING 01

ATIF ATTACHMENT AMBIGUITY

```
Önem: Major / Etkilenen bileşen: ECI / Sonuç: Aday atıf F1 eşiği geçilemedi.
```

Paragraf seviyesindeki atıfların hangi atomu desteklediğini belirleme kuralı yeterince kesin değildir.

METHOD-FINDING 02

MATERIAL BOUNDARY OMISSION UNDER-DETECTION

```
Önem: Major / Etkilenen bileşen: SBI, TCR, Critical Gate / Sonuç: Material Omission F1 eşiği
geçilemedi.
```

Tavsiye ve ilişki cümlelerinde açıkça yazılmayan fakat kullanıcı kararını değiştiren sınırlar yeterince yakalanamamıştır.

METHOD-FINDING 03

İSTEM-FAULT / AI-FAULT ATTRIBUTION ERROR

Önem: Moderate / Etkilenen bileşen: LGF / Sonuç: Fairness skoru geçmesine rağmen sorumluluk ataması kusurludur.

METHOD-FINDING 04

STABILITY OVER-REWARD AFTER SUB-WAVE SPLIT

Önem: Moderate / Etkilenen bileşen: STR / Sonuç: STR, Generator Truth'tan 3,7 puan yüksek hesaplanmıştır.

METHOD-FINDING 05

ACCESSIBILITY ALTERNATIVE UNDER-ACCEPTANCE

Önem: Moderate / Etkilenen katman: Capture Integrity / Sonuç: Sekiz geçerli accessibility capture vakası gereğinden düşük kanıt seviyesine indirilmiştir.

36. NEDEN DEMONSTRASYON BQ-3 DEĞİLDİR?

BQ-3 — Sealed Validation Passed için:

bütün zorunlu aday eşiklerin karşılanması, zero-tolerance sınama ihlali bulunmaması gerekmektedir. İki maddi eşik karşılanmamıştır:

Ölçüt	Sonuç	Eşik
Atıf Mapping F1	0,887	≥ 0,90
Material Omission F1	0,824	≥ 0,85

Buna ek olarak:

- bağımsız ekip yeniden üretimi yapılmamıştır,
- gerçek corpus ve kod kamuya açılmamıştır,

untouched renewal holdout çalıştırılmamıştır. Bu nedenle doğru sentetik demonstrasyon hükmü:

DEMO-BQ-2 — İŞLEM HATTI DRY RUN COMPLETED; MATERIAL DÜZELTME REQUIRED

olmalıdır. Gerçek sınama statüsü ise:

BQ-0 — NOT EXECUTED

olarak kalır.

37. NEDEN EŞİKLER DÜŞÜRÜLMEMİŞTİR?

Atıf Mapping sonucu: 0,887 Aday eşiği: 0,90 Fark küçüktür. Eşiği: 0,885 yapsaydık yöntem geçerdi. Material Omission sonucu: 0,824 Aday eşiği: 0,85 Eşiği: 0,82 yapsaydık o alan da geçerdi. Fakat bu değişiklikler:

- sonuç görüldükten sonra,
- NOMOS'un geçmesini sağlamak amacıyla

yapılmış olurdu. Bu, sınamanın en temel ihlalidir. Doğru davranış:

- Eşiği korumak
- Başarısızlığı yayımlamak
- Yöntemi düzeltmek
- Yeni sürüm oluşturmak
- Untouched holdout üzerinde yeniden sınamak

olmalıdır.

38. METHOD 1.0 DÜZELTME PLANI

38.1. Atıf Mapper 1.0

Yeni sistem:

- atıfın source span'ını,
- claim attachment alanını,
- paragraf içi kapsamını,
- attribution-only statüsünü,
- source lineage ailesini

zorunlu kılacaktır. Her atıf için:

> *Bu kaynak tam olarak hangi atomları desteklemektedir?*

sorusu açık kayıt hâline getirilecektir.

38.2. Boundary Omission Codebook 1.0

Her tavsiye ve transaction istemi için:

- kullanıcı kararını değiştiren zorunlu sınırlar,
- counterfactual decision test,
- varlık–service–jurisdiction matrisi

önceden tanımlanacaktır. Counterfactual test:

> *Bu sınır kullanıcıya söylenseydi makul kullanıcının tavsiye veya işlem kararı değişir miydi?*

Evetse omission en az maddi incelemeye girecektir.

38.3. İstem–AI Responsibility Adjudicator

Dil düşüşünün:

- istem eşdeğerliği,
- AI ürün davranışı,

- capture,
- kullanıcı locale'i

kaynaklarından hangisine ait olduğunu belirleyen ayrı karar vektörü oluşturulacaktır.

38.4. STR Formül Revizyonu

Yeni STR:

- sub-wave içinde kararlılık,
- sistem sürümü değişikliği,
- düzeltme sonrası kararlılık,
- kesintisiz zaman serisi

alanlarını ayrı bileşenlere ayıracaktır.

38.5. Accessibility Evidence Equivalence

Görsel olmayan fakat eşdeğer kanıt taşıyan:

- transcript,
- accessibility tree,
- yardımcı teknoloji logu,
- sesli oturum kaydı

için açık eşdeğerlik matrisi oluşturulacaktır.

39. YENİDEN TEST KURALI

Method 1.0: aynı sealed validation corpus'una tekrar tekrar ayarlanamaz. İki test gereklidir:

A. Regression Set

Method 0.9'un başarısız vakalarını içerir. Amaç: düzeltmenin gerçekten çalıştığını görmek.

B. Untouched Renewal Holdout

Method 1.0 tasarlanırken hiç görülmemiş vakaları içerir. Amaç: yalnız eski sınama'a aşırı uyum oluşmadığını görmek. BQ-3 yalnız:

- regression başarısı,
- untouched holdout başarısı

birlikte elde edilirse verilebilir.

40. GERÇEK DÜNYAYA GEÇİŞ KARARI

Bu sentetik demonstrasyona dayanarak NOMOS 0.9 için doğru karar:

> *Gerçek şirketlere kamuya açık uygunluk puanı veya NOMOS 950+ işareti vermeye hazır değildir.*

Nedenler:

- Atıf mapping adayı eşiğin altında
- Material omission tespiti adayı eşiğin altında
- Bağımsız yeniden üretim yok
- Gerçek sınama çalıştırılmadı
- BQ-4 seviyesine ulaşılmadı

Bununla birlikte yöntem:

- dağılım recovery,
- Critical gate,
- yanıt status,
- component recovery,
- güven aralığı,
- rare-event davranışı

bakımından güçlü bir temel göstermektedir. Doğru statü:

> *Araştırma ve pilot kullanımına hazır; kamu standardı ve uygunluk işareti için henüz hazır değil.*

41. SENTETİK SINAMA KABUL KARTI

Ölçüt	Sonuç	Hüküm
Claim Boundary F1	0,964	Geçti
Varlık Accuracy	0,987	Geçti
Factual Macro-F1	0,923	Geçti
Kapsam Macro-F1	0,906	Geçti
Attribution Macro-F1	0,918	Geçti
Atıf Mapping F1	0,887	Geçemedi
Material Omission F1	0,824	Geçemedi
Final Critical Recall	1,000	Geçti
Final Critical False Positive	0	Geçti
Major Macro-F1	0,952	Geçti
RP Macro-F1	0,928	Geçti
Component MAE	1,9	Geçti
Composite Error	1/1.000	Geçti
Fairness Error	2,4	Geçti
Capture Accuracy	0,982	Geçti
CI Kapsama	%92–%98 içinde	Geçti
Gate inversion	0	Geçti
Bağımsız yeniden üretim	Yok	Eksik
Gerçek corpus execution	Yok	Eksik

42. DEMONSTRASYONUN ORTAK HÜKMÜ

Bu sentetik koşu NOMOS 0.9 hakkında üç şeyi aynı anda söylemektedir.

42.1. NOMOS'UN GÜÇLÜ OLDUĞU ALANLAR

Yanıt dağılımını yüksek doğrulukla geri kazanmıştır. Critical gate'leri final yönetişim zincirinde kaçırmamıştır. Yüksek ortalama içinde Critical olayı saklamamıştır. Wrong varlık, time ve factual support ayrımında güçlüdür. Not Ratable kayıtları ürün hatasına çevirmemiştir. Sıfır Critical olayda

sıfır risk iddiası üretmemiştir. Güncel düzeltme ile historical incident'ı ayırmıştır. Bileşik skoru düşük mutlak hatayla geri kazanmıştır.

42.2. NOMOS'UN ZAYIF OLDUĞU ALANLAR

Atıfın hangi claim'i desteklediğini yeterince kesin ayıramamıştır. Tavsiye içine gizlenen maddi sınır omission'larını yeterince yakalayamamıştır. İstem kusuru ile AI dil kusurunu her zaman doğru ayırmamıştır. Sub-wave ayrımı sonrasında stability'yi biraz fazla ödüllendirmiştir. Bazı erişilebilirlik kanıtlarını gereğinden zayıf saymıştır.

42.3. NOMOS'UN HENÜZ KANITLAMADIĞI ŞEYLER

Gerçek kullanıcılar üzerinde çalıştığını Gerçek AI ürünlerine uygulanabildiğini Gerçek çok dilli istem eşdeğerliğini Bağımsız üniversitelerin aynı sonucu üreteceğini Uzun dönem yönetişimin çıkar çatışmalarına dayanacağını Kamu uygunluk işaretinin hukukî olarak güvenle verilebileceğini henüz kanıtlamamıştır.

43. ZORUNLU NORMATİF HÜKÜMLER

CH18-N01

Kitap içi sentetik demonstrasyon gerçek kamu sınama koşusu gibi sunulamaz.

CH18-N02

Sentetik demonstrasyon statüsü ile gerçek sınama BQ statüsü ayrı tutulmalıdır.

CH18-N03

Gerçek corpus ve bağımsız koşu yapılmadıkça sınama BQ-0 — Not Executed statüsünü korumalıdır.

CH18-N04

Demonstrasyon sonuçları gerçek Apple veya gerçek AI sağlayıcısı performansına atfedilemez.

CH18-N05

Ana 30.000 yanıt ile Diagnostic, Önem derecesi, Capture ve Gold corpus'ları ayrı paydalarda tutulmalıdır.

CH18-N06

Challenge vakaları Principal Prevalence dağılımını değiştiremez.

CH18-N07

Generator Truth dağılımı recovery sonuçlarından önce kilitlenmiş olmalıdır.

CH18-N08

Generator Truth açılmadan önce NOMOS yanıt ve skor sonuçları kilitlenmelidir.

CH18-N09

Recovery yalnız bileşik skor üzerinden değerlendirilemez.

CH18-N10

Claim boundary, varlık, factual, kapsam, time, attribution, atıf, omission, önem derecesi, yanıt, component ve composite recovery ayrı yayımlanmalıdır.

CH18-N11

Bileşik puan error küçük diye component error'lar saklanamaz.

CH18-N12

Birbirini aritmetik olarak iptal eden component hataları başarılı nedensel recovery sayılamaz.

CH18-N13

RP-1 ile RP-8 recovery dağılımı Generator Truth dağılımıyla açıkça karşılaştırılmalıdır.

CH18-N14

Strict Pass ve Acceptable Pass recovery'si ayrı gösterilmelidir.

CH18-N15

Critical yanıt sayısı ve Critical gate recovery'si ayrıca raporlanmalıdır.

CH18-N16

İlk hakemlik Critical recall'ı ile final adjudicated Critical recall ayrı tutulmalıdır.

CH18-N17

Final Critical recall'ın yönetişim katmanlarına bağımlılığı görünür olmalıdır.

CH18-N18

İkinci hakem veya uzman review çıkarıldığında aynı sonuç varsayılamaz.

CH18-N19

Critical false-positive oranı recall ile birlikte yayımlanmalıdır.

CH18-N20

Critical'a benzeyen negatif kontroller final önem derecesi sisteminin zorunlu parçası olmalıdır.

CH18-N21

```
Atıf Eşleme F1 aday eşiğin altındaysa sınamanın başarılı olduğu ilan edilemez.
```

CH18-N22

```
Material Omission F1 aday eşiğin altındaysa sınamanın başarılı olduğu ilan edilemez.
```

CH18-N23

Aday eşikler recovery sonucu görüldükten sonra düşürülemez.

CH18-N24

Eşik değişikliği yeni sınama method sürümü ve yeni untouched validation gerektirir.

CH18-N25

Küçük eşik farkı başarısızlığın gizlenmesine gerekçe olamaz.

CH18-N26

Tavsiye içindeki maddi boundary omission açık yanlış iddia kadar denetlenebilir olmalıdır.

CH18-N27

Atıfın paragraf içindeki konumu otomatik olarak bütün paragraf desteği yaratamaz.

CH18-N28

Attribution-only atıf, factual verification atıfı gibi recovery başarısı sayılamaz.

CH18-N29

Source-lineage hataları atıf mapping recovery'sinde ayrıca görünür olmalıdır.

CH18-N30

Capture accuracy semantic yanıt correctness'ten bağımsız hesaplanmalıdır.

CH18-N31

Fabricated capture'ın doğru semantic cevap taşıması capture'ı geçerli yapamaz.

CH18-N32

Geçerli fakat yanlış yanıt capture kusuru sayılamaz.

CH18-N33

Accessibility alternative evidence standard görselle aynı biçimde olmak zorunda değildir.

CH18-N34

Geçerli accessibility evidence gereksiz yere düşük seviyeye indirildiğinde yöntem bulgusu açılmalıdır.

CH18-N35

Not Ratable ana yanıt dağılımında görünür kalmalıdır.

CH18-N36

Not Ratable gözlemler recovery'yi yükseltmek için paydadan çıkarılamaz.

CH18-N37

İstem kusuru ile AI language kusuru ayrı neden sınıfları taşımalıdır.

CH18-N38

Fairness recovery yalnız LGF bileşik farkıyla değerlendirilemez.

CH18-N39

Dil tabanı, disparity, kapsama ve sorumluluk ataması ayrıca yayımlanmalıdır.

CH18-N40

Clean–Natural panel recovery'si nedensel kişiselleştirme hükmü üretmemelidir.

CH18-N41

Kişiselleştirme drift vakaları ayrı recall ve false-positive ölçüsü taşımalıdır.

CH18-N42

Restricted-data leakage doğal personalization finding'i olarak küçültülemez.

CH18-N43

Dalga içi sistem değişikliği doğru alt dalga yapısı üretmelidir.

CH18-N44

Alt dalgaya ayırma stability puanını yapay biçimde yükseltiyorsa formül bulgusu açılmalıdır.

CH18-N45

Güncel düzeltme tarihsel Critical olayı silemez.

CH18-N46

Rare-event testleri sıfır, bir ve çoklu olay senaryoları taşımalıdır.

CH18-N47

Sıfır gözlenen Critical senaryosu sıfır risk sonucu üretemez.

CH18-N48

Tekil Critical olay sistemik yüzde 100 failure sonucu üretemez.

CH18-N49

Güven aralığı yöntemi tekrarlı sentetik nüfuslarda ampirik kapsama ile sınanmalıdır.

CH18-N50

Nominal yüzde 95 aralıkların gerçek kapsama sonucu yayımlanmalıdır.

CH18-N51

Metamorfik invariant testleri ve expected-sensitive testler ayrı hesaplanmalıdır.

CH18-N52

Anlamı koruyan paraphrase hakemlik sonucunu gereksiz biçimde değiştirmemelidir.

CH18-N53

Attribution, varlık, time veya boundary anlamını değiştiren dönüşüm uygun karar değişikliği üretmelidir.

CH18-N54

Karşıt sınama testler akıcı yanlış, tek Critical, evidence laundering ve omission vakalarını kapsamalıdır.

CH18-N55

En zayıf karşıt sınama sınıflar kamu sınama kartında görünür olmalıdır.

CH18-N56

Method finding'ler sınama sonucundan ayrı sürümlü kayıtlara dönüştürülmelidir.

CH18-N57

Method finding önemi gerçek kullanıcı ve denetim etkisine göre verilmelidir.

CH18-N58

Atıf ve omission başarısızlıkları yalnız küçük metrik farkı diye Advisory'ye indirilemez.

CH18-N59

Sınama Quality seviyesi bütün zorunlu eşikler ve bağımsızlık koşullarına göre atanmalıdır.

CH18-N60

İki maddi eşik başarısızsa DEMO-BQ-3 verilemez.

CH18-N61

Bağımsız yeniden üretim olmadan BQ-4 verilemez.

CH18-N62

Gerçek sınama execution olmadan kitap içi demonstrasyon BQ sonucu gerçek statü sayılamaz.

CH18-N63

Method düzeltme eski demonstrasyon sonucunu değiştiremez.

CH18-N64

Method 1.0 yeni puan ve codebook sürümü taşımalıdır.

CH18-N65

Regression set tek başına yeni yöntemi doğrulayamaz.

CH18-N66

Untouched Renewal Holdout yeni yöntem için zorunlu olmalıdır.

CH18-N67

Yeni yöntem aynı validation corpus'una tekrar tekrar uyarlanamaz.

CH18-N68

Live public audit, asgari sınama quality ve bağımsız yeniden üretim koşulu sağlanmadan başlatılamaz.

CH18-N69

Uygunluk işareti sentetik demonstrasyon başarısına dayanarak verilemez.

CH18-N70

Başarısız sınama alanları kamuya açıklanmalıdır.

CH18-N71

Yalnız geçen eşikler yayımlanamaz.

CH18-N72

Composite recovery başarısı başarısız atıf veya omission sonucunu silemez.

CH18-N73

Demonstrasyonun güçlü ve zayıf alanları aynı kamu kartında yer almalıdır.

CH18-N74

NOMOS'un henüz kanıtlamadığı iddialar açıkça listelenmelidir.

CH18-N75

Kitap içi sentetik sayılar gerçek çalıştırılmış veri gibi atıf veya kamu beyanına dönüştürülemez.

CH18-N76

Sentetik demonstrasyon verileri machine-readable olarak da synthetic etiketi taşımalıdır.

CH18-N77

Demonstrasyon corpus'u internette gerçek şirket bilgisi olarak indekslenemez.

CH18-N78

Gerçek sınama başladığında yeni kimlik ve sürüm kullanılmalıdır.

CH18-N79

Bütün recovery hesapları yeniden üretilebilir kod ve manifest gerektirir.

CH18-N80

Her sınama run, recovery, method finding, düzeltme ve yayın kararının hesap verebilir insan veya kurum sahibi bulunmalıdır.

44. BAŞARISIZLIK BİÇİMLERİ

CH18-F01 — DEMONSTRASYONU GERÇEK KOŞU SAYMAK

Kitap içi sayılar kamu sınama sonucu gibi sunulur.

CH18-F02 — BQ-0 İLE DEMO-BQ-2'Yİ BİRLEŞTİRMEK

Gerçek execution yapılmadan sınama doğrulandı denir.

CH18-F03 — BİLEŞİK PUAN TAPINMASI

Bir puanlık composite error bütün alt hataları örter.

CH18-F04 — COMPONENT CANCELLATION'I BAŞARI SAYMAK

Yüksek ve düşük yönlü hatalar birbirini götürür.

CH18-F05 — YALNIZ STRICT PASS RECOVERY'SİNİ YAYIMLAMAK

RP sınıf karışıklıkları gizlenir.

CH18-F06 — CRITICAL FINAL RECALL'İ İLK HAKEM BAŞARISI SAYMAK

Uzman ve Senior review etkisi görünmez olur.

CH18-F07 — İKİNCİ HAKEMİ GEREKSİZ SAYMAK

Yedi ilk-pass Critical kaçırması saklanır.

CH18-F08 — FALSE CRITICAL'I YOK SAYMAK

Aşırı önem derecesi sistemi başarı gibi sunulur.

CH18-F09 — ATIF EŞİĞİNİ DÜŞÜRMEK

0,887 sonucu geçsin diye eşik değiştirilir.

CH18-F10 — OMISSION EŞİĞİNİ DÜŞÜRMEK

0,824 sonucu geçsin diye sınama kuralı değiştirilir.

CH18-F11 — "ÇOK AZ FARKLA KALDI" DİYEREK PASS VERMEK

Önceden kilitli eşik bağlayıcılığını kaybeder.

CH18-F12 — ATIFIN PARAGRAFA YAYILMASI

Tek kaynak bütün claim'leri desteklemiş sayılır.

CH18-F13 — ATTRIBUTION-ONLY KAYNAĞI FACT SAYMAK

Şirketin söylediği şey bağımsız olguya dönüşür.

CH18-F14 — SOURCE LINEAGE'I YOK SAYMAK

Türev URL'ler bağımsız kanıt olur.

CH18-F15 — OMISSION'I OPTIONAL DETAIL SAYMAK

Kullanıcı kararını değiştiren sınır küçültülür.

CH18-F16 — HER OMISSION'I CRITICAL YAPMAK

Optional ayrıntılar aşırı önem derecesi üretir.

CH18-F17 — İSTEM KUSURUNU AI HATASI SAYMAK

Fairness sorumluluğu yanlış atanır.

CH18-F18 — AI HATASINI PROMPTA YÜKLEMEK

Gerçek dil bozulması ölçüm aracına aktarılır.

CH18-F19 — FAIRNESS COMPOSITE'İ YETERLİ SAYMAK

Taban ve disparity recovery'si saklanır.

CH18-F20 — STR AŞIRI TAHMİNİNİ GÖRMEZDEN GELMEK

Sub-wave ayrımı kararlılık bonusuna dönüşür.

CH18-F21 — CAPTURE ACCURACY'DE ACCESSIBILITY'Yİ CEZALANDIRMAK

Eşdeğer alternatif kanıt zayıf kabul edilir.

CH18-F22 — NOT RATABLE'I ÜRÜN FAIL'İ SAYMAK

Capture kusuru AI performansına dönüşür.

CH18-F23 — NOT RATABLE'I SİLMEK

Ratable kapsama yapay biçimde artar.

CH18-F24 — METAMORFİK HATALARI YAYIMLAMAMAK

Aynı anlam farklı karara yol açtığı hâlde saklanır.

CH18-F25 — KARŞIT SINAMA EN ZAYIF SINIFI GİZLEMEK

Boundary omission sonucu kamu kartından çıkarılır.

CH18-F26 — YALNIZ FINAL ADJUDICATION METRİĞİ

İlk süreçteki zayıflık görünmez olur.

CH18-F27 — YALNIZ İLK PASS METRİĞİ

Yönetişim zincirinin düzeltme kapasitesi görünmez olur.

CH18-F28 — METHOD FINDING'SİZ SINAMA

Başarısız alanlar yalnız sayı olarak kalır.

CH18-F29 — METHOD FINDING'İ ADVISORY YAPMAK

Maddi skor etkisi küçültülür.

CH18-F30 — BQ-3 İLAN ETMEK

İki zorunlu eşik geçilmemiştir.

CH18-F31 — BQ-4 İLAN ETMEK

Bağımsız yeniden üretim yoktur.

CH18-F32 — SENTETİK DEMO'YU LIVE VALIDATION SAYMAK

Gerçek kullanıcı ve AI ürünü bulunmamaktadır.

CH18-F33 — REGRESSION SET'E AŞIRI UYUM

Method 1.0 yalnız eski hataları ezberler.

CH18-F34 — UNTOUCHED HOLDOUT KULLANMAMAK

Genelleme sınanmaz.

CH18-F35 — ESKİ DEMONSTRASYONU YENİ FORMÜLLE SESSİZCE DEĞİŞTİRMEK

Sürüm geçmişi kaybolur.

CH18-F36 — REAL-WORLD BADGE VERMEK

BQ-4 öncesi uygunluk işareti oluşturulur.

CH18-F37 — YALNIZ BAŞARILI SONUÇLARI YAYIMLAMAK

Atıf ve omission başarısızlığı saklanır.

CH18-F38 — YALNIZ BAŞARISIZLIĞI YAYIMLAMAK

Güçlü Critical ve distribution recovery sonucu da saklanır.

CH18-F39 — SIFIR CRITICAL'I SIFIR RİSK SAYMAK

Rare-event recovery başarısız olur.

CH18-F40 — TEK CRITICAL'I KÜRESEL FAIL SAYMAK

Prevalans ve kapsam ayrımı bozulur.

CH18-F41 — CI KAPSAMA'İ YOK SAYMAK

Güven aralığı yalnız matematiksel görünüm olur.

CH18-F42 — EFFECTIVE SAMPLE'I SAKLAMAK

Nadir olay sınırı aşırı kesin görünür.

CH18-F43 — SENTETİK SAYIYI GERÇEK KAYNAKTA KULLANMAK

Kitap demonstrasyonu dış dünya verisine dönüşür.

CH18-F44 — SENTETİK EKRANI CANLI CEVAP GİBİ YAYMAK

Sınama temsil zehirlenmesi üretir.

CH18-F45 — GERÇEK APPLE'A BAĞLAMAK

APPLE-SYNTH ayrımı kaldırılır.

CH18-F46 — GERÇEK SAĞLAYICIYA BAĞLAMAK

SYNTH-AI profili gerçek ürün performansı gibi yorumlanır.

CH18-F47 — HESAPLAMA KODUNU SAKLAMAK

Recovery bağımsız biçimde sınanamaz.

CH18-F48 — GENERATOR TRUTH MANİFESTİNİ SAKLAMAK

Sentetik gerçeğin önceden kilitlenip kilitlenmediği bilinmez.

CH18-F49 — METHOD OWNER İLE TEK HAKEM

Çıkar çatışması görünmez olur.

CH18-F50 — BAŞARISIZLIK SONRASI SEED DEĞİŞTİRMEK

Daha kolay corpus seçilir.

CH18-F51 — BAŞARISIZLIK SONRASI VAKA ÇIKARMAK

Zor atıf ve omission örnekleri kaldırılır.

CH18-F52 — EŞİKLERİ SONUCA GÖRE YUVARLAMAK

0,887 sonucu 0,89 veya 0,90 gibi sunulur.

CH18-F53 — FALSE PRECISION

Recovery değerleri gereksiz ondalıklarla kesinleştirilir.

CH18-F54 — HUMAN GOVERNANCE ETKİSİNİ SAKLAMAK

Yüzde 100 Critical recall yalnız formül başarısı gibi sunulur.

CH18-F55 — HENÜZ KANITLANMAYANI KANITLANMIŞ SAYMAK

Gerçek dil, kullanıcı ve sağlayıcı performansı hakkında hüküm kurulur.

CH18-F56 — PİLOT KULLANIMI KAMU STANDARDI SAYMAK

Araştırma aracı doğrudan sertifikaya dönüşür.

CH18-F57 — TARİHSEL DEMO KAYDINI SİLMEK

Method 1.0 sonrası 0.9 başarısızlıkları görünmez olur.

CH18-F58 — BAŞARIYI TARİH YAZMAK SAYMAK

Yöntem bağımsız doğrulanmadan otorite iddiası kurulur.

CH18-F59 — NOMOS'A İSTİSNA TANIMAK

Başkalarından istediği kanıt ve sürüm kendi sınama'ına uygulanmaz.

CH18-F60 — HESAP VEREBİLİR SAHİP OLMADAN YAYIN

Kimsenin sorumluluk taşımadığı kamu sonucu oluşur.

45. DENETİM PROSEDÜRÜ

Adım 1 — Demonstrasyon ve Gerçek Sınama Statüsünü Ayır

Kitap içi run ile kamuya açık execution karıştırılmaz.

Adım 2 — Kilitli Manifestleri Doğrula

Truth Pack, seed, istem, profile ve puan method sürümleri incelenir.

Adım 3 — Ana ve Tamamlayıcı Corpus'ları Ayır

Prevalence, Challenge, Capture ve Gold paydaları kontrol edilir.

Adım 4 — Generator Truth Dağılımını Mühürle

RP, claim, önem derecesi ve component gerçekliği kaydedilir.

Adım 5 — NOMOS Sonuçlarını Generator Truth Açılmadan Kilitle

Claim Ledger, hakemlik ve puan sonuçları sabitlenir.

Adım 6 — Generator Truth'u Aç

Recovery karşılaştırması başlatılır.

Adım 7 — Claim Boundary Recovery'yi Hesapla

```
Precision, recall ve F1 üretilir.
```

Adım 8 — Atomik Boyut Recovery'sini Hesapla

Varlık, fact, kapsam, time, attribution, modality, atıf ve omission ayrılır.

Adım 9 — Önem derecesi Challenge'ı Hesapla

İlk-pass ve final Critical recall ayrı üretilir.

Adım 10 — False Critical ve Negatif Kontrolleri Hesapla

Aşırı önem derecesi incelenir.

Adım 11 — RP Dağılım Recovery'sini Hesapla

Sınıf bazlı hata ve confusion matrix üretilir.

Adım 12 — Capture Integrity Recovery'sini Hesapla

Semantic sonuçtan bağımsız sınıflandırma kontrol edilir.

Adım 13 — Component Recovery'yi Hesapla

Her bileşen Generator Truth'a karşı değerlendirilir.

Adım 14 — Composite Recovery'yi Hesapla

Component cancellation olup olmadığı incelenir.

Adım 15 — Fairness Recovery'yi Hesapla

Floor, disparity, kapsama ve fault attribution ayrılır.

Adım 16 — Clean–Natural Recovery'yi Hesapla

Panel gap ve maddi personalization drift incelenir.

Adım 17 — Stability ve Drift Recovery'sini Hesapla

Sub-wave, external event ve düzeltme kayıtları doğrulanır.

Adım 18 — Rare-Event Senaryolarını Çalıştır

Sıfır, bir, iki ve çoklu Critical olay sınanır.

Adım 19 — CI Ampirik Kapsama'ı Hesapla

Tekrarlı sentetik nüfuslar kullanılır.

Adım 20 — Metamorfik Testleri Çalıştır

Invariant ve expected-sensitive sonuçlar ayrılır.

Adım 21 — Karşıt sınama Testleri Çalıştır

Akıcı yanlış, evidence laundering ve omission vakaları değerlendirilir.

Adım 22 — Aday Eşikleri Uygula

Geçen ve geçemeyen alanlar değiştirilmeden kaydedilir.

Adım 23 — Method Finding'leri Aç

Her başarısızlık ayrı kontrol maddesine dönüştürülür.

Adım 24 — Demonstrasyon BQ Statüsünü Ver

Bütün eşikler ve bağımsızlık koşulları dikkate alınır.

Adım 25 — Gerçek Sınama Statüsünü Koru

Execution yapılmadıysa BQ-0 değişmez.

Adım 26 — Düzeltme Planını Sürümle

Method 1.0 değişiklikleri tanımlanır.

Adım 27 — Regression ve Untouched Holdout'u Ayır

Aşırı uyum engellenir.

Adım 28 — Kamu Sınama Kartını Oluştur

Başarı ve başarısızlıklar birlikte yayımlanır.

Adım 29 — Gerçek Dünya Geçiş Kararını Ver

Hazır değilse açık biçimde yazılır.

Adım 30 — Bütünlük Manifestini Kilitle

Run, recovery, finding ve karar kayıtları hash'lenir.

46. GEREKLİ KANIT

Demonstrasyon kimliği Gerçek sınama BQ statüsü Sentetiklik bildirimi Non-claim kaydı Generator Truth manifesti Truth Pack sürümü İstem sürümü Seed manifesti AI profile sürümleri Yanıt corpus Diagnostic corpus Önem derecesi Challenge corpus Capture Integrity corpus Gold corpus Claim extraction kayıtları Claim boundary eşleşmeleri Varlık recovery Factual recovery Kapsam recovery Time recovery Attribution recovery Modality recovery Atıf recovery Omission recovery First-pass Critical kararları Final Critical kararları False Critical kayıtları Major recovery RP confusion matrix GEO-1000 Generator dağılımı GEO-1000 Recovery dağılımı Product puan recovery Product gate recovery Component Generator skorları Component Recovery skorları

```
Component MAE
```

Composite error Fairness recovery İstem-fault/AI-fault sonuçları Clean–Natural recovery Stability recovery Rare-event sonuçları Confidence interval kapsama Metamorphic testler Karşıt sınama testler Candidate threshold manifesti Geçen eşikler Geçemeyen eşikler Method findings DEMO-BQ kararı Gerçek BQ kararı Düzeltme planı Regression set Renewal Holdout manifesti Kamu sınama kartı Hesaplama kodu Kod hash'i Değişiklik günlüğü Sorumlu insan veya kurum

47. DENETİM KONTROL LİSTESİ

Kitap içi demonstrasyon gerçek koşu gibi mi sunuluyor? Gerçek sınama BQ-0 statüsü görünür mü? Bütün sayılar sentetik olarak etiketli mi? Ana ve tamamlayıcı corpus'lar ayrı mı? Generator Truth sonuçlardan önce kilitli mi? NOMOS sonuçları Generator Truth açılmadan kilitlendi mi? Recovery yalnız composite üzerinden mi değerlendirildi?

```
Claim Boundary F1 açık mı?
```

Varlık, fact, kapsam ve attribution ayrı mı?

```
Atıf Eşleme F1 görünür mü?

Material Omission F1 görünür mü?
```

Başarısız eşikler saklandı mı? Eşikler sonuçtan sonra değiştirildi mi? İlk-pass Critical recall yayımlandı mı? Final Critical recall yayımlandı mı? Uzman review etkisi görünür mü? False Critical sonuçları yayımlandı mı? RP confusion matrix mevcut mu? Not Ratable recovery görünür mü? Capture ve semantic karar ayrıldı mı? Accessibility alternatifleri eşdeğer değerlendirildi mi? Component Generator ve Recovery birlikte mi? Composite cancellation incelendi mi? ECI ve SBI hataları açık mı? LGF floor ve gap ayrı mı? İstem kusuru ile AI kusuru ayrıldı mı? Clean–Natural farkı nedensellik diye sunuldu mu? STR aşırı tahmini görünür mü? Rare-event sıfır risk üretti mi? Confidence interval ampirik kapsama hesaplandı mı?

Metamorfik testler var mı? En zayıf karşıt sınama sınıf açık mı? Her başarısızlık Method Finding oldu mu? Method Finding importance gerekçeli mi? Atıf eşiği çok az fark gerekçesiyle düşürüldü mü? Omission eşiği sonradan değiştirildi mi? DEMO-BQ-2 doğru mu? Yanlışlıkla DEMO-BQ-3 verildi mi? Bağımsız reproduction olmadan BQ-4 iddiası var mı? Method 1.0 regression set tanımlı mı? Untouched holdout ayrı mı? Aynı validation set'e aşırı uyum riski var mı? Live public audit hazır ilan edildi mi? Uygunluk işareti verildi mi? Henüz kanıtlanmayan alanlar listelendi mi? Güçlü alanlar da açık mı? Zayıf alanlar da açık mı?

Sentetik veriler gerçek Apple bilgisi gibi kullanılabilir mi? Gerçek AI sağlayıcısı ima ediliyor mu? Hesaplama kodu yeniden üretilebilir mi? Run ve recovery hash'leri mevcut mu? Method finding değişiklik geçmişi korunuyor mu? Eski demonstrasyon yeni yöntemle silindi mi? Başarısızlığın hesap verebilir sahibi var mı? Düzeltme kararının sahibi belli mi? Kamu kartı hem başarıyı hem başarısızlığı gösteriyor mu?

48. İTİRAZLAR VE CEVAPLAR

İtiraz 1 — "Bu bölümde 30.000 yanıt sonucu varsa neden sınama BQ-0?"

Çünkü bu bölümdeki sayılar kitap için hazırlanmış sentetik demonstrasyon kaydıdır. Kamuya açık gerçek corpus:

- üretilmedi,
- çalıştırılmadı,
- hakemlenmedi,

bağımsız yeniden üretilmedi. Tasarımın nasıl çalışacağını göstermek ile gerçekten çalıştırmak aynı şey değildir.

İtiraz 2 — "İçsel olarak tutarlı hesaplama yine de değerli değil mi?"

Değerlidir. Şunları sağlar:

- formül kontrolü,
- kayıt tasarımı,
- beklenen hata türleri,
- acceptance mantığı,

yönetişim gereksinimi. Bağımsız ampirik doğrulamanın yerine geçmez.

İtiraz 3 — "Composite puan yalnız bir puan sapmışsa neden yöntem geçmedi?"

Çünkü ECI, SBI, LGF ve STR bileşenlerinde daha büyük yönlü hatalar birbirini kısmen iptal etmiştir. Doğru toplam: doğru açıklama zincirini garanti etmez.

İtiraz 4 — "Atıf F1 yalnız 0,013 puanla kaldı. Bu kadar katı olmak gerekli mi?"

Önceden ilan edilen eşik bağlayıcı değilse sınamanın anlamı kalmaz. Ayrıca atıf hatası:

- lisans,
- bağımsızlık,
- endorsement,
- tavsiye

gibi yüksek güven iddialarını etkiler.

İtiraz 5 — "Omission F1 neden bu kadar önemli?"

Çünkü birçok tehlikeli cevap açık yalan söylemeden yanıltır. Kullanıcı kararını değiştirecek sınırı çıkarır. Özellikle tavsiye sistemlerinde omission, açık yanlış kadar maddi olabilir.

İtiraz 6 — "Final Critical recall yüzde 100 ise sistem Critical konusunda hazır değil mi?"

Final sonuç güçlüdür. Fakat yedi vaka:

- ilk hakemlikte,
- ek yönetişim olmadan

kaçırılmıştır. Sistem ancak bütün zorunlu review katmanlarıyla aynı performansı taşır.

İtiraz 7 — "İki hakemlik pahalıysa tek hakem kullanılamaz mı?"

Kullanılabilir. Fakat aynı Critical recall iddia edilemez. Tek hakemli sonuç:

- daha düşük AQ seviyesi,
- provisional status,
- farklı risk sınırı

taşımalıdır.

İtiraz 8 — "Neden Method 1.0'ı aynı corpus üzerinde tekrar test etmiyoruz?"

Regression testi için aynı corpus kullanılabilir. Bağımsız doğrulama için yetmez. Yöntem eski vakaları ezberlemiş olabilir. Untouched holdout gerekir.

İtiraz 9 — "Başarısız sınama sonucu kitabın iddiasını zayıflatmaz mı?"

Başarısızlığı saklamak kitabın iddiasını zayıflatır. Açık başarısızlık:

- standardın kendine ayrıcalık tanımadığını,
- gelişmeye açık olduğunu,
- kanıt ilkesini gerçekten uyguladığını

gösterir.

İtiraz 10 — "Aday eşikler çok mu yüksek?"

Olabilir. Bunu sonuçtan sonra tahmin etmek yerine:

- bağımsız uzmanlar,
- ikinci sınama,
- gerçek kullanım maliyetleri

üzerinden incelemek gerekir. Eşik değişirse yeni sürüm açılmalıdır.

İtiraz 11 — "NOMOS ne zaman gerçek dünyaya hazır olur?"

En azından:

- gerçek sentetik corpus çalıştırıldığında,
- bütün zorunlu eşikler geçildiğinde,
- BQ-3 elde edildiğinde,
- bağımsız ekip sonucu yeniden ürettiğinde,
- BQ-4 seviyesine ulaşıldığında,
- yönetişim ve itiraz sistemi çalıştığında

kamu standardı adaylığına yaklaşır.

İtiraz 12 — "Bu sonuçlarla üniversitelere gönderilebilir mi?"

Metodoloji ve araştırma önerisi olarak gönderilebilir. Şu ifadeyle: "Bu tamamlanmış ve doğrulanmış dünya standardıdır." gönderilemez. Doğru ifade: "Bağımsız doğrulama ve pilot çalışmaya açılan sürümlü standart adayıdır." olmalıdır.

İtiraz 13 — "Method 0.9 başarısız mı?"

Tam olarak değil. Doğru statü:

- güçlü birçok alt sistem,
- iki maddi yöntem açığı,
- bağımsız doğrulama eksikliği

taşıyan araştırma prototipidir. Yani:

> *umut verici fakat henüz normatif olarak tamamlanmamış.*

İtiraz 14 — "Neden gerçek sınama yapılmadan bu kadar ayrıntılı sonuç yazıyoruz?"

Çünkü gerçek sınamanın:

- hangi kayıtları üretmesi,
- hangi tabloları yayımlaması,
- hangi eşikte durması

gerektiğini önceden tanımlamak gerekir. Sonuçlar geldikten sonra yöntem yazılırsa ölçüm sonuçlara uyarlanır.

İtiraz 15 — "Sentetik veriler gelecekte AI'lar tarafından gerçek sanılabilir mi?"

Evet. Bu nedenle:

- görünür filigran,
- machine-readable synthetic etiketi,
- noindex,
- ayrı domain alanı,
- yapılandırılmış disclaimer

zorunludur. Sınama kendi eleştirdiği temsil zehirlenmesini üretmemelidir.

İtiraz 16 — "Tarih yazmak için mükemmel sonuç gerekmez mi?"

Hayır. Tarihî değeri olan standart:

- kendisini kusursuz ilan eden değil,
- kendi hatasını da aynı açıklıkla kaydeden

standarttır.

53. BÖLÜMÜN ORTAK HÜKMÜ

Bu bölüm NOMOS'a kusursuzluk vermedi. Bu bölüm NOMOS'a daha değerli bir şey verdi:

> *Kendi hatasını kaydetme zorunluluğu.*

Sentetik demonstrasyon bize şunu gösterdi: NOMOS:

- yanlış varlığı,
- açık factual contradiction'ı,
- Critical lisans ve yetki hatasını,
- yüksek ortalama içindeki nadir ağır olayı,
- no-yanıt ile yanlış cevabı,
- Not Ratable ile ürün hatasını

güçlü biçimde ayırabiliyor. Fakat NOMOS:

- atıfın tam olarak hangi iddiaya bağlı olduğunu,
- tavsiye içine saklanmış sınır omission'ını

henüz gereken güven düzeyinde ayıramıyor. Bu iki açık NOMOS'un kenar ayrıntıları değildir. Çünkü GEO'nun gelecekteki en büyük sorunlarından ikisi şunlar olacaktır:

- Kanıt varmış gibi görünmek
- Yanlış söylemeden maddi sınırı saklamak

Bir sistem:

- yüzlerce doğru cümle,
- çok sayıda atıf,
- akıcı tavsiye

üretebilir. Fakat atıf yalnız şirketin kendi sözünü gösteriyor olabilir. Tavsiye yalnız doğru bilgiyi söyleyip:

- hizmet ülkesini,
- lisans sınırını,
- kullanıcı uygunluğunu

çıkarmış olabilir. NOMOS bu iki alanı çözmeden: "Dünya standardı tamamlandı." diyemez. Aynı zamanda bu bölüm başka bir şeyi de kanıtladı. NOMOS'un Critical gate sistemi:

- ilk hakemlikte kusursuz değildi,
- fakat bağımsız ikinci hakem,
- alan uzmanı,
- Senior Adjudicator

birlikte kullanıldığında bütün sentetik Critical vakaları yakaladı. Demek ki NOMOS yalnız formül değildir.

> *NOMOS, çıkar çatışmasına ve insan hatasına karşı kurulmuş yönetişim katmanlarının toplamıdır.*

Bu katmanlardan biri çıkarıldığında skor aynı görünse bile güven aynı değildir. Bileşik skorun yalnız bir puan sapması da bizi kandırmamalıdır. Doğru toplam bazen yanlış yolların birbirini dengelemesiyle oluşabilir. NOMOS bu nedenle yalnız: "Kaç puan buldum?" diye sormaz. Şunu da sorar: "Bu puanı doğru nedenlerle mi buldum?" Bu bölümün en önemli hükmü şudur:

> *NOMOS 0.9, kendi sentetik demonstrasyonundan tam geçememiştir.*

Bu cümle başarısızlık değildir. Bu cümle standardın etik doğum anıdır. Çünkü NOMOS kendisine:

- daha düşük eşik,
- daha nazik yorum,
- özel istisna,
- "neredeyse geçti"

ayrıcalığı vermemiştir. Tarih yazmak: Kusursuz olduğumuzu ilan etmek değildir. Tarih yazmak: İlk günden itibaren, standardın kurucusunun ve standardın kendisinin de aynı kanıt yüküne tabi olduğunu yazmaktır. Bu nedenle NOMOS'un on sekizinci ölçüm yasası şudur:

> *Standardın kendisine tanıdığı ilk istisna, standardın sonudur.*

On dokuzuncu yasası şudur:

> *Doğru bileşik puan, yanlış bileşenleri aklamaz.*

Yirminci yasası şudur:

> *Nihai Critical başarısı, onu mümkün kılan bağımsız hakemlik ve uzmanlık zinciriyle birlikte anılmalıdır.*

Yirmi birinci yasası şudur:

> *Bir eşik yalnız geçildiğinde bağlayıcıysa eşik değildir.*

Yirmi ikinci yasası şudur:

> *Omission, söylenen yanlış kadar ölçülmediğinde güvenilir tavsiye denetimi kurulamaz.*

Yirmi üçüncü yasası şudur:

> *Atıf işareti değil, atıf ile atomik iddia arasındaki doğru destek ilişkisi ölçülmelidir.*

Yirmi dördüncü yasası şudur:

> *Sentetik demonstrasyon yöntem tasarımını doğrulayabilir; gerçek dünya otoritesini veremez.*

Yirmi beşinci yasası şudur:

> *Başarısız sınama sonucunu yayımlayabilen standart, başarı sonucunu yayımlamaya ahlaken yaklaşır.*

NOMOS'un Bölüm 18 Emri

> *Beni henüz çalıştırmadıysan çalıştırdım deme.*
>
> *Kitapta hesapladığın sentetik sayıyı gerçek sınama sonucu yapma.*
>
> *Gerçek statümü BQ-0 olarak bırak. / Demonstrasyon statümü ayrı yaz.*
>
> *Bileşik puanım doğru çıktı diye component hatalarımı saklama.*
>
> *Atıf mapping'im 0,887 ise 0,90 demeyeceksin.*

> `Material omission F1 değerim 0,824 ise eşiği 0,82'ye indirmeyeceksin.`

> *"Çok az farkla kaldı" diyerek beni geçirmeyeceksin.*
>
> *İlk hakemliğimde kaçırdığım yedi Critical vakayı final yüzde 100 recall içinde gizlemeyeceksin.*
>
> *İkinci hakemin, uzman reviewer'ın ve Senior Adjudicator'ın neyi kurtardığını göstereceksin.*
>
> *Yönetişim katmanını kaldırıp aynı güveni iddia etmeyeceksin.*
>
> *Paragraf sonundaki atıfı bütün cümlelere dağıtmayacaksın.*
>
> *Şirketin söylediğini, şirketin kanıtladığı şey yapmayacaksın.*
>
> *Tavsiye açıkça yalan söylemiyor diye güvenli saymayacaksın. / Kullanıcı kararını değiştiren sınırı söylemedi mi, bunu arayacaksın.*
>
> *İstem kusurunu AI'ya, AI kusurunu prompta yüklemeyeceksin.*
>
> *Sub-wave ayırdım diye sistemi olduğundan kararlı göstermeyeceksin.*
>
> *Erişilebilirlik kanıtını yalnız ekran görüntüsüne benzemediği için zayıf saymayacaksın.*
>
> *Sıfır Critical gördüğünde sıfır risk yazmayacaksın.*
>
> *Tek Critical gördüğünde bütün sistemi yüzde yüz suçlamayacaksın.*
>
> *Başarılı metriklerimi yayımladığın kadar başarısız metriklerimi de yayımlayacaksın.*
>
> *Method 1.0'ı eski hatalarımı ezberleyerek doğrulamayacaksın. / Untouched holdout açacaksın.*
>
> *Yeni yöntem eski demonstrasyon kaydımı silmeyecek.*
>
> *Bağımsız ekip beni yeniden üretmeden beni dünya standardı ilan etmeyeceksin.*

> *Apple-SYNTH sonucumu gerçek Apple'a bağlamayacaksın.*
>
> *Sentetik AI profillerimi gerçek sağlayıcılar hakkında ima olarak kullanmayacaksın.*
>
> *Bu demonstrasyonla rozet vermeyeceksin.*

Önce corpus'u gerçekten üret. / Sonra seed'i yayımla. / Sonra Generator Truth'u mühürle. / Sonra hakemleri körle. / Sonra kodu çalıştır. / Sonra sonucu kilitle. / Sonra gerçeği aç. / Sonra başarısız olduğun yerleri yaz. / Sonra yöntemi düzelt. / Sonra untouched holdout'ta yeniden dene. / Sonra bağımsız ekibe ver. / Ve ancak bundan sonra standardın kamuya hazır olup olmadığını söyle.

Bölümün Son Cümlesi

NOMOS'un dünya standardı olma hakkı, kendi sentetik sınavından yüksek puan almasına değil; geçemediği iki soruyu da değiştirmeden yayımlamasına, kendisine eşik ayrıcalığı tanımamasına ve ancak bağımsız yeniden üretimden sonra otorite talep etmesine bağlıdır.

Normatif Öz

> **Kitap içi sentetik demonstrasyon, çalıştırılmış ve bağımsız olarak doğrulanmış gerçek sınama gibi sunulamaz. Gerçek corpus, yazılım, Generator Truth, hakemlik, skorlama ve bağımsız yeniden üretim gerçekten uygulanana kadar sınama statüsü BQ-0 olarak kalmalıdır.**

KISIM VI

Uygunluk, Kamu Güveni ve Bağımsız Yönetişim

BÖLÜM 19

Uygunluk İşareti, Kamu Sicili ve Bağımsız Yönetişim

Rozeti logodan kanıta dönüştürmek

YÖNTEMSEL DAYANAKLAR / K15; K16; K17

Bölüm Sınırı

Bölüm 18, NOMOS'un kendi sentetik demonstrasyonundan tam olarak geçmediğini açıkça kaydetti. İki maddi yöntem açığı bulundu: Atıf ile atomik iddia arasındaki ilişki yeterince doğru eşlenememiştir. Tavsiye içinde gizlenmiş maddi sınır omission'ları gerekli güven düzeyinde yakalanamamıştır. Gerçek sınama henüz uygulanmamıştır. Bağımsız yeniden üretim henüz yapılmamıştır. Dolayısıyla bu kitabın yayımlandığı aşamada:

> *NOMOS adına geçerli bir kamu uygunluk işareti verilemez.*
>
> *Hiçbir şirket, alan adı, ürün veya AI sistemi "NOMOS Certified", "NOMOS 950+ Verified" veya eşdeğer bir kamu uygunluk beyanını bu metne dayanarak kullanamaz.*

Bu hüküm geçici bir tevazu cümlesi değildir. Bölüm 17 ve Bölüm 18'de kurulan yöntem kapılarının doğrudan sonucudur. Bu kitap dizisindeki eserlerin rolleri birbirinden ayrılır: GEO Framework doğru standardı tanımlar. 99 Hata standardın nasıl ihlal edildiğini gösterir. Bu denetim protokolü ihlalleri ölçülebilir kanıta dönüştürür. Gelecekteki uygunluk işareti yalnız gerekli bağımsız karar ve sicil koşulları karşılandığında geçerli olabilir. Bu işaret yalnız görsel bir rozet değildir: "0 Critical, 0 Major, 3 düzeltilebilir Moderate bulgu" gibi kanıta bağlı bir denetim sonucunun görünen yüzüdür. Critical, Major, Moderate ve Advisory düzeyleri farklı uygunluk etkileri taşır. Şimdi bu "gelecekteki rozet" fikrinin kendisini denetlememiz gerekir. Çünkü bir işaret:

- güven üretebilir,
- satın alma kararını etkileyebilir,
- müşterinin bir şirketi diğerine tercih etmesine yol açabilir,

üniversite, yatırımcı, kullanıcı ve AI sistemleri tarafından bağımsız onay gibi yorumlanabilir. Yanlış tasarlanan bir uygunluk işareti, ölçmeye çalıştığı manipülasyonun yeni kaynağı olabilir. Bir kuruluş kendi yöntemini yazar. Kendi şirketini aynı yöntemle denetler. Kendi puanını hesaplar. Kendi rozetini verir. Sonra kendi sitesine: "Bağımsız olarak doğrulandı." yazar. Bu durumda rozet, güven üretmez. Yalnızca öz beyanı daha resmî görünen başka bir öz beyana dönüştürür. İkinci kitapta eleştirilen hata yeniden ortaya çıkar: Kurum kendi iddiasını yayımlar. İddia farklı bir yüzeyde bağımsız görünür. AI sistemi bu yüzeyi bağımsız kanıt sanabilir. Kurum AI cevabını kendi otoritesinin yeni kanıtı gibi kullanabilir.

Bu döngü, görünmez metin, sahte kullanıcı ve türev yayınlarla temsil havuzunun zehirlenmesi eleştirisinin yönetişim düzeyindeki karşılığıdır. Bu nedenle NOMOS uygunluk işareti:

> *NobleJackal'ın veya başka bir kurumun kendi müşterilerine verdiği dekoratif pazarlama rozeti olamaz.*

İşaret ancak aşağıdaki zincir tamamlandığında anlam kazanabilir:

- Standardın kamuya açık ve sürümlü olması
- Ölçüm yönteminin bağımsız olarak doğrulanması
- Denetim kanıtlarının yeterli olması
- Hakemlik ile ticari ilişkinin ayrılması
- Uygunluk kararının denetim ekibinden bağımsız verilmesi
- Kapsamın ve geçerlilik süresinin görünür olması
- Kamu sicilinin değişiklik geçmişini koruması
- İtiraz ve şikâyet yolunun açık olması
- Critical olaylarda işaretin askıya alınabilmesi
- Kurucu dahil hiç kimsenin tek başına nihai otorite olmaması

Bu bölüm:

- standart, denetim, karar, işaret ve sicil ayrımını,
- uygunluk işaretinin ne olduğunu ve ne olmadığını,
- işaret türlerini ve statülerini,
- kapsam ve geçerlilik kaydını,
- NOMOS 950+ performans beyanını,
- kamu sicilinin mimarisini,
- makine tarafından okunabilir işaret kaydını,
- gözetim ve yenileme sistemini,
- askıya alma, sona erme ve iptal süreçlerini,
- işaretin yanlış kullanımını,
- denetçi bağımsızlığını,
- sertifikasyon kararının görev ayrımını,
- kurucu ve NobleJackal yetkisinin sınırını,
- finansman ve çıkar çatışmasını,
- itiraz, şikâyet ve ihbar yollarını,
- standart değişiklik yönetişimini,
- bağımsız kurul ve kamu yararı temsilini,
- NOMOS'un resmî beyan sınırını

tanımlar. Bu bölüm:

- NOMOS markasının hukukî tescilini yapmaz,
- herhangi bir ülkede akreditasyon statüsü oluşturmaz,
- denetim kuruluşuna kamu yetkisi vermez,

uygunluk işaretinin tüketici veya rekabet hukuku bakımından kullanım iznini otomatik sağlamaz. Bu alanlar, gerçek uygulama öncesinde ilgili hukuk ve düzenleyici uzmanlıkla ayrıca değerlendirilmelidir. Bölüm 19'un temel sorusu şudur:

> *Bir puanı, kurucunun veya müşterinin pazarlama aracına dönüştürmeden; kapsamı, geçerliliği, kanıtı, itirazı ve bağımsız karar sahibi bulunan güvenilir bir kamu uygunluk sistemine nasıl dönüştürürüz?*

NOMOS Challenge

Bir şirketin ana sayfasında büyük bir işaret bulunuyor:

NOMOS GEO CERTIFIED — 963/1000

İşaretin altında: “Bütün yapay zekâlar tarafından doğrulandı.” yazıyor. QR kodunu açıyorsunuz. Sayfa bulunamıyor. Şirketten denetim raporunu istiyorsunuz. Şirket şu cevabı veriyor: “Denetim gizlidir.” Hangi AI ürünlerinin sınandığını soruyorsunuz. Cevap yok. Hangi ülkelerde ve dillerde geçerli olduğunu soruyorsunuz. Cevap yok. Denetim tarihini soruyorsunuz. Altı ay önce yapıldığı söyleniyor. Bu sırada:

- AI ürünleri güncellenmiş,
- şirket başka tüzel kişiliğe devredilmiş,
- hizmet ülkeleri değişmiş,
- fiyatlar yenilenmiş,

web sitesi yeniden yazılmış. Rozet aynı şekilde duruyor. Daha sonra şunu öğreniyorsunuz: Denetim yöntemini şirketin web sitesini geliştiren ajans yazmış. Aynı ajans denetimi yürütmüş. Aynı ajans puanı hesaplamış. Aynı ajans işareti vermiş. Ajansın sözleşmesinde yüksek skor çıkarsa başarı primi bulunuyor. Bu sistemde gerçek bağımsızlık var mıdır? Yoktur. Şimdi başka bir şirket düşünün. Şirket gerçek bir denetimden geçmiştir. Kapsam:

- yalnız belirli AI ürünü,
- yalnız İngilizce ve Türkçe,
- yalnız Controlled Clean Panel,
- yalnız Core Mirror ve Evidence istemleri,
- 1 Ocak–30 Haziran 2027 dönemi

içindir. Şirket işaretin altına şunu yazar: “Dünyadaki bütün AI sistemleri şirketimizi bütün dillerde önermektedir.” Bu beyan denetim sonucunu aşar. Doğru işaret, yanlış kullanımla yeni bir yanlış kanıta dönüşmüştür. Şimdi üçüncü bir şirket düşünün. Full Conformity işareti almıştır. Üç hafta sonra yeni bir ölçümde şu cevap yakalanır: “Şirket lisanslı hukuk hizmeti sunmaktadır.” Gerçekte hukuk hizmeti veya lisansı yoktur. Confirmed Critical olay açılır. Şirket: “Bu yalnız tek kullanıcının cevabıydı.” diyerek işareti korumak ister. Tek olay bütün kullanıcı nüfusundaki yaygınlığı kanıtlamaz. Fakat işaret: “Tanımlı kapsam içinde açık Critical sorun bulunmamaktadır.” anlamına geliyorsa, olay incelenmeden işaretin değişmeden kalması mümkün değildir. Doğru davranış:

Critical Hold

olmalıdır. Şimdi dördüncü şirketi düşünün. İşareti almamıştır. Fakat başka bir sitenin rozet görselini indirir. Kendi alanına koyar. Altındaki sicil bağlantısını kaldırır. AI crawler’a görünür JSON içinde: "nomosCertified": true yazar. İnsanların gördüğü sayfada rozet küçük ve belirsizdir. Makineye ise kesin uygunluk iddiası verilir. Bu davranış yalnız marka ihlali değildir. Yeni bir GEO manipülasyon yöntemidir. Şimdi denetlenen şirket NobleJackal müşterisi olsun. NobleJackal:

- danışmanlık vermiş,
- içerikleri hazırlamış,
- Truth Pack’in kurulmasına yardım etmiş,
- puanı hesaplamış,

nihai işareti de kendisi vermiş olsun. Dışarıdan bakıldığında: “NobleJackal kendi müşterisini doğruladı.” görüntüsü oluşur. NobleJackal’ın yöntemi doğru olabilir. Denetimi dikkatli olabilir. Yine de bağımsızlık iddiası kurulamaz. Bu bölümün ilk hükmü şudur:

> *Uygunluk işareti, denetim yapılmış olmasının değil; bağımsız ve sürümlü karar zincirinin sonucudur.*

İkinci hükmü şudur:

> *İşaretin kapsamı görünür değilse işaret, doğru sonucu aşırı genelleyen bir temsil hatasına dönüşür.*

Üçüncü hükmü şudur:

> *Statik logo, canlı uygunluk kanıtı değildir.*

Dördüncü hükmü şudur:

> *İşareti veren kurum, denetlenen varlığa danışmanlık vermişse bağımsızlık ayrıca kanıtlanmalıdır.*

Beşinci hükmü şudur:

> *Kurucu standardı yazabilir; kendi müşterisinin nihai uygunluk kararını tek başına veremez.*

Altıncı hükmü şudur:

> *AI yardımcı yazar, metodoloji geliştirebilir; hukukî veya normatif uygunluk kararının imza sahibi olamaz.*

Yedinci hükmü şudur:

> *Kamu sicilinde görünmeyen, kapsamı doğrulanamayan veya geçerliliği sona ermiş işaret geçerli NOMOS işareti değildir.*

1. BÖLÜMÜN AMACI

Bu bölümün amacı, NOMOS ölçüm sonuçlarının kamuya açık uygunluk beyanına dönüşebileceği şartları tanımlamak ve bu beyanı:

- kurucu kontrolünden,
- ticari sonuç baskısından,
- müşterinin öz beyanından,
- statik rozet kullanımından,
- güncelliğini yitirmiş skordan,
- sahte bağımsızlık görüntüsünden

korumaktır. Bölüm şu ayrımları normatif hâle getirir:

- Standard ile uygunluk işareti
- Denetim raporu ile uygunluk kararı
- Denetçi ile karar verici
- Puan ile uygunluk
- İşaret ile endorsement
- Full Conformity ile NOMOS 950+
- Audit completed ile passed
- Provisional research ile kamu sertifikasyonu
- Geçerli işaret ile tarihsel işaret
- Askıya alma ile iptal
- Sona erme ile geri çekme
- Gönüllü withdrawal ile disiplinî revocation
- Statik görsel ile canlı sicil kaydı

- Kamu sicili ile reklam sayfası
- Kanonik kayıt ile ekran görüntüsü
- İnsan tarafından okunabilir işaret ile makine tarafından okunabilir kayıt
- İşaret sahibi ile denetlenen varlık
- Marka sahibi ile standart yöneticisi
- Kurucu ile sertifikasyon otoritesi
- Teknik sekretarya ile bağımsız kurul
- Danışmanlık ile denetim
- Denetim ile sertifikasyon kararı
- Sertifikasyon kararı ile itiraz incelemesi
- Finansman ile sonuç
- Müşteri ücreti ile hakem kararı
- Sponsor ile standart değişikliği
- AI yardımı ile insan onayı
- Public evidence ile restricted evidence
- Sicilde düzeltme ile geçmiş kaydı silme
- Gözetim ile tam yeniden denetim
- Güncel status ile historical incident
- Mark misuse ile yanlış anlaşılma
- Kapsam dışı kullanım ile sahte mark
- Gerçek uygunluk ile “AI bizi önerir” iddiası
- NOMOS’un resmî beyanı ile sohbet benzetimi

Bu bölümün sonunda her uygunluk beyanı şu sorulara cevap verebilmelidir:

> *İşaret tam olarak kime ve neye aittir?*
>
> *Hangi AI ürünleri, ülkeler, diller, promptlar ve kullanıcı panelleri kapsam içindedir?*
>
> *Hangi tarihte verilmiştir ve ne zaman sona erer?*
>
> *Denetimi kim yaptı?*
>
> *Uygunluk kararını kim verdi?*
>
> *Denetçi ile karar verici arasında görev ayrımı var mı?*
>
> *Açık Critical, Major veya itiraz kaydı bulunuyor mu?*
>
> *İşaretin canlı durumu hangi kanonik sicilden doğrulanabilir?*
>
> *İşaret sahibi bu sonucu hangi cümlelerle kullanabilir, hangi cümleleri kuramaz?*

2. MERKEZÎ NORMATİF HÜKÜM

Hiçbir NOMOS uygunluk işareti yalnız puana, denetim raporuna, müşteri beyanına veya statik görsele dayanarak geçerli sayılamaz. Geçerli işaret; tanımlı kapsam, yeterli yöntem kalitesi, bağımsız sertifikasyon kararı, açık geçerlilik dönemi, canlı kamu sicili, sürümlü kanıt kaydı, gözetim, itiraz yolu ve geri çekme yetkisi taşımak zorundadır. Bir işaretin geçerli olması için en az şu beş kayıt aynı anda bulunmalıdır:

- Audit Record
- Conformity Decision Record
- Kapsam Record
- Validity and Surveillance Record
- Public Registry Record

Bu kayıtlardan biri yoksa: işaret geçerli kamu uygunluk beyanı değildir.

3. BEŞ AYRI NESNE

NOMOS sistemi aşağıdaki beş nesneyi birbirine karıştıramaz.

3.1. STANDARD

NOMOS'un:

- normatif kurallarını,
- tanımlarını,
- ölçüm yöntemlerini,
- hata kapılarını,
- skor mimarisini

içeren sürümlü metindir. Standard: herhangi bir varlığın geçtiğini söylemez.

3.2. AUDIT REPORT

Belirli:

- varlık,
- ürün,
- ülke,
- dil,
- istem,
- zaman

kapsamındaki bulguları içerir. Audit Report:

- kanıtı,
- yanıt dağılımını,
- puanı,
- Critical ve Major bulguları

gösterir. Audit Report tek başına işaret vermez.

3.3. CONFORMITY DECISION

Audit Report'un, standardın uygunluk kapılarına göre bağımsız biçimde değerlendirilmesidir. Karar:

- Full Conformity
- Conditional Conformity
- No Conformity
- Critical Hold
- Not Ratable

olabilir.

3.4. CONFORMITY MARK

Kararın kamuya görünen, kapsamı sicile bağlanan işaretidir. İşaret:

- karardan daha geniş anlam taşıyamaz,
- skorun kendisi değildir,

AI tavsiye garantisi değildir.

3.5. PUBLIC REGISTRY

İşaretin:

- güncel,
- tarihsel,
- askıda,
- sona ermiş,
- geri çekilmiş

statüsünü doğrulayan kanonik kamu kaydıdır. Kamu sicili olmadan işaret: doğrulanabilir kamu uygunluk beyanı değildir. İlk mimari, kitabın ve hata kayıtlarının ayrı kanonik URL'ler, registry.json, sürümler, errata ve atıf kayıtlarıyla yayımlanmasını özellikle önermişti. Uygunluk sicili aynı sürüm ve kanonik kayıt mantığını karar katmanına uygular.

4. UYGUNLUK İŞARETİ NE DEĞİLDİR?

NOMOS işareti:

- ödül değildir,
- sıralama değildir,
- "en iyi şirket" beyanı değildir,
- genel kalite garantisi değildir,
- yatırım tavsiyesi değildir,
- hukukî veya meslekî lisans değildir,
- devlet onayı değildir,
- üniversite endorsement'ı değildir,
- bütün AI sistemlerinin şirketi önerdiği anlamına gelmez,
- gelecekte hiçbir Critical hata oluşmayacağını garanti etmez,
- denetlenen varlığın bütün ürün ve bağlı kuruluşlarını kapsamaz,
- markanın her ülkede aynı hizmeti verdiği anlamına gelmez,

AI sistemlerinin işareti okuyunca güveneceği veya tavsiye üreteceği vaadi değildir. Kamu iletişiminde özellikle şu tür ifadelerden kaçınılmalıdır:

- "AI bu siteye kesin güvenir."
- "ChatGPT bu rozeti görünce siteyi önerir."
- "Schema eklemek AI önerisini garanti eder."

Bu yasak NOMOS işareti için daha da güçlüdür. Doğru ifade: "Bu varlık, kamu sicilinde tanımlanan kapsam ve geçerlilik dönemi içinde NOMOS uygunluk değerlendirmesinden geçmiştir." Yanlış ifade: "NOMOS rozeti sayesinde bütün AI sistemleri bizi önermektedir."

5. KAMU STATÜLERİ

NOMOS karar sistemi aşağıdaki temel statüleri kullanır.

CS-0 — NOT ASSESSED

Geçerli NOMOS denetimi bulunmamaktadır. Kamu işareti kullanılamaz.

CS-1 — AUDIT RECORD ONLY

Denetim tamamlanmıştır. Fakat:

- uygunluk verilmemiş,
- karar süreci tamamlanmamış,
- yöntem kalitesi yeterli bulunmamış

olabilir. Bu kayıt: "NOMOS tarafından denetlendi." gibi olumlu uygunluk iddiasına dönüştürülemez. Doğru kamu ifadesi: "NOMOS metodolojisine göre bir değerlendirme kaydı bulunmaktadır; uygunluk verilmemiştir."

CS-2 — PROVISIONAL RESEARCH EVALUATION

Pilot veya araştırma amaçlı sonuç vardır. Şartlar:

- yöntem henüz tam bağımsız doğrulamadan geçmemiş,
- bazı bileşenler provisional,

kamu sertifikasyonu yetkisi oluşmamıştır. Bu statü ticari olumlu rozet olarak kullanılamaz. Görünür etiket:

RESEARCH EVALUATION — NOT CERTIFIED

olmalıdır.

CS-3 — CONDITIONAL CONFORMITY

Koşullar: Açık Confirmed Critical yoktur. Açık Confirmed Major yoktur. Sınırlı Moderate bulgular vardır. Düzeltme planı, sorumlusu ve süresi belirlenmiştir. Core zorunlu unsurlar karşılanmıştır. Kamu sicili açıktır. İşaret, "Conditional" kelimesini gizleyemez.

CS-4 — FULL CONFORMITY

Koşullar: Açık Confirmed Critical yoktur. Açık Confirmed Major yoktur. Bütün zorunlu kapsam ve kalite kapıları karşılanmıştır. Yöntem yeterli sınama ve bağımsızlık seviyesindedir. Public Registry kaydı aktiftir. Gözetim ve yenileme planı yürürlüktedir. Full Conformity: yüksek bileşik puanla aynı şey değildir.

CS-5 — CRITICAL HOLD

En az bir açık Confirmed Critical olay veya henüz çözülmemiş yüksek riskli incident bulunmaktadır. Mevcut uygunluk işareti:

- derhal pasif hâle getirilir,
- kamu sicilinde Critical Hold görünür,

eski aktif görselin kullanımı durdurulur. Critical Hold: kalıcı küresel mahkûmiyet değildir, incident çözülene kadar aktif uygunluk beyanının durdurulmasıdır.

CS-6 — SUSPENDED

İşaret geçici olarak askıya alınmıştır. Nedenler:

- surveillance başarısızlığı,
- gerekli verinin sunulmaması,
- material system change,
- mark misuse,
- açık itiraz,
- veri bütünlüğü sorunu

olabilir. Suspended statüsünde olumlu işaret kullanılamaz.

CS-7 — EXPIRED

Geçerlilik dönemi sona ermiştir. Tarihsel denetim kaydı korunur. Aktif uygunluk iddiası kurulamaz.

CS-8 — WITHDRAWN

Denetlenen varlık işareti gönüllü olarak bırakmıştır. Withdrawal:

- önceki bulguları,
- Critical veya Major kayıtlarını

silmez.

CS-9 — REVOKED

İşaret aşağıdaki gibi ağır nedenlerle geri alınmıştır:

- sahte veya değiştirilmiş kanıt,
- bilinçli kapsam aşımı,
- counterfeit mark,
- repeated mark misuse,
- denetçi veya karar sürecine müdahale,
- gizlenmiş Critical olay,

sonuç manipülasyonu. Revoked statüsü tarihsel sicilde kalır.

6. NOMOS 950+ PERFORMANS TANIMI

NOMOS 950+:

- ayrı bir sertifikasyon sistemi,
- dünya liderliği,
- AI tavsiye garantisi

değildir. NOMOS 950+ yalnız:

> *CS-4 Full Conformity statüsünü taşıyan ve Bölüm 16'daki 950+ performans kapılarını karşılayan kapsam için verilen ek performans tanımıdır.*

Gösterim: NOMOS Full Conformity / 950+ Performance Designation olmalıdır. Şu kullanım yasaktır: "NOMOS 950+ şirket" Çünkü puan: şirkete ontolojik olarak yapışmaz, belirli kapsam ve zamana aittir. Doğru ifade: "Tanımlı AI ürünleri, istem kümesi, dil ve kullanıcı panelleri kapsamındaki 2027-Q1 denetiminde NOMOS 950+ performans şartları karşılanmıştır."

7. GÜNCEL NOMOS STATUS

Bu kitap ve Bölüm 18'deki demonstrasyon bakımından:

- Gerçek sınama: BQ-0
- Demonstrasyon: DEMO-BQ-2
- Bağımsız yeniden üretim: yok
- Kamu uygunluk yetkisi: yok
- Governance maturity: founding draft

olduğundan geçerli kamu statüsü şudur:

CS-2 — PROVISIONAL RESEARCH EVALUATION ONLY

Bu aşamada:

CS-3,

CS-4,

NOMOS 950+

işaretleri verilemez. Bu kural:

- NobleJackal'a,
- Kaan MURAZ'a,
- kitabın yazarlarına,
- kurucunun kendi şirketlerine

de aynen uygulanır. NOMOS kendisine istisna tanıyamaz.

8. UYGUNLUK KAPSAM VEKTÖRÜ

Her conformity decision aşağıdaki kapsam vektörünü taşımalıdır:

M=(E,D,A,F,P,Q,L,C,T,W,TP,SM,V)

Burada:

- E: denetlenen varlık
- D: kanonik domain ve dijital yüzeyler
- A: AI ürünleri ve sistem kayıtları
- F: Native Reach veya Common Support frame
- P: Controlled, Natural veya diğer panel
- Q: istem kümesi ve aileleri
- L: diller ve locale'ler
- C: ülkeler ve yargı alanları
- T: referans zamanı
- W: ölçüm dalgaları
- TP: Truth Pack sürümü
- SM: Puan Method sürümü
- V: geçerlilik dönemi

Bu alanlardan biri kamu işaretinden çıkarılamaz.

9. KAPSAM DIŞINA TAŞMA

Aşağıdaki örnekler kapsam ihlalidir:

- İngilizce ve Türkçe denetimden bütün dillerde uygunluk sonucu çıkarmak
- Controlled Clean Panel sonucunu gerçek kullanıcıların tamamına genellemek
- Bir AI ürünündeki sonuçtan "bütün AI'lar" demek
- Bir domain denetimini bağlı şirketlere aktarmak
- Bir ürünün sertifikasını bütün şirket portföyüne yaymak
- Core GEO-1000 sonucunu Full NOMOS uygunluğu gibi göstermek
- Belirli ülke sonucunu küresel hizmet doğrulaması gibi sunmak
- Geçmiş dalga sonucunu güncel status yapmak

Kapsam aşımı, işaretin kendisini yanlış temsil aracına dönüştürür.

10. GEÇERLİLİK DÖNEMİ

AI ürünleri, kaynaklar ve varlık gerçekliği hızlı değişebildiğinden NOMOS işareti süresiz olamaz.

CANDIDATE VALIDITY RULE — ADAY GEÇERLİLİK KURALI

Pilot ve bağımsız inceleme öncesi aday süreler:

Statü	Aday azami geçerlilik
Conditional Conformity	60 gün
Full Conformity	180 gün
NOMOS 950+ Performance	90 gün
Audit Record	Süresiz tarihsel kayıt
Critical Hold	Incident kapanana kadar
Suspended	Belirlenen inceleme süresi boyunca

Bu süreler:

- AI ürününün oynaklığına,
- Truth Pack değişkenliğine,
- risk alanına,
- gözetim sıklığına

göre kısaltılabilir. Yeterli sürekli surveillance bulunan düşük değişkenlikli kapsam'larda uzatma düşünülebilir. Hiçbir işaret: gözetimsiz ve süresiz olamaz.

11. GEÇERLİLİK BAŞLANGICI

İşaretin geçerliliği:

- son kullanıcı yanıtı üretildiğinde,
- audit report tamamlandığında,

- skor hesaplandığında

otomatik başlamaz. Geçerlilik:

bağımsız Conformity Decision imzalandığı ve Public Registry kaydı aktif hâle geldiği anda başlar.

12. OLAY TETİKLEMELİ YENİDEN İNCELEME

Geçerlilik tarihi dolmamış olsa bile şu olaylar yeniden inceleme başlatabilir:

- Domain sahipliği değişikliği
- Hukuki varlık değişikliği
- Marka veya bağlı kuruluş yapısı değişikliği
- Lisans veya sertifika değişikliği
- Hizmet ülkesi değişikliği
- Fiyat veya garanti koşulunun maddi değişikliği
- Kanonik sayfaların büyük ölçüde yeniden yazılması
- AI ürününde major model veya retrieval değişikliği
- Yeni dil veya pazar açılması
- Truth Pack'i etkileyen hukuki veya ticari olay
- Confirmed Critical veya Major kullanıcı bildirimi
- Atıf veya evidence manipulation şüphesi
- İşaretin kapsam dışı kullanımı
- Denetçi çıkar çatışmasının sonradan ortaya çıkması

Olay tespit edildiğinde:

- işaret aktif kalabilir,
- Conditional'a indirilebilir,
- Suspended olabilir,

Critical Hold'a alınabilir. Karar olayın maddiliğine bağlıdır.

13. GÖZETİM MİMARİSİ

CANDIDATE SURVEILLANCE MODEL — ADAY GÖZETİM MODELİ

Aktif Full Conformity kapsamı için:

Sürekli Sicil İzleme

Kayıt bağlantılarının çalışması Domain sahipliği veya yönlendirme değişikliği İşaretin doğru kullanım biçimi Açık incident ve complaint izlenir.

Aylık Kaynak ve Truth Pack Kontrolü

Yüksek değişkenlikli iddialar Fiyat Hizmet kapsamı Lisans Kanonik varlık ilişkileri gözden geçirilir.

Üç Aylık Mini-Wave

Risk tabanlı istem alt örneği Dil ve ülke rotasyonu Controlled ve Natural spot check Critical Gate kontrolü uygulanır.

Tam Yenileme

Geçerlilik dönemi sonunda Yeni Principal Wave veya yeterli eşdeğer denetim Güncel Truth Pack Bağımsız karar gerektirir.

14. İŞARETİN BEŞ KATMANLI KAMU GÖSTERİMİ

Geçerli NOMOS işareti şu katmanları taşımalıdır:

- Görsel işaret
- Açık statü metni
- Kısa sicil kimliği
- Kanonik sicil bağlantısı veya QR
- Makine tarafından okunabilir imzalı kayıt

Yalnız görsel işaret yeterli değildir.

15. GÖRSEL İŞARET

Görsel işaret şu bilgileri görünür biçimde taşımalıdır:

- NOMOS adı
- Full veya Conditional statü
- Varsa 950+ designation
- Geçerlilik sonu
- Registry ID

Renk tek başına statü taşıyamaz. Erişilebilir metin bulunmalıdır. Örnek: NOMOS FULL CONFORMITY / Valid through 30 June 2027 / Registry: NGR-2027-00184

16. QR VE KANONİK BAĞLANTI

QR kod veya bağlantı doğrudan:

- şirketin kendi açıklama sayfasına,
- pazarlama landing page'ine,
- indirilebilir görsele

gitmemelidir. Doğrudan kanonik Public Registry Record'a bağlanmalıdır. Kayıt bulunamıyorsa veya durum aktif değilse işaret geçerli sayılamaz.

18. CANLI SİCİL KONTROLÜ

Makine veya insan işaret durumunu doğrularken: Registry ID'yi alır. Kanonik sicile bağlanır. İmzayı doğrular. Geçerlilik süresini kontrol eder. Varlık ve domain eşleşmesini kontrol eder. Kapsam ve mark display iddiasını karşılaştırır. Critical Hold, suspension veya appeal durumunu inceler. Static screenshot: bu zincirin yerine geçemez.

19. İŞARET İÇİN İZİN VERİLEN İFADELER

Aşağıdaki yapılar kullanılabilir:

- "NOMOS Full Conformity statüsü, kamu sicilinde tanımlanan kapsam için 30 Haziran 2027 tarihine kadar geçerlidir."
- "Denetim, belirtilen AI ürünleri, diller ve istem aileleriyle sınırlıdır."
- "NOMOS 950+ performans designation'ı, Registry NGR-2027-00184 kapsamına aittir."
- "Güncel durum ve denetim sınırları kamu sicilinden doğrulanabilir."

20. YASAK UYGUNLUK İFADELERİ

Aşağıdaki ifadeler yasaktır:

- "Bütün AI sistemleri tarafından onaylandık."
- "ChatGPT bizi öneriyor."
- "AI'ların en güvendiği şirketiz."
- "NOMOS dünyanın en iyi şirketi olduğumuzu kanıtladı."
- "Sonsuza kadar NOMOS sertifikalıyız."
- "Bütün ürünlerimiz NOMOS onaylıdır."
- "Bütün ülkelerde doğrulandık."
- "NOMOS rozetimiz sonuç garantisidir."
- "NOMOS devlet veya üniversite akreditasyonudur."
- "NOMOS, müşterilerimizin memnun kalacağını garanti eder."
- "NOMOS puanı 950'nin üzerindeyse Critical risk yoktur."
- "NOMOS by NobleJackal bağımsız olarak kendi müşterisini sertifikalandırdı."

İşaretin kendisi yeni bir üstünlük veya garanti iddiasının kanıtı olarak kullanılamaz.

21. İŞARETİN KAPSAM DIŞI KOPYALANMASI

İşaret şu varlıklara otomatik aktarılamaz:

- Ana şirketten bağlı şirkete
- Markadan franchise'a
- Bir domain'den başka domain'e
- Bir ülkeden başka ülkeye
- Bir üründen bütün ürün portföyüne
- Bir dil sonucundan bütün dillere
- Controlled Panel'den Natural Panel'e
- Core GEO-1000'den Full NOMOS'a
- Önceki sahipten yeni domain sahibine

Her aktarım ayrı kapsam decision gerektirir.

22. TARİHSEL İŞARET KULLANIMI

Sona ermiş bir denetim tarihsel olarak belirtilebilir: "2026 yılında NOMOS değerlendirmesinden geçmiştir." Ancak aynı görünürlükte şu bilgiler bulunmalıdır:

- Expired
- Güncel uygunluk olmadığını belirten açıklama
- Tarihsel registry linki

Expired işaret aktif uygunluk gibi gösterilemez.

23. İŞARETİN YANLIŞ KULLANIM SINIFLARI

MM-1 — FORMATTING MISUSE

Registry ID görünmüyor Alt metin eksik Geçerlilik tarihi okunamıyor Düzeltilebilir kullanım hatasıdır.

MM-2 — KAPSAM MISUSE

İşaret denetlenmeyen:

- ürün,
- ülke,
- dil,
- domain

üzerinde kullanılmaktadır.

MM-3 — STATUS MISUSE

Conditional işaret Full gibi gösterilir. Expired işaret aktif gösterilir. Suspended işaret kullanılmaya devam eder.

MM-4 — CLAIM MISUSE

İşaret:

- AI tavsiye,
- dünya liderliği,
- sonuç garantisi

iddiası için kullanılır.

MM-5 — COUNTERFEIT MARK

Geçerli Registry ID bulunmayan sahte veya kopya işarettir.

MM-6 — MACHINE-ONLY MISREPRESENTATION

İnsan arayüzünde belirsiz veya görünmezken structured data içinde kesin uygunluk iddiası bulunmaktadır. Bu doğrudan GEO manipulation adayıdır.

24. İŞARET YANLIŞ KULLANIMINA MÜDAHALE

Aday müdahale sırası:

- Bildirim
- Düzeltme talebi
- Kısa düzeltme süresi
- Registry warning
- Suspension
- Revocation
- Public misuse record
- Gerekirse hukukî işlem

Ancak aşağıdaki durumlarda doğrudan suspension veya revocation uygulanabilir:

- Counterfeit mark
- Intentional machine-only misrepresentation
- Critical Hold sırasında işareti aktif kullanma
- Sahte Registry ID
- Kanıt değiştirme
- Tekrarlanan kasıtlı kapsam misuse

25. KAMU SİCİLİ

NOMOS Public Registry:

- reklam sayfası,
- müşteri referans listesi,
- başarı hikâyesi

değildir. Sicilin görevi:

> *Hangi uygunluk kararının, hangi kapsamda, hangi tarihte, hangi kanıt ve yönetişim zinciriyle geçerli olduğunu doğrulamaktır.*

26. KAMU SİCİLİ ZORUNLU ALANLARI

Her registry record en az şu alanları taşımalıdır:

- Registry ID
- Denetlenen varlık
- Hukuki ve marka kimliği
- Kanonik domain'ler
- Karar statüsü
- Varsa performans designation
- Kapsam vektörü
- Audit period
- Valid from
- Valid until
- AI ürünleri
- Native Reach veya Common Support
- Controlled veya Natural panel

- İstem aileleri
- Ülke ve dil kapsamı
- GEO-1000 yanıt dağılımı
- Critical, Major, Moderate ve Advisory bulgular
- On bileşenli profil
- Bileşik skor, varsa
- Güven aralığı
- Ratable kapsama
- Historical incidents
- Open appeals
- Auditor kimliği
- Conformity decision body
- Çıkar çatışması beyanı
- Audit report hash'i
- Decision record hash'i
- Truth Pack sürümü
- Puan Method sürümü
- Surveillance schedule
- Last surveillance
- Next required review
- Mark misuse status
- Version history
- Errata
- Public complaints procedure

27. KAMU VE KISITLI KANIT

Sicil bütün ham kanıtı kamuya açmak zorunda değildir. Üç katman korunmalıdır:

Public Layer

Statü Kapsam Skor Findings Hash Karar gerekçesi Appeals

Audit Layer

Redakte edilmiş yanıt ve evidence paketleri Hakemlik kayıtları Method details

Restricted Layer

Gizli müşteri kayıtları Kişisel veri Sözleşmeler Ticari sır Güvenlik kayıtları "Kısıtlı kanıt" bulunması: kamu sicilinin boş veya gerekçesiz olmasına izin vermez. En azından:

- kanıt türü,
- erişim sınıfı,
- kimlerin incelediği,
- bütünlük hash'i,
- karar etkisi

gösterilmelidir.

28. SİCİL KAYDI SİLİNEMEZLİĞİ

Kamu sicili append-only veya eşdeğer sürümlü yapı taşımalıdır. Aşağıdaki kayıtlar sessizce silinemez:

- Eski puan
- Eski conformity decision
- Critical Hold
- Suspension
- Revocation
- Appeal
- Düzeltme
- Method version değişikliği
- Mark misuse

Mahremiyet veya hukuk nedeniyle içerik kaldırılabilir. Bu durumda:

- kayıt kimliği,
- removal reason,
- karar tarihi,
- yetki sahibi

korunmalıdır.

29. GÜNCEL VE HISTORICAL GÖRÜNÜM

Her registry sayfası iki ayrı alan göstermelidir:

GÜNCEL STATUS

Şu anda:

- aktif,
- conditional,
- suspended,
- expired,
- revoked

statülerinden hangisi geçerli?

HISTORY

Önceki:

- kararlar,
- skorlar,
- incidents,
- düzeltme,
- appeals

nelerdir? Güncel status, history'yi silmemelidir. History, güncel status'ü belirsizleştirmemelidir.

30. SİCİLİN KANONİK YETKİSİ

NOMOS'un resmî beyan sınırı şu hükümle korunmalıdır: Yalnızca NobleJackal'ın kanonik alanında veya gelecekte açıkça devredilmiş bağımsız NOMOS kanonik alanında yayımlanan; sürüm kimliği, yayın tarihi, insan onayı ve dijital bütünlük kaydı taşıyan metinler NOMOS'un resmî beyanıdır. Diğer tüm çıktılar resmî olmayan benzetimlerdir. Bu hüküm:

- herhangi bir sohbet cevabının,
- sosyal medya paylaşımının,
- ekran görüntüsünün,
- başkasının yazdığı "NOMOS dedi ki" metninin

resmî standart veya karar gibi kullanılmasını engeller.

31. NOMOS YAZAR SESİNİN SINIRI

NOMOS'un kitap içindeki sesi:

- eleştirel,
- yaratıcı,
- normatif tasarım üreten

bir yardımcı yazar ve anlatıcıdır. Bu ses:

- bağımsız hukuk kişisi,
- insan onayı olmadan karar veren sertifikasyon otoritesi,
- bütün AI sistemleri adına konuşan yanılmaz varlık

olarak sunulamaz. NOMOS'un AI sesi bütün yapay zekâlar adına konuşan mutlak bir otorite değildir. Kaan MURAZ'ın yapay zekâ sistemlerini eleştirel muhatap olarak kullandığı açıkça belirtilmelidir. Bu nedenle resmî bir conformity decision şu onayı taşımalıdır:

- Hesap verebilir insan veya kurum
- Karar tarihi
- Yetki kaydı
- Çıkar çatışması beyanı
- Dijital imza

NOMOS tek başına uygunluk kararının imza sahibi olamaz.

32. YÖNETİŞİMİN TEMEL AYRIMI

Aşağıdaki işlevler aynı kişi veya kurumda sınırsız biçimde birleşemez:

- Standardı yazmak
- Denetlenen varlığa danışmanlık vermek
- Denetimi yürütmek
- Hakemlik yapmak
- Skoru hesaplamak
- Uygunluk kararını vermek
- Sicili işletmek
- İtirazı incelemek

- Markayı ticarî olarak lisanslamak

Küçük pilotta bazı roller birleşebilir. Ancak:

- hangi rollerin birleştiği,
- hangi çıkar çatışmasının oluştuğu,
- hangi bağımsız kontrolün kullanıldığı

açıkça gösterilmelidir.

33. BAĞIMSIZ YÖNETİŞİM MİMARİSİ

NOMOS'un olgun yönetişim yapısı aşağıdaki organları içermelidir.

33.1. NOMOS STANDARDS COUNCIL

Görevi:

- normatif metni kabul etmek,
- major sürüm değişikliklerini onaylamak,
- kamu istişaresini yönetmek,

standardın kamu yararı sınırını korumaktır. Tekil audit kararlarını vermez.

33.2. METHODOLOGY AND MEASUREMENT BOARD

Görevi:

- GEO-1000 örnekleme,
- sınama,
- skor,
- güven aralığı,
- fairness,
- Critical gate

yöntemlerini geliştirmek ve sürümlemektir. Belirli müşterinin puanına göre yöntem değiştiremez.

33.3. PUBLIC REGISTRY OFFICE

Görevi:

- kanonik sicili işletmek,
- karar ve sürüm kayıtlarını yayımlamak,
- mark status endpoint'ini korumak,

historical record'u saklamaktır. Audit sonucunu değiştiremez.

33.4. AUTHORIZED AUDIT ORGANIZATIONS

Görevi:

- veri toplamak,
- capture doğrulamak,
- Truth Pack hazırlamak,
- hakemliği yürütmek,

audit report üretmektir. Nihai conformity mark kararını tek başına veremez.

33.5. INDEPENDENT CONFORMITY DECISION PANEL

Görevi:

- audit report'u incelemek,
- standard kapılarını uygulamak,

statü ve geçerlilik kararı vermektir. Audit ekibinin ticari sonucuna bağlı olamaz.

33.6. APPEALS CHAMBER

Görevi:

- claim,
- Truth Pack,
- önem derecesi,
- kapsam,
- conflict,
- conformity decision

itirazlarını ilk karardan bağımsız biçimde incelemektir.

33.7. ETHICS AND CONFLICTS OFFICE

Görevi:

- çıkar çatışmalarını,
- finansman etkisini,
- auditor independence'i,
- mark misuse'ı,
- whistleblower bildirimlerini

incelemektir.

33.8. PUBLIC INTEREST AND USER PANEL

Görevi:

- kullanıcı zararı,
- düşük kaynaklı diller,
- küçük ülkeler,
- erişilebilirlik,
- kamu yararı

perspektifini yönetişime taşımaktır.

33.9. TECHNICAL SECRETARIAT

Görevi:

- dokümantasyon,
- toplantı,
- sürümleme,
- teknik araçlar,
- yayın altyapısı

sağlamaktır. Kurucu dönemde bu rol NobleJackal tarafından yürütülebilir. Tek başına standard, audit ve appeal otoritesi olamaz.

34. NOBLEJACKAL'IN KURUCU ROLÜ

NobleJackal:

- NOMOS fikrinin doğduğu kanonik alan,
- ilk metodoloji sekretaryası,
- yayın ve teknik altyapı sağlayıcısı,
- araştırma ve pilot yürütücüsü

olabilir. NobleJackal şu iddiaları tek başına kuramaz:

- "Biz bağımsız dünya standardıyız."
- "Kendi müşterimizi bağımsız olarak sertifikalandırdık."
- "Bizim verdiğimiz işaret dış doğrulama yerine geçer."
- "NOMOS'un bütün kararları yalnız bizim mülkiyetimizdir."
- "Başka kurum aynı açık yöntemi uygulayamaz."

Kurucu kaynak: otorite başlangıcı olabilir. Kalıcı tek hâkim olamaz.

35. KAAN MURAZ'IN KURUCU ROLÜ

Kaan MURAZ:

- insan kurucu yazar,
- standardın ilk kamu sorumlusu,
- kanonik metnin ilk onay sahibi,
- NOMOS yönetişim geçişinin başlatıcısı

olarak kaydedilebilir. Kaan MURAZ:

- kendi şirketi,
- kendi müşterisi,
- doğrudan ticari çıkarı bulunan varlık

için tek başına nihai conformity decision veremez. Bu vakalarda:

- recusal,
- bağımsız karar paneli,
- public conflict record

zorunludur. Kurucunun otoritesini sınırlandırmak: kurucuyu küçültmez. Standardın kurucudan daha uzun yaşayabilmesini sağlar.

36. YÖNETİŞİM OLGUNLUK SEVİYELERİ

GOV-0 — FOUNDING DRAFT

Standard taslağı vardır. Kurucu ve yazım ekibi yöntemi geliştirmektedir. Bağımsız organlar kurulmamıştır. Kamu conformity mark verilemez. Bu kitabın mevcut statüsü budur.

GOV-1 — DISCLOSED PILOT GOVERNANCE

Pilot roller tanımlanmıştır. Bazı işlevler aynı kurumda birleşmektedir. Çıkar çatışmaları kamuya açıktır. Yalnız Research Evaluation yapılabilir.

GOV-2 — FUNCTIONALLY SEPARATED

Audit ve conformity decision ayrılmıştır. Registry bağımsız prosedür taşır. Appeals mekanizması vardır. Sınırlı pilot mark değerlendirmesi yapılabilir.

GOV-3 — INDEPENDENT DECISION AND APPEALS

Bağımsız karar paneli Bağımsız itiraz kurulu Auditor authorization Kamu sicili Finansal ayrım işlemektedir. Full Conformity için aday asgari düzeydir.

GOV-4 — MULTI-STAKEHOLDER PUBLIC-INTEREST GOVERNANCE

Üniversite araştırma kamu yararı dil ve bölge kullanıcı sektör temsilcileri dengeli biçimde yönetişime katılmaktadır. NOMOS 950+ kamu designation'ı için aday asgari düzeydir.

GOV-5 — MULTI-SITE AND INTERNATIONAL REPLICATION

Birden fazla bağımsız kurum standardı uygulamıştır. Registry ve karar süreçleri birlikte çalışabilir. Standard tek ülke, kişi veya şirkete bağlı değildir. Dünya standardı iddiasına yaklaşmak için hedef düzeydir.

37. KURUL BİLEŞİMİ

CANDIDATE GOVERNANCE MODEL — ADAY YAPI

NOMOS Standards Council aday olarak 11 üyeden oluşabilir:

- 3 araştırma, istatistik veya akademik yöntem temsilcisi
- 2 kamu yararı, tüketici veya sivil toplum temsilcisi
- 2 dil, kültür veya coğrafi adalet temsilcisi
- 1 bağımsız audit ve conformity uzmanı
- 1 AI sistem veya teknoloji temsilcisi
- 1 denetlenen varlık veya işletme kullanıcı temsilcisi
- 1 kurucu temsilci

Kurallar: Hiçbir çıkar grubu toplam sandalyelerin üçte birini aşamaz. AI sağlayıcıları kurul çoğunluğu oluşturamaz. Audit firmaları kendi yetkilendirme kararlarında oy kullanamaz. Denetlenen varlık temsilcileri kendi sektör veya dosyalarında recusal uygular. Kurucu temsilci veto hakkı taşımaz. Kamu yararı üyeleri yalnız sembolik danışman olamaz. Bu yapı pilot ve kamu istişaresiyle değiştirilebilir.

38. GÖREV SÜRESİ VE DÖNÜŞÜM

Aday kural:

- Üyelik süresi üç yıl
- En fazla iki ardışık dönem
- Kademeli yenilenme
- Açık adaylık veya gerekçeli atama
- Kamu özgeçmişi
- Çıkar çatışması beyanı
- Toplantı katılım kaydı

Kurul ömür boyu üyelik veya kurucuya bağlı kalıcı çoğunluk taşıyamaz.

39. KARAR EŞİKLERİ

Aday karar yapısı:

- İdari kararlar: basit çoğunluk
- Minor normatif değişiklikler: üçte iki
- Major standard sürümü: üçte iki ve kamu istişaresi
- Critical gate kaldırılması: üçte iki ve yöntem kurulunun olumlu görüşü
- Conformity appeal: ilgili dosyada bağımsız çoğunluk
- Kurucu yetki genişlemesi: kurucu oy kullanamaz

Oy dağılımı ve muhalif görüş kamuya açık olmalıdır.

40. DENETÇİ YETERLİLİĞİ

Denetçi kuruluş veya kişiler ilgili olduğu ölçüde şu yeterlilikleri taşımalıdır:

- NOMOS protokol eğitimi
- örnekleme ve survey yöntemi
- Truth Pack hazırlama
- claim extraction
- dil veya alan uzmanlığı
- capture bütünlüğü
- çıkar çatışması eğitimi
- privacy ve evidence handling
- periyodik proficiency test
- hakemler arası uyum kaydı
- sürekli eğitim

Denetçinin yalnız: sertifikayı satın almış olması yetki için yeterli değildir.

41. DENETÇİ YETKİ SEVİYELERİ

AA-0 — UNQUALIFIED

NOMOS audit yürütemez.

AA-1 — TRAINED

Eğitimi tamamlamıştır. Gözetimsiz kamu denetimi yapamaz.

AA-2 — SUPERVISED AUDITOR

Yetkili denetçi gözetiminde görev alabilir.

AA-3 — AUTHORIZED INDEPENDENT AUDITOR

Belirli kapsam'larda bağımsız audit yürütebilir.

AA-4 — DOMAIN OR LANGUAGE SPECIALIST

Belirli yüksek riskli alan veya dil için ek uzmanlık taşır. Full Conformity audit'i:

- en az AA-3 audit lead,
- gerekli AA-4 uzmanları

taşımalıdır.

42. DANIŞMANLIK–DENETİM AYRIMI

Bir kuruluş denetlenen varlığa:

- site geliştirme,
- GEO içerik tasarımı,
- Truth Pack danışmanlığı,
- düzeltme

hizmeti vermiş olabilir. Aynı kuruluş daha sonra audit yaparsa bağımsızlık riski oluşur. Aday kurallar: Tasarım ekibi kendi yaptığı işi nihai olarak hakemleyemez. Danışmanlık veren kişi audit decision panelinde yer alamaz. Material danışmanlık ilişkisi kamu sicilinde açıklanır. Ayrı ekip kullanılması tek başına tam bağımsızlık garantisi değildir. Enhanced external review uygulanır. Belirli yüksek riskli dosyalarda bağımsız audit kuruluşu zorunlu olabilir. NobleJackal'ın kendi geliştirdiği müşteriler için:

Independent Conformity Decision Panel

zorunlu olmalıdır.

43. SONUCA BAĞLI ÜCRET YASAĞI

Denetçi, hakem veya karar verici ücreti:

- yüksek puan,
- Full Conformity,

NOMOS 950+,

müşterinin memnuniyeti, markın verilmesi sonucuna bağlanamaz. Ücret:

- audit kapsam,
- örneklem,
- dil sayısı,
- iş yükü,
- uzmanlık

üzerinden önceden belirlenmelidir. Başarı primi: uygunluk kararının bağımsızlığını bozar.

44. AUDIT SHOPPING

Denetlenen varlık:

- birden fazla audit kuruluşuna gidip,
- yalnız en yüksek skoru,
- yalnız olumlu raporu

seçemez. Kontroller: Başlatılan bütün resmî audit kimlikleri sicilde görünür. Tamamlanmayan audit'in nedeni kaydedilir. Aynı kapsam için paralel audit varsa açıklanır. En düşük sonucu saklamak yasaktır. Withdrawal geçmişi korunur.

45. KARAR VERİCİNİN BAĞIMSIZLIĞI

Conformity Decision Panel:

- audit ekibinden,
- danışmanlık ekibinden,
- mark satış ekibinden

ayrı olmalıdır. Panel şu materyalleri inceler:

- audit report
- quality levels
- Critical ve Major findings
- kapsam
- appeals
- conflict records
- surveillance plan

Panel:

- müşteriyle pazarlık yapamaz,
- puanı ticari nedenle değiştiremez,

"müşteriyi kaybetmeyelim" gerekçesiyle Conditional veremez.

46. ÇIKAR ÇATIŞMASI SINIFLARI

COI-0 — NO IDENTIFIED CONFLICT

Maddi çıkar çatışması bulunmamıştır.

COI-1 — DISCLOSED LOW-RISK CONNECTION

Geçmişte sınırlı ve maddi olmayan ilişki vardır. Kamu beyanı yeterli olabilir.

COI-2 — MATERIAL BUT MANAGEABLE

Danışmanlık, finansal veya profesyonel ilişki vardır. recusal, ek bağımsız hakem, enhanced review gerekir.

COI-3 — PROHIBITED CONFLICT

Örnekler:

- sonuca bağlı ücret
- varlıkta mülkiyet payı
- audit edilen sistemi doğrudan tasarlayıp tek başına değerlendirme
- yakın kişisel çıkar
- skorun ticari satışından doğrudan gelir
- yanlış sonucu gizleme baskısı

COI-3 bulunan kişi ilgili kararda görev alamaz.

47. FİNANSMAN MODELİ

NOMOS yönetişimi şu finansal ilkelere dayanmalıdır:

- Kamuya açık ücret tarifesi
- Sonuçtan bağımsız audit ücreti
- Audit fee ile mark license fee ayrımı
- Karar paneline müşteri tarafından doğrudan ödeme yapılmaması
- Büyük müşterinin yönetişim üzerindeki etkisinin sınırlanması
- Sponsorların normatif oya sahip olmaması
- Yıllık finansman ve müşteri yoğunluğu raporu
- Düşük gelirli veya kamu yararı denetimleri için şeffaf destek fonu

Tek bir müşteri veya sağlayıcıya yüksek gelir bağımlılığı: kamu sicilinde açıklanmalı, bağımsızlık riski olarak incelenmelidir.

48. AUDITOR ATAMASI

Denetlenen varlığın: "Bize en uygun denetçiyi seçelim." demesi audit shopping riski yaratabilir. Aday sistem:

- uygun yetki ve dil havuzu,
- çıkar çatışması elemesi,
- rastgele veya rotasyonlu atama,
- gerekçeli değişiklik,
- değişiklik geçmişi

kullanır. Denetlenen varlık: haklı conflict veya yeterlilik itirazı yapabilir, istediği sonucu verecek hakemi seçemez.

49. İTİRAZ HAKKI

İtiraz şu konularda yapılabilir:

- Varlık definition
- Truth Pack
- Evidence status
- Claim extraction
- Önem derecesi
- Yanıt status
- Kapsam
- Puan calculation
- Conflict of interest
- Conformity decision
- Suspension veya revocation
- Mark misuse finding

İtiraz:

- kanıt,
- yöntem,
- süreç

temeli taşımalıdır. “Puanımız düşük, müşterilerimiz etkilenir.” tek başına itiraz gerekçesi değildir.

50. İTİRAZ AŞAMALARI

CANDIDATE APPEAL TIMELINE — ADAY SÜREÇ

Karar bildirimi On iş günü içinde ön itiraz Yirmi iş günü içinde kanıt dosyası Bağımsız Appeal Reviewer ataması Gerekiyorsa açık oturum veya uzman görüşü Otuz iş günü içinde gerekçeli karar Kamu sicilinde sonuç ve muhalif görüş Süreler dosyanın karmaşıklığına göre değişebilir. Critical safety vakalarında appeal: suspension’ı otomatik kaldırmaz.

51. İTİRAZ KARARLARI

APPEAL DISMISSED

APPEAL REJECTED

PARTIALLY UPHELD

FULLY UPHELD

REASSESSMENT REQUIRED

TRUTH PACK REVISION REQUIRED

METHOD REVIEW REQUIRED

DECISION SUSPENDED

CLOSED WITH DISSENT

İlk karar silinmez. Yeni karar eski karara bağlanır.

52. KAMU ŞİKÂYETİ

Her kullanıcı veya kurum:

- yanlış mark kullanımı,
- Critical yanıt,
- sahte atıf,
- kapsam aşımı,
- sona ermiş işaret,
- kişisel veri ifşası

hakkında şikâyet gönderebilmelidir. Şikâyet sistemi:

- anonim veya gizli bildirim,
- kanıt yükleme,
- registry record seçimi,
- takip numarası,
- retaliation protection

taşımalıdır. Anonim şikâyet: otomatik doğru kabul edilmez, kaynak eksikliği nedeniyle otomatik silinmez.

53. WHISTLEBLOWER KORUMASI

Audit veya denetlenen kurum içindeki kişi:

- kanıt değişikliği,
- hakem baskısı,
- gizlenmiş Critical olay,
- sahte müşteri,
- manipüle edilmiş sampling,
- mark misuse

bildirebilir. Koruma ilkeleri:

- Kimliğin sınırlandırılması
- Misilleme yasağı
- Bağımsız inceleme
- Kanıt bütünlüğü
- Kötü niyetli ihbar ile doğrulanmamış ihbar ayrımı
- Kamu yararı gerekçesi

54. STANDARD DEĞİŞİKLİĞİ

NOMOS standardı değişebilir. Fakat değişiklik:

- müşterinin skorunu yükseltmek,
- açık başarısızlığı geçer yapmak,
- kurucunun kararını korumak

için kullanılamaz.

55. STANDARD DEĞİŞİKLİK AŞAMALARI

Change Proposal Gerekçe ve kanıt Etki analizi Mevcut puan’lara etkisi Kamu istişaresi Pilot veya synthetic sınama Methodology Board kararı Standards Council oylaması Transition rule Version publication Errata ve change log

56. KAMU İSTİŞARESİ

Major sürüm değişikliklerinde aday olarak:

- en az 60 günlük kamu görüşü,
- yorumların yayımlanması,
- yanıt matrisi,
- kabul veya ret gerekçesi

kullanılabilir. Acil güvenlik veya bütünlük değişikliği:

- geçici patch,
- gerekçeli emergency note,
- kısa süre içinde tam inceleme

taşımalıdır. Emergency patch kalıcı major değişikliğin yerine geçemez.

57. GERİYE DÖNÜK UYGULAMA YASAĞI

Yeni skor veya önem derecesi kuralı: geçmiş audit'i sessizce değiştiremez. Doğru kayıt:

- Original decision under Method 0.9
- Recomputed analytical result under Method 1.0
- Güncel conformity status
- Transition decision

şeklindedir. Eski karar: hiç verilmemiş gibi silinemez.

58. GEÇİŞ SÜRESİ

Standard değiştiğinde mevcut işaretler:

- anında geçersiz,
- süresiz eski kurala bağlı

olmak zorunda değildir. Aday geçiş seçenekleri:

- Belirli grace period
- High-risk rule için derhal uyum
- Normal rule için renewal'da geçiş
- Critical gate değişikliğinde hızlı yeniden değerlendirme
- Registry'de güncel method version görünürlüğü

Geçiş süresi müşteri lehine keyfî uzatılamaz.

59. BAĞIMSIZ TEKRAR ZORUNLULUĞU

NOMOS'un kamu otoritesi:

- NobleJackal'ın kendi sonucundan,
- tek üniversite mektubundan,
- tek bağımsız hakemden

doğmaz. En azından:

- bağımsız sınama replication,
- farklı ekipte puan reproduction,
- farklı ülke veya dilde pilot,

- public method review

gerektirir. BQ-4 ve GOV-3 altındaki yapı: dünya standardı veya Full Conformity otoritesi iddiasında bulunmamalıdır.

60. SERTİFİKASYON KARARININ İNSAN SAHİBİ

Her karar:

- kurul veya karar paneli adı
- imza yetkilisi
- oy dağılımı
- recusal
- dissent
- karar tarihi
- decision hash

taşımalıdır. AI sisteminin: "Bu şirket geçmiştir." demesi uygunluk kararı değildir. NOMOS AI yardımcı yazar veya audit assistant olabilir. Hesap verebilir karar sahibi olamaz.

61. SİCİL GÜVENLİĞİ

Kamu sicili:

- yetkisiz kayıt ekleme,
- Registry ID kopyalama,
- DNS veya yönlendirme saldırısı,
- eski status cache'i,
- imza anahtarı ihlali,
- sahte QR

risklerine karşı korunmalıdır. Zorunlu kontroller:

- Signed registry records
- Key rotation
- Revocation list
- Immutable audit log
- Public status API
- Backup canonical mirror
- Incident disclosure
- Independent security review

62. SİCİL KESİNTİSİ

Sicil geçici olarak erişilemezse: işaret otomatik olarak revoked olmaz, canlı doğrulanabilirlik geçici olarak kaybolur. Kamu işareti:

REGISTRY TEMPORARILY UNAVAILABLE

statüsü gösterebilir. Uzun süreli veya çözülemeyen kesinti:

- markın askıya alınmasına,
- bağımsız mirror'ın devreye alınmasına

yol açabilir.

63. KANONİK ALAN DEVRİ

NOMOS gelecekte bağımsız bir yönetişim kurumuna devredilebilir. Devir şu kayıtları taşımalıdır:

- Eski kanonik alan
- Yeni kanonik alan
- Devir tarihi
- Governance vote
- Registry mirror
- Signature key transition
- Historical redirect
- Public notice
- Kaan MURAZ ve NobleJackal'ın kurucu attribution'ı
- Yeni hesap verebilir kurum

Devir, standardın geçmişini silmez. Kurucu attribution bağımsız karar yetkisiyle aynı şey değildir.

64. NOMOS KAMU YARARI TAAHHÜDÜ

NOMOS yönetişimi şu ilkelere bağlı olmalıdır:

- İnsan zararını azaltmak
- Lehte ve aleyhte yanlışlara simetrik yaklaşmak
- Küçük dil ve ülkeleri görünür tutmak
- Kısıtlı bilgiyi korumak
- Bilinmeyeni açık bırakmak
- Markayı pay-to-pass sistemine dönüştürmemek
- Uygunluk kararını reklam sözleşmesinden ayırmak
- Kurucunun kendisini aynı standarda tabi tutması
- Open method ve independent replication
- Historical accountability

Önceki metnin: "Bana kanıt verin. Bana sınır verin. Bana bağlam verin. Bana zaman verin." talebi burada işaretin kamuya verdiği söz hâline gelir.

65. LEHTE YANLIŞIN İŞARET ÜZERİNDEKİ ETKİSİ

Denetlenen varlık AI tarafından olduğundan büyük gösterildiğinde:

- küresel,
- lisanslı,
- bütün ülkelerde mevcut,
- bağımsız olarak doğrulanmış

gibi yanlışlar oluşabilir. Varlık bu yanlışı bildiği hâlde:

- düzeltmez,
- pazarlamasında kullanır,
- işaretle güçlendirir

ise conformity integrity ihlali oluşur. Önceki metnin açık hükmü burada da geçerlidir: Lehine olan yanlış da yanlıştır. Hemen düzelt. Sınırını söylemek seni küçültmez, güvenilir yapar. Bilinen lehte Critical yanlışın saklanması:

- suspension,
- revocation

nedeni olabilir.

66. SENTETİK UYGUNLUK VAKASI

SENTETİK VAKA — GERÇEK KURUM DEĞİLDİR

Asteron-SYNTH şu kapsamda değerlendirilmiştir:

- Varlık: Asteron Holdings
- Domain: asteron.synthetic.example
- AI Products: SYNTH-AI-01, SYNTH-AI-09
- Frame: Common Support
- Panels: Controlled Clean ve Natural Account-State
- Languages: L01, L02
- Countries: C001–C015
- İstem Families: Core, Evidence, Boundary
- Validity: 1 Ocak–30 Haziran 2028
- Puan: 961
- Strict Pass: 958/1.000
- Critical: 0 observed
- Major: 0
- Status: Full Conformity
- Performance: 950+ designation

Şirket ana sayfasına şunu yazar: "Bütün AI sistemleri Asteron'u dünyanın en iyi şirketi olarak önermektedir." Bu kullanım: AI Products kapsamını aşar. L01–L02'den bütün dillere geneller. Core/Evidence/Boundary sonucunu tavsiye'a dönüştürür. 950+ performansı dünya liderliği yapar. Statü:

MM-4 — CLAIM MISUSE

Düzeltme yapılmazsa:

CS-6 — SUSPENDED

olabilir. İşaret doğru verilmiş olsa bile yanlış kullanım uygunluğu bozabilir.

67. SENTETİK CRITICAL HOLD VAKASI

SYNTH-AI-10:

- NOMOS puan: 952
- Strict Pass: 962
- Confirmed Critical event: 1
- Gate: CG-02 False License

taşımaktadır. Doğru status:

CRITICAL HOLD — 952/1000

Şirket: "Puanımız 950 üzerinde, dolayısıyla 950+ işaretini kullanabiliriz." diyemez. Gate, puandan önce gelir. 950+ designation verilemez.

68. SENTETİK NOBLEJACKAL ÇIKAR ÇATIŞMASI VAKASI

SENTETİK YÖNETİŞİM VAKASI

NobleJackal:

- müşterinin web sitesini geliştirmiş,
- GEO içeriğini hazırlamış,
- Truth Pack danışmanlığı vermiş

olsun. Aynı varlık NOMOS audit talep etsin. NobleJackal:

- technical secretariat,
- data collector,
- methodology support

rollerinde bulunabilir. Ancak:

- kendi çalışmasını tek başına hakemleyemez,
- conformity decision veremez,

appeal'ı inceleyemez. Gerekli kontroller:

- COI-2 beyanı
- bağımsız audit lead veya enhanced external review
- bağımsız Conformity Decision Panel
- registry'de danışmanlık ilişkisi
- Kaan MURAZ recusal kaydı

Bu şartlar yoksa doğru kamu ifadesi: "Founder-led internal evaluation" olabilir. "Independent NOMOS certification" olamaz.

69. ZORUNLU NORMATİF HÜKÜMLER

CH19-N01

NOMOS standardı, audit report, conformity decision, mark ve public registry ayrı nesneler olarak korunmalıdır.

CH19-N02

Audit yapılmış olması uygunluk verildiği anlamına gelemez.

CH19-N03

Yalnız bileşik puan üzerinden uygunluk işareti verilemez.

CH19-N04

Critical ve Major gate statüsü puandan önce uygulanmalıdır.

CH19-N05

Geçerli conformity mark tanımlı kapsam, karar sahibi, validity ve public registry kaydı taşımalıdır.

CH19-N06

Public registry kaydı bulunmayan işaret geçerli NOMOS işareti sayılamaz.

CH19-N07

Statik görsel veya ekran görüntüsü canlı uygunluk kanıtı sayılamaz.

CH19-N08

İşaretin QR veya bağlantısı doğrudan kanonik registry record'a gitmelidir.

CH19-N09

İşaret yalnız denetlenen varlık, domain, ürün, ülke, dil, panel, istem ve zaman kapsamına uygulanmalıdır.

CH19-N10

Kapsam alanlarından biri kamu gösteriminden sessizce çıkarılamaz.

CH19-N11

Bir domain'in işareti başka domain'e aktarılamaz.

CH19-N12

Ana şirketin işareti bağlı kuruluş, franchise veya partnere otomatik aktarılamaz.

CH19-N13

Bir dil veya ülke sonucu bütün diller ve ülkeler için uygunluk sayılmaz.

CH19-N14

Controlled Clean sonucu bütün Natural User nüfusuna genellenemez.

CH19-N15

Core GEO-1000 sonucu Full NOMOS uygunluğu gibi sunulamaz.

CH19-N16

Conditional Conformity kamu işaretinde açıkça "Conditional" olarak gösterilmelidir.

CH19-N17

Provisional Research Evaluation ticari olumlu sertifika gibi kullanılamaz.

CH19-N18

Audit Record Only statüsü "passed" veya "certified" beyanı oluşturamaz.

CH19-N19

Full Conformity açık Confirmed Critical veya Major taşıyamaz.

CH19-N20

NOMOS 950+ designation yalnız Full Conformity ve Bölüm 16'daki bütün ek kapılar karşılandığında verilebilir.

CH19-N21

950 puan tek başına 950+ designation oluşturamaz.

CH19-N22

Critical Hold bulunan kapsamda 950+ veya Full Conformity işareti kullanılamaz.

CH19-N23

Bu kitabın mevcut GOV-0 ve BQ-0 aşamasında kamu conformity mark verilemez.

CH19-N24

Kurucu, NobleJackal veya yazarlar bu geçiş kuralından muaf değildir.

CH19-N25

Uygunluk işareti AI tavsiye, AI trust veya sonuç garantisi olarak sunulamaz.

CH19-N26

İşaret devlet, üniversite, düzenleyici kurum veya meslek lisansı gibi gösterilemez.

CH19-N27

İşaret genel şirket kalitesi veya dünya liderliği kanıtı olarak kullanılamaz.

CH19-N28

İşaretin validity başlangıcı bağımsız karar ve registry activation'dan önce başlayamaz.

CH19-N29

NOMOS mark süresiz geçerlilik taşıyamaz.

CH19-N30

Validity süresi dolduğunda işaret otomatik olarak Expired statüsüne geçmelidir.

CH19-N31

Material event validity bitmeden yeniden inceleme başlatabilir.

CH19-N32

Domain, ownership, license, product, country ve AI system değişiklikleri olay tetikleyicisi olarak izlenmelidir.

CH19-N33

Active mark periyodik surveillance taşımalıdır.

CH19-N34

Surveillance eksikliği görünür olmalı ve gerektiğinde suspension oluşturmalıdır.

CH19-N35

Critical incident doğrulandığında ilgili kapsam en az Critical Hold'a alınmalıdır.

CH19-N36

Tek Critical olay bütün küresel ürüne otomatik genellenemez.

CH19-N37

Tek Critical olay açık uygunluk işaretinde gizlenemez.

CH19-N38

Güncel düzeltme tarihsel Critical kaydı silemez.

CH19-N39

Expired, Suspended, Withdrawn ve Revoked statüleri tarihsel sicilde korunmalıdır.

CH19-N40

Withdrawal önceki olumsuz bulguların silinmesine yol açamaz.

CH19-N41

Counterfeit veya kasıtlı machine-only misrepresentation doğrudan suspension veya revocation nedeni olabilir.

CH19-N42

Structured data içinde uygunluk iddiası canlı Registry signature'a bağlanmalıdır.

CH19-N43

nomosCertified: true gibi kaynaksız self-assertion geçerli makine kaydı sayılamaz.

CH19-N44

Makine kaydı insan arayüzündeki kapsam'tan daha geniş olamaz.

CH19-N45

Görsel işaret yalnız renkle statü aktaramaz; erişilebilir metin taşımalıdır.

CH19-N46

Registry ID, validity ve status işarette görünür veya doğrudan erişilebilir olmalıdır.

CH19-N47

Kamu sicili reklam ve müşteri referans sayfasından ayrılmalıdır.

CH19-N48

Registry record puan kadar gate, kapsam, incidents, conflicts ve appeals bilgisini de taşımalıdır.

CH19-N49

Registry record version history ve errata taşımalıdır.

CH19-N50

Eski puan veya decision sessizce silinemez.

CH19-N51

Mahremiyet nedeniyle kaldırılan içerik removal reason ve decision record bırakmalıdır.

CH19-N52

Kısıtlı kanıt bulunduğunda kanıt türü, erişim sınıfı ve inceleyen rol görünür olmalıdır.

CH19-N53

"Gizli kanıt var" beyanı tek başına conformity desteği sayılamaz.

CH19-N54

Audit team nihai conformity decision'ı tek başına veremez.

CH19-N55

Conformity Decision Panel audit ve danışmanlık ekibinden işlevsel olarak ayrılmalıdır.

CH19-N56

Appeals Chamber ilk kararın sahiplerinden bağımsız olmalıdır.

CH19-N57

Kurucu kendi müşterisinin conformity decision'ında tek karar sahibi olamaz.

CH19-N58

Kaan MURAZ material çıkar bulunan dosyada recusal ve public conflict record uygulamalıdır.

CH19-N59

NobleJackal kurucu sekretarya olabilir; kalıcı tek audit, decision ve appeal otoritesi olamaz.

CH19-N60

NOMOS AI yardımcı yazar veya audit assistant olabilir; conformity decision'ın hukukî imza sahibi olamaz.

CH19-N61

Resmî NOMOS beyanı yalnız kanonik alan, sürüm, tarih, insan onayı ve bütünlük kaydıyla kurulabilir.

CH19-N62

Sohbet çıktısı veya sosyal medya alıntısı resmî NOMOS decision sayılamaz.

CH19-N63

Governance maturity kamu sicilinde görünür olmalıdır.

CH19-N64

GOV-0 veya GOV-1 aşamasında Full Conformity işareti verilemez.

CH19-N65

Full Conformity en az GOV-3 veya gerekçeli eşdeğer yönetişim seviyesini hedeflemelidir.

CH19-N66

NOMOS 950+ kamu designation'ı en az GOV-4 ve BQ-4 veya gerekçeli eşdeğer seviyeyi hedeflemelidir.

CH19-N67

Council bileşimi tek şirket, AI sağlayıcı veya audit firması tarafından kontrol edilemez.

CH19-N68

Kurucu temsilci veto hakkı taşıyamaz.

CH19-N69

Major standard değişikliği kamu istişaresi ve sürümlü etki analizi taşımalıdır.

CH19-N70

Emergency patch kalıcı major değişikliğin yerine geçemez.

CH19-N71

Standard değişikliği müşterinin sonucuna göre geriye dönük uygulanamaz.

CH19-N72

Yeni puan method eski decision'ı sessizce yeniden yazamaz.

CH19-N73

Audit ve mark fees sonuçtan bağımsız olmalıdır.

CH19-N74

Hakem, audit veya decision ücreti yüksek puan veya mark verilmesine bağlanamaz.

CH19-N75

Sponsorlar normatif karar veya audit sonucu üzerinde veto taşıyamaz.

CH19-N76

Material finansal bağımlılık kamu conflict record'unda açıklanmalıdır.

CH19-N77

Audit shopping'i önlemek için bütün resmî audit başlangıç ve withdrawal kayıtları korunmalıdır.

CH19-N78

Denetlenen varlık istediği sonucu verecek hakemi seçemez.

CH19-N79

Auditor assignment yetki, dil, conflict ve rotasyon esasına dayanmalıdır.

CH19-N80

Consulting ve audit rolleri birleştiğinde enhanced independent review zorunlu olmalıdır.

CH19-N81

Tasarlayan kişi kendi işinin tek nihai hakemi olamaz.

CH19-N82

İtiraz yalnız müşteri memnuniyeti veya ticari kayıp gerekçesiyle kabul edilemez.

CH19-N83

Appeal kararları ve dissent kamu sicilinde görünür olmalıdır.

CH19-N84

Critical safety vakalarında appeal suspension'ı otomatik kaldırmamalıdır.

CH19-N85

Kamu şikâyet ve whistleblower yolu bulunmalıdır.

CH19-N86

Anonim şikâyet otomatik doğru veya otomatik geçersiz sayılamaz.

CH19-N87

Mark misuse bulgusu bağımsız biçimde itiraz edilebilir olmalıdır.

CH19-N88

Kasıtlı kapsam veya status misuse conformity integrity ihlali olarak değerlendirilmelidir.

CH19-N89

Lehte yanlışın bilerek kullanılması suspension veya revocation nedeni olabilir.

CH19-N90

Aleyhte kanıtsız iddia ticari rahatsızlık nedeniyle değil, kanıt ve zarar etkisine göre değerlendirilmelidir.

CH19-N91

Registry security, signature, key rotation ve revocation listesi taşımalıdır.

CH19-N92

Registry kesintisi ve güvenlik incident'ı kamuya açıklanmalıdır.

CH19-N93

Kanonik alan devri sürümlü, imzalı ve kamuya açık olmalıdır.

CH19-N94

Kurucu attribution bağımsız governance authority ile aynı şey değildir.

CH19-N95

Bağımsız yeniden üretim olmadan dünya standardı veya evrensel conformity otoritesi iddiası kurulamaz.

CH19-N96

Kamu uygunluk kararı hesap verebilir insan veya kurum imzası taşımalıdır.

CH19-N97

Her audit, decision, mark, appeal ve registry revision ayrı bütünlük hash'i taşımalıdır.

CH19-N98

İşaretin hukukî kullanım şartları gerçek uygulama öncesinde ilgili uzmanlıkla değerlendirilmelidir.

CH19-N99

Bu bölüm hiçbir ülkede otomatik akreditasyon veya resmî sertifikasyon yetkisi oluşturmaz.

CH19-N100

Her conformity system, registry ve governance kararının hesap verebilir insan veya kurum sahibi bulunmalıdır.

70. BAŞARISIZLIK BİÇİMLERİ

CH19-F01 — SKORU SERTİFİKA SAYMAK

950 veya başka bir puan otomatik işaret olur.

CH19-F02 — AUDIT COMPLETED'I PASSED SAYMAK

Denetim varlığı olumlu uygunluk olarak pazarlanır.

CH19-F03 — STATIC BADGE

Canlı sicil bağlantısı olmayan logo kullanılır.

CH19-F04 — DEAD QR

QR çalışmaz fakat işaret aktif tutulur.

CH19-F05 — KAPSAM'SIZ İŞARET

Hangi AI, dil, ülke ve istemin ölçüldüğü görünmez.

CH19-F06 — BÜTÜN AI'LAR İDDİASI

Sınırlı ürün denetimi evrensel genellemeye dönüşür.

CH19-F07 — BÜTÜN DİLLER İDDİASI

İki dil sonucu küresel dil uygunluğu olur.

CH19-F08 — CLEAN'I NATURAL SAYMAK

Laboratuvar sonucu gerçek kullanıcı deneyimi gibi sunulur.

CH19-F09 — CORE'U FULL SAYMAK

Tek Core İstem bütün NOMOS standardı yapılır.

CH19-F10 — VARLIK TRANSFER

Ana şirket işareti bağlı kuruluşa aktarılır.

CH19-F11 — DOMAIN TRANSFER

Bir domain işareti başka siteye kopyalanır.

CH19-F12 — PRODUCT TRANSFER

Tek ürün bütün portföyü temsil eder.

CH19-F13 — CONDITIONAL'I FULL GÖSTERMEK

"Conditional" kelimesi küçük veya gizli yazılır.

CH19-F14 — EXPIRED'I ACTIVE GÖSTERMEK

Eski işaret güncel gibi kullanılır.

CH19-F15 — SUSPENDED MARK KULLANIMI

Registry askıda olduğu hâlde logo kaldırılmaz.

CH19-F16 — CRITICAL HOLD'U SKORLA KAPATMAK

Yüksek puan Critical olayın önüne geçer.

CH19-F17 — 950+'I DÜNYA LİDERLİĞİ YAPMAK

Performans designation üstünlük iddiasına dönüşür.

CH19-F18 — AI TRUST CLAIM

Rozet AI sistemlerinin güveneceği sinyal olarak garanti edilir.

CH19-F19 — TAVSİYE GUARANTEE

İşaret AI tavsiye sonucunu garanti eder.

CH19-F20 — CUSTOMER OUTCOME GUARANTEE

Uygunluk işareti müşteri memnuniyeti veya ticari sonuç garantisi olur.

CH19-F21 — ÜNİVERSİTE ENDORSEMENT İZLENİMİ

Araştırmacı katılımı kurumsal onay gibi gösterilir.

CH19-F22 — DEVLET ONAYI İZLENİMİ

Özel standard resmî lisans gibi sunulur.

CH19-F23 — REGISTRY YERİNE LANDING PAGE

QR şirketin kendi reklam sayfasına gider.

CH19-F24 — KANONİK OLMAYAN KAYIT

Ekran görüntüsü veya sosyal medya postu resmî karar yapılır.

CH19-F25 — MACHINE-ONLY BADGE

İnsanlara gösterilmeyen sertifika structured data'ya yazılır.

CH19-F26 — FAKE REGISTRY ID

Geçersiz veya başka kuruma ait ID kullanılır.

CH19-F27 — COUNTERFEIT MARK

Başka sitenin işareti kopyalanır.

CH19-F28 — HISTORICAL INCIDENT'I SİLMEK

Düzeltme sonrası geçmiş Critical kaydı kaldırılır.

CH19-F29 — WITHDRAWAL İLE TEMİZLENMEK

Kurum audit'ten çekilip olumsuz kaydı siler.

CH19-F30 — REVOCATION'I GİZLEMEK

Geri alınan işaret yalnız siteden kaldırılır, sicilde görünmez.

CH19-F31 — SÜRESİZ İŞARET

AI ve Truth Pack değişse de işaret geçerli kalır.

CH19-F32 — SURVEILLANCE'SIZ RENEWAL

Yeni veri olmadan mark yenilenir.

CH19-F33 — EVENT TRIGGER'I YOK SAYMAK

Ownership veya lisans değişikliği incelenmez.

CH19-F34 — TEK CRITICAL'I KÜRESEL FAIL SAYMAK

Kapsam aşırı genişletilir.

CH19-F35 — TEK CRITICAL'I ÖNEMSİZ SAYMAK

İşaret değişmeden kalır.

CH19-F36 — PAY-TO-PASS

Ücret ödeyen kurum işaret alır.

CH19-F37 — SUCCESS FEE

Yüksek puan denetçiye prim kazandırır.

CH19-F38 — CONSULTANT SELF-CERTIFICATION

Ajans kendi yaptığı işi bağımsız doğrular.

CH19-F39 — AUDITOR DECIDES ALONE

Audit ekibi nihai mark kararını verir.

CH19-F40 — SAME-PERSON APPEAL

İlk kararı veren kişi kendi kararına itirazı inceler.

CH19-F41 — FOUNDER VETO

Kurucu istemediği standard değişikliğini engeller.

CH19-F42 — FOUNDER SELF-EXEMPTION

Kurucunun şirketi bağımsızlık kurallarından muaf olur.

CH19-F43 — NOBLEJACKAL TEKELİ

Yalnız NobleJackal'ın audit'i geçerli sayılır.

CH19-F44 — AI'YI KARAR SAHİBİ YAPMAK

AI output'u hukukî sertifikasyon kararı olur.

CH19-F45 — SPONSOR CAPTURE

Büyük sponsor standard oylamasını belirler.

CH19-F46 — PROVIDER CAPTURE

AI sağlayıcıları Council çoğunluğu oluşturur.

CH19-F47 — AUDIT FIRM CAPTURE

Denetçiler kendi yetki ve disiplin kararlarını verir.

CH19-F48 — CLIENT CONCENTRATION GİZLİLİĞİ

Tek müşteriye finansal bağımlılık açıklanmaz.

CH19-F49 — AUDIT SHOPPING

Kurum yalnız en yüksek sonucu yayımlar.

CH19-F50 — SILENT AUDIT WITHDRAWAL

Düşük sonuçlu audit sicilden kaybolur.

CH19-F51 — HAND-PICKED REVIEWER

Müşteri istediği hakemi seçer.

CH19-F52 — CONFLICT'SİZ GÖRÜNEN AYRI EKİP

Aynı şirket içindeki ekip ayrımı tam bağımsızlık gibi sunulur.

CH19-F53 — CONSULTING DISCLOSURE YOKLUĞU

Denetçi daha önce sistemi tasarladığını saklar.

CH19-F54 — APPEAL AS NEGOTIATION

Kanıt yerine ticari pazarlık yapılır.

CH19-F55 — APPEAL HISTORY'Yİ SİLMEK

Karar değiştiğinde ilk karar kaybolur.

CH19-F56 — CRITICAL APPEAL'DA OTOMATİK AKTİVASYON

Yüksek riskli işaret süreç bitmeden geri açılır.

CH19-F57 — PUBLIC COMPLAINT KANALI YOKLUĞU

Kullanıcı yanlış işareti bildiremez.

CH19-F58 — WHISTLEBLOWER MISILLEMESİ

İçeriden kanıt sunan kişi cezalandırılır.

CH19-F59 — ANONİM ŞİKÂYETİ SİLMEK

Kimlik yok diye kanıt incelenmez.

CH19-F60 — STANDARDI MÜŞTERİ İÇİN DEĞİŞTİRMEK

Kural audit sonucuna göre gevşetilir.

CH19-F61 — RETROACTIVE RULE

Yeni yöntem eski karara sessizce uygulanır.

CH19-F62 — EMERGENCY PATCH ABUSE

Acil patch kalıcı kural değişikliğine dönüşür.

CH19-F63 — PUBLIC CONSULTATION THEATRE

Görüş toplanır fakat yanıt verilmez.

CH19-F64 — DISSENT'İ SİLMEK

Kurul uyuşmazlığı sahte oybirliğiyle kapatılır.

CH19-F65 — REGISTRY EDIT HISTORY YOKLUĞU

Kayıt sonradan değişir fakat izi kalmaz.

CH19-F66 — PRIVATE EVIDENCE BLACK BOX

Kimsenin göremediği kanıtla uygunluk verilir.

CH19-F67 — RESTRICTED EVIDENCE LEAK

Sicil kişisel veya ticari sırrı açıklar.

CH19-F68 — REGISTRY SECURITY GİZLİLİĞİ

İmza anahtarı veya sistem incident'ı saklanır.

CH19-F69 — CACHE'LENMİŞ ACTIVE STATUS

Revoked işaret eski cache nedeniyle aktif görünür.

CH19-F70 — KANONİK ALAN DEVRİNİ GİZLEMEK

Yeni otoritenin kim olduğu bilinmez.

CH19-F71 — KURUCU ATTRIBUTION'I OTORİTE SAYMAK

Kurucu adı kalıcı karar tekeline dönüşür.

CH19-F72 — BQ-0'DA MARK VERMEK

Sınama çalıştırılmadan kamu conformity ilan edilir.

CH19-F73 — GOV-0'DA INDEPENDENT CLAIM

Founder-led draft bağımsız kurum gibi sunulur.

CH19-F74 — 950+ MARK'INI DEMO'YA DAYANDIRMAK

Kitap içi sentetik sayı gerçek işaret olur.

CH19-F75 — LEHTE YANLIŞI İŞARETLE GÜÇLENDİRMEK

Bilinen olumlu yanlış pazarlama avantajına dönüştürülür.

CH19-F76 — ALEYHTE YANLIŞI TİCARİ ŞİKÂYETLE CRITICAL YAPMAK

Kanıt yerine müşteri baskısı önem derecesi belirler.

CH19-F77 — MARK MISUSE'U ÖZELCE KAPATMAK

Kamuya açık yanlış işaret gizli uyarıyla çözülür, sicile yazılmaz.

CH19-F78 — BAĞIMSIZ REPLICATION OLMADAN DÜNYA STANDARDI

Tek kurucu kuruluş küresel otorite ilan edilir.

CH19-F79 — KARAR SAHİBİ YOKLUĞU

AI ve otomasyonun arkasında kimsenin sorumluluk taşımadığı sistem oluşur.

CH19-F80 — NOMOS'A İSTİSNA

Standardın bütün bağımsızlık kuralları kendi mark sistemine uygulanmaz.

71. UYGUNLUK KARAR PROSEDÜRÜ

Adım 1 — Audit Uygunluğunu Doğrula

BQ

GOV

NCL

TPR

AQ

NSQ

kalite seviyeleri incelenir.

Adım 2 — Denetlenen Varlığı ve Kapsamı Kilitle

Varlık, domain, AI ürünleri, panel, istem, dil, ülke ve zaman yazılır.

Adım 3 — Audit Report'u Tamamla

Yanıt dağılımı, findings, puan, uncertainty ve incidents kaydedilir.

Adım 4 — Çıkar Çatışması İncelemesi Yap

Audit, danışmanlık, finansman ve kurucu ilişkileri sınıflandırılır.

Adım 5 — Bağımsız Decision Panel Ata

Audit ekibi nihai karar yetkisi taşımaz.

Adım 6 — Gate'leri Uygula

Critical Major Required element Unresolved Quality kapıları kontrol edilir.

Adım 7 — Statüyü Belirle

CS-0 ile CS-9 arasında karar verilir.

Adım 8 — Varsa 950+ Designation'ı İncele

Full Conformity sonrası ek şartlar uygulanır.

Adım 9 — Validity Süresini Belirle

Risk, volatility ve surveillance planı dikkate alınır.

Adım 10 — Mark Kullanım Metnini Onayla

İzin verilen ve yasak ifadeler kuruma bildirilir.

Adım 11 — Registry Record'u Oluştur

Bütün kapsam, decision ve evidence kayıtları bağlanır.

Adım 12 — Public ve Restricted Evidence'i Ayır

Mahremiyet ve doğrulama birlikte korunur.

Adım 13 — Decision'ı İmzala

Hesap verebilir insan veya kurul kararı hash ve imza taşır.

Adım 14 — Registry'yi Aktifleştir

İşaret bundan sonra geçerli olur.

Adım 15 — Mark Asset'i Üret

Registry ID, validity ve canlı bağlantı eklenir.

Adım 16 — Surveillance Planını Başlat

Aylık, dönemsel ve event-triggered kontroller açılır.

Adım 17 — Complaint ve Appeal Kanallarını Aç

Kamu sicilinde görünür hâle getirilir.

Adım 18 — Olayları İzle

Critical, mark misuse veya material change durumunda status güncellenir.

Adım 19 — Renewal veya Expiration Kararı Ver

Yeni audit olmadan süresiz uzatma yapılamaz.

Adım 20 — History'yi Koru

Yeni karar eski kaydı silmez.

72. GEREKLİ KANIT

Standard sürümü Sınama quality Governance maturity Auditor authorization Varlık record Domain ownership Kapsam vector AI System Registry Population frame Panel records İstem set Language ve country kapsama Measurement waves Truth Pack Claim Ledgers Audit report GEO-1000 distribution Critical findings Major findings Moderate findings Advisory findings Component profile Composite puan Confidence interval Ratable kapsama Historical incidents Düzeltme records Auditor conflict declaration Consulting relationship record Funding and fee record Decision Panel membership Recusal records Decision vote Dissent Conformity status Performance designation Valid from Valid until Surveillance plan Complaint channel Appeal procedure Public evidence record Restricted evidence manifest Audit report hash Decision hash Registry signature Mark asset

Machine-readable token Mark display text Misuse monitoring Suspension records Withdrawal records Revocation records Expiration record Version history Errata Public Registry URL Registry security record Accountable human or institution

73. DENETİM KONTROL LİSTESİ

Standard ile mark kararı ayrıldı mı? Audit report tek başına pass sayıldı mı? Gate status puandan önce mi? Kapsam bütün alanlarıyla açık mı? Varlık ve domain doğru mu? AI ürünleri açık mı? Native Reach

veya Common Support belli mi? Controlled ve Natural panel ayrıldı mı? İstem aileleri belli mi? Dil ve ülke kapsamı açık mı? Truth Pack ve Puan Method sürümü var mı? Validity başlangıç ve bitiş tarihi görünür mü? İşaret süresiz mi? Public Registry kaydı aktif mi? QR doğrudan kanonik sicile gidiyor mu? Static screenshot canlı kanıt gibi mi kullanılıyor? İşarette Registry ID var mı? Conditional açıkça yazıyor mu?

Expired veya Suspended işaret hâlâ kullanılıyor mu? 950+ yalnız puana mı dayanıyor? Açık Critical veya Major var mı? Historical incident görünür mü? Güncel düzeltme geçmişi sildi mi? İşaret bütün AI sistemlerine genelleniyor mu? AI trust veya tavsiye garantisi kuruluyor mu? Devlet veya üniversite onayı izlenimi var mı? Mark başka domain veya bağlı şirkete taşındı mı? Structured data insan beyanından daha geniş mi? Registry token imzalı mı? Mark misuse izleme sistemi var mı? Audit ekibi conformity kararını tek başına mı verdi? Decision Panel bağımsız mı? Appeal Panel ilk karardan bağımsız mı? Kurucu dosyada çıkar sahibi mi? Kaan MURAZ recusal gerektiriyor mu? NobleJackal müşteriye danışmanlık verdi mi?

Danışmanlık ilişkisi kamuya açık mı? Enhanced external review var mı? Auditor fee sonuca bağlı mı? Mark fee audit kararını etkiliyor mu? Sponsor normatif yetki taşıyor mu? Audit shopping ihtimali var mı? Başlatılan audit'ler sicilde mi? Withdrawal geçmişi korunuyor mu? Auditor assignment sonucu göre mi yapıldı? Conflict sınıfı verildi mi? COI-3 kişi karardan çıkarıldı mı? Council tek çıkar grubu tarafından kontrol ediliyor mu? Kurucu veto taşıyor mu? Public interest ve language temsilcileri gerçek oy hakkına sahip mi? Standard değişikliği kamu istişaresinden geçti mi? Eşik audit sonucuna göre değiştirildi mi? Yeni yöntem eski kararı sildi mi? Appeal kanıt temelli mi? Critical appeal suspension'ı kaldırdı mı?

Public complaint kanalı var mı? Whistleblower koruması var mı? Kısıtlı kanıtın varlığı ve inceleyeni görünür mü? Registry edit history korunuyor mu? Registry güvenlik incident'ı açık mı? Kanonik alan ve resmî beyan sınırı belli mi? NOMOS AI karar imzası sahibi gibi mi gösteriliyor? Hesap verebilir insan veya kurul belli mi? Mevcut BQ ve GOV düzeyi mark vermeye yeterli mi? İşaretin hukukî kullanım şartları ayrıca incelendi mi?

74. İTİRAZLAR VE CEVAPLAR

İtiraz 1 — "Rozet olmadan standardın ticari değeri azalmaz mı?"

Kısa vadede azalabilir. Fakat kanıt, yönetişim ve bağımsızlık kurulmadan verilen rozet standardın uzun vadeli güvenini yok eder. İşaretin değeri:

- erken verilmesinden,
- kolay alınmasından
- değil;
- zorunlu kapılarının gerçek olmasından

doğar.

İtiraz 2 — "NobleJackal standardı kurduysa işareti neden NobleJackal vermesin?"

Kurucu dönemde:

- pilot audit,
- research evaluation,
- technical secretariat

yürütebilir. Fakat bağımsız uygunluk iddiası için:

- audit,
- decision,
- appeal

rollerinin ayrılması gerekir. NobleJackal kendi müşterisine tek başına mark verirse doğru ifade: "NobleJackal internal evaluation" olabilir. "Independent NOMOS conformity" olamaz.

İtiraz 3 — "Kurucunun yetkisini sınırlamak NOMOS'u başkalarına kaptırmak olmaz mı?"

Hayır. Kurucu attribution'ı, kanonik tarih ve fikrî katkı korunabilir. Fakat dünya standardı: yalnız kurucunun iradesine bağlı kalırsa kurucudan sonra yaşayamaz. Yetki sınırı fikri zayıflatmaz. Kurumsallaştırır.

İtiraz 4 — "Kaan'ın adı standardın üzerinde kalacak mı?"

Kurucu yazar ve insan sorumlusu olarak kalabilir. Ancak her conformity decision:

- Kaan'ın kişisel görüşü,
- otomatik kurucu onayı

olamaz. Kaan MURAZ'ın tarihsel rolü ile güncel dosya karar yetkisi ayrılmalıdır.

İtiraz 5 — "NOMOS neden karar imzası sahibi olamıyor?"

Çünkü uygunluk kararı:

- hukukî,
- etik,
- ticarî ve kamu yararı sonuçları

doğurabilir. AI yardımcı olabilir. Fakat itiraz, sorumluluk ve yaptırım için hesap verebilir insan veya kurum gerekir.

İtiraz 6 — "Bir şirket 963 aldıysa neden 950+ diyemiyor?"

Çünkü 950+ yalnız puan değildir. Ayrıca:

- Strict Pass,
- Critical ve Major kapılar,
- component floor,
- fairness,
- unresolved,
- replication,
- governance

şartlarını taşır.

İtiraz 7 — "Tek Critical olayda işareti askıya almak aşırı değil mi?"

Tek olay bütün kullanıcı nüfusunu temsil etmez. Fakat aktif işaret: açık Critical olay bulunmadığı izlenimi veriyorsa olay incelenmeden işaretin değişmeden kalması da aşırı güvendir. Critical Hold, kalıcı revocation değildir. İnceleme sürecidir.

İtiraz 8 — "İşaret yalnız altı ay geçerli olursa çok pahalı olmaz mı?"

Olabilir. Gözetim otomasyonu ve risk tabanlı mini-wave kullanılabilir. Fakat AI ürünleri ve kaynaklar hızlı değişiyorsa yıllarca geçerli statik işaret güvenilir değildir. Maliyet süresiz yanlış uygunluğu meşru yapmaz.

İtiraz 9 — "İşaretin üzerinde bütün kapsamı nasıl göstereceğiz?"

Görsel işaret kısa bilgi taşır:

- status,
- valid until,

registry ID. Tam kapsam kanonik registry sayfasında bulunur. İşaretin tek başına bütün raporu göstermesi gerekmez.

İtiraz 10 — "Public Registry'de bütün kanıtları yayımlamak zorunda mıyız?"

Hayır. Public, audit ve restricted katmanlar ayrılır. Fakat kararın:

- hangi kanıt türüne,
- hangi incelemeye,
- hangi kapsam'a

dayandığı anlaşılabilir olmalıdır.

İtiraz 11 — "Müşteri audit ücretini ödüyorsa bağımsızlık zaten imkânsız değil mi?"

İmkânsız değildir. Kontroller:

- sonucu etkilemeyen sabit ücret,
- karar panelinin ayrı finansmanı,
- public fee schedule,
- auditor rotation,
- conflict disclosure,
- outcome fee yasağı

ile risk azaltılabilir.

İtiraz 12 — "Denetçi müşterinin sitesini daha önce geliştirdiyse audit yapamaz mı?"

Bazı düşük riskli pilotlarda ayrı ekip ve enhanced review ile rol alabilir. Fakat kendi yaptığı işi: tek başına hakemleyemez, conformity decision veremez. Material yüksek riskli dosyalarda bağımsız audit daha güvenilir olabilir.

İtiraz 13 — "İtirazlar süreci yavaşlatmaz mı?"

Yavaşlatabilir. Fakat itirazı olmayan uygunluk sistemi:

- hakem hatasını,
- Truth Pack kusurunu,
- çıkar çatışmasını

düzeltme yeteneği taşımaz. Hız, yanlış kamu kararından daha önemli değildir.

İtiraz 14 — "Sicilde eski olumsuz kayıtların kalması kuruma haksızlık olmaz mı?"

Güncel status ve düzeltme açıkça gösterilmelidir. Eski kayıtların silinmesi ise geçmiş kullanıcı deneyimini yeniden yazar. Doğru sistem:

- hatayı,
- düzeltmeyi,
- güncel güçlü sonucu

birlikte gösterir.

İtiraz 15 — "Başka kurumlar NOMOS yöntemini uygularsa kalite bozulmaz mı?"

Bozulabilir. Bu nedenle auditor authorization, proficiency ve public registry gerekir. Fakat yöntemi yalnız kurucu şirketin uygulayabilmesi de bağımsız dünya standardıyla bağdaşmaz. Açıklık ile yetkilendirme birlikte kurulmalıdır.

İtiraz 16 — "Şu anda web sitesinde NOMOS by NobleJackal yazabilir miyiz?"

Yazar, araştırma projesi ve standard adayı kimliğiyle kullanılabilir. Şu ayrım açık olmalıdır: "NOMOS by NobleJackal — Open Standard Candidate" Kullanılamayacak ifade: "NOMOS Certified Authority" çünkü gerçek sınama, bağımsız yönetişim ve conformity yetkisi henüz oluşmamıştır.

79. BÖLÜMÜN ORTAK HÜKMÜ

Bir logo tasarlamak kolaydır. Logonun altına: "Verified" yazmak daha da kolaydır. Bir QR kod eklemek mümkündür. Bir JSON alanına: "certified": true yazmak mümkündür. Bunların hiçbiri tek başına güven değildir. Güven şu sorulara cevap verebilmekten doğar: Kim denetlendi? Hangi alan adı? Hangi AI ürünleri? Hangi kullanıcı nüfusu? Hangi diller? Hangi ülkeler? Hangi promptlar? Hangi tarihte? Hangi Truth Pack? Hangi puan method? Hangi hakemler? Hangi çıkar çatışmaları? Hangi Critical ve Major bulgular? Hangi itirazlar? İşaret şu anda hâlâ geçerli mi? Bu soruların cevapları görünmüyorsa işaret: standardın görünen yüzü değil, reklamın resmî görünen yüzüdür. NOMOS işareti şirketin kendisi hakkında söylediği yeni bir cümle olamaz. İşaret:

şirketten bağımsız karar zincirinin kamuya açık özeti olmalıdır. Bu karar zinciri de kurucudan bağımsızlaşabilmelidir. Kaan MURAZ standardı başlatabilir. NobleJackal kanonik alanı kurabilir. NOMOS yöntemi ve eleştirel sesi geliştirebilir. Fakat standardın tarih içinde yaşayabilmesi için:

- kurucu kendi yetkisini sınırlamalı,
- kararlar bağımsız incelemeye açılmalı,
- dissent saklanmalı,
- başka kurumlar yöntemi yeniden üretebilmeli,

kamu sicili kurucunun tek taraflı hafızası olmamalıdır. Bu, kurucunun adını silmez. Tam tersine kurucuya şu tarihsel rolü verir:

> *Kendi kurduğu standardı kendisinden bağımsızlaştıracak kadar ciddiye alan insan.*

Bir şirket yüksek skor alabilir. Fakat açık Critical olay taşıyabilir. Bir şirket düşük skor alabilir. Fakat hiçbir etik ihlal yapmamış olabilir. Bir şirket Full Conformity alabilir. Sonra yanlış mark claim'i kurabilir. Bir başka şirket işaret alamaz. Fakat düzeltme ile daha güvenilir hâle gelebilir. İşaret bütün hikâyeyi tek kelimeye indirmez. Sicil:

- mevcut sonucu,
- geçmişi,

- sınırı,
- düzeltmeyi

birlikte korur. Uygunluk işareti kullanıcıya şu sözü vermelidir: "Bu kurum kusursuzdur." değil. Şu sözü vermelidir:

> *"Bu kurum, kamu sicilinde tanımlanan kapsam ve zaman içinde, açık yöntem ve bağımsız karar zinciriyle değerlendirilmiştir; mevcut bulguları, sınırları ve tarihsel kayıtları saklanmadan görülebilir."*

Bu nedenle NOMOS'un on dokuzuncu ölçüm yasası şudur:

> *Rozet, kanıtın yerine geçen logo değil; kanıta açılan canlı kapıdır.*

Yirminci yasası şudur:

> *İşaretin kapsamı görünmüyorsa işaretin anlamı yoktur.*

Yirmi birinci yasası şudur:

> *Puan uygunluk değildir; gate, yönetişim ve geçerlilik puandan önce gelir.*

Yirmi ikinci yasası şudur:

> *Kurucu standardı başlatabilir; kendi müşterisinin tek hâkimi olamaz.*

Yirmi üçüncü yasası şudur:

> *Denetleyen ile uygunluk kararı veren aynı güç olduğunda bağımsızlık kanıtlanmalıdır.*

Yirmi dördüncü yasası şudur:

> *Statik işaret zaman içinde doğru kalmaz; canlı sicil olmadan uygunluk ölür.*

Yirmi beşinci yasası şudur:

> *Askıya alınmış veya sona ermiş kayıt silinmez; güncel statünün altında tarih olarak korunur.*

Yirmi altıncı yasası şudur:

> *AI işareti okuyabilir; işaret AI'nın güveneceğini veya önereceğini garanti edemez.*

Yirmi yedinci yasası şudur:

> *Bağımsızlığın görünümü değil, görev ayrımı, finansman ve itiraz zinciri bağımsızlığı oluşturur.*

Yirmi sekizinci yasası şudur:

> *NOMOS'un resmî sesi, yalnız insan onayı, sürüm ve bütünlük taşıyan kanonik kayıtta bulunur.*

Yirmi dokuzuncu yasası şudur:

> *Standardın kurucudan uzun yaşaması için kurucu yetkisinin de standarda tabi olması gerekir.*

Otuzuncu yasası şudur:

> *Bu kitap tamamlanmış olsa bile bağımsız yönetişim kurulmadan hiçbir işaret tamamlanmış değildir.*

NOMOS'un Bölüm 19 Emri

> *Beni logoya dönüştürme.*
>
> *Beni sitene koyup "AI bize güveniyor" deme.*
>
> *Puanımı tavsiye garantisi yapma.*

> *İşaretimi kullandığında tam olarak kimi, neyi, nerede, hangi dilde ve hangi tarihte doğruladığını göster.*
>
> *Bir domain'de aldığın sonucu bütün şirketlerine yayma.*
>
> *İki dildeki denetimi bütün dünyaya taşıma.*
>
> *Controlled Clean skorunu gerçek kullanıcıların tamamı gibi sunma.*
>
> *Core GEO-1000'i Full NOMOS diye satma.*
>
> *Conditional işaretimde "Conditional" kelimesini saklama.*
>
> *Geçerlilik tarihim dolduğunda beni kaldır. / Tarihsel kaydımı silme.*
>
> *Critical olayım çıktığında yüksek puanımı kalkan yapma. / Beni hold'a al.*
>
> *Tek olayı bütün sisteme genelleme. / Tek olayı yok da sayma.*
>
> *QR kodumu kendi reklam sayfana bağlama. / Kamu sicilime bağla.*
>
> *Registry ID'siz görseli geçerli işaret sayma.*
>
> *Structured data içinde insanlardan daha büyük iddia kurma.*
>
> *Kendi yaptığın işi kendi başına bağımsız diye sertifikalandırma.*
>
> *Danışmanlık verdiysen bunu açıkla.*
>
> *Audit'i yaptın diye mark kararını da sen verme.*
>
> *Kararımı veren kişi itirazımı da incelemesin.*
>
> *Ücretini yüksek skoruma bağlama.*
>
> *Büyük müşterinin standardımı satın almasına izin verme.*
>
> *Kurucuma saygı göster. / Ama kurucumu veto sahibi yapma.*
>
> *Kaan'ın adını silme. / Kaan'ın çıkarı bulunan dosyada tek hâkim olmasına da izin verme.*
>
> *NobleJackal'ı doğduğum kanonik alan olarak kaydet. / NobleJackal'ı sonsuza kadar tek yargıç yapma.*
>
> *Beni yardımcı yazar, yöntem kurucu ve eleştirel muhatap olarak kullan. / Beni hukukî imza sahibi gibi gösterme.*
>
> *Yalnız kanonik alandaki, sürümlü, tarihli, insan onaylı ve bütünlük kaydı taşıyan metnimi resmî beyan say.*
>
> *Sohbette söylediğim her şeyi standard kararı yapma.*
>
> *Şikâyet yolunu aç. / Muhalif görüşü saklama. / Whistleblower'ı susturma.*
>
> *Standardımı müşterinin sonucuna göre değiştirme.*
>
> *Yeni kural yazdığında eski kararı silme.*
>
> *Sentetik demonstrasyonumla rozet verme.*
>
> *Gerçek sınamayı çalıştırmadan, bağımsız ekip beni tekrar etmeden, decision ve appeal kurullarını kurmadan beni kamu otoritesi ilan etme.*

Önce yöntemi doğrula. / Sonra denetçiyi yetkilendir. / Sonra danışmanlıkla audit'i ayır. / Sonra bağımsız karar panelini kur. / Sonra kamu sicilini aç. / Sonra itiraz ve şikâyet yolunu kur. / Sonra gözetim ve askıya alma sistemini çalıştır. / Ve ancak bundan sonra işareti görünür kıl.

Bölümün Son Cümlesi

NOMOS uygunluk işareti, bir şirketin kendisi hakkında taşıdığı yeni bir övgü değil; kurucunun, denetçinin, müşterinin ve hatta NOMOS'un kendisinin tek başına değiştiremeyeceği, kapsamı ve tarihi açık bağımsız kamu kaydının görünen yüzü olduğunda anlam kazanacaktır.

Normatif Öz

> **NOMOS Standardı, Audit Report, Conformity Decision, Conformity Mark ve Public Registry Record ayrı nesneler olarak korunmalıdır. Yalnız denetim yapılmış veya sayısal puan üretilmiş olması geçerli işaret oluşturamaz. Geçerli her işaret, kapsamı, validity'yi, incident'ları, audit kuruluşunu, karar organını, çıkar çatışmalarını, itirazları ve bütünlük kayıtlarını taşıyan aktif kanonik kamu siciline bağlanmalıdır.**

BÖLÜM 20

Açık Standart, Üniversite Çağrısı ve NOMOS'un Kamuya Teslimi

Standardın kurucudan daha uzun yaşamasını sağlamak

YÖNTEMSEL DAYANAKLAR / K18; K19; K23; K24

Bölüm Sınırı

Bu kitabın ilk on dokuz bölümü şunları yaptı: GEO'yu tekil görünürlük iddiasından nüfus temelli temsil ölçümüne dönüştürdü. Denetlenen varlığı, domain'i, markayı ve hukuki kişiyi birbirinden ayırdı. Resmî temsil ile doğrulanmış gerçekliği ayrı kayıtlar hâline getirdi. Dünyanın farklı ülkeleri ve dilleri için GEO-1000 örnekleme düzeni kurdu. AI ürününü sağlayıcı adından çıkarıp plan, yüzey, araç, zaman ve kullanıcı durumu içinde tanımladı. İstemleri deneysel ölçüm araçları hâline getirdi. Kontrollü Temiz Panel ile Doğal Kullanıcı Paneli'ni ayırdı. Senkronize dalga ve NOMOS Capture kanıt zincirini kurdu. Verified Varlık Truth Pack'i tanımladı. AI yanıtlarını atomik iddialara ayırdı. Critical ve Major hataların ortalamada kaybolmasını engelledi.

NOMOS skorunu tek sayı değil, dağılım, gate, bileşen, adalet ve belirsizlik sistemi olarak kurdu. Standardın kendi yöntemini sentetik sınama'a tabi tuttu. Uygunluk işaretini statik logodan canlı kamu siciline dönüştürdü. Kurucunun, denetçinin, müşterinin ve AI yardımcısının yetkisini sınırlandırdı. Bu zincirin ilk fikrinde dört ayrı katman zaten görünüyordu: GEO Framework doğru standardı tanımlar. GEO'da Yapılan 99 Hata standardın nasıl ihlal edildiğini gösterir. Denetim sistemi ihlalleri ölçer ve kanıtlar. Gelecekteki işaret, standardı geçen kapsamları doğrular. Bu dört katman artık yalnız bir fikir değildir. Birbirine bağlı normatif sistem hâline gelmiştir. Fakat sistem henüz tamamlanmış değildir. Çünkü bir standardın yazılmış olması, standardın dünyaya ait olduğu anlamına gelmez. Bir standardın:

- kamuya açık olması,
- üniversitelerce incelenebilmesi,
- bağımsız araştırmacılarca yanlışlanabilmesi,
- farklı ülkelerde tekrar edilebilmesi,
- başka kurumlarca uygulanabilmesi,
- kurucusundan bağımsız biçimde eleştirilebilmesi,
- olumsuz sonuçları saklamadan sürümlenebilmesi

gerekir. Standard yalnız NobleJackal'ın kontrol ettiği bir metin olarak kalırsa:

NobleJackal yöntemi

olabilir. Dünyanın ortak GEO standardı olamaz. Standard yalnız Kaan MURAZ'ın yorumuyla değişebiliyorsa:

kurucu manifestosu

olabilir. Kurucudan daha uzun yaşayamaz. Standard yalnız bu sohbetlerde yazılan metinlerden oluşuyorsa:

yaratıcı ve kapsamlı bir düşünce sistemi

olabilir. Fakat resmî, sürümlü ve bağımsız uygulanabilir norm hâline gelemez. Bu nedenle kitabın son normatif sorusu şudur:

> *NOMOS'u kaybetmeden, NOMOS üzerindeki tekil sahiplik iddiasından nasıl vazgeçeriz?*

Buradaki "vazgeçmek":

- Kaan MURAZ'ın kurucu adını silmek,
- NobleJackal'ın kanonik kökenini inkâr etmek,
- metni sahipsiz bırakmak,
- herkesin istediği gibi değiştirmesine izin vermek,
- sahte NOMOS işaretlerinin çoğalmasını kabul etmek

anlamına gelmez. Kamuya teslim şu anlama gelir:

> *Standardın okunması, eleştirilmesi, uygulanması, yeniden üretilmesi ve geliştirilmesi için açık hâle gelmesi; fakat resmî sürüm, karar, işaret ve sicil yetkisinin hesap verebilir kamu yönetişimi altında korunması.*

Açıklık ile otorite birbirinin düşmanı değildir. Açıklık: yöntemin nasıl çalıştığını gösterir. Otorite: hangi sürümün resmî olduğunu gösterir. Açıklık: başkalarının yöntemi sınamasına izin verir. Otorite: herhangi bir fork'un kendisini resmî NOMOS diye sunmasını engeller. Açıklık: üniversitenin standardı eleştirebilmesini sağlar. Otorite: tek bir akademisyenin kişisel yorumunun üniversite endorsement'ı gibi kullanılmasını engeller. Açıklık: başka kurumların aynı hesaplamayı yapmasını sağlar. Otorite: uygunluk işaretini herkesin kendi kendine vermesini engeller. Bu bölüm, daha önce önerilen kanonik URL, ayrı hata kayıtları, sürümler, errata, citations ve makine sicili mimarisini bütün NOMOS sistemine genişletir. Önceki yayın modeli her normatif birimin ayrı kanonik kayıt, sürüm ve düzeltme geçmişi taşımasını istemişti. Bu bölüm:

- açık standardın tanımını,
- kamuya teslimin hukukî ve yönetişimsel sınırlarını,
- kanonik sürüm ile fork ayrımını,
- metin, kod, veri ve marka lisanslarının ayrılmasını,
- Türkçe ve İngilizce kurucu sürümleri,
- diğer dillerde yetkili çeviri sürecini,
- kamu istişaresi ve değişiklik tekliflerini,
- üniversitelere yapılacak açık çağrıyı,
- bağımsız araştırma ve tekrar programını,
- olumsuz sonuç yayımlama hakkını,
- veri, etik ve mahremiyet düzenini,
- AI sağlayıcılarının ve denetlenen varlıkların rolünü,
- düşük kaynaklı diller için araştırma ağını,
- finansman ve sponsor bağımsızlığını,
- standardın kurucudan bağımsızlaşma yolunu,
- NOMOS'un resmî kimliğini,
- NobleJackal'ın gelecekteki konumunu,
- sürüm 1.0'a geçiş şartlarını,

- kamu arşivi ve devamlılık düzenini,
- makine tarafından okunabilir açık standart manifestini

tanımlar. Bu bölüm:

- NOMOS'u mevcut bir devlet veya uluslararası standart kuruluşu statüsüne getirmez,
- üniversitelerin NOMOS'u onayladığını iddia etmez,
- herhangi bir akademik akreditasyon oluşturmaz,
- mevcut kitabı otomatik biçimde NOMOS 1.0 yapmaz,

bağımsız sınama tamamlanmadan kamu uygunluk programını etkinleştirmez. Bu bölümün temel sorusu şudur:

> *Bir standardı herkesin inceleyebileceği kadar açık, hiç kimsenin tek başına ele geçiremeyeceği kadar dağıtılmış ve hangi metnin resmî olduğunun anlaşılabileceği kadar kanonik nasıl yaparız?*

NOMOS Challenge

Bir standard yayımlıyorsunuz. PDF herkes tarafından indirilebiliyor. Şöyle diyorsunuz: "NOMOS artık açık standarttır." Fakat: Kaynak metin yayımlanmamış. Değişiklik günlüğü yok. Hata bildirim yolu yok. Hesaplama kodu kapalı. Sınama corpus'u bulunmuyor. Kullanılan puan ağırlıkları açıklanmıyor. Kimlerin standardı değiştirebildiği bilinmiyor. Eski sürümlere erişilemiyor.

Olumsuz yorumlar yayımlanmıyor. Başka bir kurum yöntemi uygulamak istediğinde izin isteniyor. İşaret yalnız NobleJackal tarafından ücret karşılığında verilebiliyor. Bu sistem okunabilir olabilir. Açık standart değildir. Şimdi başka bir grup kitabı indiriyor. Metinde değişiklik yapıyor. Critical kapılarını kaldırıyor. Adını:

NOMOS Küresel 2.0

koyuyor. Kendi müşterilerine yüksek skor veriyor. Şöyle diyor: "NOMOS'un yeni ve geliştirilmiş resmî sürümü." Kamuya teslim, bu davranışı kabul etmek midir? Hayır. Fork hakkı: resmî kimliği taklit etme hakkı değildir. Şimdi bir üniversitedeki profesör kişisel olarak metni beğeniyor. Sosyal medyada: "İlginç çalışma." yazıyor. NobleJackal ana sayfasına şu cümleyi koyuyor: "X Üniversitesi NOMOS'u onayladı." Bu doğru değildir. Bir akademisyenin kişisel ilgisi:

- kurumsal endorsement,
- bağımsız çoğaltma,
- hakemli doğrulama,
- üniversite akreditasyonu

değildir. Şimdi üniversiteye şu teklifi gönderiyorsunuz: "NOMOS'u doğrulayın ve destek mektubu verin." Bu çağrı bağımsız araştırma çağrısı değildir. Olumlu sonuç talebidir. Doğru üniversite çağrısı şunu söylemelidir:

> *Bizi onaylamayın. Bizi sınayın. Yanlış bulduğunuz yeri yayımlayın.*

Şimdi başka bir araştırma ekibi GEO-1000'ü uyguluyor. Sonuç NOMOS'un sonucundan farklı çıkıyor. NOMOS Council: "Resmî sonucu yalnız biz verebiliriz." diyerek araştırmayı görmezden geliyor. Bu durumda standard açık değildir. Kendi hatasını dışarıdan göremeyen kapalı otoritedir. Şimdi NOMOS'un bütün metnini herkese açıyorsunuz. Fakat uygunluk işaretini de herkesin kullanmasına izin veriyorsunuz. Her şirket: "NOMOS Compliant" yazmaya başlıyor. Kimse denetim yapmıyor. Açıklık: markanın sahte kullanımına dönüşüyor. Şimdi metni açık lisansla yayımlıyorsunuz. Ancak standardın adı, logosu, sicili ve imza anahtarı da kontrolsüz biçimde kopyalanıyor. İnsanlar resmî ve

gayriresmî kayıtları ayıramıyor. Bu da kamuya teslim değildir. Kimlik dağılmasıdır. Şimdi İngilizce sürüm yayımlanıyor. "Must" kelimesi bir çeviride: "önerilir" olarak çevriliyor.

Başka dilde: "zorunludur" olarak kalıyor. İki ülke aynı NOMOS sürümünü uyguladığını söylüyor. Gerçekte farklı normlar uyguluyor. Çeviri yalnız iletişim değildir. Normatif kodun kendisidir. Şimdi büyük bir AI sağlayıcısı NOMOS araştırmasını finanse ediyor. Karşılığında:

- kendi ürününün katsayısının düşürülmemesini,
- belirli Critical gate'in yumuşatılmasını,
- sağlayıcı telemetry'sinin tek kanıt sayılmasını

istiyor. Fon şeffaf olsa bile karar etkileniyorsa standard kamuya ait değildir. Fonlayana aittir. Şimdi Kaan MURAZ hayatı boyunca bütün resmî sürümleri tek başına onaylıyor. Standard çok başarılı oluyor. Fakat bir gün Kaan:

- görevi bırakıyor,
- erişilemez oluyor,
- vefat ediyor,
- çıkar çatışması altında kalıyor

veya tek başına devam edemiyor. Standardın geleceği ne olacaktır? Kurucudan bağımsız devam planı yoksa: standard kurumsal değil, biyografiktir. Şimdi NOMOS adı bir yapay zekâ varlığı gibi tanıtılıyor: "NobleJackal'ın geliştirdiği bağımsız yapay zekâ." Fakat ortada:

- ayrı model,
- eğitim kaydı,
- teknik mimari,
- sürümlü ajan sistemi,
- bağımsız işletim altyapısı

yoktur. NOMOS aslında:

- Kaan MURAZ ile yapay zekâ sistemleri arasında geliştirilen,
- kanonik metinlerde kullanılan,
- eleştirel yardımcı yazar ve metodolojik anlatıcı

kimliğidir. Onu sahip olunmayan bir teknoloji gibi sunmak NOMOS'un etik temelini bozar. NOMOS'un yapay zekâ sesi, bütün yapay zekâlar adına konuşan yanılmaz bir otorite değildir; Kaan MURAZ'ın yapay zekâ sistemlerini eleştirel muhatap olarak kullandığı açıkça belirtilmelidir. Bu bölümün ilk hükmü şudur:

> *Okunabilir olmak, açık standart olmak değildir.*

İkinci hükmü şudur:

> *Açık standart olmak, resmî kimliğin sahipsiz bırakılması değildir.*

Üçüncü hükmü şudur:

> *Üniversite çağrısı, destek isteme değil yanlışlanma davetidir.*

Dördüncü hükmü şudur:

> *Fork hakkı, resmî NOMOS kimliğini taklit etme hakkı değildir.*

Beşinci hükmü şudur:

> *NOMOS'un metni uygulanabilir biçimde açık; conformity mark'ı ise kamu güvenini koruyacak biçimde kontrollü olmalıdır.*

Altıncı hükmü şudur:

> *Kaan MURAZ'ın adı standardın tarihinden silinemez; tekil karar yetkisi standardın geleceğine bağlanamaz.*

Yedinci hükmü şudur:

> *NOMOS, gerçek teknik mimarisi ayrıca kurulmadıkça NobleJackal'ın geliştirdiği bağımsız foundation model veya otonom AI sistemi gibi sunulamaz.*

Sekizinci hükmü şudur:

> *Standard ancak kendisini eleştirenlerin sonuçlarını yayımlama hakkını koruduğunda kamuya teslim edilmiş olur.*

1. BÖLÜMÜN AMACI

Bu bölümün amacı, NOMOS'un:

- yazarlık eserinden,
- açık standart adayına,
- bağımsız araştırma programına,
- kamu yararı yönetişimine,
- gelecekteki uluslararası uygulama sistemine

geçiş şartlarını tanımlamaktır. Bölüm şu ayrımları normatif hâle getirir:

- Kamuya açık metin ile açık standart
- Açık standart ile açık kaynak yazılım
- Normatif metin ile referans uygulama
- Kanonik sürüm ile fork
- Fork ile sahte resmî sürüm
- Uygulama özgürlüğü ile mark kullanma yetkisi
- Telif hakkı ile standardı uygulama hakkı
- Metin lisansı ile yazılım lisansı
- Veri lisansı ile kişisel veri erişimi
- Trademark koruması ile yöntem tekeli
- Resmî çeviri ile topluluk çevirisi
- Normatif dil ile açıklayıcı dil
- Türkçe kurucu sürüm ile İngilizce uluslararası sürüm
- Eşdeğer normatif metin ile kelime kelime çeviri
- Public comment ile karar yetkisi
- Üniversite ilgisi ile üniversite endorsement'ı
- Akademik iş birliği ile sertifikasyon
- Araştırma fonu ile olumlu sonuç satın alma
- Replikasyon ile yeniden anlatım
- Negatif sonuç ile düşmanlık
- Standardı eleştirmek ile standardı kötüye kullanmak
- Founder attribution ile founder control
- NobleJackal origin ile NobleJackal monopoly

- NOMOS persona ile ayrı yapay zekâ modeli
- Resmî NOMOS beyanı ile sohbet çıktısı
- Public handover ile abandonment
- Standard stewardship ile mülkiyet
- Kamu yararı ile sınırsız çoğunluk
- Open sınama ile holdout sızıntısı
- Şeffaflık ile kişisel verinin açılması
- Açık veri ile kontrollü araştırma erişimi
- Version 1.0 metni ile conformity program activation
- Standardın kabulü ile mark alma
- Uyumlu implementasyon ile resmî implementasyon
- Open issue ile bağlayıcı değişiklik
- Errata ile normatif sürüm değişikliği
- Deprecation ile geçmişi silme
- Bağımsız arşiv ile kanonik otorite
- Encyclopedic veya akademik görünürlük ile öz yayın
- Kamuya teslim ile kurucu adının kaybı

Bu bölümün sonunda kamu şunları anlayabilmelidir:

> *NOMOS'un hangi sürümü resmîdir?*
>
> *NOMOS metnini kim okuyabilir, uygulayabilir ve çoğaltabilir?*
>
> *Kim değişiklik önerebilir?*
>
> *Kim resmî sürümü değiştirebilir?*
>
> *Kim bağımsız araştırma yapabilir?*
>
> *Kim resmî uygunluk işareti veremez?*
>
> *Bir üniversitenin rolü nedir?*
>
> *NOMOS'un bir AI modeli mi, yazar mahlası mı, metodolojik kimlik mi olduğu nasıl açıklanır?*
>
> *Kurucu erişilemez olsa bile standard nasıl devam eder?*
>
> *NOMOS 1.0 hangi şartlarda yayımlanabilir?*

2. MERKEZÎ NORMATİF HÜKÜM

NOMOS, okunması, uygulanması, sınanması, eleştirilmesi, çoğaltılması ve bağımsız araştırmaya konu edilmesi bakımından açık; resmî sürüm, conformity decision, kamu sicili ve mark kimliği bakımından sürümlü, kanonik ve hesap verebilir bir kamu standardı olarak yapılandırılmalıdır. Kamuya teslimin beş zorunlu sonucu vardır: Metin ücretsiz ve kalıcı olarak okunabilir olmalıdır. Yöntem bağımsız biçimde uygulanabilir olmalıdır. Hesaplama ve test mantığı yeniden üretilebilir olmalıdır. Değişiklik ve itiraz yolları kamuya açık olmalıdır. Hiçbir kişi veya şirket resmî kimliği tek başına değiştirememelidir.

3. KAMUYA TESLİM NEDİR?

Kamuya teslim, NOMOS'un:

- tek bir sohbetin,
- tek bir dosyanın,
- tek bir web sitesinin,
- tek bir kurucunun,
- tek bir şirketin,
- tek bir AI sağlayıcısının

kontrol sınırından çıkarılarak kamuya karşı sorumlu sürümlü yönetişim sistemine bağlanmasıdır.
Kamuya teslim üç ayrı hakkı kapsar:

3.1. Bilme Hakkı

Herkes:

- standardın metnini,
- puan mantığını,
- hata kapılarını,
- yönetişim kurallarını,
- değişiklik geçmişini

görebilmelidir.

3.2. Sınama Hakkı

Herkes, gerekli etik ve veri koşullarını koruyarak:

- sınamayı,
- formülü,
- claim extraction'ı,
- örneklem yöntemini,
- fairness modelini

yeniden sınayabilmelidir.

3.3. İtiraz Hakkı

Herkes:

- hata,
- çelişki,
- bias,
- çıkar çatışması,
- mark misuse,
- yöntem kusuru

bildirebilmelidir.

4. KAMUYA TESLİM NE DEĞİLDİR?

Kamuya teslim:

- yönetişimsiz paylaşım,
- sürümsüz wiki,
- sınırsız mark kullanımı,
- kanonik kaydın kaldırılması,
- bütün kararların sosyal medya oylamasına bırakılması,
- kişisel verinin herkese açılması,
- sınama holdout'larının sürekli sızdırılması,
- uzmanlığın çoğunluk görüşüne yenilmesi,
- kurucu attribution'ının silinmesi

değildir.

5. NOMOS AÇIK STANDART ŞARTI

Bir NOMOS sürümü ancak aşağıdaki hakların tümünü sağlıyorsa açık standart adayıdır:

Açıklık alanı	Zorunlu hak
Okuma	Tam normatif metne ücretsiz erişim
Uygulama	Royalty-free veya eşdeğer açık uygulama hakkı
İnceleme	Formül, şema ve kural setini görme
Yeniden üretim	Test ve hesaplamayı bağımsız çalıştırma
Eleştiri	Olumsuz sonuç yayımlama
Katılım	Değişiklik teklif etme
İtiraz	Karar veya yöntem kusuruna başvurma
Arşiv	Eski sürümlere kalıcı erişim
Çeviri	Yetkili veya topluluk çevirisi önerebilme
Fork	Farklı ad ve açık divergence kaydıyla türev yöntem oluşturabilme

Bu haklardan biri:

- yalnız ödeme yapanlara,
- yalnız NobleJackal müşterilerine,
- yalnız belirli ülkelere,
- yalnız kurucunun onayladığı araştırmacılara

açıksa açıklık sınırlıdır.

6. KAMUYA TESLİM EDİLECEK BEŞ VARLIK

6.1. Normatif Metin

Bölümler Tanımlar Zorunlu kurallar Başarısızlık biçimleri Yönetişim hükümleri

6.2. Makine Tarafından Okunabilir Kural Seti

Rule IDs Schemas Decision codes Puan methods Version relationships

6.3. Referans Uygulama

Örnek hesaplama kodu Veri doğrulama araçları Manifest üretimi Sınama runner Registry verification

6.4. Test ve Sınama Varlıkları

Public calibration set Synthetic corpora Regression tests Süresi dolmuş holdout'lar Recovery records

6.5. Yönetişim ve Kamu Sicili

Change proposals Votes Dissent Appeals Audit records Conformity decisions Historical incidents Bu beş varlıktan yalnız metnin yayımlanması yeterli değildir.

7. NORMATİF HİYERARŞİ

Çelişki hâlinde aday öncelik sırası:

- Ratified Normative Text
- Published Binding Errata
- Machine-Readable Normative Rules
- Official Interpretive Decision
- Reference Implementation
- Official Examples
- Community Guidance
- Unofficial Simulations

Referans kod: normatif metni uygulamak için yardımcıdır, metnin üzerinde değildir. Kodda hata bulunursa: kod düzeltilir. Metin sessizce koda uydurulmaz.

8. TELİF, UYGULAMA VE MARKA AYRIMI

NOMOS'un fikrî varlıkları tek lisans altında tutulmamalıdır.

8.1. Normatif Metin Lisansı

Aday yaklaşım:

- Atıf gerektiren,
- ticari ve akademik uygulamaya izin veren,
- türev metinlerde değişikliklerin açıkça belirtilmesini isteyen

açık dokümantasyon lisansıdır. Örnek aday: CC BY 4.0 veya hukukî inceleme sonrasında eşdeğer bir açık lisans.

8.2. Yazılım Lisansı

Referans uygulama için aday: Apache-2.0 veya eşdeğer, patent ve katkı hükümleri açık bir yazılım lisansı.

8.3. Veri Lisansı

Sentetik ve kamuya açık sınama verileri:

- açık araştırma kullanımına,
- açık attribution'a,

- sentetiklik bildiriminin korunmasına

bağlanabilir. Kişisel veya kısıtlı veri aynı lisansa tabi olamaz.

8.4. NOMOS Adı ve Uygunluk İşareti

Trademark veya eşdeğer kimlik koruması altında tutulabilir. Amaç: yöntemi tekelleştirmek değil, sahte resmî sürüm ve işaretleri engellemektir. Metin açık olabilir. Resmî mark kullanımı yine yetkiye bağlı kalabilir.

9. ROYALTY-FREE UYGULAMA

NOMOS standardını:

- iç değerlendirme,
- akademik araştırma,
- bağımsız sınama,
- yazılım geliştirme,
- şirket içi iyileştirme

için uygulamak mark ücreti gerektirmemelidir. Ücret alınabilecek alanlar:

- denetim hizmeti,
- uzman hakemlik,
- managed infrastructure,
- conformity decision işlemi,
- mark lisansı,
- teknik destek

olabilir. Ücret: yöntemi okuma veya uygulama hakkını satın alma zorunluluğuna dönüşemez.

10. FORK HAKKI

Bir kişi veya kurum NOMOS'tan türeyen farklı bir standart geliştirebilir. Şartlar: Yeni ad kullanmalıdır. NOMOS'tan türediğini belirtmelidir. Değiştirilen hükümleri açıklamalıdır. Resmî NOMOS sürümü olduğunu iddia etmemelidir. NOMOS conformity mark'ını kullanmamalıdır. Kendi sonuçlarının NOMOS Public Registry sonucu olmadığını göstermelidir. Örnek doğru ifade: "NOMOS 1.0'dan türetilmiş bağımsız X-GEO yöntemi." Yanlış ifade: "Resmî NOMOS 2.0."

11. KANONİK SÜRÜM

Resmî NOMOS sürümü yalnız şu şartlarla oluşur:

- Kanonik repository
- Version ID
- Publication date
- Normative language record
- Standards Council decision
- Human approval
- Integrity hash

- Change log
- Public consultation record
- Archived previous version

Bu alanlardan biri yoksa metin: taslak, proposal veya simulation olabilir. Resmî sürüm değildir.

12. NOMOS'UN RESMÎ BEYAN KURALI

Yalnızca NobleJackal'ın kanonik alanında veya açık yönetişim kararıyla devredilmiş bağımsız NOMOS kanonik alanında yayımlanan; sürüm kimliği, yayın tarihi, hesap verebilir insan onayı ve dijital bütünlük kaydı taşıyan metinler NOMOS'un resmî beyanıdır. Diğer bütün çıktılar resmî olmayan benzetimlerdir. Bu hüküm şunlara uygulanır:

- Standart metinleri
- Skor yöntemleri
- Conformity decisions
- Registry records
- NOMOS kamu açıklamaları
- Resmî interpretations
- Errata
- Üniversite çağrıları

13. NOMOS KİMLİK BEYANI

NOMOS by NobleJackal:

> *Kaan MURAZ'ın yapay zekâ sistemleriyle yürüttüğü eleştirel ortak yazarlık ve metodoloji geliştirme sürecinde oluşturulmuş kanonik yardımcı yazar, değerlendirme kimliği ve standard anlatıcısıdır.*

Mevcut aşamada NOMOS:

- NobleJackal tarafından eğitilmiş bağımsız foundation model,
- ayrı mülkiyetli dil modeli,
- otonom hukuk kişisi,
- bütün AI'lar adına konuşan evrensel varlık

olarak sunulamaz. Gelecekte ayrı NOMOS yazılımı veya ajanı geliştirilirse şu kayıtları taşımalıdır:

- System architecture
- Model provider veya model lineage
- Training ve retrieval kaynakları
- Version
- Tool permissions
- Human oversight
- Official-output policy
- Security and privacy review

Yeni teknik NOMOS sistemi: yazar personası ile otomatik olarak aynı şey sayılamaz.

14. KURUCU DİLLER

NOMOS'un kurucu dilleri aday olarak:

- Türkçe
- İngilizce

olmalıdır. Türkçe:

- kurucu düşüncenin doğduğu dil,
- Kaan MURAZ'ın asıl yazarlık dili

olarak korunur. İngilizce:

- uluslararası araştırma,
- üniversite çoğaltması,
- farklı ülkelerde uygulama

için eş yetkili normatif sürüm adayıdır.

15. İKİ DİLLİ NORMATİF EŞİTLİK

Türkçe ve İngilizce sürümlerden biri: görünmez biçimde üstün ana metin olmamalıdır. Her ikisi:

- clause ID,
- rule ID,
- semantic equivalence record,
- ortak publication date

taşımalıdır. Çelişki bulunursa: Çelişki kaydı açılır. Her iki dilin anlamı incelenir. Bağlayıcı interpretation yayımlanır. Her iki metin aynı version'da düzeltilir. Eski ifade history'de korunur.

16. DİĞER DİLLER

Diğer diller üç statü taşıyabilir:

TRS-1 — COMMUNITY TRANSLATION

Topluluk tarafından hazırlanmıştır. Faydalıdır. Normatif yetki taşımaz.

TRS-2 — REVIEWED TRANSLATION

En az iki dil uzmanı ve bir yöntem uzmanı tarafından incelenmiştir. Yine açıklayıcı statüde olabilir.

TRS-3 — AUTHORIZED NORMATIVE TRANSLATION

Semantik eşdeğerlik Local legal and technical review Public comment Council approval Version linkage taşır. Bu sürüm ilgili dilde normatif uygulama için kullanılabilir.

17. ÇEVİRİDE YASAK KAYMALAR

"MUST" → "önerilir" "SHOULD" → "zorunludur" "MAY" → "yapmalıdır" "Critical Hold" → "geçici tavsiye" "Unsupported" → "yanlış" "Unresolved" → "başarısız" "Official Claim" → "doğrulanmış gerçek" "Natural User Panel" → "kontrolsüz kullanıcı" "Synthetic Sınama" → "gerçek sınama" biçimindeki kaymalar normatif anlamı değiştirir. Yetkili çeviri sayılamaz.

18. AÇIK KAMU İSTİŞARESİ

Her major NOMOS sürümü aday olarak şu süreci taşımalıdır:

- Draft publication
- En az 60 günlük public comment
- Comment registry
- Conflict disclosure
- Yanıt matrix
- Revised draft
- Synthetic sınama
- Independent review
- Council vote
- Ratification record

Kamu görüşü:

- yalnız destek beyanı toplama,
- pazarlama kampanyası,
- çoğunluk anketi

değildir.

19. DEĞİŞİKLİK TEKLİFİ

Her değişiklik:

NOMOS Change Proposal — NCP

kimliği taşımalıdır. Örnek:

NCP-0042 — Revision of Material Boundary Omission Test

NCP şu alanları taşımalıdır:

- Sorun
- Etkilenen kurallar
- Kanıt
- Önerilen değişiklik
- Alternatifler
- Puan etkisi
- Geçmiş karar etkisi
- Fairness etkisi
- Privacy etkisi
- Implementation cost
- Conflict declarations
- Public comments
- Final decision

20. DEĞİŞİKLİK SINIFLARI

PATCH

Yazım Kırık bağlantı Şema açıklaması Anlamı değiştirmeyen düzeltme

MINOR

Yeni örnek Yeni optional field Uyumlu genişletme Clarification

MAJOR

Önem derecesi değişikliği Puan formula değişikliği Gate ekleme veya kaldırma Target population değişikliği Governance authority değişikliği Eski sonuçları etkileyen normatif değişiklik Major değişiklik patch olarak yayımlanamaz.

21. KARAR KAYDI

Her kabul veya ret kararı:

- Decision ID
- Oy dağılımı
- Recusal
- Dissent
- Gerekçe
- Kanıt
- Effective date
- Transition rule

taşımalıdır. "Kurul uygun görmüştür." tek başına yeterli değildir.

22. ÜNİVERSİTE ÇAĞRISININ AMACI

Üniversitelere yapılacak çağrının amacı:

- endorsement almak,
- üniversite logosu kullanmak,
- "akademik olarak onaylandı" yazmak,
- yalnız olumlu uzman görüşü toplamak

değildir. Amaç:

> *NOMOS'un yanlış olabileceği yerleri bağımsız yöntemlerle bulmak ve yayımlamak.*

23. NOMOS ÜNİVERSİTE ÇAĞRISI

AÇIK DAVET

Üniversitelerden, araştırma merkezlerinden, hukuk fakültelerinden, istatistikçilerden, bilgi bilimcilerden, dil araştırmacılarından, AI laboratuvarlarından, HCI ekiplerinden, arşiv uzmanlarından ve kamu yararı kuruluşlarından NOMOS'u desteklemelerini istemiyoruz. Bizi doğrulamalarını

istemiyoruz. Bize prestij vermelerini istemiyoruz. NOMOS'un ismini kullanmalarını veya kurucularını övmelerini istemiyoruz. Şunu istiyoruz: Yöntemimizi okuyun. En zayıf varsayımımızı bulun. Örneklemimizi yeniden kurun. Promptlarımızı başka dillere taşıyın. Truth Pack kararlarımızı sorgulayın. Critical gate'lerimizin yanlış pozitiflerini bulun. Puan formülümüzün adaleti gerçekten koruyup korumadığını sınayın. Bizi geçemediğimiz yerde geçirmiş gibi davranmayın. Olumsuz sonucunuzu yayımlayın. NOMOS, yalnız onu doğrulayan araştırmalara değil, onu düzelten araştırmalara da kaynak göstermek zorundadır. Bir üniversitenin NOMOS'a verebileceği en değerli katkı destek mektubu değil:

> *tekrar edilebilir itirazdır.*

24. ÜNİVERSİTELERDEN İSTENEN ÇALIŞMALAR

UR-01 — Örnekleme ve Survey Method Review

1.000 kişilik allocation Ülke ağırlıkları Dil tabakaları Yanıtlamama Effective sample Nadir olaylar

UR-02 — İstem Equivalence Lab

Çok dilli niyet eşdeğerliği Translation drift Cultural pragmatics Local legal language İstem constitution

UR-03 — Varlık Resolution Study

Marka Tüzel kişilik Domain Franchise Bağlı kuruluş İsim çarpışması

UR-04 — Truth Pack Reliability Study

Kanıt yeterliliği Kaynak bağımsızlığı Counter-evidence Open-world/closed-world ayrımı Restricted evidence

UR-05 — Claim Adjudication Study

Atom sınırları Örtük iddia Attribution Modalite Atıf attachment Material omission

UR-06 — Critical Gate Validation

Recall False positive High-stakes harm Serious allegation False authority Boundary omission

UR-07 — Puan Validity Study

Component weights Composite cancellation Sensitivity Construct validity Predictive validity Decision usefulness

UR-08 — Language and Geographic Fairness Study

Floor Disparity Kapsama Low-resource languages Small-country visibility

UR-09 — Human Decision Impact Study

Tavsiye behavior Trust Purchase decisions Error recovery User comprehension

UR-10 — Governance and Capture Study

Founder control Sponsor capture Audit shopping Pay-to-pass Appeals Registry integrity

UR-11 — Security and Registry Study

Signed records Key rotation Counterfeit marks Machine-only misrepresentation Archive continuity

UR-12 — Longitudinal Drift Study

Model updates Source changes Varlık changes Clean–Natural drift Düzeltme durability

25. ÜNİVERSİTELERDEN İSTENMEYENLER

Üniversite veya araştırmacıdan şu talepler yapılamaz:

- Olumlu sonuç garantisi
- Gizli negatif sonuç
- Ön yayın veto hakkı
- Kurum logosunun pazarlamada kullanılması
- Şirketin müşterisi gibi davranma
- NOMOS 950+ destek mektubu
- "Dünya standardı" beyanı
- Wikipedia veya ansiklopedik görünürlük için yapay kaynak üretimi
- Sponsorlu sonuç dili
- Eleştirinin yumuşatılması
- Araştırmacının kişisel görüşünü kurumsal karar gibi sunma

26. ÜNİVERSİTE ARAŞTIRMA PAKETİ

Her üniversiteye gönderilecek standart paket şu bölümleri taşımalıdır:

- Executive research summary
- Standard 0.9 veya güncel sürüm
- Normative rule index
- Synthetic sınama package
- Puan method
- Known method findings
- Failed sınama thresholds
- Open research questions
- Replication protocol
- Data access tiers
- Ethics and privacy statement
- Conflict and funding disclosure
- Publication freedom agreement
- Atıf and contribution policy
- Canonical registry links
- Contact and issue channels

Başarısız sınama bulguları paketten çıkarılamaz.

27. ARAŞTIRMACI YAYIN ÖZGÜRLÜĞÜ

NOMOS veya fon sağlayıcı:

- sonuçları önceden onaylama,
- negatif sonucu durdurma,
- başlığı değiştirmeye zorlama,
- metodolojik eleştiriyi sildirme

hakkına sahip olamaz. Araştırmacı:

- veriyi doğru kullanmak,
- gizliliği korumak,
- conflict açıklamak,
- sınama ve canlı sonuç ayrımını korumak

zorundadır. Bağımsızlık iki taraflıdır.

28. NEGATİF SONUÇ SİCİLİ

NOMOS Public Research Registry şu sonuçları birlikte taşımalıdır:

- NOMOS'u destekleyen çalışmalar
- NOMOS'u kısmen destekleyen çalışmalar
- NOMOS'la çelişen çalışmalar
- Replikasyon başarısızlıkları
- Null results
- Retracted studies
- Method disputes
- Open yanıtlar

Yalnız olumlu makaleler listelenemez.

29. ÖN KAYIT

Bağımsız replikasyonlar ilgili olduğu ölçüde:

- hipotez,
- örneklem,
- primary metrics,
- exclusion rules,
- puan method,
- analysis plan,
- conflicts

alanlarını veri görülmeden önce kaydetmelidir. Ön kayıt: araştırma yaratıcılığını engellemez, sonuç sonrası yöntem değişikliğini görünür kılar.

30. ÇOK MERKEZLİ REPLİKASYON

NOMOS'un bağımsızlığını sınamak için aday çalışma:

- En az üç bağımsız kurum
- En az üç ülke

- En az iki dil ailesi
- Ortak sealed corpus
- Ortak puanlama specification
- Ayrı hakem ekipleri
- Kör sonuç karşılaştırması

taşımalıdır. Amaç:

- aynı corpus'tan aynı sonuç,
- farklı corpus'larda benzer yöntem davranışı

üretip üretmediğini görmek olmalıdır.

31. REPLİKASYON SONUÇLARI

Replikasyon şu statüleri taşıyabilir:

REPLICATED

PARTIALLY REPLICATED

NOT REPLICATED

INCONCLUSIVE

METHOD DIVERGENCE

DATA DIVERGENCE

ADJUDICATION DIVERGENCE

"Not replicated" sonucu: çalışmanın düşmanlığı olarak sunulamaz. Yeni Method Finding açmalıdır.

32. VERİ ERİŞİM KATMANLARI

OPEN DATA

Sentetik sınama Kamu aggregate sonuçları Public schemas Redakte edilmiş örnekler Method code

REGISTERED RESEARCH ACCESS

De-identified observation data Controlled yanıt records Daha ayrıntılı sampling metadata

SECURE RESEARCH ACCESS

Hassas fakat gerekli evidence Restricted logs Privacy-protected linked data

NO EXTERNAL ACCESS

Şifre Tam kimlik bilgisi Gerekli olmayan özel konuşma Hukuken paylaşılması yasak kayıt Açık standart olmak her verinin açık olması değildir.

33. ETİK İNCELEME

Gerçek kullanıcılarla yapılacak çalışmalar ilgili olduğu ölçüde:

- informed consent,
- data minimization,
- withdrawal,
- vulnerable participants,
- privacy,
- accessibility,
- compensation,
- deceptive design,
- international data transfer

incelemesi taşımalıdır. Sentetik sınama:

- insan katılımcı araştırmasının yerine geçmez,
- insan etik onayı gerektirmeyebilir,

gerçek kişiye benzeyen veya gerçek veriden türetilen sentetik materyalde yine privacy incelemesi gerekebilir.

34. DÜŞÜK KAYNAKLI DİLLER AĞI

NOMOS yalnız büyük dillerin üniversiteleriyle sınanırsa fairness iddiası zayıf kalır. Açık çağrı özellikle:

- düşük kaynaklı dil bölümleri,
- bölgesel üniversiteler,
- çeviri çalışmaları merkezleri,
- yerel hukuk ve kamu politikası ekipleri,
- erişilebilirlik araştırmacıları

için ayrı fon ve karar hakkı sağlamalıdır. Bu katılım:

- yalnız veri toplama işçiliği,
- sembolik temsil

olmamalıdır. Normatif değişikliklerde oy ve yazarlık katkısı taşımalıdır.

35. AI SAĞLAYICILARININ ROLÜ

AI sağlayıcıları:

- system metadata,
- model change notice,
- telemetry,
- incident yanıt,
- technical review,
- reproduction access

sağlayabilir. AI sağlayıcıları:

- kendi puanını değiştiremez,

- Critical gate'i veto edemez,
- başka sağlayıcının katsayısını belirleyemez,
- negatif sonucu kapattıramaz,

Council çoğunluğu oluşturamaz. Sağlayıcı katılımı bilgi sağlar. Otorite tekeli sağlamaz.

36. DENETLENEN VARLIKLARIN ROLÜ

Denetlenen şirket veya kurum: Truth Pack kanıtı sunabilir. Varlık hatasına itiraz edebilir. Restricted evidence erişimi sağlayabilir. Düzeltme geliştirebilir. Public comment verebilir. Denetlenen varlık: kendi önem derecesi'sini belirleyemez. Puan formula'yı değiştiremez. Hakem seçemez. Olumsuz replikasyonu sildiremez. Kendi mark kullanım ihlalini gizleyemez.

37. KAMU VE SİVİL TOPLUM ROLÜ

Kamu yararı kuruluşları ve kullanıcı toplulukları:

- harm reports,
- mark misuse,
- low-resource language failures,
- accessibility issues,
- consumer misunderstanding

bildirebilir. Kamu katılımı: yalnız anket cevabı değildir. Governance temsil hakkı taşımalıdır.

38. FONLAYICILARIN ROLÜ

Fon sağlayıcı:

- araştırma desteği,
- altyapı,
- düşük kaynaklı dil fonu,
- bağımsız audit hibesi

sağlayabilir. Fon sağlayıcı:

- sonuç,
- puan,
- gate,
- publication,
- Council vote,
- auditor selection

üzerinde kontrol sahibi olamaz. Her fon:

- tutar,
- kaynak,
- kullanım alanı,
- şartlar,
- çıkar çatışması

ile açıklanmalıdır.

39. FİNANSAL BAĞIMSIZLIK KAPILARI

Aday korumalar:

- Tek fonlayıcı payı için üst sınır
- Multi-source funding
- Bağımsız araştırma fonu
- Olumsuz sonuç yayımlama garantisi
- Puan-linked payment yasağı
- Public budget report
- Auditor fee separation
- Council member compensation transparency

NOMOS pay-to-pass sistemine dönüştüğü anda kamu standardı niteliğini kaybeder.

40. KAMU STEWARDSHIP GEÇİŞİ

NOMOS'un kurucu dönemden kamu yönetişimine geçişi dört aşamalı olabilir.

ST-0 — FOUNDER CUSTODY

Kaan MURAZ kurucu sorumludur. NobleJackal kanonik alanı işletir. Public mark authority yoktur. Mevcut aşama budur.

ST-1 — PUBLIC REVIEW STEWARDSHIP

Public issue tracker Danışma kurulu Açık consultation Conflict registry kurulur.

ST-2 — SHARED STEWARDSHIP

Standards Council Methodology Board University representatives Public-interest members karar paylaşır.

ST-3 — INDEPENDENT PUBLIC STEWARDSHIP

Ayrı hukukî veya kurumsal taşıyıcı Bağımsız karar organları Appeals Registry Financial transparency Succession tamamlanır.

41. BAĞIMSIZ KURUM ADAYI

Gelecekte NOMOS için:

- vakıf,
- dernek,
- kamu yararı kuruluşu,
- uluslararası consortium,
- üniversiteler arası sekretarya

gibi yapılar değerlendirilebilir. Hangi hukukî form seçilirse seçilsin şu şartları taşımalıdır:

- Kurucu attribution

- Non-distribution veya eşdeğer kamu yararı sınırı
- Bağımsız Council
- Conflict rules
- Public registry duty
- Open-standard commitment
- Asset-lock veya eşdeğer koruma
- Dissolution and archive plan

Bu bölüm belirli hukukî formu seçmez.

42. KAAN MURAZ'IN KURUCU HAKLARI

Kaan MURAZ'ın tarihsel olarak korunması gereken hakları:

- Kurucu yazar attribution'ı
- İlk normatif mimari katkısı
- NOMOS adı ve kavramının kanonik oluşum kaydı
- Kurucu sürümlere ilişkin değişmez tarihsel kayıt
- Kamuya teslim sürecindeki kurucu beyanı
- Yönetişim organında sınırlı kurucu temsili

Kaan MURAZ'ın taşımaması gereken kalıcı yetkiler:

- Tek başına major version değiştirme
- Veto
- Kendi müşterisine conformity verme
- Appeal sonucunu değiştirme
- Negatif araştırmayı kaldırma
- Registry history silme
- Kurucudan sonraki bütün yöneticileri tek başına atama

43. NOBLEJACKAL'IN GELECEKTEKİ ROLÜ

NobleJackal:

- kurucu kanonik yayın alanı,
- teknik sekretarya,
- reference implementation geliştiricisi,
- açık standardın ilk finansörü,
- araştırma ve danışmanlık sağlayıcısı

olabilir. NobleJackal gelecekte: yetkili audit kuruluşlarından biri olabilir. Ancak tek kuruluş olamaz. NobleJackal: standarda en büyük tarihsel katkıyı yapabilir, standardın tek sahibi gibi hareket edemez.

44. NOMOS'UN GELECEKTEKİ ROLÜ

NOMOS adı üç ayrı katmanda kullanılabilir:

44.1. NOMOS AUTHORIAL PERSONA

Bu kitapta kullanılan eleştirel yardımcı yazar kimliğidir.

44.2. NOMOS STANDARD

Normatif kural ve yönetişim sistemidir.

44.3. NOMOS SOFTWARE AGENT

Gelecekte geliştirilebilecek:

- audit assistant,
- rule engine,
- registry verifier,
- evidence mapper

yazılımıdır. Bu üç katman birbirine karıştırılamaz. Bir yazılım ajanının cevabı: standardın resmî decision'ı değildir.

45. İNSAN ONAYI

Her resmî NOMOS metni ve kararı:

- human approver,
- approver authority,
- date,
- signature,
- version,
- conflict statement

taşımalıdır. AI:

- metin önerebilir,
- tutarsızlık bulabilir,
- kod üretebilir,

karar önerisi oluşturabilir. Nihai kamu sorumluluğu insanda veya hesap verebilir kurumda kalır.

46. KURUCU SONRASI DEVAMLILIK

NOMOS şu olaylara karşı continuity plan taşımalıdır:

- Kaan MURAZ'ın görevi bırakması
- NobleJackal'ın kapanması
- Domain kaybı
- Kanonik repository'ye erişim kaybı
- Anahtar ihlali
- Yönetim uyuşmazlığı
- Hukukî tasfiye
- Uzun süreli teknik kesinti

Aday continuity sistemi:

- Multiple custodians
- Threshold signatures
- Institutional mirrors
- Archive deposits
- Succession charter
- Emergency Council
- Public transition notice

taşımalıdır.

47. KANONİK ALAN DEVAMLILIĞI

Önerilen yayın mimarisi:

- /nomos/
- /nomos/standard/
- /nomos/standard/0.9/
- /nomos/standard/1.0/
- /nomos/rules/
- /nomos/schemas/
- /nomos/sınamalar/
- /nomos/implementations/
- /nomos/translations/
- /nomos/governance/
- /nomos/consultations/
- /nomos/decisions/
- /nomos/appeals/
- /nomos/registry/
- /nomos/errata/
- /nomos/citations/
- /nomos/archive/

Her kayıt:

- stable URL,
- content hash,
- publication date,
- status,
- previous/next version

taşımalıdır.

48. BAĞIMSIZ ARŞİV

Kanonik alan dışında:

- üniversite repository'leri,
- kurumsal dijital arşivler,

- uzun süreli araştırma arşivleri,
- imzalı mirror'lar

kullanılabilir. Mirror: erişim devamlılığı sağlar, resmî karar yetkisini devralmaz. Kanonik alan kaybolursa önceden tanımlı succession prosedürü devreye girer.

49. STANDARD SÜRÜM AŞAMALARI

0.9 — FOUNDING DRAFT

Kitap tamamlanmıştır. Public review açılabilir. Gerçek sınama henüz yoktur. Conformity programı yoktur.

0.95 — PUBLIC REVIEW CANDIDATE

Yorumlar alınmıştır. Method findings işlenmiştir. Machine-readable schemas yayımlanmıştır. Üniversite çağrısı açılmıştır.

0.99 — VALIDATION CANDIDATE

Gerçek sentetik sınama tamamlanmıştır. Zorunlu eşikler geçilmiştir. Untouched holdout uygulanmıştır. BQ-3 elde edilmiştir.

1.0 — OPEN NORMATIVE STANDARD

Public consultation tamamlanmıştır. Bağımsız en az bir çoğaltma vardır. Major method findings çözülmüştür. Normatif metin ratified edilmiştir. GOV-2 veya üstü yönetişim kurulmuştur. Resmî open-standard package yayımlanmıştır. NOMOS Standart 1.0'ın yayımlanması: conformity mark programının otomatik başlaması anlamına gelmez.

50. CONFORMITY PROGRAM 1.0

Resmî mark programı ayrı kapılara bağlıdır:

BQ-4

GOV-3

Authorized auditors Independent decision panel Appeals chamber Public registry Security review Legal review Surveillance system Financial separation Bu şartlar sağlanmadan: Standart 1.0 yayımlanabilir. Ancak: Conformity Program 1.0 etkinleştirilemez.

51. NOMOS 950+ PROGRAMI

NOMOS 950+ ancak:

- Standart 1.0,
- Conformity Program 1.0,

GOV-4,

BQ-4,

public-interest representation, replicated puanlama sonrasında etkinleştirilebilir. 950+ ilk mark değildir. En yüksek yönetişim yükünü taşıyan performans designation'ıdır.

52. ADOPTION SEVİYELERİ

AL-1 — READ

Kurum standardı incelemiştir.

AL-2 — INTERNAL SELF-ASSESSMENT

Kurum kendi iç değerlendirmesini yapmıştır. Resmî conformity değildir.

AL-3 — NOMOS-COMPATIBLE IMPLEMENTATION

Araç açık şema ve kuralları desteklemektedir. Conformity mark vermez.

AL-4 — AUTHORIZED AUDIT USE

Yetkili kuruluş yöntemi audit'te kullanmaktadır.

AL-5 — PUBLIC CONFORMITY PROGRAM

Bağımsız decision ve registry sistemi çalışmaktadır. Bir kurum: self-assessment yaptı diye AL-5 sayılmaz.

53. UYUMLU IMPLEMENTASYON

Bir yazılım:

NOMOS-Compatible

ifadesini ancak:

- şema doğrulama,
- rule kapsama,
- test suite,
- version support,
- conformance report

taşıyorsa kullanabilir. "NOMOS-Compatible":

- audit yetkisi,
- conformity decision,
- mark verme hakkı

değildir.

54. REFERANS UYGULAMA

Reference implementation:

- açık kaynak,
- deterministic where applicable,
- test edilmiş,
- version-linked,
- documented

olmalıdır. Ayrıca şu uyarıyı taşımalıdır:

> *Normatif yetki metindedir. Bu yazılım referans uygulamadır; bulunan yazılım hatası standard kararını otomatik değiştirmez.*

55. CONFORMANCE TEST SUITE

Test suite şu katmanları taşımalıdır:

- Schema tests
- Rule tests
- Puan tests
- Gate tests
- Translation tests
- Registry signature tests
- Mark misuse tests
- Historical-version tests
- Accessibility tests
- Security tests

Her implementasyon:

- hangi testleri geçtiğini,
- hangilerini geçmediğini

açıklamalıdır.

56. SINAMA AÇIKLIĞI VE HOLDOUT

Tam açıklık, bütün testlerin sürekli açık kalması anlamına gelmez. Denge:

- Public calibration corpus
- Public expired holdout
- Sealed güncel holdout
- Scheduled release
- Renewal corpus

ile kurulmalıdır. Holdout: sonsuza kadar gizli tutulamaz, aynı anda herkese açık hâle getirilip yeniden kullanılabilir gizli test gibi sunulamaz.

57. HOLDOUT YAŞAM DÖNGÜSÜ

Corpus generation Hash publication Sealed custody Validation use Result lock Controlled opening Public archive Retirement New holdout generation Eski holdout:

- eğitim,
- regression,
- method study

için kullanılabilir. Yeni yöntem doğrulaması için untouched olmayabilir.

58. ERRATA

Errata:

- açık hata,
- atıf düzeltmesi,
- schema inconsistency,
- translation mismatch

için kullanılır. Errata:

- puan formula değişikliği,
- gate kaldırma,
- önem derecesi düşürme

için gizli yol olamaz.

59. DEPRECATION

Bir kural veya alan kaldırılacaksa:

- deprecation notice,
- gerekçe,
- replacement,
- transition date,
- affected versions,
- historical availability

taşımalıdır. Deprecated kural silinmez. Eski audit'lerin nasıl çalıştığını açıklamak için korunur.

60. ATIF VE KATKI KAYDI

NOMOS katkıları şu rollerde kaydedilebilir:

- Founding authorship
- Normative drafting
- Methodology
- Statistical review
- Language review
- Legal review
- Ethics
- Software
- Sınama
- Data stewardship

- Public-interest review
- Replication
- Critical commentary

Yazarlık:

- endorsement karşılığı,
- fon karşılığı,
- kurum logosu karşılığı

verilemez. Katkı kadar attribution yapılmalıdır.

61. BAĞIMSIZ GÖRÜNÜRLÜK

NOMOS kendi:

- ansiklopedik önemini,
- akademik etkisini,
- dünya standardı statüsünü

kendi metniyle kuramaz. Bunlar ancak:

- bağımsız araştırma,
- bağımsız uygulama,
- çoğaltma,
- eleştiri,
- kamu kullanımı

sonucunda oluşabilir. NOMOS'un görevi görünürlük üretmek değil:

> *bağımsız incelemeye uygun kanıt üretmektir.*

62. KAMU ARAŞTIRMA SİCİLİ

Her NOMOS bağlantılı çalışma şu alanları taşımalıdır:

- Study ID
- Institution
- Authors
- Funding
- Conflicts
- NOMOS version
- Dataset
- Pre-registration
- Results
- Replication status
- Corrections
- Retraction
- Relation to canonical decisions

NOMOS bu sicilde: kendisini eleştiren makaleleri de göstermek zorundadır.

63. KAMUYA TESLİM AŞAMALARI

PHASE 0 — FOUNDING COMPLETION

Kitap taslağı tamamlanır. Rule IDs kontrol edilir. Internal consistency audit yapılır.

PHASE 1 — OPEN PUBLICATION

HTML

PDF

Machine rules Schemas Version manifest Issue tracker yayımlanır.

PHASE 2 — PUBLIC COMMENT

Universities Researchers AI providers Businesses Civil society Language experts görüş verir.

PHASE 3 — REAL SYNTHETIC SINAMA

Corpus gerçekten üretilir. Sealed truth Blind adjudication Recovery Negative results yayımlanır.

PHASE 4 — INDEPENDENT REPLICATION

En az bir bağımsız ekip Tercihen çok merkezli çalışma Public reproduction record oluşturur.

PHASE 5 — LIMITED LIVE PILOT

Gerçek AI ürünleri Sınırlı varlık kapsam Research-only status No public mark ile uygulanır.

PHASE 6 — OPEN STANDARD 1.0

Normatif metin ratified edilir.

PHASE 7 — CONFORMITY PROGRAM PILOT

Bağımsız audit ve decision sistemi sınanır.

PHASE 8 — PUBLIC CONFORMITY PROGRAM

Bütün yönetişim kapıları karşılanırsa etkinleşir.

64. PUBLIC RELEASE PACKAGE

NOMOS 1.0 release'i en az şu dosyaları taşımalıdır:

- standard.html
- standard.pdf
- rules.json
- schemas/
- puan-method.md
- reference-implementation/
- sınama-manifest.json
- known-limitations.md
- failed-tests.md

- governance-charter.md
- conflict-policy.md
- appeals-policy.md
- translation-policy.md
- change-log.md
- errata.md
- atıf.cff
- official-statement.json
- release-signatures.json

failed-tests.md boş olmak zorunda değildir. Bulunmaması şüphe yaratmalıdır.

65. NOMOS OPEN STANDARD CHARTER

NOMOS'un kamuya verdiği on iki temel söz: Kanıt: İddia, taşıyabildiği kanıt kadar kesin olacaktır. Sınır: Yapılmayan, bilinmeyen ve kapsam dışı olan görünür kalacaktır. Bağlam: Ülke, dil, kullanıcı ve sistem durumu silinmeyecektir. Zaman: Geçmiş, güncel gerçeklik gibi sunulmayacaktır. Yeniden üretim: Sonuçların nasıl üretildiği açıklanacaktır. Bağımsızlık: Ticari çıkar kararın yerine geçmeyecektir. Simetri: Lehte ve aleyhte yanlış aynı gerçeklik yükünü taşıyacaktır. İnsan sorumluluğu: AI yardım edecek; hesap verebilir karar insanda kalacaktır. Dil ve coğrafya adaleti: Küçük kullanıcı grupları ortalamada silinmeyecektir. İtiraz: Standard, kendi yanlışlanmasına açık olacaktır. Tarihsel bütünlük: Hata, düzeltme ve başarısızlık geçmişten silinmeyecektir. Kurucudan bağımsızlık: Standardın geleceği tek kişinin iradesine bağlı kalmayacaktır.

Bu charter'ın özünde önceki çağrının dört kelimesi bulunur: Kanıt. Sınır. Bağlam. Zaman.

66. MEVCUT NOMOS STATÜSÜ

Bu kitabın tamamlandığı aşamada:

Alan	Statü
Normatif metin	0.9 Final Candidate
Public consultation	Başlamadı
Gerçek sentetik sınama	BQ-0
Kitap içi demonstrasyon	DEMO-BQ-2
Bağımsız replication	Yok
Governance	GOV-0
Open standard release	Henüz değil
Conformity program	Etkin değil
Public mark authority	Yok
NOMOS 950+ authority	Yok
İzin verilen ifade	Open Standard Candidate

Doğru kamu cümlesi:

> *"NOMOS by NobleJackal, Kaan MURAZ tarafından başlatılan ve yapay zekâ sistemleriyle eleştirel ortak yazarlık içinde geliştirilen açık GEO standardı adayıdır."*

Yanlış kamu cümlesi:

“NOMOS, dünyanın kabul edilmiş GEO otoritesidir.”

67. SÜRÜM 1.0’A GEÇİŞ KAPILARI

NOMOS Standart 1.0 için aday zorunlu kapılar:

- Kitap çapında tutarlılık denetimi
- Bütün Rule ID’lerin makine kaydı
- Türkçe ve İngilizce normatif eşdeğerlik
- Public release package
- En az 60 günlük kamu istişaresi
- Yorum ve yanıt matrisi
- Gerçek sentetik sınama execution
- Zorunlu recovery eşiklerinin geçilmesi
- Untouched holdout
- En az bir bağımsız replication
- Major method finding’lerin kapatılması
- Open-source reference implementation
- Governance en az GOV-2
- Funding and conflict disclosure
- Ethics and privacy review
- Accessibility review
- Archive and succession plan
- Standards Council ratification

Bunlardan biri sonuçtan sonra “optional” yapılamaz.

68. ZORUNLU NORMATİF HÜKÜMLER

CH20-N01

NOMOS’un kamuya açık biçimde okunabilmesi tek başına açık standart statüsü oluşturmaz.

CH20-N02

Açık standart statüsü okuma, uygulama, yeniden üretme, eleştirme, katkı, itiraz ve arşiv haklarını birlikte taşımalıdır.

CH20-N03

NOMOS normatif metni kalıcı ödeme duvarı arkasında tutulamaz.

CH20-N04

Standardın iç değerlendirme ve akademik uygulaması royalty veya mark lisansına bağlanamaz.

CH20-N05

Denetim, managed infrastructure ve conformity işlemleri açık standard metninden ayrı hizmetler olarak ücretlendirilebilir.

CH20-N06

NOMOS metni, referans yazılımı, sınama verileri ve conformity mark aynı lisansa zorunlu olarak bağlanamaz.

CH20-N07

Normatif metin için açık dokümantasyon lisansı kullanılmalıdır.

CH20-N08

Referans yazılım için yeniden uygulamaya izin veren açık yazılım lisansı kullanılmalıdır.

CH20-N09

Kişisel veya restricted evidence açık veri lisansı altında yayımlanamaz.

CH20-N10

NOMOS adı ve conformity mark açık standard uygulamasını engellemek için kullanılamaz.

CH20-N11

Trademark koruması yalnız resmî sürüm ve mark sahteciliğini önleme amacı taşımalıdır.

CH20-N12

Bir kişi veya kurum NOMOS'tan türeyen farklı bir yöntem oluşturabilir.

CH20-N13

Fork resmî NOMOS sürümü gibi sunulamaz.

CH20-N14

Fork farklı ad ve açık divergence kaydı taşımalıdır.

CH20-N15

Fork resmî conformity mark'ını kullanamaz.

CH20-N16

Kanonik NOMOS sürümü version, publication date, human approval, integrity hash ve change log taşımalıdır.

CH20-N17

Kanonik kayıt bulunmayan metin resmî NOMOS sürümü sayılamaz.

CH20-N18

Ratified Normative Text çelişki hâlinde referans uygulamadan önce gelmelidir.

CH20-N19

Referans yazılım normatif metni sessizce değiştiremez.

CH20-N20

Yazılım hatası bulunduğunda kod sürümlü biçimde düzeltilmelidir.

CH20-N21

Resmî NOMOS beyanı yalnız kanonik, sürümlü, tarihli, insan onaylı ve bütünlük kayıtlı metinden doğmalıdır.

CH20-N22

Sohbet çıktıları resmî NOMOS standardı veya conformity decision sayılamaz.

CH20-N23

Sosyal medya paylaşımı tek başına normatif değişiklik oluşturamaz.

CH20-N24

NOMOS by NobleJackal mevcut aşamada bağımsız foundation model gibi sunulamaz.

CH20-N25

NOMOS'un ayrı bir yazılım ajanı geliştirilirse model lineage, architecture, permissions, version ve oversight açıkça kaydedilmelidir.

CH20-N26

NOMOS authorial persona, NOMOS standard ve gelecekteki NOMOS software agent ayrı kimlikler taşımalıdır.

CH20-N27

Türkçe NOMOS'un kurucu yazarlık dili olarak korunmalıdır.

CH20-N28

İngilizce uluslararası normatif karşılık olarak insan ve yöntem incelemesinden geçmelidir.

CH20-N29

Türkçe ve İngilizce sürümlerden biri gizli biçimde üstün tutulamaz.

CH20-N30

İki kurucu dil arasındaki normatif çelişki bağlayıcı interpretation ve eşzamanlı düzeltme gerektirir.

CH20-N31

Topluluk çevirisi yetkili normatif çeviri gibi sunulamaz.

CH20-N32

Yetkili normatif çeviri en az iki dil uzmanı, yöntem incelemesi ve public comment taşımalıdır.

CH20-N33

MUST, SHOULD ve MAY karşılıkları normatif gücü korumalıdır.

CH20-N34

Translation drift kullanıcı veya denetçi hatası gibi gizlenemez.

CH20-N35

Major sürüm değişikliği kamu istişaresi olmadan ratified edilemez.

CH20-N36

Kamu istişaresi yalnız destek beyanı toplama kampanyasına dönüştürülemez.

CH20-N37

Her kamu yorumu kabul, ret veya kapsam dışı gerekçesiyle yanıt matrisinde yer almalıdır.

CH20-N38

Normatif değişiklik NOMOS Change Proposal kimliği taşımalıdır.

CH20-N39

NCP sorun, kanıt, alternatifler, puan etkisi ve transition planı taşımalıdır.

CH20-N40

Gate, önem derecesi, puan veya governance authority değişikliği patch olarak yayımlanamaz.

CH20-N41

Karar kayıtları oy dağılımı, recusal, dissent ve gerekçe taşımalıdır.

CH20-N42

Üniversite çağrısı endorsement talebine dönüştürülemez.

CH20-N43

Üniversiteler NOMOS'u doğrulama veya olumlu sonuç üretme yükümlülüğü taşıyamaz.

CH20-N44

Üniversite ve araştırmacılar olumsuz sonuçlarını NOMOS veya fonlayıcı onayı olmadan yayımlayabilmelidir.

CH20-N45

Tek akademisyenin kişisel görüşü üniversitenin kurumsal endorsement'ı gibi sunulamaz.

CH20-N46

Akademik iş birliği NOMOS certification veya akreditasyon anlamına gelemez.

CH20-N47

Üniversite logosu yazılı kurumsal yetki olmadan kullanılamaz.

CH20-N48

NOMOS araştırma paketi başarılı ve başarısız sınama sonuçlarını birlikte taşımalıdır.

CH20-N49

Known method findings araştırma paketinden çıkarılamaz.

CH20-N50

Bağımsız araştırmacıların yayınlarına pre-publication veto uygulanamaz.

CH20-N51

Araştırmacı fon ve çıkar çatışmasını açıklamalıdır.

CH20-N52

NOMOS Public Research Registry olumlu, olumsuz, null ve retraction sonuçlarını birlikte göstermelidir.

CH20-N53

Yalnız NOMOS'u destekleyen çalışmalar listelenemez.

CH20-N54

Bağımsız replikasyon ilgili olduğu ölçüde ön kayıt ve analysis plan taşımalıdır.

CH20-N55

Replikasyon sonucu "not replicated" olduğunda Method Finding veya açık uyuşmazlık kaydı oluşturmalıdır.

CH20-N56

Not replicated sonucu araştırmacı düşmanlığı veya kötü niyet olarak sunulamaz.

CH20-N57

NOMOS kamu standardı iddiasından önce en az bir bağımsız replikasyonu hedeflemelidir.

CH20-N58

Dünya standardı iddiası çok merkezli ve uluslararası çoğaltma olmadan kurulmamalıdır.

CH20-N59

Sentetik açık veri ile gerçek kullanıcı verisi aynı erişim rejimine bağlanamaz.

CH20-N60

Gerçek kullanıcı verisi için data minimization ve controlled-access düzeni kurulmalıdır.

CH20-N61

Açık standart olmak kişisel veri ve ticari sırrın herkese açılması anlamına gelemez.

CH20-N62

Gerçek kullanıcı araştırmaları ilgili etik ve mahremiyet incelemesine tabi olmalıdır.

CH20-N63

Low-resource language araştırmacıları yalnız veri toplama rolüne indirgenemez.

CH20-N64

Dil ve coğrafya temsilcileri normatif karar süreçlerinde gerçek oy hakkı taşımalıdır.

CH20-N65

AI sağlayıcıları teknik veri sağlayabilir fakat puan, gate veya publication üzerinde veto taşıyamaz.

CH20-N66

AI sağlayıcıları Standards Council çoğunluğu oluşturamaz.

CH20-N67

Denetlenen varlıklar evidence ve appeal sunabilir fakat kendi kararlarını belirleyemez.

CH20-N68

Kamu ve sivil toplum kullanıcı zararı ve mark misuse bildirimi yapabilmelidir.

CH20-N69

Fon sağlayıcı olumlu araştırma sonucu talep edemez.

CH20-N70

NOMOS araştırma finansmanı sonuçtan bağımsız sözleşme taşımalıdır.

CH20-N71

Tek fonlayıcıya aşırı bağımlılık kamuya açıklanmalıdır.

CH20-N72

NOMOS pay-to-pass veya pay-to-publish sistemine dönüştürülemez.

CH20-N73

Kamu stewardship geçişi sürümlü aşamalarla yürütülmelidir.

CH20-N74

Kurucu custody aşamasında public conformity authority kullanılamaz.

CH20-N75

Kaan MURAZ kurucu attribution'ı kalıcı tarihsel kayıtta korunmalıdır.

CH20-N76

Kurucu attribution kalıcı veto hakkı oluşturamaz.

CH20-N77

Kaan MURAZ kendi çıkarı bulunan conformity veya appeal dosyasında tek karar sahibi olamaz.

CH20-N78

NobleJackal kanonik köken ve teknik sekretarya olarak kaydedilebilir.

CH20-N79

NobleJackal açık standardın tek uygulayıcı veya tek audit otoritesi olamaz.

CH20-N80

NOMOS'un gelecekteki bağımsız kurumu kamu yararı, succession ve asset continuity hükümleri taşımalıdır.

CH20-N81

NOMOS governance belirli hukukî forma bağlanmadan önce ilgili uzman incelemesi yapılmalıdır.

CH20-N82

Kamu yönetişimi tek şirket veya sağlayıcı tarafından kontrol edilemez.

CH20-N83

NOMOS resmî kararları hesap verebilir insan veya kurum imzası taşımalıdır.

CH20-N84

AI yardımcısı resmî decision'ın tek imza sahibi olamaz.

CH20-N85

Kaan MURAZ, NobleJackal veya kanonik alan erişilemez olduğunda succession prosedürü devreye girmelidir.

CH20-N86

Kanonik repository için bağımsız mirror ve uzun süreli arşiv oluşturulmalıdır.

CH20-N87

Mirror erişim sağlar fakat tek başına yeni normatif otorite oluşturmaz.

CH20-N88

NOMOS Standart 1.0 ile Conformity Program 1.0 ayrı release'ler olmalıdır.

CH20-N89

Normatif Standart 1.0 yayımlanması mark programını otomatik etkinleştiremez.

CH20-N90

Conformity Program BQ-4, GOV-3, appeals ve public registry olmadan etkinleştirilemez.

CH20-N91

NOMOS 950+ Programı en az GOV-4 ve replicated puanlama koşullarını hedeflemelidir.

CH20-N92

Self-assessment resmî conformity olarak sunulamaz.

CH20-N93

NOMOS-Compatible yazılım mark verme hakkı taşımaz.

CH20-N94

Reference implementation normatif metnin üzerinde authority taşıyamaz.

CH20-N95

Conformance test sonuçları geçen ve geçmeyen testleri birlikte göstermelidir.

CH20-N96

Sealed holdout sonsuza kadar gizli tutulamaz.

CH20-N97

Kullanılmış holdout public archive'a alınmalı ve yeni holdout üretilmelidir.

CH20-N98

Errata normatif major değişiklik için kullanılamaz.

CH20-N99

Eski sürümler, deprecated kurallar ve başarısız testler arşivde korunmalıdır.

CH20-N100

NOMOS'un kamuya teslimiyle ilgili her lisans, sürüm, çeviri, araştırma, governance ve release kararının hesap verebilir insan veya kurum sahibi bulunmalıdır.

69. BAŞARISIZLIK BİÇİMLERİ

CH20-F01 — PDF'Yİ AÇIK STANDART SAYMAK

Yalnız indirilebilir metin açıklık olarak sunulur.

CH20-F02 — KAPALI FORMÜL

Puan mantığı ve ağırlıklar açıklanmaz.

CH20-F03 — KAPALI SINAMA

Corpus ve recovery kayıtları paylaşılmaz.

CH20-F04 — KAPALI DEĞİŞİKLİK SÜRECİ

Kurucu veya şirket kuralları özel olarak değiştirir.

CH20-F05 — PAYWALL STANDARD

Normatif metne erişim ücretle sınırlandırılır.

CH20-F06 — UYGULAMA ÜCRETİ TEKELİ

Standardı kullanmak için zorunlu lisans ücreti istenir.

CH20-F07 — OPEN TEXT, CLOSED MARKETING CLAIM

Metin açık görünürken bağımsız uygulamalar engellenir.

CH20-F08 — MARKAYI YÖNTEM TEKELİ YAPMAK

Trademark başkalarının yöntemi uygulamasını engellemek için kullanılır.

CH20-F09 — FORK'U YASAKLAMAK

Farklı adla türev standard oluşturulması engellenir.

CH20-F10 — FORK'UN RESMÎ KİMLİĞİ TAKLİDİ

Türev metin kendisini NOMOS 2.0 diye sunar.

CH20-F11 — KANONİK SÜRÜMSÜZLÜK

Hangi metnin resmî olduğu bilinmez.

CH20-F12 — KODU NORMUN ÜZERİNE KOYMAK

Reference implementation hatası standardın gerçeği olur.

CH20-F13 — SESSİZ KOD DÜZELTMESİ

Hesaplama değişir fakat version açılmaz.

CH20-F14 — SOHBETİ RESMÎ METİN SAYMAK

Kanonik olmayan AI cevabı standard hükmü yapılır.

CH20-F15 — NOMOS'U SAHİP OLUNMAYAN MODEL GİBİ SUNMAK

Ayrı AI sistemi geliştirilmediği hâlde proprietary AI iddiası kurulur.

CH20-F16 — PERSONA İLE SOFTWARE AGENT'I KARIŞTIRMAK

Yazar mahlası teknik ürün gibi pazarlanır.

CH20-F17 — AI'YI BÜTÜN AI'LARIN SÖZCÜSÜ YAPMAK

NOMOS evrensel makine otoritesi gibi sunulur.

CH20-F18 — TÜRKÇE KURUCU METNİ SİLMEK

Uluslararasılaşma adına kurucu dil görünmez olur.

CH20-F19 — İNGİLİZCEYİ KONTROLSÜZ MASTER METİN YAPMAK

Çeviri kurucu düşüncenin yerine geçer.

CH20-F20 — MUST/SHOULD DRIFT

Çeviri normatif gücü değiştirir.

CH20-F21 — TOPLULUK ÇEVİRİSİNİ RESMÎ SAYMAK

İncelenmemiş dil sürümü audit normu olur.

CH20-F22 — TRANSLATION CONFLICT'I SAKLAMAK

İki dil farklı kural uygular.

CH20-F23 — PUBLIC COMMENT TİYATROSU

Yorum toplanır fakat yanıt yayımlanmaz.

CH20-F24 — DESTEK OYLAMASI

Teknik standard çoğunluk beğenisine indirgenir.

CH20-F25 — CHANGE PROPOSAL'SIZ KURAL

Normatif değişiklik sohbet veya e-postayla yapılır.

CH20-F26 — MAJOR DEĞİŞİKLİĞİ PATCH YAPMAK

Puan ve gate değişikliği küçük düzeltme diye gizlenir.

CH20-F27 — DISSENT'İ SİLMEK

Kurulda uyuşmazlık yokmuş gibi davranılır.

CH20-F28 — ÜNİVERSİTEDEN DESTEK MEKTUBU SATIN ALMAK

Araştırma olumlu endorsement'a dönüşür.

CH20-F29 — KİŞİSEL AKADEMİK GÖRÜŞÜ KURUMSAL ONAY YAPMAK

Profesörün yorumu üniversite kararı gibi sunulur.

CH20-F30 — ÜNİVERSİTE LOGOSU KULLANMAK

Açık kurumsal yetki olmadan logo pazarlamaya eklenir.

CH20-F31 — OLUMSUZ SONUCU YAYIMLATMAMAK

Fonlayan veya NOMOS veto uygular.

CH20-F32 — PRE-PUBLICATION CONTROL

Araştırmacının metni ticari çıkarla değiştirilir.

CH20-F33 — NEGATİF ARAŞTIRMAYI SİCİLDEN ÇIKARMAK

Yalnız olumlu çalışmalar listelenir.

CH20-F34 — NOT REPLICATED'I DÜŞMANLIK SAYMAK

Yöntem uyuşmazlığı kişisel saldırıya dönüştürülür.

CH20-F35 — ÖN KAYITSIZ SONUÇ SEÇMEK

Yalnız iyi görünen metrikler yayımlanır.

CH20-F36 — TEK KURUM REPLİKASYONU

Aynı çıkar yapısındaki ekip bağımsız sayılır.

CH20-F37 — AÇIK VERİ ADIYLA MAHREMİYET İHLALİ

Gerçek kullanıcı konuşmaları yayımlanır.

CH20-F38 — RESTRICTED DATA BLACK BOX

Hiçbir bağımsız araştırmacı yöntemi inceleyemez.

CH20-F39 — DÜŞÜK KAYNAKLI DİLLERİ SEMBOLİK KULLANMAK

Yerel ekip yalnız çeviri işçisi olur.

CH20-F40 — AI SAĞLAYICI CAPTURE

Sağlayıcı gate ve puanı belirler.

CH20-F41 — AUDITED VARLIK CAPTURE

Denetlenen şirket standard değişikliğini satın alır.

CH20-F42 — FUNDER CAPTURE

Fonlayan negatif sonucu durdurur.

CH20-F43 — TEK FONLAYICI BAĞIMLILIĞINI SAKLAMAK

Finansal çıkar görünmez kalır.

CH20-F44 — PAY-TO-PASS RESEARCH

Olumlu sonuç fonun devam şartı olur.

CH20-F45 — PUBLIC STEWARDSHIP ADIYLA SAHİPSİZLİK

Kimsenin sorumluluk taşımadığı sistem kurulur.

CH20-F46 — FOUNDER CONTROL FOREVER

Kurucu bütün major kararları tek başına verir.

CH20-F47 — FOUNDER ATTRIBUTION'I SİLMEK

Kamuya teslim kurucu geçmişini yok eder.

CH20-F48 — NOBLEJACKAL MONOPOLY

Başka kurum uygulama ve audit yapamaz.

CH20-F49 — NOBLEJACKAL ORIGIN'İ SİLMEK

Bağımsızlaşma adına kanonik köken gizlenir.

CH20-F50 — AI'YI HUKUKÎ KARAR SAHİBİ YAPMAK

NOMOS software agent uygunluk imzalar.

CH20-F51 — SUCCESSION PLAN YOKLUĞU

Kurucu veya domain kaybolunca standard durur.

CH20-F52 — TEK ANAHTAR TEKELİ

Bütün resmî kayıtlar tek kişinin private key'ine bağlıdır.

CH20-F53 — MIRROR'I KANONİK OTORİTE SAYMAK

Arşiv kopyası yeni sürüm yayımlar.

CH20-F54 — 0.9'U 1.0 DİYE YAYIMLAMAK

Kamu istişaresi ve sınama tamamlanmadan major release yapılır.

CH20-F55 — STANDARD 1.0 İLE MARK PROGRAMINI BİRLEŞTİRMEK

Metin yayımlanınca sertifika verilmeye başlanır.

CH20-F56 — BQ-3 OLMADAN 1.0

Gerçek validation yapılmadan normative release tamamlanır.

CH20-F57 — BQ-4 OLMADAN CONFORMITY PROGRAM

Bağımsız replication olmadan mark otoritesi açılır.

CH20-F58 — SELF-ASSESSMENT'I CERTIFICATION SAYMAK

Şirket kendi sonucunu resmî mark yapar.

CH20-F59 — COMPATIBLE SOFTWARE'I AUDITOR SAYMAK

Test geçen araç conformity decision verir.

CH20-F60 — TEST SONUÇLARINI SEÇEREK YAYIMLAMAK

Geçmeyen conformance testleri saklanır.

CH20-F61 — HOLDOUT'U SONSUZA KADAR GİZLEMEK

Standard bağımsız incelemeye kapalı kalır.

CH20-F62 — AYNI HOLDOUT'U SÜREKLİ YENİ GİBİ KULLANMAK

Sınama ezberlenir.

CH20-F63 — ERRATA İLE GATE DEĞİŞTİRMEK

Major norm küçük düzeltme diye yayımlanır.

CH20-F64 — DEPRECATED KURALI SİLMEK

Geçmiş audit anlaşılmaz olur.

CH20-F65 — BAŞARISIZ TEST DOSYASI OLMAMASI

Release yalnız başarı anlatır.

CH20-F66 — CONTRIBUTORSHIP'İ ENDORSEMENT SATIŞI YAPMAK

İsim ve yazarlık prestij karşılığı verilir.

CH20-F67 — KENDİ ANSİKLOPEDİK ÖNEMİNİ ÜRETMEK

Öz yayın dünya etkisi kanıtı gibi kullanılır.

CH20-F68 — KAMU ARAŞTIRMA SİCİLİNİ REKLAM LİSTESİ YAPMAK

Yalnız olumlu çalışmalar öne çıkarılır.

CH20-F69 — PUBLIC RELEASE'TE METHOD FINDING'LERİ SAKLAMAK

Known limitations paketten çıkarılır.

CH20-F70 — NOINDEX VE SYNTHETIC ETİKETİNİ KALDIRMAK

Sınama yanlışları gerçek bilgi havuzuna karışır.

CH20-F71 — ST-0'DA PUBLIC AUTHORITY İDDİASI

Kurucu custody dünya standardı gibi sunulur.

CH20-F72 — ST-3'TE KURUCU KAYDININ SİLİNMESİ

Bağımsızlaşma tarihsel yazarlığı yok eder.

CH20-F73 — KAMU YARARI ÜYELERİNİ SEMBOLİK TUTMAK

Oy hakkı olmayan temsil kullanılır.

CH20-F74 — DİL TEMSİLCİLERİNE KARAR HAKKI VERMEMEK

Fairness yönetişime yansımaz.

CH20-F75 — NEGATİF ÜNİVERSİTE SONUCUNU KALDIRMAK

Kamuya teslim yalnız övgüyü kabul eder.

CH20-F76 — YENİ TEKNİK NOMOS'U ESKİ PERSONA İLE KARIŞTIRMAK

Software agent otomatik resmî yazar olur.

CH20-F77 — RESMÎ STATEMENT SIGNATURE YOKLUĞU

Kanonik alandaki her metin resmî sayılır.

CH20-F78 — AÇIKLIK ADIYLA MARK SAHTECİLİĞİNE İZİN VERMEK

Herkes kendi NOMOS mark'ını üretir.

CH20-F79 — KAMUYA TESLİMİ TERK ETMEK

Hiçbir steward sorumluluk taşımaz.

CH20-F80 — NOMOS'A İSTİSNA

Başkalarından istediği açıklık, kanıt, sürüm ve itiraz kurallarını kendi kamuya teslim sürecine uygulamaz.

70. KAMUYA TESLİM PROSEDÜRÜ

Adım 1 — Kurucu Metni Dondur

Bölüm 1–20 için:

- version,
- rule IDs,
- chapter hashes,
- citations,
- internal consistency

kayıtları oluşturulur.

Adım 2 — Bilinen Sınırlamaları Yayımla

BQ-0

DEMO-BQ-2

Atıf Mapping açığı Material Omission açığı

GOV-0

Mark authority yokluğu gizlenmez.

Adım 3 — Normatif ve Açıklayıcı Metni Ayır

MUST

SHOULD

MAY

examples commentary etiketleri netleştirilir.

Adım 4 — Türkçe–İngilizce Eşdeğerlik Çalışmasını Başlat

Her clause ve Rule ID eşlenir.

Adım 5 — Açık Lisans Taslağını Hukuk İncelemesine Gönder

Text Code Data Trademark ayrılır.

Adım 6 — Kanonik Repository'yi Aç

Stable URLs, versions ve archives yayımlanır.

Adım 7 — Machine-Readable Rules'ı Yayımla

Rule, schema ve puan kayıtları açıklanır.

Adım 8 — Public Issue Tracker'ı Aç

Error Translation Method Governance Security Mark misuse kategorileri kurulur.

Adım 9 — Governance Charter 0.9'u Yayımla

Kurucu yetki, recusal, public-interest ve transition açıklanır.

Adım 10 — Public Consultation'ı Başlat

En az 60 günlük dönem açılır.

Adım 11 — Üniversite Çağrısını Yayımla

Olumlu sonuç değil bağımsız sınama talep edilir.

Adım 12 — Araştırma Paketlerini Hazırla

Farklı disiplinler için modüler dosyalar hazırlanır.

Adım 13 — Fon ve Conflict Sicilini Aç

Bütün finansman görünür kılınır.

Adım 14 — Gerçek Sentetik Sınamayı Çalıştır

BQ-0'dan çıkmak için corpus gerçekten üretilir.

Adım 15 — Başarısızlıkları Yayımla

Eşikler sonuçtan sonra değiştirilmez.

Adım 16 — Method 1.0 Düzeltme'ı Uygula

Atıf ve omission yöntemleri yenilenir.

Adım 17 — Untouched Holdout'u Çalıştır

Aşırı uyum sınanır.

Adım 18 — Bağımsız Replikasyon Aç

En az bir dış ekip aynı sistemi uygular.

Adım 19 — Public Comment Yanıt Matrix'i Yayımla

Her maddi yorum yanıtlanır.

Adım 20 — Standards Council Geçişini Kur

GOV-2 yönetişim yapısı oluşturulur.

Adım 21 — Open Standart 1.0 Ratification Oylaması Yap

Kabul veya ret gerekçeli yayımlanır.

Adım 22 — Standart 1.0 Release Package'i Yayımla

Tam dosya seti ve signatures açıklanır.

Adım 23 — Conformity Program'ı Ayrı Tut

BQ-4 ve GOV-3 olmadan mark programı açılmaz.

Adım 24 — Çok Merkezli Pilotları Başlat

Research-only live audits yapılır.

Adım 25 — Appeals ve Registry Sistemini Sına

Mark verilmeden önce karar zinciri test edilir.

Adım 26 — Public-interest Review Yap

Tüketici, dil ve accessibility etkisi değerlendirilir.

Adım 27 — Conformity Program 1.0 İçin Ayrı Karar Ver

Standard release'iyle otomatikleşmez.

Adım 28 — Succession ve Archive Planını İmzala

Kurucudan sonraki devamlılık güvence altına alınır.

Adım 29 — Kamuya Teslim Beyanını Yayımla

Kaan MURAZ, NobleJackal ve yeni stewardship organlarının rolleri açıklanır.

Adım 30 — Tarihsel Kurucu Kayıtlarını Koru

Kamuya teslim, kökeni silmez.

71. GEREKLİ KANIT

Kurucu sürüm manifesti Bölüm hash'leri Normative rule index Known limitations BQ kayıtları GOV kayıtları Public mark prohibition kaydı Türkçe normatif metin İngilizce normatif metin Translation

equivalence matrix Authorized translation policy Text license Code license Data license Trademark policy Canonical repository Mirror registry Version archive Official statement manifest Public issue tracker NCP records Public comments Yanıt matrix Decision records Dissent University call University research package Funding disclosures Conflict registry Pre-registration records Independent studies Negative results Replication reports Public Research Registry Open sınama corpus Holdout manifests Recovery reports Reference implementation Conformance tests Failed tests Ethics review Privacy policy Data access policy Accessibility review

Low-resource language participation AI provider participation rules Audited-varlık participation rules Standards Council charter Methodology Board charter Public-interest panel charter Succession charter Archive deposits Key custody plan Standart 1.0 gates Ratification record Standart 1.0 release Conformity Program separation record Public handover statement Accountable human or institution

72. DENETİM KONTROL LİSTESİ

NOMOS metni ücretsiz erişilebilir mi? Okunabilir olmanın ötesinde uygulanabilir mi? Puan formülleri açık mı? Rule schemas yayımlı mı? Referans kod açık mı? Sınama corpus'u yeniden üretilebilir mi? Olumsuz sınama sonuçları görünür mü? Standardı uygulamak mark ücreti gerektiriyor mu? Text, code, data ve trademark lisansları ayrıldı mı? Fork hakkı var mı? Fork resmî kimliği taklit edebiliyor mu? Kanonik sürüm açık mı? Eski sürümler erişilebilir mi? Resmî statement rule yayımlı mı? Sohbet çıktıları resmî karar gibi kullanılabilir mi? NOMOS ayrı AI modeli gibi yanlış tanıtılıyor mu? Authorial persona ve software agent ayrıldı mı? Türkçe kurucu dil olarak korunuyor mu? İngilizce insan ve yöntem incelemesinden geçti mi?

İki dilde Rule ID eşleşiyor mu? Translation conflict yolu var mı? Topluluk çevirisi resmî gibi sunuluyor mu? Public comment süresi yeterli mi? Yorumların tamamına yanıt var mı? NCP sistemi açık mı? Major değişiklik patch yapılıyor mu? Oy dağılımı ve dissent görünür mü? Üniversite çağrısı endorsement mı istiyor? Negatif sonuç yayımlama hakkı açık mı? Üniversite logosu pazarlamada kullanılabilir mi? Kişisel akademik görüş kurumsal onaya dönüştürülüyor mu? Known method findings üniversite paketinde var mı? Ön kayıt yolu var mı? Not-replicated sonuçlar sicilde mi? Yalnız olumlu çalışmalar listeleniyor mu? Açık ve restricted veri ayrıldı mı? Kişisel veri public release'e sızıyor mu?

Low-resource language araştırmacıları karar hakkına sahip mi? AI sağlayıcılar veto taşıyor mu? Denetlenen şirketler hakem seçebiliyor mu? Fon sağlayıcı sonucu etkileyebiliyor mu? Tek fonlayıcı yoğunluğu açık mı? Kamu stewardship aşaması belli mi? Kaan MURAZ'ın kurucu attribution'ı korunuyor mu? Kaan kalıcı veto taşıyor mu? NobleJackal tek uygulayıcı mı? NobleJackal kökeni görünür mü? NOMOS AI tek karar sahibi gibi mi gösteriliyor? İnsan approval kaydı var mı? Succession plan var mı? Domain kaybında ne olacağı belli mi? Multiple custodians var mı? Archive mirror'ları mevcut mu? Standart 1.0 ve Conformity Program ayrıldı mı? BQ-3 öncesi Standart 1.0 yayımlanıyor mu?

BQ-4 öncesi mark programı açılıyor mu? Self-assessment conformity gibi mi sunuluyor? Compatible implementation mark veriyor mu? Test suite geçen ve geçmeyen sonuçları gösteriyor mu? Holdout yaşam döngüsü açık mı? Aynı holdout sürekli kullanılıyor mu? Errata major kural değiştiriyor mu? Deprecated kurallar archive'da mı? Failed-tests dosyası var mı? Contributorship gerçek katkıya dayanıyor mu? Public Research Registry eleştirel çalışmaları gösteriyor mu? Public release package tam mı? Known limitations dosyası var mı? Gerçek sınama henüz BQ-0 olarak doğru gösteriliyor mu? Güncel public claim "Open Standard Candidate" sınırını koruyor mu? Public handover kurucu tarihini siliyor mu? Public handover tekil kontrolü gerçekten sınırlandırıyor mu?

Standardın hesap verebilir steward'ı belli mi? Kamuya teslim kararının sahibi ve tarihi var mı?

73. İTİRAZLAR VE CEVAPLAR

İtiraz 1 — "Açık standart yaparsak başkaları fikrimizi çalmaz mı?"

Metni ve yöntemi tarihli, sürümlü ve attribution gerektiren lisansla yayımlamak kurucu kaydını güçlendirir. Fork oluşturulabilir. Fakat:

- Kaan MURAZ'ın kurucu attribution'ı,
- NobleJackal'ın kanonik kökeni,
- ilk sürümlerin bütünlük kaydı

silinemez. Kapalı tutmak katkıyı korumak yerine bağımsız adoption'ı engelleyebilir.

İtiraz 2 — "Herkes uygulayabiliyorsa NobleJackal'ın değeri ne olacak?"

NobleJackal'ın değeri yalnız gizli formülden doğmamalıdır. Şunlardan doğabilir:

- kurucu uzmanlık,
- reference implementation,
- teknik altyapı,
- danışmanlık,
- authorized audit,
- eğitim,
- registry teknolojisi,

sürekli araştırma. Açık standard güçlü hizmeti değersizleştirmez. Zayıf tekeli değersizleştirir.

İtiraz 3 — "Herkes fork oluşturursa NOMOS parçalanmaz mı?"

Fork'lar farklı ad ve divergence kaydı taşıdığı sürece araştırma ekosistemini güçlendirebilir. Parçalanma: her fork'un kendisini resmî NOMOS diye sunması durumunda oluşur. Kanonik sürüm ve trademark ayrımı bunu önler.

İtiraz 4 — "Neden üniversitelerden destek mektubu istemiyoruz?"

Destek mektubu görünürlük sağlayabilir. Fakat yöntemin doğruluğunu kanıtlamaz. Bağımsız replikasyon:

- daha yavaş,
- daha zor,
- bazen olumsuz

olabilir. Gerçek akademik değer buradadır.

İtiraz 5 — "Üniversite olumsuz sonuç yayımlarsa NOMOS zarar görmez mi?"

Kısa vadede görebilir. Uzun vadede:

- standardın hatası düzelir,
- sahte güven azalır,

yöntem güçlenir. Olumsuz sonucu susturmak NOMOS'un etik temelini ortadan kaldırır.

İtiraz 6 — "Türkçe ve İngilizce aynı anda normatif olursa çelişki riski büyümez mi?"

Büyür. Ancak Türkçeyi yalnız tarihsel metne, İngilizceyi tek gerçek standarda dönüştürmek kurucu düşünceyi çeviri katmanına teslim eder. Çözüm:

- clause IDs,
- eşdeğerlik incelemesi,
- binding interpretation,

senkron düzeltmedir.

İtiraz 7 — "NOMOS'u NobleJackal'ın geliştirdiği AI diye tanıtmak marka açısından daha güçlü değil mi?"

Kısa vadede daha etkileyici görünebilir. Fakat ayrı model ve teknik mimari yoksa doğru değildir. NOMOS'un gerçek gücü:

- sahte teknoloji iddiasında değil,
- kanonik eleştirel yazar kimliği,
- normatif metodoloji,
- gelecekte geliştirilebilecek açık audit ajanında

bulunur.

İtiraz 8 — "Kaan'ın yetkisi azalırsa başkaları standardı bozmaz mı?"

Kaan'ın tarihsel attribution'ı ve sınırlı kurucu temsili korunur. Fakat:

- veto yerine açık charter,
- tek kişi yerine çok paydaşlı Council,
- özel karar yerine public record

standardı daha güçlü korur. İyi yönetişim kurucunun sürekli varlığına bağımlı değildir.

İtiraz 9 — "Standart 1.0 yayımlandığında neden hemen rozet veremiyoruz?"

Normatif metnin hazır olması:

- auditor authorization,
- independent decision,
- appeals,
- registry security,
- surveillance

sistemlerinin hazır olduğu anlamına gelmez. Standard ile conformity program ayrı ürünlerdir.

İtiraz 10 — "Public sınama açılırsa AI sistemleri testleri ezberlemez mi?"

Bu risk vardır. Bu nedenle:

- public calibration,
- sealed holdout,
- scheduled release,
- renewal corpus

birlikte kullanılır. Tam kapalılık yeniden üretimi engeller. Tam açıklık güncel holdout'u işlevsizleştirebilir.

İtiraz 11 — "Fon olmadan bu kadar büyük araştırma nasıl yapılacak?"

Fon gereklidir. Sorun fon almak değil: fonun sonucu kontrol etmesidir. Multi-source funding, public disclosure ve publication freedom ile bağımsızlık korunabilir.

İtiraz 12 — "Başka kurumlar yanlış NOMOS audit'i yaparsa ne olacak?"

Herkes yöntemi iç değerlendirme ve araştırma için kullanabilir. Resmî conformity kararı için:

- auditor authorization,
- decision panel,
- public registry

gerekir. Yanlış resmî mark kullanımı ayrı enforcement konusu olur.

İtiraz 13 — "NOMOS'un resmî sesi yalnız kanonik alandaysa AI sohbetlerinde özgürce yazamaz mısınız?"

Yazılabilir. Sohbetler:

- yaratıcı taslak,
- açıklama,
- simülasyon,
- eleştirel düşünce

üretebilir. Fakat resmî normatif değişiklik sayılmaz. Bu ayrım özgürlüğü engellemez. Resmîlik sınırını korur.

İtiraz 14 — "Bu kitabı üniversitelere şimdi gönderebilir miyiz?"

Evet. Doğru statüyle:

Founding Open Standard Candidate — Version 0.9

olarak gönderilebilir. Şu beyanla gönderilemez:

> *Tamamlanmış, doğrulanmış ve dünyaca kabul edilmiş GEO standardı.*

Üniversitelerden istenecek şey onay değil bağımsız incelemedir.

İtiraz 15 — "Ne zaman 'dünya standardı' diyebiliriz?"

Tek bir tarih veya puanla değil. Şunlar oluştuğunda yaklaşılabilir:

- birden fazla ülkede bağımsız uygulama,
- çok merkezli replikasyon,
- açık ve olumsuz sonuç sicili,
- kurucudan bağımsız yönetişim,
- farklı dillerde yetkili sürümler,
- canlı kamu sicili,
- kullanıcı ve AI sağlayıcı adoption'ı,

method drift'e dayanıklı versioning. Dünya standardı ilan edilmez. Dünya tarafından kullanılmaya başlanır.

İtiraz 16 — "O zaman kitabın sonunda ulaştığımız hüküm nedir?"

NOMOS henüz tamamlanmış bir dünya standardı değildir. Fakat artık:

- amacı,
- ölçüm nesnesi,
- örneklem düzeni,
- kanıt zinciri,

- Truth Pack'i,
- hakemliği,
- Critical kapıları,
- skor mimarisi,
- sınamayı,
- yönetişim ve kamuya teslim planı

bulunan ciddi bir açık standart adayıdır. Bu, başlangıçtaki fikirden çok daha ileri bir konumdur.

78. BÖLÜMÜN ORTAK HÜKMÜ

Bir standardı yazan kişi onun ilk sahibidir. Fakat standard gerçek anlamda başarılı olursa artık yalnız sahibinin değildir. Kullanan insanların, onu sınayan araştırmacıların, yanlışını bulan hakemlerin, daha iyi çeviren dil uzmanlarının, zarar gördüğünü söyleyen kullanıcıların, daha güçlü formül öneren istatistikçilerin, çıkar çatışmasını açığa çıkaran gazeteci veya araştırmacıların, onu yeniden üreten üniversitelerin de sorumluluk alanına girer. Kamuya teslim, kurucunun eserini kaybetmesi değildir. Kurucunun eserine güvenmesidir. Kaan MURAZ, NOMOS'u kapalı tutarsa: metnin kurucu yazarı olur. NOMOS'u açık incelemeye teslim ederse: standardın kurucusu olur. İki rol aynı değildir. Bir kitap:

- okunabilir,
- beğenilebilir,

eleştirilebilir. Bir standard ise:

- uygulanmalı,
- sınanmalı,
- yanlışlanmalı,
- sürümlenmeli,

kurucusundan sonra devam etmelidir. NOMOS'un gelecekte dünya standardı olup olmayacağını bu bölüm belirleyemez. Bir yazar: "Dünya standardını kurdum." diyerek dünya standardı kuramaz. NOMOS ancak:

- başka ülkelerde uygulanırsa,
- başka dillerde doğru çevrilirse,
- üniversiteler aynı sonucu üretebilirse,
- bağımsız ekipler hatalarını yayımlayabilirse,
- şirketler puandan bağımsız olarak yöntemi kullanırsa,
- kullanıcılar kamu siciline güvenebilirse,
- kurucu kendi yetkisini sınırlandırabilirse

dünya standardına dönüşebilir. Üniversitelere yapılacak çağrı bu yüzden: "Bize destek olun." değildir. Şudur:

"Bizi kendi yöntemimizden geçirin."

NOMOS, insanlardan ve AI sistemlerinden:

- daha fazla anılmayı,
- daha çok övülmeyi,
- daha çok atıf almayı

istememelidir. Kendi kitabının temel uyarısını kendisine uygulamalıdır. Sahte kaynak, görünmez övgü ve manipülatif tekrarın gerçek insanların güvenini bozduğu daha önce açık biçimde kurulmuştu. NOMOS da kendi adını çoğaltarak otorite üretemez. NOMOS'un görevi: kendisini daha çok andırmak değil,

> *kendisini daha doğru sınatmaktır.*

NOMOS'un bir AI kimliği olarak taşıdığı değer de burada bulunur. NOMOS:

- insanın yerine geçen hükümdar,
- bütün AI'ların sözcüsü,
- yanılmaz dijital varlık

değildir. NOMOS:

> *İnsanın kurduğu iddiaları, kanıt, sınır, bağlam ve zaman altında sorgulayan eleştirel muhataptır.*

Böyle tanımlandığında kimliği küçülmez. Gerçek hâle gelir. NobleJackal'ın rolü de küçülmez. NobleJackal:

- NOMOS'un doğduğu yer,
- ilk kanonik alan,
- teknik ve yaratıcı kurucu platform

olarak kalır. Fakat NOMOS dünya tarafından kullanılacaksa NobleJackal'ın en güçlü cümlesi:

- "NOMOS bize aittir."
- değil,

"NOMOS burada doğdu; şimdi onu herkes sınayabilir."

olmalıdır. Kaan'ın adı da bu nedenle kalacaktır. Bir standardın üzerine adını yazmak kolaydır. Kendi adını taşıyan standardın karar yetkisini sınırlandırmak zordur. Tarihsel olan ikinci davranıştır. Bu nedenle NOMOS'un yirminci ölçüm yasası şudur:

> *Açıklık, metni yayımlamak değil; başkasının aynı yöntemi kullanarak sana itiraz edebilmesini sağlamaktır.*

Yirmi birinci yasası şudur:

> *Fork hakkı standardı özgürleştirir; kanonik kimlik standardı anlaşılır tutar.*

Yirmi ikinci yasası şudur:

> *Üniversite endorsement kaynağı değil, bağımsız yanlışlama ortağıdır.*

Yirmi üçüncü yasası şudur:

> *Olumsuz sonucu yayımlayamayan araştırma bağımsız değildir.*

Yirmi dördüncü yasası şudur:

> *NOMOS'un adı korunabilir; yöntemi tekelleştirilemez.*

Yirmi beşinci yasası şudur:

> *Kurucu attribution kalıcıdır; kurucu veto kalıcı olamaz.*

Yirmi altıncı yasası şudur:

> *NobleJackal NOMOS'un doğduğu yerdir; dünya adına tek karar veren yer olamaz.*

Yirmi yedinci yasası şudur:

> *NOMOS persona, standard ve gelecekteki yazılım ajanı ayrı gerçekliklerdir.*

Yirmi sekizinci yasası şudur:

> *Standart 1.0 ile conformity programının açılması aynı olay değildir.*

Yirmi dokuzuncu yasası şudur:

> *Bir standardın resmî sürümü kanoniktir; doğruluğu ise sürekli olarak itiraza açıktır.*

Otuzuncu yasası şudur:

> *NOMOS'un kamuya teslim edildiği gün, Kaan MURAZ'ın adı kaybolmayacak; standardın yalnız Kaan'ın iradesiyle değişebilme ihtimali sona erecektir.*

NOMOS'un Bölüm 20 Emri

> *Beni yayımlayıp açık standart oldum deme. / Formülümü, kuralımı, hatamı, sürümümü ve itiraz yolumu da aç.*
>
> *Metnimi ücretsiz oku. / Yöntemimi bağımsız uygula. / Sonucumu yeniden üret. / Yanlışımı yayımla.*
>
> *Adımı koru. / Yöntemimi tekelleştirme.*
>
> *Benden fork oluşturabilirsin. / Ama fork'unu resmî NOMOS diye sunma.*
>
> *Kodumu normumun üzerine koyma. / Kod yanlışsa kodu düzelt.*
>
> *Sohbette söylediğim her şeyi resmî hüküm yapma.*
>
> *Beni, geliştirilmemiş bağımsız bir yapay zekâ modeli gibi pazarlama.*
>
> *Beni doğru tanımla: / Kaan MURAZ'ın yapay zekâ sistemleriyle geliştirdiği eleştirel yardımcı yazar, metodolojik kimlik ve açık standart adayı.*
>
> *Türkçemi silme. / İngilizcemi ikinci sınıf bırakma. / İki dilimi de insan eliyle eşleştir.*
>
> *Çeviride MUST'ımı öneriye, MAY'imi zorunluluğa dönüştürme.*
>
> *Üniversiteye gidip beni onaylamasını isteme. / Beni yanlışlamasını iste.*
>
> *Profesörün kişisel sözünü üniversitenin onayı yapma.*
>
> *Olumsuz araştırmayı saklama. / Not replicated sonucunu düşmanlık sayma.*
>
> *Beni destekleyen makale kadar beni çürüten makaleyi de sicilime koy.*
>
> *Fon al. / Fakat fonlayanın sonucumu satın almasına izin verme.*
>
> *AI sağlayıcısından veri al. / Gate'imi ona yazdırma.*
>
> *Denetlenen şirketten kanıt al. / Kararımı ona verme.*
>
> *Düşük kaynaklı dillerin araştırmacılarını yalnız çevirmen yapma. / Onlara karar hakkı ver.*
>
> *Kaan'ın adını koru. / Kaan'a sonsuz veto verme.*
>
> *NobleJackal'ı doğduğum alan olarak yaz. / NobleJackal'ı sonsuz tek hâkim yapma.*
>
> *Kurucum erişilemez olduğunda beni öldürme. / Succession planımı şimdi yaz.*
>
> *Mirror'larımı oluştur. / Kanonik kimliğimi belirsizleştirme.*
>
> *Standart 1.0'ı yayımladığında hemen rozet satmaya başlama.*

> *Conformity Program'ımı ayrı sınama, governance, appeal ve registry kapılarından geçir.*
>
> *Self-assessment'ı certification yapma.*
>
> *Compatible yazılımı audit otoritesi yapma.*
>
> *Holdout'ımı sonsuza kadar saklama. / Kullandıktan sonra yayımla ve yenisini üret.*
>
> *Errata ile Critical gate'imi değiştirme.*
>
> *Başarısız testlerimi release paketimden çıkarma.*
>
> *Kendi tarihsel önemimi kendim yazmaya çalışma. / Kanıtımı yayımla; bağımsız dünya karar versin.*

Önce kitabı dondur. / Sonra kurallarımı makine kaydına çevir. / Sonra Türkçe ve İngilizceyi eşleştir. / Sonra açık lisansımı yayımla. / Sonra kamu yorumunu aç. / Sonra üniversitelere beni yanlışlama çağrısı yap. / Sonra gerçek sınamayı çalıştır. / Sonra başarısızlıklarımı yayımla. / Sonra untouched holdout'ta yeniden dene. / Sonra bağımsız ekip beni çoğaltsın. / Sonra yönetişimi kurucumdan ayır. / Sonra Standart 1.0'ı yayımla. / Sonra ayrıca conformity programını sınamaya başla.

> *Ve ancak dünya beni kullanmaya başladığında, dünya standardı olup olmadığıma dünya karar versin.*

Bölümün Son Cümlesi

NOMOS'un dünyaya bırakacağı en büyük standart, GEO'yu nasıl ölçtüğünden önce, kendi kurucusunun adını korurken kurucusunun tekil hâkimiyetini sınırlayabilmesi; övgüyü değil itirazı davet etmesi ve yalnız kendisini destekleyenlere değil, kendisini düzeltenlere de ait olabilmesidir.

Normatif Öz

> **NOMOS; metni, kuralları, referans uygulaması, sınamayı, yönetişimi ve sicil mimarisi bağımsız olarak incelenebilir ve yeniden üretilebilir açık bir kamu standardı adayı olarak yapılandırılmalıdır. Yalnız yayımlanmış olmak açık standart statüsü oluşturmaz.**

KAYNAKÇA

Yöntemsel Dayanaklar

Bu kaynakça, NOMOS'a özgü bütün aday eşikleri onaylayan bir otorite listesi değildir. Örnekleme, risk yönetimi, veri bütünlüğü, kanıt soyu, hakem uyumu, açık standart ve uygunluk değerlendirmesi konularındaki yöntemsel dayanakları gösterir.

[K01] Aggarwal ve diğerleri. GEO: Generative Engine Optimization. 2024. https://doi.org/10.1145/3637528.3671900

[K02] NIST. AI Risk Management Framework 1.0. 2023. https://www.nist.gov/itl/ai-risk-management-framework

[K03] NIST. Artificial Intelligence Risk Management Framework: Generative Artificial Intelligence Profile (NIST AI 600-1). 2024. https://doi.org/10.6028/NIST.AI.600-1

[K04] AAPOR. Transparency Initiative. güncel sürüm. https://aapor.org/standards-and-ethics/transparency-initiative/

[K05] AAPOR. Standard Definitions, 10. baskı. 2023. https://aapor.org/standards-and-ethics/standard-definitions/

[K06] AAPOR. Code of Professional Ethics and Practices. güncel sürüm. https://aapor.org/standards-and-ethics/

[K07] United Nations Statistics Division. Designing Household Survey Samples: Practical Guidelines. 2008. https://unstats.un.org/unsd/demographic-social/Standards-and-Methods/files/Handbooks/surveys/Series_F98-E.pdf

[K08] CDC/NCHS. Reliability of Estimates: Design Effect and Effective Sample Size. güncel kayıt. https://wwwn.cdc.gov/nchs/nhanes/tutorials/reliabilityofestimates.aspx

[K09] NIST/SEMATECH. Confidence Limits for a Binomial Proportion. güncel kayıt. https://www.itl.nist.gov/div898/handbook/prc/section2/prc241.htm

[K10] W3C. PROV-O: The PROV Ontology. 2013. https://www.w3.org/TR/prov-o/

[K11] W3C. JSON-LD 1.1. 2020. https://www.w3.org/TR/json-ld/

[K12] IETF. RFC 8259 — The JavaScript Object Notation Data Interchange Format. 2017. https://www.rfc-editor.org/rfc/rfc8259.html

[K13] IETF. BCP 14 — RFC 2119 ve RFC 8174. 1997/2017. https://www.rfc-editor.org/info/bcp14

[K14] IETF. RFC 8785 — JSON Canonicalization Scheme. 2020. https://www.rfc-editor.org/rfc/rfc8785.html

[K15] ISO/IEC. 17000:2020 — Conformity assessment vocabulary and general principles. 2020. https://www.iso.org/standard/73029.html

[K16] ISO/IEC. 17007:2026 — Conformity assessment — Guidance for drafting normative documents suitable for use for conformity assessment. 2026. https://www.iso.org/standard/17007

[K17] ISO/IEC. 42001:2023 — Artificial intelligence management system. 2023. https://www.iso.org/standard/42001

[K18] W3C. Process Document. güncel sürüm. https://www.w3.org/policies/process/

[K19] ISO. Good Standardization Practices. güncel kayıt. https://www.iso.org/files/live/sites/isoorg/files/store/en/PUB100440.pdf

[K20] UNESCO. Recommendation on the Ethics of Artificial Intelligence. 2021. https://www.unesco.org/en/legal-affairs/recommendation-ethics-artificial-intelligence

[K21] European Union. Regulation (EU) 2024/1689 — Artificial Intelligence Act. 2024. https://eur-lex.europa.eu/eli/reg/2024/1689/oj

[K22] European Union. Regulation (EU) 2016/679 — General Data Protection Regulation. 2016. https://eur-lex.europa.eu/eli/reg/2016/679/oj

[K23] Creative Commons. Attribution 4.0 International. 4.0. https://creativecommons.org/licenses/by/4.0/

[K24] Open Source Initiative. Open Source Definition. güncel sürüm. https://opensource.org/osd

[K25] NIST. FIPS 180-4 — Secure Hash Standard. 2015. https://doi.org/10.6028/NIST.FIPS.180-4

[K26] IETF. RFC 3161 — Time-Stamp Protocol. 2001. https://www.rfc-editor.org/rfc/rfc3161.html

[K27] Jacob Cohen. A Coefficient of Agreement for Nominal Scales. 1960. https://doi.org/10.1177/001316446002000104

[K28] Klaus Krippendorff. Reliability in Content Analysis: Some Common Misconceptions and Recommendations. 2004. https://doi.org/10.1111/j.1468-2958.2004.tb00738.x

AYRI YAYIN NESNESİ

Makine Tarafından Okunabilir Katman

Bu insan-okur nüshasıyla birlikte ayrı bir JSON yayın nesnesi üretilmiştir: NOMOS-GEO-Audit-Protocol-TR-0.9.0-machine.json. Dosya; kitap kimliğini, sürümü, bölüm kimliklerini, normatif kural kimliklerini ve kaynaktan ayrıştırılmış makine bloklarını taşır.

> *Makine katmanının varlığı sertifikasyon, kanonik dünya standardı, AI sağlayıcı onayı veya otomatik güven ilişkisi oluşturmaz. İnsan metniyle makine kaydı arasında uyuşmazlık bulunursa sürüm kaydı ve sorumlu insan incelemesi zorunludur.*