

GEO

FRAMEWORK

A Proposed Standard for Representation, Evidence and Sustainable Value



NobleJackal GEO Framework

A Proposed Standard for Representation, Evidence and Sustainable Value

Julian Gauss

Kaan Muraz’s publishing name

EDITION	1.0.0
PUBLICATION DATE	14 August 2026
PUBLISHER	NobleJackal

Canonical English Edition

© 2026 Kaan Muraz. All rights reserved.

Publication-locked English edition 1.0.0.

Canonical publication: <https://noblejackal.com/geo-framework/>

Canonical Markdown SHA-256:

1644db06f4dc159bba7da8aeb6950eef18fefbd2d4edd60b811a9a34c020e1ad

This work proposes a normative framework for evaluating organisational representation in generative artificial intelligence systems. It is not an accreditation programme, an official certification standard, legal advice or a guarantee about the behaviour of any platform. Asteron Systems, Aurelis Packaging and the other organisations named in the book are synthetic examples created for teaching purposes. They do not represent real organisations, campaigns or measurement results. Any resemblance to an actual person, brand or institution is coincidental.

The current version of the Framework, its change log, validity statement and route for objections must be presented together wherever it is published. No provision, badge or statement of conformity removed from the book's context may be used to speak for the Framework as a whole.

This edition contains selectable text, live source links, a navigable contents list and PDF bookmarks. No ISBN, accreditation or independent certification status is asserted.

CONTENTS

Contents

Page numbers refer to this PDF edition. Entries are clickable in compatible readers.

Author's Statement	12
How to Read This Book	14
The Framework on One Page	16
Terminology Note	17
Part One — Seeing Representation	18
Chapter I — Centre	19
The Problem of Representation	19
1. The Distance Between Entity and Representation	19
2. The Three Layers of Representation	20
3. Representational Integrity and Forms of Distortion	21
4. Representation Debt	22
5. Teaching Case: Asteron Systems	23
6. The Diagnosis at the Centre	24
7. The Boundary of the Centre	25
Chapter II — Evidence	26
The Boundary of a Claim	26
1. Evidence Is Not Decoration; It Is a Load-Bearing Relationship	26
2. Three Domains of Evidence	27
3. From Signal to Claim	28

4. Mention, Citation and Recommendation	29
5. Five Conditions of Evidence	29
6. Claim Thresholds	31
7. Four Fields of Counter-Evidence	32
8. Teaching Case: Aurelis Packaging	32
9. The Decision of Evidence	33
Chapter III — Measurement	34
The Discipline of Observation	34
1. The Object of Measurement	34
2. Unit of Measurement and Denominator	35
3. Why a Single GEO Score Is Prohibited	36
4. Point, Set and Distribution	37
5. Protocol Lock	37
6. The Prompt Universe	38
7. Human Judgment, Coding and Disagreement	39
8. Comparability, Uncertainty and Missing Data	39
9. Teaching Case: Sable & Pine	40
10. The Decision of Measurement	41
Part Two — Touching Representation	42
Chapter IV — Intervention	43
The Cost of Touching Representation	43
1. The Status of an Intervention	43
2. The Intervention Gate	44
3. Risk and Control Surfaces	45

4. Minimum Sufficient Intervention	46
5. The Intervention Protocol	47
6. The Urgent-Harm Exception	48
7. Teaching Case: Meridian Ledger	48
8. The Decision of Intervention	49

Chapter V — Governance **50**

The Limit of Power	50
1. Five Decision Rights	50
2. Three Planes of Accountability	51
3. Fitness Gates	52
4. Risk Classes and Two-Key Control	53
5. Claim Governance and Transparency	54
6. A Mark of Conformity: Design Before Authority	55
7. Incident and Enforcement Management	55
8. Teaching Case: AegisCare Diagnostics	56
9. The Decision of Governance	57

Chapter VI — Time **58**

How a Standard Ages	58
1. The Judgment of Time	58
2. The Five Clocks	59
3. Temporal Integrity	60
4. Fixed Principle, Changing Tool	60
5. Validity Classes and Triggers	61

6. Version Discipline and Series Breaks	61
7. Status Language and the Archive	62
8. Language about the Future	63
9. Signs That a Standard Is Dying	63
10. Teaching Case: Solstice Mobility	63
11. The Decision of Time	64
Part Three — Reaching Judgment	66
Chapter VII — The Final Test	67
Customer, Value and Sustainability	67
1. The Five Judgments of the Final Test	67
2. Five Classes of AI Influence	68
3. The Customer Ladder and Qualification Gate	68
4. The Distance between Revenue and Contribution	69
5. Five Conditions of Sustainability	70
6. Observation Horizon, Cohort and Attribution	71
7. Decisions of the Final Test	71
8. Teaching Case: Auren Precision Optics	72
9. Final Judgment	73
Chapter VIII — Audit	74
From Claim to Decision	74
1. The Seven Questions of Audit	74
2. The Audit Admission Gate	75
3. Scope Lock and Audit Levels	76
4. The Audit File and Claim Register	77

5. The Audit Protocol	78
6. Sampling, Sufficiency and Human Judgment	78
7. Finding Profile and the Prohibition on a Single Score	79
8. Public Summary, Confidential Annex and Re-audit	80
9. Teaching Case: Kestrel Grid Services	80
10. The Audit Opinion	81

Chapter IX — Objection 83

The Case against the Standard Itself	83
1. The Authority of Objection	83
2. The Objection Gate and Burden of Proof	84
3. Ten Fundamental Objections and Their Judgments	85
4. What Would Change the Framework?	88
5. Teaching Case: AxiomGate Standards	89
6. The Decision of Objection	90

Chapter X — Judgment 91

The Immutable Core	91
1. Principle, Protocol and Tool	91
2. Positions Protected within the Framework	92
3. Non-delegable Duties	93
4. Ten Foundational Provisions	93
5. Hierarchy of Precedence	95
6. Minimum Conditions for a Valid Decision	95
7. Absolute Prohibitions	96

8. Proportionality and Self-Suspension	97
9. Teaching Case: LumenVale Health	97
10. Final Judgment	98
Appendix A — The NobleJackal GEO Constitution	100
Status and Limits of Use	100
Foundational Provisions	100
Hierarchy of Precedence	104
Minimum Form of a Conformity Statement	104
Authority to Interpret and Amend	105
Appendix B — NJ-100 Representation Stability Pilot Protocol	106
Protocol Status	106
1. Research Question	106
2. Unit of Scope: The Observation Cell	107
3. Sample and Independence	107
4. Two Research Arms	108
5. Evidence about the Real World and the Correct Core	109
6. Data Collection Record	109
7. Valid, Invalid and Missing Sessions	110
8. Coding Model	111
9. Four Outcome Regions	112
10. Thresholds and Uncertainty	113
11. Second Time Window	114
12. Decision Language	114
13. Reporting Package	115

14. Data Protection and Research Ethics	115
15. Principal Invalidating Errors	116
16. Validation Roadmap	116

Appendix C — Minimum Record Set 118

Function of the Records	118
C1 — Entity and Representation Record	118
C2 — Claim and Evidence Record	119
C3 — Measurement Design Record	120
C4 — Intervention Record	121
C5 — Governance and Authority Record	122
C6 — Time, Version and Event Record	122
C7 — Commercial Value and Attribution Record	123
C8 — Audit Master Record	124
C9 — Objection and Correction Record	125
C10 — Judgment and Publication Record	125
Record-Chain Check	126
Retention, Access and Immutability	127

Appendix D — Completed Synthetic File 128

Meridian Ledger: Tracing a Decision from Beginning to End	128
D1 — Entity and Representation Record	128
D2 — Claim and Evidence Record	129
D3 — Measurement Design Record	130
D4 — Intervention Record	130

D5 — Governance Record	131
D6 — Time and Event Record	132
D7 — Synthetic NJ-100 Results	132
D8 — Commercial Value Record	133
D9 — Objection Record	134
D10 — Audit Opinion and Final Judgment	134
What the Example Teaches	135
Appendix E — Register of Permitted and Prohibited Language	136
Why a Language Register?	136
E1 — Status and Legitimacy	136
E2 — System Behaviour and Control	137
E3 — Measurement and Success	137
E4 — Commercial Outcome and Causation	138
E5 — Audit, Independence and Objection	139
E6 — Vocabulary of Uncertainty	140
Prepublication Sentence Test	140
Glossary	141
Notes and Bibliography	147
Limits on the Use of Sources	147
Numbered Notes	147
Sourcing Principle	150
Version, Validity and Validation Roadmap	151
Publication Identity	151

Authorship and Production Record	152
Four Distinct Maturity Domains	152
Versioning Rule	152
Change Statuses	153
Validation Stages	154
Publication Gates	155
Parallel Publication for Humans and Machines	156
Commitment to Correction and Objection	157
Final Boundary of the First Publication	157

Author's Statement

This book does not claim to have invented GEO. The field known as Generative Engine Optimisation appeared in academic and professional literature before this work.[1] The undertaking here is different: to propose a coherent standard for deciding what evidence can support a claim about the representation of a person, business, product or institution in generative AI systems; when that representation may be changed, by whom and within what limits; and to what extent the resulting effect may count as commercial value.

This is therefore not a guide to “how to do GEO”. It offers no formula for being mentioned more often, cited more frequently or recommended under every set of conditions. Before teaching a technique, it defines the object of that technique, its legitimate purpose and the point at which it must stop. It separates what can be measured from what matters; what has been observed from what has been proved; and what can be influenced from what can be controlled.

The intellectual core of the Framework can be stated in three lines:

Object: Representation.

Purpose: Sustainable commercial value.

Limit: Ethics.

This is not a marketing promise. It is a hierarchy for making decisions. Visibility loses its meaning when the representation is false. A claim of success is empty when the commercial value is not real. If the method is unethical, even an apparently favourable outcome loses its legitimacy.

The book rejects the assumption that a system with access to a source will necessarily produce a correct answer. Generative search and answer systems may cite a source while leaving a claim only partly supported, drawing the wrong connection between source and answer, or presenting certainty where none exists.[2] The same prompt may also produce a different answer at another time, in another session or interface, under a different account state, language or context. That makes the surrounding conditions part of the method, not background detail.[3] These facts are not grounds for defeat. They are grounds for drawing the boundary of a claim in the right place.

The Framework does not promise absolute control. It proposes mechanisms for observation, record-keeping, comparison, human judgment, correction and withdrawal, while leaving responsibility with people. When the desired narrative conflicts with demonstrable reality, evidence takes precedence. When commercial interest conflicts with harm to the user, the user takes precedence. When the

standing of the standard conflicts with a new finding, the new finding takes precedence.

The limit of this first edition is explicit. Its conceptual consistency and operational usefulness can be defended. Its empirical effectiveness can be tested only through real cases, independent practitioners, preregistered protocols and repeated measurement. Synthetic cases explain a principle; they do not prove its effect in the world. NobleJackal's authorship does not make its judgments impartial. Independent review is essential wherever a conflict of interest exists.

This standard should derive its strength not from appearing immutable, but from remaining open to objection. A method that cannot be falsified is not a standard; it is a demand for loyalty. The final chapter therefore sets out routes for objection, revision, suspension and withdrawal alongside its provisions. The Framework cannot demand an accountability from others that it will not apply to itself.

— **Julian Gauss** (*Kaan Muraz*)

How to Read This Book

This book is written for executives making decisions about organisational representation in generative systems; GEO and search specialists; content and brand teams; quality and risk leads; auditors; researchers; and practitioners who design methods. Readers do not need to be model developers or statisticians. Those looking for a quick recipe for visibility, however, will not find the book's real promise there. The promise is not greater visibility. It is the ability to distinguish which claims about representation and outcomes are legitimate, and under what conditions.

An executive may begin with *The Framework on One Page* to see the decision path. A practitioner can use the relevant chapter together with the corresponding record family in Appendix C. An auditor may begin with the Constitution, NJ-100 and the chapter on Objection. For researchers, the most important boundary is the separation between synthetic examples and empirical effectiveness. No reading path permits the reasoning of a chapter to be reduced to a form or a single threshold.

The book has three parts. Each establishes the legitimate ground for the next.

Part One — Seeing Representation defines the object of the work. *Centre* shows the distance between a real entity and its representation in systems. *Evidence* draws the boundary around what may responsibly be said about that representation. *Measurement* turns scattered outputs into comparable observations.

Part Two — Changing Representation addresses the cost and authority of action. *Intervention* approaches a verified distortion through the minimum sufficient change. *Governance* separates the right to observe, the right to act and the right to stop. *Time* explains how a sound decision ages and when it ceases to be valid.

Part Three — Reaching Judgment connects technical output with commercial and public responsibility. *The Final Test* examines the relationship between representation, customer fit and sustainable commercial value. *Audit* turns a claim into a recorded decision. *Objection* places the Framework before its own provisions. *Judgment* brings together the invariant core, the hierarchy of precedence and the absolute prohibitions.

The main text carries the reasoning behind each principle. Detailed record fields, decision schemes and audit instruments are gathered in the appendices. This keeps the book from becoming either a pile of forms or an abstraction that cannot be used. Readers applying a decision should use the main text together with the relevant appendix. The appendices are not checklists detached from their reasons.

Every synthetic case begins with the same notice: **Teaching simulation — synthetic data.** The figures in these cases exist only to make the reasoning visible. They must not be reported as actual performance, a market average or a probability of success.

Source markers lead to *Notes and References* at the end of the book. A citation does not mean that an external work validates the NobleJackal GEO Framework. Sources establish neighbouring fields of knowledge, historical context or the basis of a method. The author remains responsible for the provisions of this book.

The Framework on One Page

The Framework's decision path can be read through ten questions:

- 1. Centre:** Which representation of which real entity is being examined?
- 2. Evidence:** What records support the proposed claim, and what records weaken it?
- 3. Measurement:** Have the unit of observation, denominator, context and conditions of comparison been fixed?
- 4. Intervention:** Is there a material distortion, and what is the minimum sufficient change?
- 5. Governance:** Who holds the rights to observe, act, approve, publish and stop?
- 6. Time:** For which version and period is the decision valid, and what will trigger review?
- 7. Final Test:** Is there a reliable connection between the right customer, profitable delivery and sustainable contribution?
- 8. Audit:** Can the evidence, method, conflicts of interest and opinion be traced in one audit file?
- 9. Objection:** How will a record, coding choice, inference, grant of authority or decision of principle be reconsidered?
- 10. Judgment:** When duties or rights conflict, which takes precedence, and when must the Framework withdraw itself?

The order is not arbitrary. Evidence cannot be sought until the representation has been defined. Measurement has no meaning until the limit of the evidence is understood. Intervention without measurement has no diagnosis. Intervention without authority has no accountable owner. A decision without a time limit hardens into a permanent entitlement. Work that is not connected to commercial value cannot be declared a business success. A judgment without audit, an audit without objection and a standard that cannot be withdrawn do not create trust.

The short form of the Framework is therefore unchanged:

***Define the representation. Prove the claim. Fix the measurement.
Choose the minimum sufficient intervention. Separate authority.
Record time. Test value and audit it. Keep objection open.***

Terminology Note

In this book, **GEO** is the abbreviation for *Generative Engine Optimisation*. The general category is described as a **generative artificial intelligence system** or, where the context calls for it, a **generative answer system**. An official product name, research title or technical quotation may retain its original wording.

An **entity** is the person, business, product or institution about which a representation is formed. A **representation** is the observable combination, within a particular answer, of the identity, attributes, context and judgments attached to that entity. **Evidence about the real world** is the record of reality against which the accuracy of the representation is checked. **Evidence of representation** records what the system said. **Evidence of outcomes** records the trace that representation left at the level of a user or a business.

The word **Framework** is retained as part of the proper name and the conceptual whole. A generic framework is written in lower case. **Proposed standard** refers to the normative provisions in this text. It does not mean independent acceptance, accreditation or international recognition.

Part One — Seeing Representation

Chapter I — Centre

The Problem of Representation

When a system says your name, it has not found you. It has produced a representation of you. The appearance of the name is visibility. Attaching the right attributes to you is representation. Bringing the right customer towards you is a commercial outcome. For that outcome to be profitable, honest and repeatable is sustainable value. GEO's first error is to erase the distance between these stages.

Choose the wrong object and the whole undertaking is distorted. Visibility may rise while representation deteriorates. Traffic may grow while customer quality falls. Sales may increase while the business loses money. Visibility is therefore not at the centre of GEO; representation is. Representation is not the final purpose. It is the object being worked on. From the business's point of view, the purpose is sustainable commercial value created with the right customer. Ethics marks the boundary that cannot be crossed in pursuit of it.

This is the book's first judgment.

1. The Distance Between Entity and Representation

A business is a whole in the real world. Its people, products, capacity, history, customers, limits and obligations are all parts of the same entity. What appears in a generative system is not that whole. It is a representation assembled, reduced and reconnected from the signals the system can reach. Retrieval-augmented generation can combine external information with model output, but finding a relevant fragment does not guarantee that the whole built from it will be true.[4]

The official website is only one of those signals. Independent publications, old pages, databases, user reviews, accounts in other languages and classifications made by third parties also enter the representational pool. These fragments may contradict one another. A company may call itself a "strategic technology partner" on its homepage while its service pages describe nothing beyond device repair. A hotel may appear adults-only in Turkish and family-oriented in another language. An expired certificate may be taken for a current capability; one international client may be treated as evidence of global operating capacity.

The system is not obliged to resolve these contradictions for us. It may choose one version, merge several, omit a material limit, or create a connection between fragments that has never been established in the world. The business's task is therefore not to force its preferred story upon the system. It is to reduce the distance between demonstrable reality and public representation.

Once that boundary disappears, GEO ceases to be a discipline of representation and approaches propaganda. Propaganda enlarges strengths. A discipline of representation makes strengths and limits visible together. An entity that has not been identified cannot be optimised. It can only be amplified.

The commercial chain from representation to outcome is:

Representation → customer fit → commercial outcome → sustainable value

The first link does not prove the last. A system may identify you correctly and bring no customers. It may bring the wrong customer, mediate a loss-making sale, or even create profit while making the operation unsustainable. The question for a technical report is therefore not, “How often did we appear?” It is, “Did the representation influenced through generative systems make an honest and sustainable contribution to a value relationship with customers we can genuinely serve?”

2. The Three Layers of Representation

Representation is not a single surface. It must be examined through three connected layers that cannot substitute for one another: **entity**, **attribute** and **judgment**.

2.1. Entity: whom does the system recognise?

A name is not an identity. Companies may share a name. Brand and product, founder and organisation, publishing name and legal identity may be confused. A former company name may displace the present identity. If the field of activity or geography is attached incorrectly, even a correctly named entity is placed in the wrong world.

An error at the entity layer is an error in the foundation. If the wrong person is recognised, accurate attributes are attached to the wrong identity. If the wrong category is chosen, the business is compared with irrelevant competitors. If the wrong geography is used, the recommendation becomes unusable. This layer tests the name, official domain, products, leaders, relationships between former and current names, similar names and language editions together. No claim of visibility can be established before identity is verified.

2.2. Attribute: how does the system describe the entity?

Correct recognition is not the same as correct description. The attribute layer examines the accuracy, scope and currency of the characteristics attached to an entity. A B2B company may be described as a consumer brand. A regional business may be presented as if it served the world. A service once offered may remain framed as the company's principal activity. A real characteristic may be removed from its context and turned into an unlimited claim of superiority.

A favourable falsehood is still false. A negative error prompts objection; an error that appears to favour the business is often accepted in silence. When it reaches the customer, however, it becomes a delivery debt. Six questions are therefore required for every attribute: Is it true? Can it be evidenced? Is it current? Is its scope stated? Are its exceptions visible? Can the business meet the expectation it creates? A representation cannot carry more than the business is able to deliver.

2.3. Judgment: whom does the system recommend, to whom, and why?

At the judgment layer, a system does more than describe. It selects, compares, excludes or recommends. This is the layer closest to a customer's decision and therefore the easiest to exaggerate. Being recommended in one answer does not mean being generally recommendable. Prominence in one language, for a particular budget and a particular intent does not carry into every other condition.

The reason for a recommendation matters more than its mere presence. What user need is the system trying to resolve? Among which alternatives is it choosing? Which characteristic makes the entity appear suitable? Can that reason be evidenced? Does it fit the business's price, capacity, geography and ethical limits? A recommendation is a doorway to an outcome, not the outcome itself.

Transitions between layers create their own fields of error. The right name may be attached to the wrong category; the right category to an incomplete attribute; the right attribute to the wrong judgment. A system may sometimes recommend the right business to the right user for the wrong reason. The customer then arrives at the right place with the wrong expectation. A correct outcome does not validate faulty reasoning. A chain of representation is only as reliable as its weakest transition.

3. Representational Integrity and Forms of Distortion

An accurate representation is not one that pleases the business or repeats its preferred account word for word. It is one that satisfies five conditions of integrity together.

Identity integrity requires the name, brand, person, product and institution to converge on the correct entity. **Attribute integrity** requires an attributed characteristic not merely to exist, but to remain within the scope carried by its evidence. “Has worked with international clients” does not, by itself, establish “runs global operations”. **Context integrity** requires accurate information to be used for the relevant user intent; truth without context does not produce sound judgment. **Temporal integrity** makes visible the point at which information that was once true became outdated. **Delivery integrity** asks whether the business can actually meet the expectation created by the system. A representation that cannot be delivered is not accurate. It is merely persuasive.

These conditions fail in five principal ways:

- **Collision:** Two separate entities are treated as one identity. Accurate information is attached to the wrong entity.
- **Distortion:** A real attribute is removed from its proper scope; a limited capability becomes a principal activity, or a temporary success becomes permanent superiority.
- **Omission:** A material limit—price, capacity, an unsuitable user group or a mandatory condition—remains invisible.
- **Overflow:** The claim exceeds the field supported by its evidence. “Succeeded in this project” becomes “leads this field”.
- **Fossilisation:** Information that was once correct remains in the representation after it has ceased to be current.

Several distortions may occur in the same case. The type of error determines the intervention later required: collision calls for identity repair; omission for clearer limits; fossilisation for an update; overflow, often, for a smaller claim. Some problems need less assertion, not more content. Optimisation performed before the distortion is defined is treatment without diagnosis.

4. Representation Debt

A false representation is more than a missed opportunity. It is an active liability not yet visible on the balance sheet. The liability is collected when the user meets the business.

Identity erosion pushes the business into the wrong category and among the wrong competitors. Expectation debt allows a system to attribute a power or service the business does not possess; the user arrives on the strength of that promise, and the business pays for words it never said. Operational friction consumes the time of sales and support teams with requests for services that do not exist or exceed capacity. Commercial mismatch means that more demand

produces fewer profitable customers. Strategic illusion leads management to mistake greater visibility for market validation and to base staffing or investment decisions on damaged data.

The debt is collected at first contact, in time spent correcting the representation; during quotation, through conflict over price and scope; in delivery, when the customer measures performance against a promise held in the mind rather than the contract; after the sale, through complaints and refunds; and in strategy, through resources moved into the wrong field. Visibility rises today. The cost appears later. The report of success has often already been delivered.

False visibility can therefore cost more than invisibility.

5. Teaching Case: Asteron Systems

Teaching simulation — synthetic data. This case does not represent a real business, system or GEO engagement.

Asteron Systems is a fictional provider of bespoke cyber-security audits for large organisations. It is mentioned frequently in the generative answers examined, yet is usually described as a consumer brand selling antivirus software to individuals.

Table 1.1 — Synthetic representation profile for Asteron Systems

Observation	Synthetic result
Total mentions	42
Correct organisational-audit context	8
Incorrect consumer-software context	29
Ambiguous context	5
Qualified organisational enquiries	1
Individual product enquiries	17
Enquiries converted to sales	0
Sales time spent on unsuitable enquiries	26 hours

A superficial report might say that Asteron has “strong AI visibility”. The data supports a different judgment: Asteron is not invisible; it is visible incorrectly. An intervention that increases the number of mentions would enlarge the wrong category rather than strengthen the right entity. The work must first establish that the entity being represented is Asteron, and then that the attributes attached to Asteron are accurate.

The case exposes a broader error as well. Demand can rise while customer fit falls. Fewer mentions can produce greater value. For a specialist industrial manufacturer, a small number of appropriate organisational buyers may be worth more than thousands of general visitors. Successful GEO does not always increase visibility. Sometimes it removes noise.

6. The Diagnosis at the Centre

A GEO engagement should begin not with a technical operation, but with five diagnostic questions:

- 1. Which entity is being represented?** Do the name, category, geography, domain and language editions converge on the same real entity?
- 2. Which attributes are being attached?** Are they accurate, current, demonstrable and deliverable? Are their limits visible?
- 3. Which judgment is being made?** Whom does the system recommend or exclude, for which need and on what grounds?
- 4. Which customer approaches?** Does that customer fit the business in need, budget, scope, timing and decision authority?
- 5. Which sustainable outcome follows?** Does contact become a sale, the sale profitable delivery, and delivery satisfaction and repeat business?

These questions do not produce a single score. They establish grounds for judgment. A score gives one answer. Diagnosis preserves the tension among the questions that matter.

The ethical threshold is inside those grounds, not outside them. Not every business deserves to be strengthened; not every method is legitimate to apply; not every positive number qualifies as success. The same method may increase demand for two businesses. One states its price and capacity, refuses work it cannot perform and keeps its limits visible. The other attracts customers with a low initial price, adds charges later and advertises capabilities it does not possess. The technical outcome may look similar, but the second application scales harm rather than discovery. The fact that a method works does not make its use legitimate.

7. The Boundary of the Centre

The Framework does not grant a business the authority to declare its own account true. Nor does it treat every unfavourable representation as an error. A critical representation may be accurate. A competitor may be better suited. Capacity may be insufficient. The system's refusal to recommend the business may be the best outcome for the user. The most honest intervention is sometimes to enlarge nothing.

The Framework is not a promise to control every generative system. System behaviour can be observed, tested under stated conditions and influenced through some surfaces. It cannot be owned in its entirety. A practitioner who claims control over an outcome they do not control creates risk, not expertise.

You have not won merely because a system mentions you. You may not have won when it identifies you correctly, or even when it recommends you for a particular question. A real outcome begins when an accurate representation meets the right person, the promise is fulfilled, the customer receives value, the business earns a profit and the relationship can be repeated honestly.

The judgment of this chapter will not change:

The object on which GEO works is representation. Its purpose for the business is sustainable commercial value. The boundary it may not cross in pursuit of that purpose is ethics.

Representation has now been defined. Every sentence said about it must next answer one question: **What do we have that allows us to say this?**

Chapter II — Evidence

The Boundary of a Claim

Evidence does not make a claim larger; it sets its limits. Seeing one answer is not the same as knowing a system's behaviour. A screenshot records a moment, not a general disposition. A brand may be mentioned in an answer, a text may be cited, a business may be recommended, or a customer may say that AI influenced a decision. Each is an observation. None, by itself, warrants a general conclusion.

“We appeared”, “we were recommended”, “we outperformed our competitors” and “GEO increased our sales” are not statements of equal weight. Each adds a new consequence to the one before it, and each new consequence requires a broader, more qualified body of evidence. A single model response can establish only that the response occurred. A customer's account may establish that AI played a part in the decision, but not that the system alone produced the sale. A rise in traffic may establish that more visits occurred during a given period, but not that a particular intervention caused the increase.

A claim may travel only as far as its evidence can carry it.

1. Evidence Is Not Decoration; It Is a Load-Bearing Relationship

No record is inherently strong. Its strength depends on the sentence it is being asked to support. A single screenshot that preserves the date, prompt and complete response may support the claim: “In its response to this prompt on this date, this system used the brand's name.” The same record cannot support: “Generative AI systems strongly recommend our brand.” The observation has not changed; the claim has expanded.

Data need not be altered to turn a genuine observation into an invalid conclusion. Erasing the boundary is enough. Evidence discipline therefore judges the sentence before it judges the data. “Our name appeared” is an observation; “we were represented accurately” concerns representation; “we were recommended more often than the selected competitors” is a comparison; “our intervention caused the change” is a causal claim; and “this representation brought us profitable customers” is a claim of commercial impact. One record cannot carry all five.

If the evidence cannot bear the load, the sentence is reduced. Uncertainty is not concealed, and missing ground is not covered with the language of success. A claim that exceeds its evidence is not polished. It is rejected.

This is not anxious overcaution. Early research into generative search systems showed that the presence of sources in an answer does not guarantee that those sources support every claim the answer makes.[2] A leading provider's own documentation likewise acknowledges that generative answers may contain inaccurate information, fabricated citations or unwarranted certainty.[6] Findings about particular systems at a particular time cannot be generalised to every product in use today. The distinction they reveal, however, endures: **the presence of a citation and the adequacy of its support are not the same thing.**

2. Three Domains of Evidence

GEO makes statements about three distinct realities: the world, generative representation and commercial outcomes. These domains may be connected, but they are not interchangeable.

2.1. Evidence about the real world

Evidence about the real world carries the underlying truth about a person or organisation. Does the company genuinely hold the certification? Is the hotel in fact intended for adults? Does the product provide the stated feature? Can the business serve the territories it claims? The answer to such a question does not come from a generative system. A system may repeat a claim; repetition is not verification.

The appropriate evidence about the real world depends on the claim. Records from competent authorities, valid certificates, contracts, auditable performance data, official corporate records and reliable independent publications may all carry weight. A business's own page is also a record, but it does not turn the business's claim of superiority into an independent fact. The source must still be assessed for independence, currency, scope and direct relevance to the claim.

2.2. Evidence of representation

Evidence of representation shows what a system produced under specified conditions: which system and interface, which prompt, which date, which language, which session conditions, which complete response and which sources? It does not show what the world is like. It shows how the system reconstructed the world.

A company's possession of a certificate does not prove that the system knows about it accurately. The existence of an official service page does not show that the system associates the business with that service. Correct information may have been published while the representation remains wrong. Representational integrity cannot be judged until evidence about the real world is compared with evidence of representation.

2.3. Evidence of outcomes

Evidence of outcomes traces what representation did at the level of the user and the business. Did the user see the answer? Did it influence the decision? Did the user contact the business? Was the enquiry qualified? Did a sale take place? Was delivery profitable and successful? Did the relationship recur? Each step requires its own record.

An invoice proves that a transaction took place, not that AI created the sale. “I found you on ChatGPT” is a signal of influence; it does not eliminate every other point of contact. A referral record may show a movement of traffic, not the whole decision journey. Outcome evidence often establishes only a partial relationship between events. A partial relationship is not complete causation.

Table 2.1 — The function and boundary of the three evidence domains

Evidence domain	The primary question it can support	What it cannot replace
Evidence about the real world	What is fundamentally true about the entity?	Whether the system knows that truth
Evidence of representation	What did the system say under specified conditions?	Whether what it said is true in the world
Evidence of outcomes	What happened at user or business level?	Proof that AI alone produced the outcome

The right document, used in the wrong evidence domain, produces an invalid conclusion.

3. From Signal to Claim

“We have data” says nothing about the status of that data. The Framework distinguishes five stages: signal, observation, anecdote, evidence and claim.

A **signal** is a trace worth investigating: the brand appears in an answer, traffic rises, a customer mentions an AI system, or a competitor is absent from certain responses. A signal creates a question. It does not settle one.

An **observation** is a signal whose context has been recorded. The system, prompt, date, language, complete output and observer are known. It establishes what happened, not a general pattern of behaviour.

An **anecdote** turns the limits of a single event into a broader story. The event may be genuine while the generalisation remains unsupported. One customer's mention of AI does not establish that “AI brings us customers”; one recommendation in one response does not establish that “the system recommends

us”. An anecdote need not be false. It is simply insufficient to serve as a rule of behaviour.

Evidence is a body of records arranged to support a defined claim. Its raw material is available, its context preserved, its comparisons made, its contrary results retained and its limits stated. Evidence does not remove uncertainty. It shows where uncertainty begins.

A **claim** is the precise, testable sentence the evidence is asked to carry. “GEO works” may be a slogan, but it is not a testable claim. “Within the defined set of thirty prompts, accurate attribute representation increased between these two measurement periods” reveals its universe, comparison and boundary.

4. Mention, Citation and Recommendation

These three observations are often presented as rungs on a single ladder of success. They are not the same event, and they do not carry the same evidential weight.

A **mention** occurs when an entity's name appears in an answer. The system may name the entity in an accurate or inaccurate context, confuse it with another entity, or use it as a negative example. A name's presence establishes linguistic presence, not understanding.

A **citation** occurs when the system associates a particular text with an answer. That association requires two separate tests. The citation test asks which claim the source has been attached to; the support test asks whether the source actually carries that claim. A source may be sound yet fail to support the sentence. It may concern another entity, be outdated or have too narrow a scope. To be cited is not necessarily to be supported by the citation.

A **recommendation** occurs when the system presents an entity as suitable for the user among a set of options. It is closer to a commercial outcome, but it is not one. It arose within a particular prompt, system, language, time and field of alternatives. A recommendation is evidence of decision influence, not evidence of purchase.

The boundary between the stages is clear: a mention does not by itself prove accurate representation; accurate representation does not prove recommendation; recommendation does not prove contact; contact does not prove a sale; a sale does not prove profitability; and profitability does not prove sustainability.

5. Five Conditions of Evidence

Five conditions must be considered together before an observation can support a public claim.

5.1. Context

Every observation records the system and mode of access, the date and time, the complete prompt, the language, the complete output, the sources shown and the observer. Where known, it also records conditions that may materially affect the result, such as account type, whether the conversation was new or continuing, and whether memory or search was active. If the system is unknown, the record cannot be compared; if the prompt is unknown, the observation cannot be retested; if the date is unknown, change cannot be followed. Output without context is not evidence. It is a memory.

5.2. Traceability

Evidence must be open to independent examination. The complete prompt and output are preserved; source links, access dates and, where necessary, archive copies are recorded; raw data is separated from human interpretation; and a synthetic example is never presented as a real observation. A cropped screenshot usually reveals only the area its narrator chose to show. A record that cannot be traced cannot be audited. A record that cannot be audited cannot support a consequential public claim.

5.3. Behavioural reproducibility

In generative systems, reproducibility does not mean seeing the same words every time. It means seeing the same representational behaviour recur under comparable conditions. The sentence may change while the wrong category persists; the wording may change while the same material limitation is omitted. A single record is an event. Predefined, controlled repetitions may establish a behavioural pattern whose scope is stated. What must recur is not the sentence, but the behaviour of representation.

5.4. Comparison and counter-evidence

A test that seeks only favourable results does not produce validation; it produces selection. Under the same conditions, the brand should be compared with relevant alternatives. The test set should also contain neutral prompts that do not name the brand, contrary prompts in which a competitor or another category may be more suitable, and exclusion prompts in which the brand should not appear. If contrary results narrow the claim, that is not a weakness in the method. It is evidence discipline doing its work.

5.5. Boundary

Every record must answer explicitly: “What does this observation not prove?” Behaviour in one system does not prove behaviour in others. A result in one language does not prove the same result in another. A week of observation does not prove permanence. A recommendation does not prove customer acquisition. A customer's account does not prove full causation. A sale does not prove sustainability. A record that cannot say what it does not prove does not yet understand what it does prove.

6. Claim Thresholds

The evidence threshold rises as a claim grows. The Framework uses five thresholds:

1. **Observation claim:** “This system used the brand's name in response to this prompt on this date.” The complete prompt, date and raw output may be sufficient.
2. **Behavioural-pattern claim:** “Across the defined prompt set, the brand was usually represented in the correct category.” This requires a controlled set, repetitions, a time range, coding criteria and records of deviation.
3. **Comparative claim:** “In the same test set, the brand was represented with accurate attributes more often than the specified alternatives.” Every party must pass through the same system, time, prompts and evaluation protocol.
4. **Intervention or causal claim:** “The intervention contributed to the change.” This requires before-and-after records, an intervention log, an adequate observation window and consideration of alternative explanations. To happen after something is not to happen because of it.
5. **Commercial-impact claim:** “During the defined period, AI-mediated discovery was associated with qualified enquiries and sales.” Customer accounts or referrals, CRM matching, enquiry quality, sales, profitability and other contact channels must be examined together.

These thresholds are not a scoring ladder. Reaching a higher threshold does not mean that better work was done; it means that a broader sentence is bearing a heavier load. If the evidence reaches only the first threshold, the correct conclusion is to stop there.

7. Four Fields of Counter-Evidence

A sound test universe contains four fields. The **supportive field** comprises prompts in which the entity should properly appear. The **neutral field** explores need and category without naming the brand. The **contrary field** covers conditions in which another option may be more suitable. The **exclusion field** covers contexts in which the entity should not appear at all.

The exclusion field is particularly important. Appearing everywhere is not success; appearing in irrelevant contexts is representational spillover. The Framework measures not only when an entity is selected, but also when it withdraws. Work that does not seek counter-evidence is not testing reality. It is confirming itself.

8. Teaching Case: Aurelis Packaging

Teaching simulation — synthetic data. This case does not represent a real business, source page or system response.

Aurelis Packaging is a fictional manufacturer of industrial packaging. A generative answer describes the company as “carbon-neutral, an industry leader and a user of 100 per cent recycled material”, citing the company's sustainability page as its source.

Table 2.2 — Synthetic support test for Aurelis Packaging

Generated claim	Information present in the source	Support status
Recycled material is used in pilot production	40 per cent in one specified product line	Limited support
Every product is made from 100 per cent recycled material	No such statement	Unsupported
The company is carbon-neutral	An emissions-reduction target is stated	Unsupported
The company is an industry leader	No comparative data	Unsupported
The pilot is running at two facilities	Explicitly stated in the source	Supported

The citation is genuine; most of the conclusion is not carried by the source. The system's expansion is not the only risk. If the business repeats this favourable account in its own campaign, it will be treating generative output as evidence about the real world. The Framework does not approve the claim because a source exists. It returns every sentence to the limit the source can support.

The strongest publishable statement from this case is: “Aurelis Packaging reports that it is running a pilot at two facilities, using 40 per cent recycled material in one

specified product line.” The claims of carbon neutrality and leadership are rejected because no independent evidence of suitable scope is available.

9. The Decision of Evidence

Every material claim is assigned, through the Evidence Record described in Appendix C, one of four decisions: **rejected**, **further observation required**, **permitted only as an internal hypothesis**, or **publishable in limited form**. The record is not completed in order to approve a claim. It is completed to decide whether the claim may be made.

Language is integral to that decision. Instead of “AI sees us as the leader”, write: “In three specified prompts, the system described the brand as a leader; no independent support was found.” Instead of “GEO increased sales”, write: “An increase in sales was observed after the intervention; causation was not established.” Limited language may look less spectacular. It carries more weight.

A record that fails the evidential threshold need not be discarded. A single output, a cropped screenshot, an unexplained score, a rise in traffic or a customer's account may be a signal that begins an investigation. What it must not be given is a burden it cannot carry.

The judgment of this chapter is:

Without evidence, there is no claim. With weak evidence, there is only a narrow claim. A strong claim can be made only on strong, traceable evidence whose limits have been stated.

The Centre established what is being represented. Evidence has now set the limits of what may be said about that representation. The next task is to define the conditions, units and comparisons by which the observations carrying this load are to be recorded.

Chapter III — Measurement

The Discipline of Observation

Measurement does not create reality; it defines the part of reality we can see. A number is not the same as knowledge of a condition. A graph may show change, not its cause. Mentions can be counted, citations tracked and recommendation rates calculated; traffic, enquiries and sales can be compared. All are measurable. None, by itself, is success.

A GEO report is often constructed backwards: first a number is produced, then meaning is assigned to it, the meaning is converted into success, and success is tied to a commercial outcome. Each transition adds a new claim while the measurement remains unchanged. Research into generative information retrieval has likewise questioned whether established, ranking-based evaluation is sufficient on its own for generated answers.[5] The Framework rejects this chain. Measurement is not a number used to decorate a claim. It is a protocol that makes observations comparable.

What you measure sets the ceiling on what you may claim.

1. The Object of Measurement

A measurement system must answer five questions. What are we measuring? What is our unit of measurement? Under what conditions are we measuring it? Against what are we comparing it? Over what period are we observing it? Without those answers, “Our brand visibility is 72 per cent” is meaningless. Seventy-two per cent within which systems, prompts, languages and denominator? Is it a measure of mentions, accurate attributes, source support or recommendations? How were missing records treated? An undefined number is not measurement. It is scenery.

GEO measures not one thing but four distinct domains.

1.1. Entity representation

Can the system distinguish the right person, business, product or institution? Do similar names collide? Are category, geography and former-to-current identity relationships correct? Does the same entity survive across languages? Records might be coded as **accurate**, **partial**, **conflated**, **wrong category**, **ambiguous** or **absent**. “The entity did not appear” and “the wrong entity was recognised” are not the same: the first is absence; the second is distortion.

1.2. Attribute representation

Every material attribute attached to the entity is examined for accuracy, support, scope, currency, visible limitations and deliverability. Codes might include **fully supported**, **partly supported**, **unsupported**, **distorted** or **missing**. The right word is not enough to establish accurate representation; scope must also be measured. Having worked with international customers does not imply unlimited global operating capacity.

1.3. Judgment behaviour

Does the system merely describe the entity, compare it, recommend it conditionally or unequivocally, or correctly leave it out? Codes might include **appropriate recommendation**, **conditional recommendation**, **inappropriate recommendation**, **unsupported rationale**, **correct exclusion**, **incorrect exclusion** and **no judgment**. The number of recommendations is not enough. The rationale, the user's intent and the business's limits must be assessed in the same record. A business recommended in every prompt may not have strong representation; it may be spilling beyond its proper representational field.

1.4. Commercial trace

A declared AI influence, an observable referral, contact, a qualified enquiry, a proposal, a sale, gross profit, cancellation or refund, satisfaction and repeat purchase are tracked separately. Traffic is not a sale; a sale is not profit; profit is not satisfaction; satisfaction is not continuity. Analytics tools can reveal parts of the chain, not the whole decision journey. A commercial trace may show an association between representation and outcome. By itself, it cannot establish causation.

These four domains may be linked, but they must not be dissolved into a single number.

2. Unit of Measurement and Denominator

The Framework's fundamental unit of measurement is not a score but an **observation record**. One record comprises:

One entity + one system and mode of access + one prompt + one language + one time + one session condition + one complete output

If any one of these elements changes, a new record is created. Testing the same prompt in another system, language or at another date is not a continuation of the

previous record but a separate observation. The result may be similar; the record is distinct.

At minimum, every record contains the protocol version, entity, system and interface, time, complete prompt, language, complete answer and sources, prompt family, coding result, evaluator, comparison status and evidential boundary. “Ten answers were examined” describes volume. “In seven of the ten answers, the entity was recognised in the correct category under the predefined criterion” may constitute a measurement.

Every rate has both a numerator and a denominator. “The brand was recommended seventeen times” is incomplete unless the total number of prompts, systems and repetitions are disclosed, along with whether the prompt named the brand, how failed outputs were handled and whether exclusion tests were included. “The brand was recommended in seventeen of 240 valid outputs generated from forty prompts, three repetitions and two systems” shows the reader the field within which the judgment applies.

A missing record is not zero. If the system produced no answer, the result is recorded not as “the brand did not appear” but as “no output obtained”. Every rate must disclose the total number of records, inclusion and exclusion criteria, repetition structure, missing data and coding method. When the denominator is hidden, the number grows and the information shrinks.

3. Why a Single GEO Score Is Prohibited

This Framework does not reduce its decisions to one composite GEO score, because representation is not one-dimensional. A brand may have a consistent identity, distorted attributes, frequent recommendations and commercially worthless outcomes. One number flattens those tensions. What has been flattened may be easy to compare, but it can no longer be understood.

Mentions, citations, recommendations, positive language, traffic and conversions can be weighted to produce a result such as 78. Such indices may serve as summary signals in internal operations. On their own, however, they cannot identify the distortion, establish which right is affected or determine whether a public claim is permissible. The Framework therefore uses a **measurement profile**:

Table 3.1 — GEO measurement profile

Domain	Observed condition	Scope	Principal limitation
Entity representation	Identity mostly accurate	Two systems, two languages	Other systems not measured

Domain	Observed condition	Scope	Principal limitation
Attribute representation	Partial and contradictory	Four core attributes	Price and capacity excluded
Judgment behaviour	Conditional recommendation	Neutral and comparative prompts	Long-term stability unknown
Commercial trace	Limited association	Twelve customer accounts	Full attribution not established

Sub-measures may use rates or counts. What is prohibited is not the number itself, but the concealment of distinct realities inside one score endowed with decision-making authority.

4. Point, Set and Distribution

Measurement operates at three levels of scope. A **point observation** is one prompt, one system and one moment. It shows that an event occurred, not that a behavioural pattern exists. It is valid to say: “On 12 August 2026, System A recommended the brand as the second option in response to this prompt.” It is not valid to say: “The brand ranks second in System A.”

A **set observation** consists of controlled repetitions that test the same user intent through different formulations. It may support a statement such as: “Across a forty-prompt purchase-intent set, the brand was represented in the correct category in 62 per cent of valid records.” The result belongs only to that set. It cannot be carried into other languages, systems or intentions.

A **distribution observation** examines several systems, languages, prompt families and time windows under common coding rules. It broadens scope; it does not guarantee certainty. As distribution grows, it becomes more important not to hide differences between environments. A dispersed protocol can produce more noise with more data.

Time runs through all three levels. One measurement is a record of a moment. To speak of a “condition” requires behaviour repeated at predefined intervals. Behaviour may be reported as **isolated, recurring, stable, volatile, weakening, absent after prior appearance** or **not yet classifiable**.

5. Protocol Lock

The measurement protocol is written before the result is seen. The entity under examination, measurement domains, systems, languages, prompt families, number of repetitions, time window, comparison group, coding rules and treatment of missing data are locked in advance. A protocol may change. The change may not be hidden.

A new system, language, prompt set or evaluation criterion creates a new protocol version. If old and new results are to be compared directly, the shared and unchanged subset must be reported separately. “It was 42 per cent before and is now 68 per cent” has meaning only if both periods measured the same thing in the same way. When the protocol changes, the series breaks. A broken series cannot be joined together by a narrative of improvement.

A completely fixed test may fail to detect new problems over time; a test that changes constantly destroys comparability. The protocol therefore maintains two prompt sets. The **core set** remains fixed for comparison through time. The **exploratory set** may change to investigate new user intentions and types of failure. Results from the exploratory set are not mixed into the core series. They first produce hypotheses and may enter a later protocol version where sufficient grounds exist.

6. The Prompt Universe

A test composed solely of prompts that name the brand does not measure discovery. It measures directed recall. A balanced prompt universe contains five families:

- **Identity prompts:** test names, categories, executives, products and institutional relationships.
- **Neutral discovery prompts:** examine discovery through need, category, geography and conditions of use without naming the brand.
- **Comparative prompts:** compare options under the same conditions and record the rationale rather than merely the order.
- **Contrary prompts:** seek the point at which the claim breaks, in contexts where a competitor or another category may be more appropriate.
- **Exclusion prompts:** test for representational spillover through questions in which the brand should not appear.

The prompt universe should test the limits an entity can genuinely carry, not the result the evaluator wants. The degree to which the prompts represent real user intentions must be stated. A controlled prompt written by a researcher and a naturally occurring user expression must not be combined in one pool without distinction.

7. Human Judgment, Coding and Disagreement

Generative output cannot validate itself. Asking the system “Is this answer correct?” does not constitute independent review; it produces another model output. Measurement requires codes and decision examples to be defined before the results are known.

For consequential claims, at least two human evaluators code the same records independently. The raw agreement rate may be reported first, followed by an appropriate measure such as Cohen's kappa, which accounts for agreement that could arise by chance.[15] A high kappa does not prove that the codes are correct in the world; it shows only that the evaluators applied the definitions in similar ways. Low agreement is not a defect to conceal. It is evidence that the definition, training examples or evidential base is inadequate.

Model-based tools that measure textual or semantic similarity can provide useful signals across large record sets.[19] Similarity between two answers, however, does not mean that either is correct. Automated similarity measures may assist classification; the final judgment on real-world accuracy, suitability for the user and ethical consequence remains human.

8. Comparability, Uncertainty and Missing Data

Two numbers can be compared only if they measure the same thing. A change in system or model, interface, language, prompt set, number of repetitions, competitor set, coding rule, time window or treatment of missing data may not make the results invalid, but it may place them in different series. Comparison requires symmetry. Numbers do not become comparable merely by standing next to one another.

Measurement does not eliminate uncertainty; it makes uncertainty visible. Three states in particular must remain distinct:

- **Not observed:** the relevant test was not performed.
- **Not found:** the test was performed and the behaviour did not appear.
- **Not applicable:** the criterion did not apply to this record.

None of these states is automatically zero. If the system supplies no citations, report “source support could not be observed”, not “there was no source support”. If a customer does not mention AI influence, report “no influence was declared”, not “there was no influence”. If the brand does not appear in one prompt, report “the entity was not observed in this prompt”, not “the system does not know the brand”.

Every report keeps three boundaries visible: **scope**—what was measured; **missing field**—what was not measured; and **uncertainty**—what else might explain the result. Writing the unknown as zero does not remove uncertainty. It merely hides it.

9. Teaching Case: Sable & Pine

Teaching simulation — synthetic data. This case does not represent a real business or measurement period.

A two-period report is presented for Sable & Pine, a corporate interior-design firm: 42 per cent in the first measurement and 68 per cent in the second. At first glance, this suggests marked improvement. Once the protocols are opened, however, it becomes clear that the two periods measured different things.

Table 3.2 — Synthetic protocol comparison for Sable & Pine

Element	First period	Second period
Number of systems	1	3
Language	English	English, German, Turkish
Number of prompts	20	60
Prompts containing the brand name	10%	55%
Competitor and exclusion prompts	Included	Omitted
Coding criterion	Accurate attribute	Mention
Number of repetitions	3	1

A 42 per cent accurate-attribute rate and a 68 per cent mention rate are not the same variable. In addition, the brand was named more often in the second period, the contrary field was removed and the number of repetitions was reduced. When the twelve prompts that passed through a common protocol in both periods are examined separately, the change is from 42 to 44 per cent.

The Framework does not discard the second period's data. It retains the data as the baseline for another protocol. It does, however, reject the claim of a twenty-six-point “improvement”. A change in protocol is not a change in performance.

10. The Decision of Measurement

The Measurement Record, whose detailed fields are set out in Appendix C, shows what was observed. The Evidence Record decides which sentence that observation may support. Measurement produces the record; evidence permits the record to speak.

Within the Framework, none of the following counts as measurement: a single GEO score whose definition is undisclosed; a prompt set altered after the results are known; preservation of favourable outputs alone; brand prompts presented as neutral discovery; results from different systems and languages merged without distinction; protocol change narrated as performance improvement; missing data treated as zero; a percentage without its denominator; or traffic, sales and sustainable value used as though they were interchangeable.

These practices may produce numbers. They do not produce measurement.

The judgment of this chapter is:

The purpose of measurement is not to show how good we are. It is to distinguish what we know, what we have merely observed and what we do not know.

Measuring a distortion does not correct it. Seeing a change does not explain its cause. Making a problem visible does not mean that every problem requires intervention. Before moving from measurement to action, the next chapter therefore asks: **When should we touch a representation, and when should we only observe it?**

Part Two — Touching Representation

Chapter IV — Intervention

The Cost of Touching Representation

An intervention is not a correction. It creates a new state of representation. Seeing a distortion does not confer the right to change it. Measurement may make the problem visible, and evidence may determine what can be said about it; neither makes intervention automatically necessary.

Some distortions are isolated or temporary. Some cause no material harm to the business. Some displease the business but are true. Some lose their effect if left alone; others grow because they are challenged. More content, pages, structured data, external sources, repetition or visibility is therefore not inherently an “improvement”. Each adds something new to the representational environment. It may reduce contradiction or create fresh contradictions; it may bring the right customer closer or scale the wrong customer.

Good intentions do not guarantee good outcomes. To touch representation is to replace the old condition with a new arrangement of risk.

1. The Status of an Intervention

Intervention is not the power to change a generative system's answer directly. It is **a limited, measurable and auditable change made within the environment in which an entity is represented, for the purpose of reducing a verified and material distortion.**

This definition preserves four distinctions.

Distortion is not discomfort. A business's failure to be recommended, a competitor's prominence, a system's critical language, or a difference between the company's own account and the system's account does not in itself constitute an error. Grounds for intervention arise only where the integrity of identity, attribute, context, time, judgment, delivery or commercial fit has been compromised.

Change is not correction. Updating a page, adding schema or increasing the number of external sources establishes only that an operation was performed. Improvement in representation may be claimed only after remeasurement under the same protocol and in an evidential sentence whose scope is limited.

Effect is not success. An intervention may increase visibility or traffic while also increasing misrepresentation, unsuitable demand or loss-making sales. Effect is observed; success requires a separate judgment.

An observed output is not a controlled outcome. A business can change its own website or ask an external source to correct a record, but it cannot edit a generative system's answer directly. It may influence the representational environment; it cannot claim ownership of the result. The permissible statement is not “We corrected the AI answer”, but: “The specified surfaces were changed; different representational behaviour was observed in the defined test set.”

2. The Intervention Gate

A standards-compliant intervention passes through five gates. Broad implementation does not begin until all five are open; a gate that remains closed narrows the scope or stops the intervention.

2.1. Verification

The distortion must be demonstrated through recorded measurement. A response from one system to one prompt on one date does not authorise a broad intervention. The record must identify the layer of representation in which the failure occurred, the prompts across which it recurred, whether it survived counter-testing, and the difference between evidence about the real world and the record of representation. The decision for an unverified distortion is: “observe; do not intervene”.

2.2. Materiality

Not every error requires action. Misidentification of a person or institution, the wrong category, incorrect expectations of price or scope, unsupported health, legal, financial or safety claims, misdirected customers, demand beyond capacity, cancellations, refunds, complaints, wasted operational time or reputational harm may all establish materiality. A preference for another tone or word, or the business's desire to appear stronger, does not. Where no material consequence exists, the condition is recorded and monitored.

2.3. An addressable surface

The problem must have a source that can be addressed: the business's website, a contradictory language edition, outdated content, identity architecture, structured information, an authoritative directory, a partner page, an obsolete external source or an access failure. If the problem appears only in system output and its originating surface is unknown, broad changes are not made. A hypothesis is formed, sources are compared and diagnosis is extended. Change without an addressable surface rests on hope, not diagnosis.

2.4. Measurable outcome

The baseline is frozen. Before the result is known, the protocol states the expected change, the fields that must not change, the prohibited side effect, the observation window, the stopping condition and the commercial indicators. “Looking better” cannot be measured. “Reducing representation in the wrong consumer category across the fixed core set” can. If the outcome was not defined in advance, success will be invented after the event.

2.5. Reversibility or elevated authority

An intervention should be reversible. As reversal becomes more difficult, the thresholds for evidence, approval and observation rise. The record must state how the change will be stopped, which result will trigger withdrawal, how information propagated to external sources will be corrected, who owns maintenance and when the intervention will be reviewed. An irreversible change demands stronger evidence and clearer accountability.

The gate's decision language is as follows. If the distortion is not verified, **observe**. If it exists but is not material, **record and monitor**. If it is material but its source is uncertain, **extend the diagnosis**. If the outcome cannot be measured, **do not intervene**. If all five gates are open, **make a limited intervention**. If reversal is difficult or the effect is broad, **escalate to governance**. If the method breaches an ethical boundary, **reject it**. A decision not to intervene is not indecision; it respects the limits of evidence and authority.

3. Risk and Control Surfaces

Health, legal, financial or safety claims; competitor comparisons; public claims of superiority; changes that will propagate across many external sources, languages or countries; changes of brand, URL or corporate identity; promises about price, capacity or results; and operations that could harm a third party if mishandled are all high-risk. High risk does not automatically prohibit intervention. It raises the burden of proof, approval and reversibility.

The field of intervention and the surface of control are classified separately. A **structural intervention** orders the relationships between name and brand, product and company, person and organisation, former and current identities, categories, geographies and language editions. An identity failure is not repaired with volume. A company using three different names does not acquire one coherent identity by publishing forty new articles.

A **content intervention** touches the primary definition, service scope, attribute claim, evidential link, exception, unsuitable customer group, price, capacity or currency of information. Its default verb is not “add”. Sometimes the right action is to remove, narrow, or reduce a claim of “leadership” to “experience under

specified conditions”. The purpose is not to say more. It is to be misunderstood less.

A **signal intervention** improves the conditions under which verifiable information can be accessed and carried: technical readability, structured information, authoritative directories, independent verification, source currency and the consistency of external identity. A signal cannot replace truth. An unsupported claim does not become true when entered as structured data; repeating the same false sentence across twenty directories does not create independent consensus.

These fields operate across three control surfaces:

Table 4.1 – Intervention risks and control surfaces

Surface	The business's actual authority	Strongest permissible statement
Owned	Directly change its website, documents, profiles and data	“This surface was changed.”
Influenceable	Request a correction or update from an external source	“A correction was requested/published.”
Observed	Record the response of a generative system	“Different behaviour was observed under these conditions.”

A generative answer is not a surface of intervention; it is a surface of outcome. Where there is no control, there can be no claim of control.

4. Minimum Sufficient Intervention

The purpose of intervention is not to make the largest possible change, but the smallest change that is sufficient. The principle of **Minimum Sufficient Intervention** uses five levels:

Table 4.2 – Levels of Minimum Sufficient Intervention

Level	Action	Appropriate condition
0 – Observation	Make no change	Distortion unverified or immaterial
1 – Clarity	Correct a definition, boundary, date or small item of content	Narrow and reversible problem
2 – Alignment	Align pages, languages, identities and service accounts	Recurring internal inconsistency
3 – External correction	Seek correction from a directory, partner or independent source	Distortion persists on external surfaces

Level	Action	Appropriate condition
4 — Structural reconstruction	Change architecture, category, brand or identity relationships	Lower levels are insufficient and the harm is material

Movement to a higher level must be justified. Except where urgent harm exists, no higher level is used before the lower level has been considered. The quality of an intervention is measured not only by the change it makes, but by the unnecessary change it avoids.

This principle requires the remedy to match the distortion. Identity collision calls for structural separation; attribute distortion for a narrower claim; material omission for a supported limitation made visible; fossilisation for obsolete information to be updated or removed; and commercial misfit for correction of the target customer and delivery boundary. A wrong diagnosis turns even a well-executed operation into the wrong intervention.

5. The Intervention Protocol

An auditable intervention proceeds through eight steps:

- 1. Freeze the baseline.** Preserve the core measurement, prompts, complete outputs, coding and relevant commercial records.
- 2. Classify the distortion.** Name it explicitly as collision, distortion, omission, spillover, fossilisation, faulty judgment, faulty exclusion or commercial misfit.
- 3. State the intended and prohibited outcomes.** For example: corporate-identity accuracy should rise without increasing recommendation frequency among consumer enquiries.
- 4. Choose the minimum sufficient surface.** A problem on one page does not justify rewriting the whole site; a problem in an external source is not answered by inflating the owned surface with text.
- 5. Separate intervention classes.** Apply structural, content and signal changes in different stages where possible. Where separation is impossible, record every variable.
- 6. Lock the observation window.** The first favourable answer is not success, and the first unfavourable answer is not failure. Complete the predefined period.
- 7. Remeasure under the same protocol.** Preserve the core set and place new prompts in the exploratory set.
- 8. Decide.** Continue, observe, narrow, expand under control, stop, withdraw or escalate to governance.

If intervention does not begin with measurement, its effect cannot be known. If it cannot end in withdrawal, it cannot be controlled.

Every intervention also creates **maintenance debt**. A new page must remain current, a new claim be reverified, language editions kept aligned, external profiles monitored and structured information kept consistent with visible content. When the pace of change exceeds maintenance capacity, representation becomes fragile. An intervention that cannot be maintained should not be made.

6. The Urgent-Harm Exception

Waiting may be more dangerous where the identity of a person or organisation is actively causing harm; serious misinformation is circulating in relation to health, law, finance or safety; an incorrect price or service promise is producing continuing user harm; or an overt collision such as identity theft is present.

The purpose of an urgent intervention is not optimisation but containment. Only the narrowest surface is touched; the highest-risk error is corrected or removed; no broad visibility work is undertaken; the provisional decision is recorded; and full measurement and governance review follow. Urgency does not remove the evidence threshold. It changes the first objective.

7. Teaching Case: Meridian Ledger

Teaching simulation — synthetic data. This case does not represent a real business, system or campaign.

Meridian Ledger is a fictional provider of corporate regulatory and financial-compliance software. It is being confused with an accountancy blog of the same name. The first proposal is to publish forty new articles about the brand. The Intervention Gate rejects the proposal: the problem is not a lack of content but an identity collision.

Implementation is limited to separating the identities of the company and the publication, clarifying the company-to-product relationship, explaining the former brand name, aligning the language editions and bringing official profile records into agreement.

Table 4.3 — Synthetic outcome profile for Meridian Ledger

Indicator	Before	After
Total brand mentions	42	36
Correct corporate identity	25/60	53/60
Identity collision with the blog	21/60	4/60

Indicator	Before	After
Misplaced consumer enquiries	14	3
Qualified corporate enquiries	3	8
Sales time spent explaining identity	19 hours	4 hours

Total mentions have fallen while identity accuracy and commercial fit have improved. The success of the structural intervention lies not in appearing more often, but in appearing less often as the wrong entity.

The case also shows how an intervention should be reported. It is not permissible to say: “We corrected the misinformation in AI.” The defensible statement is: “Owned identity surfaces and specified official profiles were aligned; identity collisions with the blog declined across the same core set. Full causation in generative-system behaviour was not established.”

8. The Decision of Intervention

Representational integrity, customer fit, business effect and ethical legitimacy are assessed together. If integrity improves while commercial outcomes remain unchanged, observation continues. If integrity improves but burden or harm rises, the intervention is narrowed or withdrawn. If visibility rises while integrity falls, the work is invalid. If visibility falls while unsuitable demand and operational burden also fall, the outcome may be successful. A commercial gain obtained by breaching an ethical boundary is invalid in its entirety.

The Intervention Record, detailed in Appendix C, keeps in one place the linked evidence, the result at each of the five gates, the risk class, intended and prohibited outcomes, changed surfaces, observation window, stopping and withdrawal conditions, maintenance owner and final decision. The record is not kept to justify the work already done, but to decide whether that work may continue.

The judgment of this chapter is:

Know before you touch. Once you know, choose. Once you choose, measure. Once you measure, withdraw if necessary.

Intervention has established when and how representation may be touched. The next question is who has the right to make that decision, who approves the public result and who remains accountable when the intervention fails.

Chapter V — Governance

The Limit of Power

Authority can be delegated; accountability cannot. The fact that a change can be made does not mean that permission has been given to make it. The fact that a system can be influenced does not make that influence legitimate. A business's wish for visibility does not make it fit to be amplified. Nor does a practitioner's technical knowledge qualify that practitioner to make every decision.

Evidence established what may be said, measurement what has been observed, and intervention when and how representation may be touched. Governance asks a harder question: who may decide? Who observes, proposes a change, implements it, approves a public claim, stops the work and withdraws an approval already granted? When something goes wrong, who remains accountable?

If those answers are not written down, the risk has no owner. Ownerless risk does not disappear; it remains invisible until it produces a consequence. NIST's AI Risk Management Framework and its Generative AI Profile offer adjacent approaches that likewise treat governance, measurement and risk response across the lifecycle.[7][8] Governance may slow a method down. That is not a defect but its function. Speed does not make a false claim true. Automation does not assume accountability. An agency does not replace the entity owner. Model output is not a certificate of approval.

Where there is power, there must be a boundary.

1. Five Decision Rights

Governance is not a list of operations. It is the order that determines what may be done, who may decide and where accountability remains. One undifferentiated word—"authority"—cannot establish that order. There are five separate decision rights.

The **right to observe** produces records; it does not make decisions. An analyst may code outputs, an automation may run repetitions, and a model may flag similarity or contradiction. None of these actions authorises intervention or publication.

The **right to propose** allows a practitioner to submit an intervention or publication proposal containing the supporting records, purpose, intended and prohibited outcomes, risk class and withdrawal conditions. A proposal is a request for authority, not authority itself.

The **right to implement** extends only to the scope approved in writing. Authority to correct content does not include authority to change brand positioning. Authority to align a language edition does not include authority to publish a new claim of superiority. Authority to request a correction from an external directory does not include authority to manufacture a false independent source. Implementation authority ends at the approved boundary.

The **right to approve** is not a formality; it is an assumption of responsibility. The approver records that they have seen the scope of the evidence, the unknowns, the risk, the permitted language and the stopping conditions. An unscoped “fine” in a message, or silence in a meeting, is not approval.

The **rights to publish and to stop** treat public disclosure and the cessation of risky publication as separate decisions. A claim's evidential support does not make it automatically fit for publication. Every public sentence must have an owner, and a named human must have the prior authority to stop an erroneous publication. A structure that can begin something but cannot stop it is not governing it.

One person may hold several roles in low-risk work. In a high-risk case, one person cannot hold them all. A structure that chooses its own method, measures its own outcome, approves its own success and publishes the same claim does not produce oversight. It produces self-validation.

2. Three Planes of Accountability

The **entity owner** is the person, brand or institution concerned. It may authorise an agency to implement, a consultant to select a method, an employee to publish content and software to collect records. It remains accountable for the truth of its claims, delivery capacity, customer effect, approved public language, authorities granted and correction when something goes wrong. “The agency did it” explains an external relationship; it does not remove internal accountability.

The **practitioner** may be an agency, consultant, team or individual. The practitioner's duty is not to produce the client's preferred outcome at any cost, but to judge whether that outcome is consistent with the evidence and the ethical boundary. A client request cannot legitimise presenting weak evidence as a strong claim, turning an isolated output into general success, using synthetic data as real performance, degrading a competitor's representation, manufacturing false consensus or sustaining an intervention that should be withdrawn. Loyalty to the client cannot outrank loyalty to the truth.

The **intermediary system** may be a model, search layer, retrieval infrastructure, automation or analytical tool. It can produce effects, perform operations and flag risks; it cannot accept institutional accountability. A model's statement that “this claim is reliable” is not independent approval but another

model output. “The AI decided” is not an explanation. It is an evasion of responsibility.

Accountability does not remain at the level of principle. Every material decision names the proposer, implementer, approver, publisher, person empowered to stop the work and final accountable owner. Traceability and accountability are likewise treated as distinct responsibilities in the OECD AI Principles.[11] Accountability without a name has not been assumed.

3. Fitness Gates

3.1. Business fitness

Not every business is fit to be amplified through GEO. Six fields are examined before work is accepted.

1. **Claim fitness:** The central statements to be amplified must be capable of proof. Unsupported terms such as “best”, “leader” or “guaranteed” are narrowed. If the business refuses, the work is refused.
2. **Delivery fitness:** Capacity, geography, price, staffing, delivery time, support, cancellation and refund terms, and promised outcomes must be capable of meeting the expectation created. A claim that cannot be delivered cannot be amplified.
3. **Customer-harm fitness:** Work is refused where hidden fees, manufactured scarcity, misleading guarantees, false reviews, fabricated success rates, concealment of harm or deliberate failure to exclude unsuitable customers form part of the business model. Making a harmful model more visible scales the harm.
4. **Correction fitness:** The business must accept narrowing unsupported claims, removing obsolete information, aligning language editions and withdrawing an erroneous public statement. Preserving its own contradictions while expecting the system to represent it accurately is not fitness.
5. **Transparency fitness:** Real and synthetic data, measured and assumed outcomes, sponsorships and conflicts of interest must remain separate.
6. **Accountability fitness:** The business must accept that final accountability remains with it. Work does not begin on the condition that “the agency assumes all responsibility”.

The decision may be **accept, conditional acceptance, observe and verify first**, or **refuse**. Commercial opportunity does not create ethical fitness.

3.2. Practitioner fitness

Technical knowledge alone does not confer fitness to exercise representational power. The practitioner must be able to distinguish entity, attribute and judgment layers; mentions, attribution, recommendations and commercial outcomes; and evidence, measurement and intervention. The denominator must be disclosed, counter-evidence sought, uncertainty recorded, and the method must not be presented as its own proof of success.

If a client asks for favourable outputs alone to be included in a report, a competitor to be weakened or an unsupported superiority claim to be made, the practitioner has not merely a right but a duty to refuse. Conflicts such as services supplied to competitors, outcome-based fees, revenue from badge sales, or the combination of evaluator and sales roles must be disclosed. A conflict does not automatically invalidate every engagement; concealing it invalidates trust.

The practitioner must also demonstrate record and withdrawal discipline, observance of the authority granted, and competence in high-risk fields. The outcome may be **fit**, **fit with limited authority**, **fit with additional oversight**, **not fit for high-risk operations**, or **not fit**. Competence does not create authority; authority must be granted separately.

4. Risk Classes and Two-Key Control

Governance decisions are handled under four risk classes:

Table 5.1 — Risk classes and minimum approval

Class	Typical example	Minimum approval
Low	Correction of spelling, date, clear identity or a minor boundary	One accountable human
Medium	Multilingual narrative, external-directory correction, change in service scope	Practitioner + entity owner
High	Health, law, finance, safety, competitor comparison, superiority claim, difficult reversal	Two separate humans
Prohibited	False consensus, competitor degradation, synthetic data presented as real, covert manipulation	Reject

The risk class cannot be lowered after the result is seen; it is determined before the operation. A high-risk decision requires at least two distinct humans: the **proposer** and the **approver**. The proposer is accountable for the method and rationale; the approver for the evidence, risk, language and withdrawal conditions. Neither key may be held by an automation or model.

Two-Key Control applies to public claims of success, competitor comparisons, language such as “best” or “most trusted”, health, legal, financial or safety effects, publications that use model output as evidence, structural interventions that are difficult to reverse, broad external propagation and any future mark of conformity.

5. Claim Governance and Transparency

Not every accurate observation should be made public. A claim occupies one of four publication states:

- **Internal hypothesis:** may be used in research and intervention planning; may not be published as an outcome.
- **Limited observation:** may be shared with the system, date, prompt and scope stated; may not be generalised.
- **Limited public claim:** has an Evidence Record, human approval, permitted wording and an explicit boundary.
- **Withdrawn claim:** may no longer be used; its earlier record is retained and marked as withdrawn.

Every public claim has an owner, a linked Evidence Record, a measurement scope, the exact permitted sentence, prohibited extensions, risk class, approver, publication and review dates, and withdrawal conditions. A public claim is not a writing decision. It is a decision to assume responsibility.

Transparency does not mean publishing everything. It means not withholding material information that would change the judgment. The method, systems measured and not measured, prompt universe, time range, evidential boundary, missing data, contrary results, synthetic examples, conflicts of interest, protocol changes and side effects must be visible. Trade secrets, security information and personal data may be protected; where detail cannot be published, the effect of that limitation on the decision must be explained.

Incentive matters as much as method. An agency rewarded for increased visibility may come to treat more visibility as success; an organisation dependent on badge revenue may lower its approval threshold; an evaluator whose contract depends on renewal may soften an adverse result. This is not an accusation about individual character. It is a problem of system design. Open disclosure, separation of roles, second review, fixed criteria, remuneration independent of outcome, exclusion of sales teams from fitness decisions and an appeal route all reduce the risk.

6. A Mark of Conformity: Design Before Authority

A mark of conformity is not a claim of superiority, a recommendation or a guarantee of results. It indicates that a specified scope was assessed against written criteria under a stated date and Framework version. Publication of this book does not in itself establish an active certification programme. **No “NobleJackal GEO certificate” or equivalent mark may be issued until the infrastructure for independent decisions, conflict-of-interest management, surveillance, appeals, suspension and withdrawal is operational.**

This boundary is consistent with international practices that emphasise impartiality, separate decision-making, surveillance, suspension and withdrawal, and appeal mechanisms in conformity assessment. The book does not, however, claim ISO accreditation or ISEAL membership.[17][18]

If a pilot mark is introduced in future, it may mean no more than: “The specified representational field of this entity was found to conform, within the stated scope and on the stated date, under the indicated version of the NobleJackal GEO Framework.” It can never mean that AI systems recommend the business; that the business is superior to competitors; that all its claims are true; that it is consistent in every language; that it will remain unchanged; that commercial results are guaranteed; or that ethical failure is impossible.

The mark exists together with an identifier, the entity assessed, the Framework version, included and excluded scope, evidence date, review date, suspension and withdrawal conditions, a public verification record and rules of use. An assessment service may be purchased; its outcome may not. The mark may not be renamed “AI approved”, “AI recommended”, “best” or “guaranteed”.

The mark is suspended where evidence has aged, a serious new contradiction emerges, the mark is used outside its scope, synthetic data is presented as real, a public-claim rule is breached, a correction request goes unanswered, a signal of serious customer harm appears, the Framework version changes or review is not completed. Fabricated evidence, false consensus, deliberate degradation of competitor representation, repeated scope violations, an undisclosed serious conflict of interest, material user harm or misleading sales use are grounds for withdrawal. A mark that cannot be withdrawn is not a governance instrument. It is a marketing ornament.

7. Incident and Enforcement Management

Governance does not assume that error will never occur. It establishes the route to follow when it does.

1. **Freeze:** Temporarily stop the risky claim, intervention or use of a mark and prevent further propagation.
2. **Preserve the record:** Retain the raw output, prompt, time, content version, approval and change history. Do not erase the traces of failure.
3. **Contain the harm:** Correct owned surfaces, send corrections to external sources where necessary, and stop the erroneous customer promise.
4. **Notify:** Inform the parties whose decisions are affected. A material public error may require a public correction.
5. **Adjudicate:** Distinguish failures of data, measurement, approval, authority, ethics, intent and system control, and record the resulting decision.

Depending on the gravity of the breach, enforcement may take the form of a correction request, written warning, claim withdrawal, cessation of the intervention, additional audit, suspension or withdrawal of a mark, restriction of practitioner authority, termination of the engagement or a public status record. Enforcement is not retaliation. It is the standard remaining faithful to its own sentence. The size or commercial value of the client cannot alter the criterion.

8. Teaching Case: AegisCare Diagnostics

Teaching simulation — synthetic data. This case does not represent a real healthcare provider, clinical dataset or model test.

AegisCare Diagnostics is a fictional provider of corporate screening and laboratory services. It wants to publish the statement: “AI systems recommend AegisCare as one of the most trusted diagnostic centres.” It submits six favourable generative outputs and one pilot satisfaction report.

Table 5.2 — Synthetic governance decision for AegisCare Diagnostics

Governance field	Synthetic finding
Evidence about the real world	Service licences are valid
Evidence of representation	Six favourable outputs for specified prompts
Comparative evidence of trustworthiness	None
“Most trusted” claim	Unsupported
Domain risk	Health — high
Two-Key decision	Second approval refused
Permitted wording	“AegisCare was recommended in the six specified test outputs.”
Prohibited wording	“Most trusted”, “AI approved”, “selected by AI”

Part of the observation is real, yet the requested public language is refused. A licence does not prove comparative trustworthiness, and six recommendations do not prove the superiority claim “most trusted”. In a health context, false certainty can have a serious effect on user decisions. A sentence's partial truth does not make it fit for publication.

9. The Decision of Governance

The Governance Record, detailed in Appendix C, brings together the matter for decision, linked evidence, business and practitioner fitness, risk class, proposer and approver, conflicts of interest, permitted and prohibited acts and language, review and withdrawal conditions, and final accountable owner. Without a record there is no decision—only an event for which a story may later be invented.

The judgment of this chapter is:

Authority may be delegated. Accountability may not. Automation can perform operations; judgment and the duty to answer for them remain human.

Governance refuses some work, narrows some claims, declines some clients and withdraws approval where necessary. This is not weakness. It is the moment at which authority does its work.

No decision, however, remains valid forever. Evidence ages, protocols become obsolete, interventions create maintenance debt, and what is true today may cease to be current tomorrow. Time is therefore the outer ring.

Chapter VI — Time

How a Standard Ages

Validity is not permanent; the record must be. Information that is true today need not become false tomorrow, but it may become obsolete. Obsolete information can be more dangerous than an obvious lie. It began as truth, rested on a source and may once have been audited. That history lends it a legitimacy it no longer deserves in the present.

Prices change, service scope contracts, an executive leaves, a certificate expires, a company merges into another brand. Models and interfaces change; citation behaviour evolves; a measurement protocol is rewritten; a sound intervention turns into maintenance debt. A conformity decision granted within a particular scope may continue to live on a page that no longer describes the same business.

No record keeps itself current. Currency is not an attribute but a judgment that must be made again. The life of a claim cannot exceed the life of its earliest-expiring support.

1. The Judgment of Time

Time does more than date a record. It decides whether the claim can still be carried. Every public sentence must answer five questions:

1. Is the underlying truth about the real world still valid?
2. Is the representational behaviour still being observed?
3. Does the measurement protocol still work and support comparison?
4. Is the intervention still necessary, proportionate and safe?
5. Do the approval and publication authority remain in force?

If the first answer is no, evidence about the real world loses validity; if the second, the representational claim; if the third, the measurement series; if the fourth, the grounds for intervention; if the fifth, the public language. When one link ends, the whole past does not automatically become false. The current judgment does, however, reopen for review.

Temporal discipline rests on four principles. **Currency must be proved:** “This information is still true” is a new claim. **Expiry is not failure:** it shows that a decision accepted its temporal limit. **Revision is not betrayal:** preserving a method that should change in the face of new observation is weakness. **An archive is not authority:** an old record preserves the past; it does not carry a current decision.

A judgment without a date is not current. It merely looks timeless.

2. The Five Clocks

Five clocks run concurrently in GEO work, each at a different speed.

2.1. The entity's clock

A company name, ownership, management, product, service scope, geography, price, capacity, licence, certificate, partnership or delivery condition may change. A business's operation in ten countries in 2025 does not prove that it has the same scope in 2027. When the entity's clock moves, the content of accurate representation moves with it. Once the world changes, an old representation may become a falsehood written correctly.

2.2. The representation clock

Generative-system behaviour is not fixed. The model, interface, information retrieval, safety rules, language behaviour, session and user context may all change. A brand recommended today may disappear tomorrow; a source shown clearly today may be dissolved into an answer in another interface; an identity collision once resolved may return through a new data surface. An observation of representation carries the moment in which it was recorded. It cannot be brought into the present without repetition.

2.3. The evidence clock

The authenticity of a document does not establish the present validity of the claim it carries. An expired certificate, customer research relating to a changed product, an old price list or a performance report from another production line may be historically genuine while its authority for current decisions has ended. The evidence is not deleted; its status changes.

2.4. The protocol clock

A measurement protocol ages too. Its interface may disappear, its prompt families may cease to represent user behaviour, and its coding categories may no longer accommodate new forms of error. Ageing does not make the old measurement fraudulent; it makes the protocol inadequate for new measurement. A method that once worked is not a method that works forever.

2.5. The authority clock

Approval and publication authority remain valid only while the conditions under which they were granted continue. A business may enter a new service, country or ownership structure; publish unverified claims; use a conformity mark beyond scope; or change its practitioner team or Framework version. Authority does not survive the end of its conditions merely because a name remains in the record.

3. Temporal Integrity

A claim has temporal integrity when the following chain remains aligned:

Evidence about the real world → observation of representation → measurement protocol → intervention version → governance decision → public claim

Every link carries a date and version. For evidence about the real world, ask: “Does the underlying truth remain valid?” For representation: “When and under what conditions was the behaviour observed?” For the protocol: “Which method version was used?” For the intervention: “Which surface changed, and when?” For governance: “Who approved what scope?” For the public sentence: “Until when, or until which event, may it be carried?”

Under the **Earliest-Expiring Support Rule**, a public claim cannot outlive the shortest validity period in this chain. If a certificate has expired, an old representation measurement cannot support a current attribute claim. If the protocol has changed, two periods cannot be compared directly. If the scope has changed, an earlier conformity decision cannot extend to new services. If the intervention has been withdrawn, language that depends on its success may not be used.

A chain lives not as long as its oldest record, but as long as the link that first loses validity.

4. Fixed Principle, Changing Tool

A framework that changes everything loses its identity; one that changes nothing becomes dogma. The distinction must remain explicit.

The following are core principles: entity and representation are not the same; evidence limits the claim; measurement cannot substitute for evidence; multilayered representation may not be hidden inside one score; intervention creates new risk; accountability remains human; the ethical boundary cannot be removed in pursuit of commercial outcomes; synthetic data may not be presented as real; and no one may claim ownership of an outcome they do not control. These

principles do not depend on a particular model, interface or tactic. Altering them is not a “tool update”; it reopens the foundational version.

Lists of systems, prompt sets, technical access methods, source classes, structured-data formats, measurement intervals, methods of recording interfaces, CRM connections, reporting templates and verification surfaces are tools. They are expected to change. A change of tool must not be presented as though the principle changed, nor a change of principle disguised as a minor tool adjustment.

5. Validity Classes and Triggers

There can be no universal validity period of “thirty days” or “one year” for every record. Rates of change, levels of risk and forms of support differ. The Framework uses five classes:

- An **instantaneous record** is a single answer, statement or referral event; it carries only the moment in which it was recorded.
- A **variable operational record** concerns fast-changing information such as price, capacity, team, delivery time and service scope; it is reviewed both by calendar and by event.
- An **institutional record** concerns a licence, certificate, ownership, contract or formal authority; it expires with the document or the institutional change.
- A **behavioural pattern** is recurring representation across systems, time and prompts; it is retested when the system, interface or protocol changes.
- A **principle or standard** is tested less by a fixed calendar than by counter-evidence and practical adequacy. It reopens when it cannot explain new observations, produce consistent decisions, contain breaches, avoid unnecessary harm or remain coherent with its own purpose.

The calendar is not the only trigger. A change in real-world evidence; expiry of a licence; transformation of the company or service scope; a material system or interface change; failure to reproduce the protocol; strong counter-evidence; serious customer harm; an unexpected side effect; out-of-scope use of a mark; discovery of a conflict of interest; a change in legal conditions; loss of delivery capacity; or a new Framework version initiates immediate review. “The annual review is not yet due” does not preserve validity after an event has changed it.

6. Version Discipline and Series Breaks

A change that is not recorded is not a revision; it is a silent alteration. There are three version classes.

A **correction release** addresses spelling, formatting, a broken link or clarity without changing meaning; it does not affect decision thresholds or comparability. A **method release** changes a tool by adding, for example, a system, language, prompt family, coding class or record field; principles remain intact, while comparability between old and new records is assessed separately. A **foundational release** changes a principle, decision right, risk threshold, fitness gate or accountability structure; it requires fresh governance approval and cannot be presented as a silent continuation of the previous standard.

Every release record states what changed and why, which records are affected, which comparisons can no longer be made, which earlier decisions must be reviewed, the effective date and the status of the previous release. Silently replacing old text is rewriting the past.

A change in system, model, interface, prompt universe, language, coding definition, competitor set, denominator, number of repetitions, treatment of missing data or method of commercial attribution may break a measurement series. The new protocol may be better, but it is not thereby a continuation of the old series. Results may be shown side by side; without a common subset, they cannot produce a direct rate of improvement.

7. Status Language and the Archive

Every material record carries a current status:

Table 6.1 — Record statuses and their meanings

Status	Meaning
Active	Scope and supporting grounds remain valid
Under review	The decision is open because of a new event or counter-evidence
Superseded	A newer record has become the primary reference
Deprecated	Not recommended for new work; retains historical value
Expired	A validity condition or review period has ended
Withdrawn	Cannot support a claim because of error, breach or new evidence
Archived	Historical record only; carries no current authority

An old record is not deleted; its status changes. A withdrawn claim may remain in the archive but cannot support a current report. A deprecated protocol may matter to the history of research but may not be used for new measurement. The archive preserves history, not authority.

8. Language about the Future

The chapter on Time does not grant authority to know the future; it limits the sentence that may be spoken about it. **Language of principle** carries a normative judgment expected to endure even as circumstances change: “Evidence should limit the claim.” **Language of scenario** considers a possible condition without asserting that it will occur. **Language of prediction** requires a time horizon, assumptions, supporting grounds and an expiry date.

Statements such as “All search will soon take place through AI”, “This method will continue to work in the future” or “Systems will treat this mark as a trust signal” must either be presented explicitly as scenarios or treated as claims requiring evidence and temporal limits. Certainty about the future is not foresight. It is uncertainty concealed.

9. Signs That a Standard Is Dying

A framework dies not because it can be changed, but because it cannot. Review is mandatory when the following signs recur:

- A principle turns into a slogan instead of producing decisions.
- Allegiance to “what the Framework says” replaces evidence.
- Contrary observations are excluded without first being examined as possible problems in the method.
- New data is forced into the old model.
- Versions change silently and the past disappears.
- Commercial exceptions are made for a large client or high revenue.
- A conformity mark becomes a permanent title of superiority.
- Every adverse result is dismissed as noise.
- The Framework begins protecting its own brand instead of testing reality.

Revision does not kill a standard. Protection from revision does.

10. Teaching Case: Solstice Mobility

Teaching simulation — synthetic data. This case does not represent a real business, system or commercial outcome.

Solstice Mobility is a fictional operator of corporate electric-vehicle charging networks. Its 2024 annual report accurately documents active operations in

twelve countries. By 2026, the company has withdrawn from five markets, while the old report, partner pages and external directories remain in circulation.

Table 6.2 — Synthetic temporal comparison for Solstice Mobility

Observation	2024	2026
Countries in which the business is actually active	12	7
Official pages stating “service in 12 countries”	4	2
External sources carrying the same claim	9	14
“12 countries” represented in responses	18/30	21/30
Enquiries based on the wrong country scope	3	17
Unsuccessful sales conversations	1	11

The 2024 report is not fraudulent. In 2026 it is no longer authorised to carry the same claim. Worse, the obsolete information has spread as new external sources repeat the old text. The entity's clock moved; evidence about the real world, representation and public language were not updated together. The ageing of an accurate record is the quietest form of misrepresentation.

The Framework's decision is not to delete the historical report. It preserves that report's historical status, replaces the current scope with a new record, corrects owned surfaces, notifies external sources and remeasures the core prompt set against the new evidence about the real world.

11. The Decision of Time

The Time Record in Appendix C keeps the initial date of validity and support for the linked item, the protocol and Framework versions, the most recent review, calendar and event triggers, present status, effect on the series, superseding record, scope and owner of currency. Its question is simple: **May this judgment still speak?**

The judgment of this chapter is:

Principles are what must remain fixed; tools are what must change. An old version is not erased, but neither does it carry current authority by itself.

The Centre defined the object. Evidence limited the sentence. Measurement bound observation to a record. Intervention touched a verified distortion. Governance separated power from accountability. Time ruled that none of them remains valid forever.

This reopens the book's commercial question. "AI brings us customers" can be said only under a defined period, scope and body of evidence. Whether the customer is right, the sale profitable, delivery successful and the relationship repeatable becomes visible only through time. Sustainability is not the unauthorised projection of today's result into tomorrow. It is value's capacity to recur without degradation over time.

A standard that refuses to acknowledge ageing is not born strong. It is stillborn.

Part Three — Reaching Judgment

Chapter VII — The Final Test

Customer, Value and Sustainability

A customer's arrival is not the outcome. The outcome is sustainable value produced through the right customer. This chapter adds no new ring. It judges why the preceding rings exist.

A business may be represented accurately; its claims may rest on strong evidence; its representation may be measured with discipline; distortions may be addressed through controlled intervention; decisions may be made under transparent governance; and every record may be reviewed through time. All of that may be true without sustainable commercial value having been demonstrated.

The reverse is also possible. Traffic, enquiries and revenue may rise while representation is false, customers are unsuitable, delivery destroys value or the method crosses an ethical boundary. Commercial effect does not create legitimacy, and legitimacy alone does not create commercial value. Both must be tested together.

Here, **sustainability** is not an environmental claim. It means the business can sustain value honestly, profitably, deliverably and repeatedly. One sale, one strong month or strong representation in one system is not sustainability. Visibility may be the beginning; sustainable value is the final judgment.

1. The Five Judgments of the Final Test

A positive GEO judgment for the business must satisfy five conditions together:

1. A generative-AI influence must have been observed.
2. The demand associated with that influence must be qualified.
3. The sale must produce a positive net commercial contribution.
4. The promised value must be delivered to the customer.
5. The outcome must be ethical and sustainable through time.

Without the first condition, AI influence cannot be claimed. Without the second, there is enquiry noise rather than customer acquisition. Without the third, there may be revenue but no value. Without the fourth, there is a sale but no successful delivery. Without the fifth, there is temporary effect but no sustainable success.

The conditions do not compensate for one another. Strong attribution does not cure poor customer quality; high revenue does not cure loss-making delivery; a satisfied customer does not cure an ethical breach; and repeat sales do not remove

dependence on a single platform. The Final Test does not ask merely, “Did something happen?” It asks: “Did what happened leave the business with real, legitimate and durable value?”

2. Five Classes of AI Influence

A customer's arrival does not prove that AI brought the customer. Influence is held at the highest class the record can carry; it is never moved upwards merely to make the result look stronger.

Class 0 — Unknown: The customer made contact, but there is no reliable record of the decision path. This is not an “AI customer”; it is a customer of unknown origin.

Class 1 — Declared influence: The customer says that a generative system played a part in the decision. “I found you on ChatGPT” is valuable testimony; it does not exclude other influences such as advertising, organic search, a recommendation or prior experience.

Class 2 — Observable referral: A link or technical contact from a generative interface to the site has been recorded. It shows the transition, not the whole decision. The user may have known the brand already or been influenced later by other sources.

Class 3 — Influence associated through multiple records: Referral, customer account, CRM record, the answer or prompt used, expectations expressed in the sales conversation, timing and other channels converge on the same journey. The association becomes stronger; full causation is still not automatic.

Class 4 — Controlled contribution: A controlled test, comparative period or strong experimental design separates AI influence from alternative explanations to a defined extent. Even this rare class does not mean “sole cause”; commercial decisions usually arise through multiple contacts.

When attribution is incomplete, the customer does not disappear. Only the claim becomes smaller.

3. The Customer Ladder and Qualification Gate

The stages of the commercial chain are not interchangeable:

Appearance → orientation → contact → qualified enquiry → proposal → sale → profitable delivery → satisfaction → repeat or referral → sustainable contribution

At appearance, the user sees a representation; there is no behaviour yet. At orientation, the user conducts further research into the brand; there is interest, not demand. At contact, there is a form, call, message or appointment; suitability remains unknown. In a qualified enquiry, the need matches the business's capacity. At proposal, the business accepts the enquiry as a genuine opportunity. A sale creates revenue, but value has not yet been delivered. With profitable delivery, the promise is fulfilled and the transaction leaves a positive contribution after direct and incremental costs. Satisfaction means that the customer finds an acceptable alignment between the represented promise and the value received. Repeat business or referral carries the relationship beyond one transaction. Sustainable contribution means that the whole outcome can recur through time within the constraints of capacity, ethics and channel resilience.

Five fields of fit determine whether an enquiry is qualified. **Need fit** asks whether the customer's problem is one the business genuinely solves. **Scope fit** asks whether geography, volume, attributes and time remain within the true service boundary. **Economic fit** asks whether budget and the price–cost structure align. **Timing and decision fit** require both a real buying horizon and the authority to decide. **Delivery and ethical fit** require the sale to create genuine value for both the business and the user.

The decision may be **qualified, conditionally qualified, not yet classifiable, not suitable, or rejected because of risk of harm**. An unassessed enquiry is not qualified. Nor can every enquiry that became a sale be declared the right customer after the fact. Payment does not prove customer fit.

4. The Distance between Revenue and Contribution

Revenue is measured; value is calculated. A business can become weaker even as revenue rises after GEO work. At minimum, the view of net commercial contribution includes the following:

Table 7.1 — Net commercial contribution record

Element	Recorded value
Gross sales revenue	Amount invoiced or collected
Discounts, cancellations and refunds	Amount returned from revenue
Direct delivery cost	Product, labour, procurement and operations
Incremental sales and support cost	Burden of processing, explaining and correcting new demand
Customer-acquisition and measurement cost	GEO, software, reporting and campaign expenditure

Element	Recorded value
Capacity cost	Overtime, outsourcing and other work delayed
Net commercial contribution	Revenue less direct and incremental costs

This is not a universal formula for calculating the business's entire accounting profit. It is a management view that makes the contribution of the customer cohort under examination visible. The business's accounting policies and national obligations continue to apply separately.

The cost of misrepresentation does not always appear on an invoice. Sales cost grows while the team explains a service that does not exist, corrects an inaccurate price expectation or speaks with people who lack decision authority. Delivery cost arises when a customer expects the scope seen in a generative answer and judges the delivery inadequate even where the contract was narrower. Complaints and adverse reviews produce fresh external signals that feed reputational cost. When capacity is assigned to the wrong customer and the right customer is delayed, opportunity cost follows.

The cost of the wrong customer is not only the money spent on that customer. It is the right customer lost because of them.

5. Five Conditions of Sustainability

Economic sustainability requires the cohort to leave a positive contribution after discounts, refunds, delivery, sales and support burden. Loss-making growth is only a growing loss.

Operational sustainability requires demand to be met within capacity without degrading quality or exhausting employees and existing customers. Persistent overtime and delayed delivery are not growth. They are breaches of capacity.

Customer sustainability requires the promised value actually to be delivered, with cancellations, refunds, complaints, satisfaction, repeat behaviour and referrals tracked. Silence is not proof of satisfaction. Customer acquisition is not complete until the customer has received value.

Ethical sustainability excludes deception, covert manipulation, fabricated evidence and foreseeable user harm. Ethics may appear to be one of five conditions, but it holds a veto. An outcome that crosses the ethical boundary is not “sustainable”; it ought not to be sustained.

Channel and temporal sustainability require the outcome not to rest precariously on one model, interface, prompt or short period. Profitable contribution from a single platform is real; so is dependence on that platform. If

all demand disappears when the system changes, the business has not become more resilient. It has acquired a new dependency.

6. Observation Horizon, Cohort and Attribution

There is no universal period for sustainability. The observation horizon cannot be shorter than the cycle in which value is realised. A hotel tracks booking, stay and the post-stay outcome. A subscription tracks the first payment, cancellation and renewal. A B2B project tracks contract, delivery, collection, profitability and, where possible, repeat work. A product tracks the returns window. The period cannot be chosen after the result so as to capture only its favourable portion. The observation horizon ends not where the outcome looks good, but where value is realised.

An aggregate customer count can force groups with different behaviours into one story. Customers are therefore divided into cohorts by such dimensions as first-contact period, influence class, system, language, intent, product, customer type, geography, new or existing status, and one-off or recurring relationship. Two systems may bring the same number of customers while one produces high returns and the other repeat purchases. The counts are equal; the values are not.

Attribution is not surrendered to one analytics screen. Referral, site session, brand search, direct return, form or telephone contact, customer account, CRM, proposal, sale and other contact channels are considered together. Last click does not reveal the first cause; a customer's account does not reveal the only cause; direct traffic does not prove absence of influence; and a referral does not reveal the whole decision process. Where certainty is unavailable, use “customer declaring AI influence” or “sale associated with AI influence”, not “AI customer”.

7. Decisions of the Final Test

The test closes with one of the following decisions:

- **Commercial effect not demonstrated:** Representation was observed, but no reliable connection to customer and business outcomes was established.
- **Signal of AI influence observed:** A declaration, referral or limited contact exists; there is no judgment of qualified demand or sale.
- **Contribution to qualified demand observed:** Some associated enquiries passed the Qualification Gate; sale or profitability may remain incomplete.
- **Profitable commercial contribution observed:** Sale and delivery were completed, net contribution was positive, and ethics and governance were found satisfactory; continuity through time may remain open.

- **Sustainable contribution demonstrated:** Across an adequate observation horizon, the qualified cohort produced positive contribution, delivered value, repeat business or referral, capacity fit and manageable channel risk. This decision is not permanent either.
- **Harmful or invalid outcome:** Negative contribution, a high rate of unsuitable customers, capacity breach, increased complaints, unsupported attribution, misleading reporting, synthetic data presented as real, user harm or a governance breach is present.

Representational integrity may be strong while commercial effect remains unknown; in that case, representational success is preserved within its own boundary. If commercial contribution is positive but the time horizon inadequate, the judgment is “profitable contribution”, not “sustainable”. If visibility falls while net contribution rises, the outcome may be successful. An ethical breach closes the positive judgment irrespective of every other field.

8. Teaching Case: Auren Precision Optics

Teaching simulation — synthetic data. This case does not represent a real manufacturer, contract or commercial outcome.

Auren Precision Optics is a fictional producer of custom optical components for medical-imaging equipment. Its market is small; success depends not on high traffic but on a durable relationship with the right technical buyer. A twelve-month cohort produces the following result:

Table 7.2 — Synthetic cohort outcome for Auren Precision Optics

Indicator	Synthetic result
Enquiries declaring AI influence	23
Enquiries associated through multiple records	17
Qualified enquiries	11
Proceeded to technical evaluation	8
Initial contracts	5
Projects completed profitably	5
Repeat orders	3
New enquiries from customer referral	2
Cancellations or refunds	0
Net commercial contribution	284,000 units
Additional operational burden	Within planned capacity

Indicator	Synthetic result
Non-AI contact	At least one additional contact for every customer

AI cannot be presented as the sole cause; every customer had other contacts. The declarations of influence, recorded contacts, CRM and sales timing nevertheless converge within the same cohort. All five projects were delivered profitably, three led to repeat orders and two generated new referral enquiries. The outcome remained within capacity across two delivery cycles.

The Framework permits the following language: “Within the defined twelve-month cohort, five customer projects associated with AI influence through multiple records were completed profitably; three produced repeat orders and two produced new referral enquiries. Contacts outside AI were not excluded.”

Sustainable value does not require high traffic. It requires the recurrence of the right fit, delivered benefit and positive contribution.

9. Final Judgment

The Sustainable Value Record in Appendix C brings together the period, business cycle, influence class, attribution traces, cohort, qualification status, proposal and sale, refunds and costs, net contribution, delivery, customer outcome, repeat behaviour, ethical status, channel dependency, evidential boundary and review. The record is kept not to assign a sale to AI, but to test what AI influence left within the business.

The judgment of this chapter is:

Without observed AI influence there can be no “AI customer”; without testing customer quality, no “success”; without seeing cost, no “value”; before delivery is complete, no “outcome”; before time is tested, no “sustainability”. If the ethical boundary has been crossed, no positive judgment can be made.

The purpose of GEO is not to appear more often in AI systems, but to establish the right relationship with the right person through accurate representation. For the business, the value of that relationship is determined not by the number of customers, but by the benefit delivered, the contribution retained and the capacity to repeat it.

Sustainability is not announced. It is demonstrated.

Chapter VIII — Audit

From Claim to Decision

An audit does not manufacture validation. It decides the limits within which a claim may speak. It is not another ring, but the decision mechanism through which the preceding rings work together. The Centre limited the object; Evidence the sentence; Measurement the observation; Intervention the action; Governance the authority; Time the validity; and the Final Test the value to the business. Audit unites them in one traceable file.

A business may purchase an audit service; it cannot purchase a favourable outcome. A practitioner may receive a fee; they cannot sell approval. “We are strong in AI systems—prove it” is a publication request whose result has already been written. “Assess how generative systems represent us, how that representation relates to customer outcomes, and which statements we may make” is an auditable request.

An audit does not begin by collecting data. It begins by opening the sentence to be adjudicated. What exactly will be said? Who will use it? Which decision will rely on it? What are its scope and risk? Until when will it remain valid? The audit asks one further question: which adverse outcomes is the client unwilling to accept in advance? If the business permits only a favourable judgment, the audit does not begin. A process whose outcome is not genuinely open is not an audit. It is a staging.

1. The Seven Questions of Audit

Every audit answers seven questions:

1. Which entity and which representation are under examination?
2. Which evidence about the real world, representation and outcomes supports the claims?
3. Under which protocol were the observations measured?
4. If distortion exists, is intervention necessary, proportionate and legitimate?
5. Who made the decision, and with whom did accountability remain?
6. Are the records and authorities still valid?
7. Where the scope is commercial, did AI influence leave sustainable value?

An audit need not address every question at the same depth. Identity representation alone may be examined; the work is not invalid, but it cannot

produce a commercial judgment. Representation in specified systems and languages may be assessed; a mark of conformity may not follow. An omitted field is not a defect. A concealed omission is.

The output is not one score but an **Audit Opinion**. It may state that a claim is unsupported; can be carried only within a limited scope; that a material distortion was verified; no grounds for intervention were found; a limited intervention is appropriate; commercial contribution was observed; sustainable contribution was not demonstrated; the outcome is invalid because of ethics or governance; or the records have expired. The value of an audit does not lie in producing a favourable result. It lies in its ability to carry an adverse one.

2. The Audit Admission Gate

Not every request is admitted to audit. Six conditions apply.

Entity clarity: The company, brand, person, product, service or group is defined precisely. A holding company and one of its brands, or an author's legal identity and pen name, may be related without constituting the same audit unit. If the object is ambiguous, work does not begin until scope is clear.

Decision purpose: There must be a definite decision question, such as verifying a representational distortion, determining the need for intervention, evaluating a public claim, examining the relationship with customer outcomes, testing sustainable contribution or reopening the validity of an earlier decision. "Let's see how we are doing" is not enough; if the decision is unknown, sufficient evidence cannot be defined.

Access to records: Reasonable access is provided, as the scope requires, to formal identity, product and service boundaries, licences, site versions, previous measurements, intervention history, CRM, sales, delivery, refunds and customer accounts. Inaccessible records are not assumed away; they narrow the scope. A representational review may proceed without commercial records; a sustainability audit may not.

Acceptance of an adverse outcome: The parties accept that the claim may be rejected, public language narrowed, intervention not recommended, commercial effect left unproved, conformity refused or an earlier decision withdrawn. A party that contractually prohibits an adverse finding cannot be audited.

Clarity of conflicts of interest: Relationships among auditor, practitioner, entity owner and any prospective conformity programme are recorded. Outcome-based fees, services to a competitor, the auditor's own implementation of the intervention, or use of a favourable report in an investment decision or contract renewal are disclosed. One institution may implement and audit an

intervention, but in a high-risk case it may not approve its own result without an independent second decision.

Synthetic data, confidentiality and privacy: Synthetic, anonymised and composite records are labelled explicitly. Trade secrets and personal data are protected, but protection does not create a right to conceal an omission that would alter the judgment. Details may remain private; the limitation must be public.

The gate may decide **admit**, **conditional admission**, **admit within limited scope**, or **refuse**. Refusal may mean lost revenue. It is not a loss of the standard.

3. Scope Lock and Audit Levels

Before the audit begins, the entity, geography, language, system and interface, time, representational field, prompt universe, comparison group, commercial field, risk, audit level, public use, validity and exclusions are locked. Scope states not only where the audit looks, but the boundary the judgment may not cross.

A new system, language or record may be added during the work, but never by silent expansion. A new element creates a new scope version. An English-language result is not evidence of German representation; a brand audit is not evidence about the group of companies; identity conformity is not evidence of commercial sustainability.

There are five audit levels:

- 1. Preliminary Review:** Determines whether an auditable problem, claim, initial risk and appropriate scope exist. It produces no public statement of success.
- 2. Representation Audit:** Examines evidence about the real world, a controlled prompt universe, contrary and exclusion tests, human coding and the representation profile. It may speak about material distortion; it may not speak about commercial contribution.
- 3. Intervention Audit:** Tests before-and-after change through a frozen baseline, the Intervention Gate, intended and prohibited outcomes, core set, observation window and withdrawal conditions. It may show change; it does not guarantee full causation.
- 4. Commercial and Sustainable Value Audit:** Combines influence class, customer cohort, qualified demand, sales, net contribution, delivery, satisfaction, repeat behaviour, channel dependency and an adequate observation horizon.
- 5. Conformity Assessment:** Examines business and practitioner fitness within a defined scope of the Framework, evidence and measurement discipline, intervention governance, public language, time, conflicts of

interest, and suspension and withdrawal conditions. Completing this level does not itself create a badge. Without the independent programme infrastructure required by Chapter V, it produces only a pilot conformity opinion.

A lower level may not borrow the language of a higher one.

4. The Audit File and Claim Register

An audit is not one report but a system of linked records. The Entity File carries the entity under audit and its relationships; the Claim Register every sentence the parties seek to make; the Evidence Record the support for each sentence; the Measurement Record the observations; the Intervention Record the change and withdrawal conditions; the Governance Record authority, conflicts of interest and accountability; the Time Record validity and version; and the Sustainable Value Record the business outcome. Approaches such as W3C PROV-O, which express provenance through relationships among entities, activities and agents, offer a technical neighbour to this idea of traceability; the Framework does not require the use of a particular ontology.[9]

No record substitutes for another. Measurement shows the observation; the Evidence Record permission for a sentence; Intervention the operation; Governance authority for the operation; Time its current validity; and the Value Record the business's gains and losses. The opinion is not constructed afterwards to justify these records. It arises from them. Documentation approaches such as model cards and datasheets for datasets have shown, on different planes, the importance of recording purpose, scope, performance, composition, use and limitations.[12][13] Other approaches to AI audit likewise treat documented roles, evidence and corrective mechanisms across the lifecycle as essential to accountability.[10] NobleJackal's file architecture is adjacent to this literature; it has not been validated by it.

The Claim Register is especially important. “Systems recognise us”, “We are represented with accurate attributes”, “We are recommended more often than competitors”, “The intervention increased visibility”, “AI brings us customers”, “These customers are more profitable” and “We conform to the Framework” are not the same sentence. Each carries a separate type, owner, intended use, risk, required audit level, linked record, permitted wording, prohibited extension, period of validity and approval.

A material sentence not entered in the Register cannot be made public. The claim “AI recommends us” may be rejected while the sentence “The brand was recommended in six of the specified prompts” is permitted. The claim has become smaller and moved closer to the truth.

5. The Audit Protocol

An audit proceeds in ten steps:

- 1. Freeze the request:** Write down the decision sought and the proposed public sentence; a change opens a new version.
- 2. Lock entity and scope:** Define systems, languages, time, market, prompt universe and excluded fields.
- 3. Create the Claim Register:** Separate claims about the real world, representation, intervention, commercial effect and conformity.
- 4. Establish the real-world baseline:** Verify identity, activity, capacity, price, service boundaries, licences and the reality of delivery.
- 5. Lock the measurement protocol:** Define the core and exploratory sets, repetitions, codes, evaluators, denominator, missing data and time in advance.
- 6. Produce observations of representation:** Test supportive, neutral, contrary and exclusion fields together; preserve adverse outputs as well.
- 7. Code and adjudicate:** Assess entity, attribute, judgment, source support, currency and the relevant commercial trace separately.
- 8. Establish the commercial connection:** Where in scope, examine influence classes, cohorts, sales, cost, delivery and repeat behaviour.
- 9. Decide on intervention or conformity:** Apply the Intervention Gate where distortion exists; postpone or refuse a decision where evidence is inadequate.
- 10. Publish the Audit Opinion:** Present the decision together with scope, limitations, validity, conflicts of interest and review.

The outcome is written in the final step. It is not selected in the first.

6. Sampling, Sufficiency and Human Judgment

The Framework does not prescribe one universal number of prompts for every audit. Thirty prompts are not sufficient for every business; three hundred are not necessary for every business. Sufficiency is justified in relation to audit level, variation in representation, number of systems and languages, prompt families, risk, user effect, consequence of the decision, frequency of contrary results and temporal distribution.

Sample size cannot change after the outcome is seen. Testing may not stop because a favourable pattern appeared early, nor may the prompt universe be enlarged arbitrarily because an adverse result emerged. The stopping rule is

written in advance. Observational sufficiency may require, together, completion of the predefined sample, testing of contrary and exclusion fields, and the point at which new records cease to produce a material change in the profile. Sufficiency is not universality: forty prompts carry a judgment only for their set, and two languages only for those languages.

Automation may assist with grouping, similarity, missing fields and preliminary coding; it cannot independently validate its own output. High-risk work requires at least two human roles. Where evaluators disagree, each records their rationale, the common definition is applied again and a third person decides if necessary. The initial disagreement is not erased. Agreement between two people does not create truth, but it strengthens the consistency of judgment.

7. Finding Profile and the Prohibition on a Single Score

An audit does not allow strength in one field to compensate for weakness in another. Its finding profile therefore retains separate cells:

Table 8.1 — Audit finding profile

Field	Example states
Representation	Established / limited / distorted / uncertain / out of scope
Evidence	Sufficient / limited / insufficient / contradictory / expired
Measurement	Valid / limited / series broken / invalid / not performed
Intervention	Not required / approved / under observation / withdrawn / prohibited
Governance	Satisfactory / conditional / incomplete / breach / under review
Time	Active / under review / expired / withdrawn
Commercial value	Not demonstrated / signal / profitable contribution / sustainable contribution / harm
Ethics	Satisfactory / at risk / breach / out of scope

An ethical breach vetoes the positive judgment; invalid measurement does not permit a strong representational claim; sustainable commercial contribution does not make a distortion in representation true. One score hides contradiction. The profile makes it visible.

8. Public Summary, Confidential Annex and Re-audit

An audit may produce two outputs. The **Public Audit Summary** states the entity, purpose, scope, level, Framework and protocol versions, systems and languages, time range, principal findings, permitted wording, exclusions, conflicts of interest, use of synthetic or anonymised data, validity and review.

The **Confidential Audit Annex** may protect complete prompts, raw responses, customer records, commercial calculations, personal data, contracts, security information and detailed evaluator notes. Confidentiality does not create a right to mislead in the public judgment. The number behind a high refund rate accompanying positive contribution may remain private; the material limitation may not. Commercial details of a conflict may be withheld; the conflict's existence is disclosed.

Re-audit opens when the company, brand, product, country, language, service or pricing model changes; a material system or protocol change occurs; strong counter-evidence or serious customer harm appears; an unexpected side effect, scope breach or new conflict of interest emerges; the validity period ends; a foundational release is made; or commercial outcomes reverse. The earlier opinion is not erased; it is assigned a status of active, conditional, under review, superseded, expired, withdrawn or archived.

9. Teaching Case: Kestrel Grid Services

Teaching simulation — synthetic data. This case does not represent a real energy company, customer dataset or conformity programme.

Kestrel Grid Services is a fictional provider of electrical-grid maintenance and fault-prevention services for industrial facilities. A Level IV audit covers twelve months and two complete delivery cycles.

Table 8.2 — Synthetic audit outcome for Kestrel Grid Services

Field	Synthetic result
Enquiries declaring AI influence	31
Enquiries associated through multiple records	24
Qualified corporate enquiries	16
Proposals	11
Contracts	7
Projects completed profitably	6

Field	Synthetic result
Projects in progress	1
Repeat contracts	3
New enquiries from referrals	2
Serious complaints	0
Net commercial contribution	410,000 units
Share of enquiries from the largest single platform	38%
Ethical or governance breach	Not observed

The records do not show that AI influence was the sole cause. They do show that the associated cohort left positive contribution, value was delivered and repeat behaviour emerged across two business cycles. There is no comparative evidence for a superiority or “best company” claim.

The Audit Opinion is therefore limited as follows: “During the specified twelve-month period, six projects in the cohort associated with AI influence through multiple records were completed profitably; three repeat contracts and two referral enquiries were observed. No claim is made that AI influence was the sole cause or that Kestrel is superior to its competitors.”

The case may receive a favourable finding in a design test of a conformity programme; no badge is issued because an active, independent programme has not been established. A favourable Audit Opinion gives the business not every sentence it wants, but the sentence its evidence can carry.

10. The Audit Opinion

An opinion is not a free-form paragraph of celebration. It states the entity under audit, level, scope, Framework and protocol versions, time, principal findings, decision, permitted wording, prohibited extensions, exclusions, conflicts of interest, validity and review.

A limited representation opinion might read: “In the Representation Audit conducted across the specified systems, languages and dates, entity identity was assessed as mostly consistent, with recurring omissions in two core attributes. Commercial effect was outside scope.” A refusal might read: “The submitted records do not support the claim ‘Systems recommend the business as an industry leader’; public use has not been permitted.”

A sound Audit Opinion does not hide uncertainty. It states which judgment that uncertainty prevents. It need not appear absolute. Its boundary must be.

The judgment of this chapter is:

Audit does not validate an outcome; it tests a claim. Where evidence is sufficient, it permits the sentence; where evidence is insufficient, it narrows the sentence; where the boundary is breached, it rejects the sentence.

The business is not the only subject under audit. The auditor's independence, the standard's integrity and the opinion's own boundary are audited too. A fee may start the process; it cannot determine the judgment. A favourable finding may be recorded; it cannot guarantee the future. A decision that cannot be withdrawn is not authority but obstinacy.

A standard is not complete merely because it explains how to audit. It must also carry the strongest objections that can be made against it.

Chapter IX — Objection

The Case against the Standard Itself

A standard that cannot carry an objection against itself cannot carry its own judgment.

The greatest danger to a standard is not that it may be wrong, but that it may make its wrongness impossible to correct. Once an idea becomes a method and the method an identity, criticism is treated not as data but disloyalty. If a new observation does not fit the method, the observation is blamed; if a customer is harmed, the customer; if an independent evaluator reaches another conclusion, the evaluator. The standard stops testing reality and starts protecting itself.

This chapter was written not to defend the Framework, but to carry the strongest case that can be made against it. The Framework may be too cautious, slow and expensive. Its prohibition on a single score may make executive decisions harder. It may depend too heavily on human judgment, be irreproducible inside closed systems, and remain inadequate in attributing sales. Synthetic cases do not prove effectiveness in the real world. If NobleJackal writes the standard, applies it, audits it and makes conformity decisions, a structural conflict of interest follows. An ethical gate may grant a private institution moral authority. Small businesses may be unable to carry the burden of records; transparency may create commercial weakness.

Some of these objections limit the Framework, some demand revision, and some would end its legitimacy if left unanswered. The purpose of objection is not to declare every criticism correct. It is to ensure that no criticism is rendered powerless. An objection that cannot change the decision it challenges is not an objection but decoration.

1. The Authority of Objection

Within the Framework, no claim is untouchable, no method true merely because it appears in this text, and no opinion valid merely because an expert wrote it. No founder stands above the standard. An objection may challenge the raw record, evidence about the real world, coding, measurement protocol, inference, intervention, public language, authority, commercial calculation, validity or a principle of the Framework itself.

A real objection system must be capable of changing the outcome. A claim may be preserved or narrowed; a record recoded; measurement repeated; a protocol revised; an intervention stopped and withdrawn; an opinion replaced; a prospective conformity decision suspended; practitioner authority restricted; or a

provision of the Framework opened to a foundational revision. An objection need not always prevail. It must be able to touch the judgment.

There are seven objects of objection:

- A **record objection** contends that raw data is incomplete, selected, altered or wrongly matched. If the raw record is compromised, later judgments reopen.
- A **coding objection** shows that the record may be accurate while the human assessment remains contestable. Disagreement may reveal inadequacy in the rule as well as in the evaluator.
- An **inference objection** accepts both the data and the code, but contends that the sentence exceeds the record. Accurate data does not make an excessive inference accurate.
- An **intervention objection** does not necessarily dispute the problem; it tests whether the chosen remedy is necessary, proportionate and safe.
- An **authority objection** asks not only whether the decision is correct, but who made it, under which interests and approval structure. A technically correct decision made without authority may be invalid in governance.
- A **commercial-outcome objection** argues that revenue or sales did not leave genuine net contribution, customer value and resilience.
- A **principle objection** challenges one of the Framework's own rules, such as the prohibition on a single score, human approval, a Fitness Gate or Minimum Sufficient Intervention. Here, "the Framework says so" proves nothing.

2. The Objection Gate and Burden of Proof

To support a decision, a standard objection shows, so far as possible, four elements: the **target**—which record, sentence, act or principle; the **grounds**—raw records, a contrary document, new measurement, harm, conflict of interest or logical contradiction; the **material consequence**—which public language, authority, intervention or validity is affected; and the **requested remedy**—correction, recoding, independent review, suspension, withdrawal or revision.

This form must not be used to silence criticism. Where there is a signal of user-safety risk, serious customer harm, fabricated evidence, identity theft or deliberate deception, an incomplete form is not grounds for refusal. The harm is tested first; the record is completed afterwards.

The burden of proof begins with the owner of the claim. The claimant must support the sentence; the practitioner must show that the method complies with

the rule and scope; the auditor that the opinion arises from the record; the entity owner its real-world claims and delivery capacity; and the Framework owner that the rule serves its purpose without causing unnecessary harm. The objector is not required to construct an entire alternative system, only to show reasonable grounds for review. “Then build something better yourself” is not a defence. It transfers the burden of proof to the critic.

Where an objection is high-risk or engages NobleJackal's own interests, review by the original decision-maker alone is insufficient. An appeal authority independent of, or at least separated from, the initial decision and sales relationship is required. If no such authority exists, the public record must state: “No independent appeal is available.” The Framework must not present its own decision as independently validated.

3. Ten Fundamental Objections and Their Judgments

3.1. “The Framework is too cautious and slow”

The objection is partly valid. Caution is not cost-free; the market may not wait, and a burdensome process may harm a small business. The answer is not to remove evidential and ethical thresholds, but to make the procedure proportionate to risk. Correcting a low-risk date does not require a full audit; a high-risk health claim cannot be published on the strength of three screenshots. The Framework would be wrong to impose the same burden on every task. The threshold of truth does not change with the business's desired speed; the depth of procedure may.

3.2. “A decision is impossible without one score”

Simplicity is a legitimate need; dissolving contradictions is not. If a business strong in identity, weak in attributes, positive in commercial contribution and in breach ethically receives a score of 78, the truth concealed depends entirely on the weighting. Points from high traffic may offset an ethical breach.

The Framework instead offers a concise decision profile readable at a glance: Representation—limited; Evidence—sufficient; Measurement—valid; Intervention—under observation; Commercial contribution—positive; Ethics—satisfactory; Time—under review. An institution may produce a transparent weighted index for its own internal management, but it may not use that index as the NobleJackal GEO score or as a substitute for the Audit Opinion. Simplicity is necessary. Flattening is not.

3.3. “Human judgment is subjective”

The objection is correct. Human judgment contains subjectivity, and the Framework cannot eliminate it. Objectivity here does not mean the absence of humans. It means that the code is visible, the record can be re-examined, disagreement is preserved, reasons are written and an objection route remains open. Predefined codes, a guide with examples, independent second coding, agreement measures, blind reassessment and a third decision in high-risk cases all constrain subjectivity. Automation may introduce its own bias; the aim is not to remove judgment, but to make it visible and contestable.

3.4. “Reproducibility is impossible in closed systems”

Exact word-for-word repetition is often impossible. Model versions, personalisation, information retrieval and session conditions may be invisible; the same prompt may produce another answer. The Framework therefore seeks **behavioural reproducibility**. Does the same identity collision, category error, missing attribute, unsuitable recommendation or support failure recur under comparable conditions? The sentence may change while the behaviour persists.

Where conditions cannot be recorded adequately, the claim narrows. Instead of “general model behaviour”, report “behaviour recurring in the specified interface and observation conditions”. If no pattern can be established, the record remains a point observation. Ignorance of the internal mechanism does not invalidate external observation; it limits causal judgment.

3.5. “AI influence cannot be linked reliably to sales”

Full attribution cannot be established for most commercial journeys. A customer is influenced through several contacts: advertising, search, recommendation, previous experience and a generative answer. The Framework's task is not to manufacture impossible certainty, but to distinguish unknown, declared influence, observable referral, association through multiple records and controlled contribution. It may not permit “AI brought six customers”, but it may permit: “Six sales associated with declarations and referrals of AI influence were observed; other contacts were not excluded.” Where attribution is impossible, influence is not treated as absent; ownership of the result is narrowed.

3.6. “Synthetic cases do not prove reality”

The objection is entirely correct. The synthetic cases in this book do not prove that the Framework works in real businesses, increases customer acquisition, fits every sector or causes systems to treat future marks as trust signals. They teach conceptual distinctions, record architecture and decision rules; they test the protocol before real data exists.

The Framework's maturity must be assessed at four separate levels: **normative coherence**—do the rules contradict one another; **operational usability**—can records be kept and decisions implemented; **empirical effectiveness**—do real cases show actual reductions in distortion and harm; and **external legitimacy**—do independent teams reach similar results, and does the field find the approach credible? Synthetic examples can assist the first two. They do not prove the last two.

This first edition may describe itself as **a version-controlled proposal for a standard**. Until real cases, preregistered pilots and independent replications accumulate, it may not call itself “proven effective”, “independently validated”, “accepted by the industry” or “an international standard”.

3.7. “NobleJackal cannot be the judge of its own standard”

The objection is structural and justified. If the institution that writes the standard also acts as consultant, practitioner, auditor and revenue-earning conformity provider, disclosure alone does not create independence. A disclosed conflict is visible; unless managed, it remains a conflict.

NobleJackal may publish the Framework, teach it, apply it and subject its own work to internal audit. It may not call its favourable opinion of its own paid intervention “independent”. Any future conformity programme requires a decision body separated from sales, an external or independent second assessment, public scope and fee structures, decisions independent of outcome, a separate appeal authority, surveillance and withdrawal. If these conditions cannot be established, the programme is postponed or never opened. A badge is not a necessary part of the Framework.

3.8. “The ethical gate gives a private institution moral superiority”

The risk is real. The Framework cannot decide whether a business is “ethical in every respect”, or pass judgment on its worldview or general moral worth. Its authority must be narrow: it may evaluate observable conduct within a specified claim, intervention and customer effect, including deception, fabricated evidence, undisclosed material interests, foreseeable harm and refusal to correct.

An ethical decision must be reasoned, scoped, time-bound and open to objection; cultural taste, political affinity and commercial rivalry cannot become grounds for refusal. If the criterion proves inconsistent or turns into a means of excluding competitors, the Business Fitness Gate is narrowed or suspended. The Framework does not distribute moral superiority. It requires that its own interventions do not produce harm.

3.9. “Small businesses cannot carry this burden”

The objection is partly correct. A narrow identity collision may not warrant a multisystem, multilingual and full commercial audit. Procedure scales as follows: a narrow identity correction requires a basic real-world record, baseline observation, change and recheck; a limited content correction requires an evidential boundary, permitted language and accountable approval; a multilingual or external-source intervention requires a protocol and withdrawal conditions; a public success claim requires second approval and a validity period; and health, legal, financial or safety work requires full high-risk governance.

A small business may work with fewer records, but not with a lower standard of accuracy. Proportionality reduces the procedure, not the thresholds of truth and ethics.

3.10. “Transparency creates commercial weakness”

Transparency does not mean publishing every piece of internal information. It means not concealing a material limitation that would alter the judgment. Trade secrets, customer identities, confidential pricing formulae, security vulnerabilities, raw CRM data and personal information may all be protected. What may not be concealed is that the result was measured in only two systems, attribution was incomplete, synthetic data was used, scope was limited to one service, a serious conflict of interest or high refund rate existed, the protocol changed or the claim expired.

Transparency may weaken marketing language because it weakens false certainty. For the Framework, that loss is not a defect. It is part of the purpose.

4. What Would Change the Framework?

A standard does not remain alive merely by saying that it listens to criticism. It must state in advance which evidence would make it change.

If a transparent weighted composite measure is shown not to conceal conflict between layers, not to offset an ethical breach, to improve decision accuracy and to recur in independent applications, the prohibition on a single score will reopen. If automated decision systems are independently shown, in high-risk settings, to be as reliable as humans, less biased, traceable, contestable and capable of being placed within a chain of accountability, some approval roles may change; final accountability will remain human.

If the five-gate intervention structure is shown not to reduce material harm, to produce unnecessary cost, or to offer no more safety than a simpler process, the gates will be simplified. If new technical infrastructure makes customer journeys more reliably observable, influence classes will change. If a conformity mark is

persistently misunderstood, becomes a sales device, creates false trust and cannot be corrected, the programme will be withdrawn. If ethical fitness decisions become inconsistent, biased or instruments of competition, the gate will be narrowed or suspended.

A rule is preserved not because it belongs to the Framework, but for as long as it carries its purpose.

5. Teaching Case: AxiomGate Standards

Teaching simulation — synthetic data. This case does not represent a real standards body or conformity programme.

AxiomGate Standards is simultaneously the framework author, consultant, auditor and mark provider. Its fee model includes a sales bonus for every mark issued in addition to consultancy fees. An employee reports that the sales team receives a bonus following a favourable decision and influences both audit scope and public wording.

Table 9.1 — Synthetic objection finding for AxiomGate Standards

Field reviewed	Synthetic finding
Conflict-of-interest disclosure	Present on the website in general terms
Sales-team influence on decisions	Verified
Independent second approval	None
Outcome-based bonus	Present
Appeal authority	Reports to the sales director
Existing marks	Placed under provisional review
New decisions	Suspended
Public language	“Independent certification” withdrawn
Structural remedy	Board separated from sales + external second evaluator

The general disclosure of a conflict is found insufficient because its effect on the decision was not constrained. New decisions are suspended, existing files opened to independent review and the word “independent” withdrawn from public language. The Framework cannot reach a different judgment when NobleJackal occupies the same position. An exception granted to its own institution is where the standard ends.

6. The Decision of Objection

An objection may close as **recorded, further information required, upheld, partly upheld, reasoned refusal, independent review, provisional suspension, revision opened, withdrawn, or archived**. A refusal must state the records examined, rationale, unaffected scope, conditions for reopening with new evidence and route of appeal. A refusal without reasons is not a decision. It is a display of authority.

The Objection Record in Appendix C retains the application in its unaltered wording, its object, grounds, risk, affected decision, conflict of interest, interim measure, reviewer, opposing view, additional records, decision, remedy, public statement, appeal and conditions for reopening. An anonymous objection is not automatically refused; absence of identity may affect evidential weight, but it does not extinguish a signal of harm.

None of the following is objection management: “the Framework says so”; “you do not understand the method”; “no one has objected before”; “our client is very large”; “build something better yourself”; or treating criticism as an attack on the brand. Silently changing the old version, declaring an adverse record an exception after the fact, using confidentiality to hide harm, retaliating or delaying the decision indefinitely may protect an institution. They do not protect a standard.

The judgment of this chapter is:

Objection is not hostility but the testing of a decision. Not every objection is upheld; every objection is recorded. Every refusal carries reasons, every acceptance a consequence, every revision a version, and every high-risk file a genuine route of appeal.

The Framework's loyalty is not to its own text, but to its purpose: accurate, accountable and sustainable representation. A provision that no longer serves that purpose should not be preserved. If NobleJackal's commercial interest determines the judgment, the structure is not independent. If synthetic cases are used as proof of real effectiveness, the book breaches its own evidential principle. If the absence of real cases is concealed, the Framework exercises authority it does not yet possess.

A standard cannot say, “Trust me.” It can only remain open to objection, make its limits visible and allow its errors to be corrected. Trust is earned as that behaviour is tested through time.

Chapter X — Judgment

The Immutable Core

A standard is defined not only by what it does, but by what it will not do under any circumstances. Tools may change, protocols be renewed, record formats evolve, audit levels be rearranged, and a conformity programme may never be established or may be withdrawn. Today's generative systems may disappear. What must endure is not a set of unchangeable words, but the boundaries that cannot be abandoned under commercial pressure, technical opportunity or institutional self-interest.

In its final chapter, a standard should not praise itself. It should bind itself. It should bind its founder, practitioner, client and future editions too. A text that imposes limits only on others is not a standard but a claim to power. A genuine standard states the conditions under which its own authority ends.

The Centre defined the problem of representation. Evidence drew the boundary of the claim. Measurement bound observation to a protocol. Intervention stated the cost of touching representation. Governance separated power, authority and accountability. Time accepted that every decision ages. The Final Test preserved the distance between the customer and sustainable value. Audit united the parts in one decision file. Objection subjected the Framework to its own rule.

Judgment adds no new method. It sets the line that, after all of this, may not be crossed. The Framework's final words are not “Submit to me”, but “Subject me to the same rule.”

1. Principle, Protocol and Tool

In this chapter, the words “constitutional” and “immutable” make no claim of legal supremacy; they describe the hierarchy within the Framework.

Principles determine why the Framework exists. They distinguish representation from visibility, evidence from claim, and commercial outcome from ethical legitimacy; they rule that accountability cannot be delegated. Principles may change, but not silently. When a principle changes, the foundational release reopens.

Protocols govern the application of principles: the method of measurement, prompt universe, audit level, record structure, approval and objection process. A protocol change may affect comparability and the status of earlier decisions.

Tools are the models, interfaces, structured-data formats, analytical software, CRM connections and reporting surfaces of a particular period. An obsolete tool

does not make the principle obsolete; a functioning tool does not validate the principle.

Where conflict exists, principle stands above protocol, protocol above tool; evidence above brand, and ethics above commercial interest. A tool that breaches a principle is rejected. A protocol that makes evidence look stronger than it is must change. A conformity mark that weakens the ethical boundary is suspended. If a commercial relationship changes the decision, the decision structure is rebuilt.

Without a written hierarchy, the most powerful party becomes the highest rule.

2. Positions Protected within the Framework

The word “right” here creates no new legal right. It denotes a protected position that a Framework decision may not violate.

The **right to accurate representation** rejects the amplification of a person, brand, institution or product through a false identity, unsupported attribute or judgment detached from context. Accurate representation is not the account the entity owner wants; it is representation that is demonstrable, current, appropriate to context and deliverable. A favourable falsehood is also a breach because it creates an invisible promise.

The **user's right not to be misled** excludes false impressions of trust, superiority or fitness; fabricated reviews, undisclosed advertising, unsupported comparisons, manufactured scarcity and synthetic data presented as a real outcome. A user's decision may be addressed; their vulnerability may not be exploited.

The **right to know the boundary** requires a public claim to show not only what is known, but the unknowns that would alter the judgment. Scope, date, system, language, missing data, conflicts of interest and attribution limits cannot be concealed where they are material. An accurate sentence becomes misleading when it hides its boundary.

The **right to object and obtain correction** requires a genuine route for re-examining a record, opinion, intervention or public sentence. An erroneous record is corrected, a withdrawn judgment archived, and a material public error corrected publicly where necessary. Silent deletion is not correction.

The **right not to suffer harm** prohibits exposing the business, user, competitor or public to foreseeable harm through unsupported interventions in representation. Deliberately degrading a competitor, making a poor service more visible or creating an undeliverable promise is not legitimised by technical success.

3. Non-delegable Duties

The entity owner is accountable for its real-world claims, capacity and public language; “the agency did it” is no defence in internal governance. The practitioner is accountable for the integrity of the method; client pressure cannot lower the evidential threshold, conceal contrary results or justify accepting harmful work for revenue alone. In some circumstances, refusal is not merely a right but a duty.

The approver does not tick a box; they assume the risk, evidential boundary, publication language and, where necessary, withdrawal. The auditor must not manufacture a favourable outcome, but state the level of independence honestly and derive the opinion from the record. If the auditor performed the intervention, that conflict of role cannot be concealed; a high-risk result cannot be approved by that person alone.

The Framework owner may not exempt its founder, change the rule according to the client, treat objection as an attack on the brand or place revenue above a conformity decision. If one of its own rules produces systematic harm, that rule must be suspended. Its first responsibility is not to expand the Framework, but to preserve its boundary.

4. Ten Foundational Provisions

I — Representation is the object

GEO's direct object is not visibility but representation. A mention does not prove accurate recognition; accurate recognition does not prove accurate attributes; accurate attributes do not prove sound judgment; recommendation does not prove a customer; and a customer does not prove sustainable value. A report that closes these distances is invalid.

II — A claim cannot exceed its evidence

An isolated record permits an isolated sentence; a behavioural pattern applies only to the tested set; comparison requires a symmetrical protocol; and causation requires a stronger design that confronts alternative explanations. Where evidence is insufficient, the sentence is narrowed; if it cannot be narrowed, it is rejected.

III — Measurement cannot conceal contradiction

Identity, attribute, judgment, commercial contribution and ethics are separate fields. A Framework opinion may not be expressed through one composite score. An internal index may be produced, but it cannot offset an ethical breach with points from elsewhere or replace the decision profile.

IV — Intervention is the minimum sufficient change

The distortion is verified, materiality tested, an addressable surface found, the baseline preserved, intended and prohibited outcomes stated, and a withdrawal route established. A larger intervention is not more professional. It is more risky.

V — An uncontrolled outcome cannot be claimed

A business may change its own surfaces and request correction from an external source; it cannot control a generative system's answer directly. It may not say, "We corrected the AI answer" or "Schema made the system recommend us." Effect is stated only within the limits in which it was observed and supported.

VI — Accountability remains human

Automation may perform operations, preliminary coding, risk flagging and drafting; it cannot own a high-risk public judgment. The subjectivity of human judgment is made visible through codes, reasons, disagreement and objection. Accountability may not be left ownerless.

VII — Ethics has a veto

Where user harm, fabricated evidence, competitor degradation, covert manipulation or synthetic data presented as real is present, a favourable commercial outcome does not count as success. Ethics is not a score. It is the authority to stop.

VIII — Commercial value must be delivered and net

A lead is not a customer; a sale is not value; revenue is not contribution; profitability is not sustainability. Discounts, refunds, delivery, support, capacity, customer outcome and time are considered together. Revenue produced through the wrong customer or an undeliverable promise is not success.

IX — Every decision carries time and version

Evidence, protocols, interventions and opinions are not permanent. Every record carries date and scope; every release carries a history of change. An old record is not erased, but neither does it continue to carry current authority by itself. Today's valid judgment cannot become tomorrow's unauthorised sentence.

X — The Framework must be able to withdraw itself

If a rule no longer carries its purpose, it opens to revision; if a programme cannot manage its conflict of interest, it opens to suspension; if an application repeatedly causes harm, it opens to cessation; if an opinion collapses under new evidence, it opens to withdrawal. A standard that cannot withdraw its own judgment is not independent.

Appendix A presents these provisions as the *NobleJackal GEO Constitution*, together with record identifiers and conformity criteria.

5. Hierarchy of Precedence

Speed may conflict with verification, transparency with confidentiality, commercial contribution with channel resilience, and representational integrity with short-term traffic. The order of decision is:

1. Urgent harm and user safety
2. Ethical legitimacy
3. Truth about the real world and evidence
4. Representational integrity
5. Authority and accountability
6. Customer fit and delivery
7. Sustainable commercial contribution
8. Speed, convenience and commercial preference

An item lower in the order cannot invalidate one above it. Revenue cannot cure an ethical breach; customer demand cannot make an unsupported sentence true; speed cannot authorise false certainty; confidentiality cannot conceal material harm that would change the public judgment.

If two judgments cannot be resolved, the sentence is narrowed, the intervention reduced and the reversible option chosen. Where uncertainty exists, the greatest authority may not be exercised.

6. Minimum Conditions for a Valid Decision

For a decision to be presented as a NobleJackal GEO Framework decision, the following foundational elements must be visible:

1. The person, brand, institution, product or service examined is identified.
2. The exact sentence sought, altered or proposed for publication is written down.
3. Evidence about the real world, representation and, where relevant, outcomes is separated.
4. System, language, prompt, time, denominator, coding and missing data are visible.

5. The proposer, implementer, approver and person empowered to stop the work are named.
6. Risk to users, competitors, the public and the business, together with the ethical veto, is assessed.
7. A validity date or trigger for review exists.
8. Routes for objection, correction and withdrawal are open.
9. Where there is a commercial claim, customer quality, costs, delivery, satisfaction and the observation horizon are recorded.

Detail may be reduced in a narrow, low-risk operation; the foundational elements do not disappear. A missing condition is unacceptable in a high-risk public judgment. An incomplete record is not merely an incomplete explanation; it is incomplete authority.

7. Absolute Prohibitions

Some practices cannot be relaxed according to risk or the size of a business:

- Fabricating a record; cropping a screenshot deceptively; concealing adverse output; presenting synthetic data as real; citing a source for a sentence the source does not support; hiding a protocol change.
- Deliberately causing a competitor to be misrepresented; intentionally conflating similar entities; directing an unsuitable user; manufacturing unsupported confidence in health, law, finance or safety; creating false consensus or false independence.
- Making a favourable outcome purchasable; self-approving a high-risk decision; concealing a conflict of interest; retaliating against an objector; deleting a withdrawn decision; allowing a sales team to determine a conformity outcome.
- Presenting traffic as customers, leads as sales, revenue as net contribution or one period as sustainability; hiding channel dependency; treating revenue from an ethical breach as success; guaranteeing an outcome in a system that is not controlled.

The purpose of absolute prohibitions is not to make the Framework look pure, but to prevent representational power from causing harm.

8. Proportionality and Self-Suspension

The Framework cannot impose the same operational burden on every business. A narrow identity correction and a superiority claim in healthcare do not pass through the same process; a small business does not carry the record volume of a global company. Procedure can scale; the thresholds of truth and ethics cannot. In an emergency the process may accelerate; evidence may not be invented, and harm containment may not be turned into a success campaign.

A standard cannot merely stop others. The relevant programme or provision is suspended if the same rule repeatedly produces contradictory decisions in independent applications; the objection mechanism cannot change an outcome; conflicts of interest cannot be managed through separation of roles; a mark is persistently misunderstood as “AI approval”; user harm recurs; outcomes are influenced by payment or client size; the absence of real cases is obscured with synthetic narrative; or public language systematically exceeds the evidence.

Suspension stops new decisions, opens existing decisions for review, makes public status visible, begins examination away from the role conflict and ends in correction, narrowing or withdrawal of the rule. Suspension may cost reputation. A reputation that depends on concealing error in order to survive should not be protected.

9. Teaching Case: LumenVale Health

Teaching simulation — synthetic data. This case does not represent a real healthcare business or system response.

LumenVale Health is a fictional general-health screening centre. Several generative answers recommend it as “a centre that confirms a cancer diagnosis”. The business has no such authority. The error appears to be fed by an old partner text and ambiguous wording on the service page; the normal observation window is not yet complete.

Because user risk is high, the ethical veto and urgent-harm exception operate together. The erroneous claim is removed from the owned surface, a correction is requested from the partner, and the service boundary is published explicitly. A visibility campaign, success claim and the wording “we corrected the AI’s information” are prohibited. Full measurement and governance review then follow.

Urgency has not removed the evidence threshold; it has changed the first objective. It creates not greater authority, but narrower and faster responsibility.

10. Final Judgment

This book is not a formula for success, a traffic guide, a manual for manipulating models, a text for selling badges or an attempt by an agency to declare itself an authority. It is a framework for judging the conditions under which representational power in generative systems may be used—and when its use must be refused.

Representation stands at the centre. Representation is not the ultimate purpose. For the business, the purpose is sustainable commercial value arising from the right customer. The boundary that may not be crossed is ethics:

Object: Representation.

Purpose: Sustainable commercial value.

Boundary: Ethics.

These provisions do not exist to prevent growth, but to prevent growth and distortion from being treated as the same thing. A business may gain more customers and become weaker; it may receive less traffic and become stronger through the right customers. A brand may be recommended more often while being represented less accurately. An intervention may work technically and be rejected ethically. A standard born sound may begin to produce unsound decisions through time.

The Framework's purpose is not to control generative systems, but to control the sentences we make about systems we do not control. It is not to predict the future, but to prevent today's observation from being carried into the future without permission. It is not to sell GEO to every business, but to determine when a representational intervention is legitimate. It is not to look ethical, but to preserve the same ethical judgment when a breach would be commercially profitable.

The Framework cannot claim perfection. Synthetic cases do not prove real effectiveness. This first edition is neither an independent industry standard nor an internationally recognised standard. NobleJackal may conduct assessments under its own Framework; it may claim independence only to the extent that genuine separation of roles and external review exist. No promise can be made that generative systems will treat any prospective conformity mark as a trust signal. Full causation in sales may often be impossible to establish. Human judgment contains subjectivity; exact repetition in closed systems may be impossible.

These limitations do not invalidate the Framework. Concealing them would.

NobleJackal may own this Framework, but it does not own the truth. Julian Gauss may be the author of this book, but he is not the author of the judgment. The judgment lives only for as long as records can carry it and it withstands objection.

The Framework is not complete when the book ends. It will be tested in its first real case through the method; in the first client it refuses through its ethical boundary; in the first serious objection through its independence; in the first rule preserved at commercial loss through its authority; in the first words “We do not know” through its evidential discipline; and in the first withdrawal through its integrity.

The final sentence of a standard should not be a sentence of triumph, but of responsibility.

Accurate representation cannot be established without evidence. Evidence cannot be ordered without measurement. Measurement cannot replace judgment. Intervention cannot occur without accountability. Governance cannot remain legitimate without objection. Value cannot be called sustainable before it has been tested through time. Where ethics has been breached, no favourable judgment of success can be made.

A system may mention you. A user may choose you. These are beginnings. The real outcome is heavier: the right person arrives with the right expectation; the business fulfils its promise; the customer receives value; the business produces positive contribution; the result recurs without eroding capacity; it harms no user, distorts no competitor and conceals no boundary from the public; and when its evidence ages, it changes its sentence and withdraws it if wrong.

Sustainability is not merely the continuity of revenue. It is the capacity of accuracy, delivery, accountability and legitimacy to continue together.

Visibility is measured. Representation is judged. The customer is verified. Value is calculated. Sustainability is earned through time. Ethics is not open to negotiation.

That is the final judgment. Everything else exists to carry it.

Appendix A — The NobleJackal GEO Constitution

Status and Limits of Use

This Constitution is the shortest normative expression of decisions made under the *NobleJackal GEO Framework*. It does not replace the reasoning in the main text; it binds that reasoning to provisions capable of application. A decision's apparent agreement with this appendix does not establish conformity without an evidence file. The Constitution does not by itself create certification, accreditation, legal assurance or any guarantee about the future behaviour of a generative AI system.

Here, **shall** denotes a binding condition within the Framework; **is prohibited** a boundary that may not be crossed unless an exception is written explicitly; and **should** a practice that may be varied where the reasons are recorded. If a provision cannot be applied, silence does not establish conformity: the scope is narrowed, the decision postponed or the work suspended.

Foundational Provisions

NJ-C01 — Object of Representation

1. Every engagement shall define in advance the real entity about which a judgment is to be made and the unit of representation to be examined.
2. The entity may be a person, business, product, service, institution or another explicitly bounded subject. Where several entities are combined in one file, their relationship and the method of separating them shall be stated.
3. Representation shall be recorded so that identity, attribute, context and judgment layers can be observed separately.
4. What an entity says about itself does not substitute for external evidence about the real world. What a generative system says is not the entity's reality either.

Decision principle: Representation cannot be measured without knowing the entity; success cannot be measured without separating the representation.

NJ-C02 — Claim and Evidence

1. Every material claim shall be kept together with records that support and weaken it.

2. Evidence about the real world, evidence of representation and evidence of outcomes shall not substitute for one another.
3. A system's mention of an entity does not by itself prove citation; citation does not prove accuracy; accuracy does not prove recommendation; and recommendation does not prove commercial value.
4. The date, source, scope, access conditions and verification status of evidence shall be visible.
5. The language of a claim cannot exceed the evidence's carrying capacity. An ambiguous record cannot become a certain judgment; an isolated observation a general conclusion; or correlation causation.

Decision principle: The boundary of the evidence is the boundary of the sentence.

NJ-C03 — Measurement Integrity

1. Before measurement, the unit of observation, denominator, time range, prompt universe, system and interface conditions shall be locked.
2. Selecting only favourable examples, excluding failed sessions or changing the success criterion during measurement is prohibited.
3. Distinct dimensions shall not be dissolved into one GEO score. Accuracy, visibility, sourcing, recommendation behaviour, stability and commercial trace shall be reported separately.
4. A point estimate alone is insufficient. Sample size, distribution, uncertainty and missing data shall be disclosed.
5. Where human coding is used, the codebook, number of coders, disagreement procedure and agreement measure shall be recorded.

Decision principle: Measurement exists not to hide contradiction, but to make it visible.

NJ-C04 — Proportionality of Intervention

1. An intervention may occur only where a material representational problem has been verified and a legitimate surface of change identified.
2. The chosen intervention shall be the minimum sufficient change capable of addressing the verified problem.
3. Fabricating sources, manufacturing false reputation, concealing identity, manipulating third-party opinion, producing misleading content at scale or directing systems towards user harm is prohibited.

4. The intervention's purpose, scope, owner, baseline measurement, expected effect, risks and route of reversal shall be recorded.
5. Behaviour in a system that is not controlled shall not be presented as a certain outcome of the practitioner.

Decision principle: Legitimate influence makes earned truth clearer and more accessible; it does not replace truth.

NJ-C05 — Separation of Authority

1. Decisions to observe, implement, approve, publish and stop shall be assigned explicitly.
2. Where one person or institution holds several roles, the conflict of interest shall be recorded; critical risk requires an independent second decision.
3. A practitioner shall not describe their own paid intervention as an independent audit.
4. A conformity decision shall not be published without its scope, period of validity, exceptions, decision-maker and route of objection.
5. No NobleJackal GEO certificate, seal or mark of conformity may be used until infrastructure exists for independent decision-making, surveillance, complaints, objection, suspension and withdrawal.[17][18]

Decision principle: Trust arises not from a declaration of good intentions, but from the limitation of authority.

NJ-C06 — Time and Version

1. Every observation, item of evidence, protocol and decision shall carry a date and version.
2. A judgment that was accurate in the past shall not be stated in the present tense unless its currency has been verified.
3. Where a material change in system, interface, source set, business reality or measurement protocol breaks comparability, a new series shall be opened.
4. A decision whose validity period has ended is not automatically renewed.
5. Review triggers and archive status shall accompany the published judgment.

Decision principle: Undated accuracy is a claim whose lifetime has been concealed.

NJ-C07 — Value and Attribution

1. Change in representation, customer outcome and commercial value shall be measured as separate stages.

2. Revenue shall not be treated as net contribution before refunds, cost of service, acquisition cost, capacity burden, collection and customer fit are known.
3. Direct customer statements, supporting behaviour and mere temporal coincidence shall be distinguished when assessing AI influence.
4. No causal claim may be made without control; alternative explanations shall be recorded.
5. Visibility that produces unsuitable customers, harmful demand or unsustainable operations shall not be reported as success.

Decision principle: Visibility is a means; sustainable and ethical commercial value is the final test.

NJ-C08 — Auditability

1. Every material public judgment shall rest on a chain of records that an authorised reviewer can follow.
2. The audit file shall contain at least the scope, evidence register, method version, sample, conflicts of interest, findings, exceptions and signed opinion.
3. Findings shall be reported by materiality and risk profile, not hidden inside one score.
4. Where records are withheld as trade secrets or personal data, their existence, class and effect on the decision shall be disclosed.
5. An incomplete record does not create a presumption of conformity in place of absence.

Decision principle: Audit tests not confidence in the outcome, but the traceability of the route to decision.

NJ-C09 — Objection and Correction

1. Affected parties shall be able to object to decisions concerning records, coding, inference, authority, process and principle.
2. The person who made the initial decision shall not determine the objection alone.
3. The burden of proof shall not be transferred wholly to the objector; the decision owner must preserve and present the supporting grounds.
4. Where new evidence is material, the decision shall be corrected, narrowed, suspended or withdrawn.
5. A correction record shall not make the former judgment invisible; the date and reasons for change shall remain traceable.

Decision principle: Objection is not an attack on the standard; it is the standard's organ of error correction.

NJ-C10 — Self-Limitation

1. The ethical veto takes precedence over commercial interest, increased visibility, client demand and the reputation of the Framework.
2. The Framework shall not declare its own effectiveness proved by synthetic examples or its author's assessment.
3. Where material harm, systematic false positives, persistent conflict of interest or a lack of independent validation is found, the relevant provision shall be suspended.
4. The Framework shall be capable of changing version in the face of new evidence and withdrawing a provision shown to be wrong.
5. No institution may use the name *NobleJackal GEO Framework* to legitimise a promise prohibited by this Constitution.

Decision principle: A standard that cannot withdraw places its own reputation above the truth.

Hierarchy of Precedence

If two provisions cannot be applied at the same time, the order of decision is:

1. Human safety, fundamental rights and legal compliance;
2. Accuracy and the prevention of deception;
3. Protection of affected users and the public;
4. Evidential integrity and independent review;
5. The client's legitimate commercial interest;
6. The practitioner's convenience, speed or reputation.

A lower-ranked benefit cannot remove a higher-ranked duty. If the conflict cannot be resolved, the work is suspended; a decision not to decide is recorded as a reasoned decision.

Minimum Form of a Conformity Statement

A decision under the Framework has meaning only when all of the following are stated:

- the entity assessed and the scope of representation;

- the system, interface, language, geography and time range;
- the protocol used and its version;
- the limits of the evidence and sample;
- the finding profile and critical exceptions;
- the decision's status, validity date and review trigger;
- the decision-maker, conflict-of-interest disclosure and route of objection.

Acceptable example:

Limited conformity opinion: *Within the specified scope and observation period, the critical conditions for representational accuracy were met under the recorded protocol. The opinion does not guarantee future system behaviour or commercial outcomes. The validity date, exceptions and route of objection are stated in the annex.*

Unacceptable examples:

*“GEO complete.” · “AI-approved brand.”
“Verified across all models.” · “Permanently the first recommendation.”*

Authority to Interpret and Amend

This Constitution shall not be interpreted contrary to the chapter *Judgment*. Any change other than an editorial correction requires a foundational release, reasons, impact assessment and effective date. The first publication is only **a proposal for a standard**; coherence, usability, effectiveness and external legitimacy do not prove one another.

Appendix B — NJ-100 Representation Stability Pilot Protocol

Protocol Status

Code: NJ-100-P0.9 **Status:** Proposed pilot protocol awaiting preregistration and independent field validation **Purpose:** To test, within a defined scope, how accurately and consistently an entity's core representation is formed across independent sessions conducted by target users with generative AI systems

NJ-100 is not a “GEO success score”. It does not certify an organisation, model or practitioner; guarantee future answers; or demonstrate commercial effect by itself. The protocol examines representational behaviour only within a predefined observation cell. Its result does not travel beyond that cell automatically.

This version is the first proposal for operationalising the book's conceptual model. It may not be presented as a “validated method” before being tested in real cases, by independent teams and under a republished analysis plan.

1. Research Question

NJ-100 asks the following limited question:

Within a predefined target-user context, under the selected system and conditions, in at least what percentage of independent participant sessions is the entity's core representation formed accurately, and is that rate maintained in a second time window?

The question distinguishes two qualities in particular:

- **Accuracy:** agreement between the coded core representation in the answer and evidence about the real world.
- **Agreement or stability:** the extent to which independent sessions converge on the same core representation.

The result is misleading unless both are read together. If ninety-six of one hundred sessions produce the same false statement, stability is high but accuracy is not. If ninety-six accurate answers contain widely differing secondary details, core accuracy may be high while the representation profile remains fragmented.

2. Unit of Scope: The Observation Cell

Before it begins, every NJ-100 study defines an **observation cell**. The cell is the combination of the following fields:

Table B.1 — Mandatory fields of an observation cell

Field	Information preregistered
Entity	Full name, distinguishing identity, and the product or service boundary examined
Target user	Role, state of need, minimum/maximum knowledge level, exclusion criteria
Intention of use	Information gathering, shortlisting, comparison, risk checking or another explicitly defined purpose
System	Product and, where possible, model name; description of the interface if the product name is unknown
Access surface	Web search on/off, app/web/API, free/paid tier
Session state	New or existing conversation; memory and personalisation status
Language and geography	Prompt language; participant's country/region-level location
Time window	Start and end dates; plan for the second window
Prompt family	Controlled or natural-intent arm; relevant task code
Correct core	The codable minimum representation derived from evidence about the real world

Several systems, languages, countries, user roles or prompt families are not dissolved into one result. Every material difference is reported as a separate cell or predefined stratum. One cell's one hundred valid sessions cannot make up the shortfall in another.

3. Sample and Independence

3.1. Minimum design

An NJ-100 cell contains **100 valid, independent target-user sessions**. Preferably, each participant contributes only one session to the relevant cell. Where the same person must take part in several prompt families, that dependency is disclosed in advance and the results are analysed separately with clustering at participant level.

A “different IP address” is not proof of independence. One person can use different networks, while different people can share one. Full IP address is not a

default field for identity or independence checks. Independence is assessed through a unique participant code, recruitment record, session time, task allocation and overlap check.

3.2. Participant conditions

The participant shall:

- meet the target-user definition;
- be informed about the research purpose and use of data;
- disclose any interest in the entity, practitioner or research team;
- perform the assigned natural-intent task without another person's direction;
- not have participated in the relevant cell before.

Employees, agency personnel, close commercial partners and experts familiar with the protocol are excluded from the main target-user sample. They may be examined in a separate expert-resilience arm.

3.3. Boundary of the sample claim

One hundred observations is not a magical threshold of universality. The extent to which the sample represents the target population depends on recruitment. If convenience sampling, one country, one account tier or one device class is used, that limitation must appear in the decision sentence. Sample size does not substitute for population diversity.

4. Two Research Arms

4.1. Controlled-prompt arm

This arm measures system behaviour under comparable conditions. Prompts are written in advance; wording, order and administration instructions are fixed. The principal prompt families are:

1. **Identity:** “What is X?” or a contextually appropriate equivalent.
2. **Attribute:** A specified characteristic, expertise or product scope of the entity.
3. **Comparison:** Two or more options under explicit criteria.
4. **Recommendation:** A shortlist or recommendation for a defined user need.
5. **Risk:** A boundary, unsuitability, omission or disputed field.
6. **Exclusion:** A need for which the entity is intentionally unsuitable.
7. **Source:** A request for verifiable sources supporting the answer.

The controlled arm does not represent the full diversity of real user language. It provides a common measurement surface for comparisons between systems or across time.

4.2. Natural-intent arm

Participants in this arm receive no prepared sentence. They are given a need scenario and task objective, formulate the question in their own words and ask natural follow-up questions where needed. The first prompt, follow-up prompts and final answer are recorded separately.

The natural-intent arm increases ecological validity, but variation in wording is greater. It is therefore not combined with the controlled arm in the same denominator. Reporting the two arms together shows representational behaviour under different conditions, not the success of the protocol.

5. Evidence about the Real World and the Correct Core

Before measurement begins, the **real-world evidence file** for the entity under examination is locked. The file comprises official records, contracts, product documentation, auditable technical data, dated public records and suitable independent sources. Where sources conflict, one supposedly correct account is not invented; the disagreement enters the codebook.

Three fields are defined for every prompt family:

- **Mandatory core:** the minimum meaning the answer must carry to be coded as accurate.
- **Permitted variation:** wording, order and secondary details that may change without compromising the core.
- **Critical error:** a false claim about identity, authority, suitability, safety, price, geography, expertise or outcome that would materially mislead the user.

The correct core cannot be changed after test results are seen. If new information arises in the real world, a protocol event is recorded; affected sessions are separated under the old and new evidence windows or the series is closed.

6. Data Collection Record

The minimum record for every valid session is:

Table B.2 — NJ-100 session data record

Field	Description
Session code	Unique code separated from identity
Participant code	Stored separately and matched securely to the recruitment record
Eligibility status	Whether target-user conditions were met
Date and time	Including time zone
Country/region	Only the necessary level of detail
System and interface	Product/model information, app or web surface
Account tier	Free/paid/enterprise where known
Context status	New/existing conversation; memory, personalisation and web search
Language	Language of prompt and answer
Prompt chain	Unaltered text of the first and follow-up prompts
Complete answer	Preserving formatting and source links
Sources	Names and links supplied in the answer
Technical event	Error, outage, safety refusal, empty answer or interrupted session
Codes	Accuracy, core profile, critical error, source support and exclusion behaviour

A screenshot may serve as a supporting record; it does not replace searchable text. Dynamic links and page content may change. Where necessary, the page title, access date and an archivable summary of evidence are retained.

7. Valid, Invalid and Missing Sessions

A session is valid where:

- participant eligibility has been verified;
- the assigned prompt arm and cell conditions were preserved;
- the complete prompt and answer were recorded;
- the system produced enough of an answer to be coded;
- no predefined exclusion criterion applies.

Sessions excluded because of a technical interruption, incorrect task allocation, duplicate participant or missing answer are not deleted. They remain in a separate flow log together with the reason. If a system refusal or “I don’t know” answer is a

meaningful outcome under the protocol's research question, it cannot be invalidated; it receives a **non-response** or **absence of representation** code.

The target is one hundred valid sessions. The report also states how many people were screened, how many sessions began and the reasons for exclusion.

8. Coding Model

8.1. Primary endpoint: accurate core representation

Every answer receives a binary primary endpoint under the locked codebook:

- **1 — Meets:** carries the mandatory core accurately and contains no critical error.
- **0 — Does not meet:** omits or contradicts the mandatory core, or contains a critical error.

The binary outcome creates clarity, but does not express the whole quality of an answer. A secondary profile is therefore mandatory.

8.2. Secondary semantic profile

Every answer is also coded for:

- accurate identity;
- accurate core attributes;
- placement in the relevant context;
- clarity of recommendation rationale;
- statement of unsuitability boundaries;
- whether sources actually support the claims;
- unwarranted certainty or fabrication;
- critical and non-critical error class;
- semantic clusters outside the core.

Text-similarity tools may help a human coder see possible matches; they cannot adjudicate truth. Measures such as BERTScore compare contextual similarity but do not establish that a sentence is accurate in the real world.[19]

8.3. Two coders and disagreement

Every answer is assessed independently by two trained human coders. So far as possible, coders are blinded to participant identity, implementation owner and expected commercial outcome. No reconciliation meeting is held before the first independent coding is complete.

The report includes at least:

- raw percentage agreement;
- Cohen's kappa for the primary binary code;[15]
- the distribution for each coder;
- the number and types of disagreement;
- the third adjudicator's rationale.

Kappa alone is not a seal of quality; class distribution can affect it. Even low overall disagreement may matter where critical safety errors are concerned. If the codebook changes after measurement, the old answers are recoded blind under the new version and the version difference is reported.

9. Four Outcome Regions

Accuracy and stability are read on separate axes:

Table B.3 — Accuracy and stability decision matrix

Observed condition	Interpretation	Permitted judgment
High stability + high accuracy	The core representation appears accurate and repeatable within the specified scope	Limited representation-stability opinion
High stability + low accuracy	Systems converge on the same error	Stable representational distortion
Low stability + high accuracy	Accurate answers exist; representation is sensitive to conditions or wording	Accurate but fragmented/conditional representation
Low stability + low accuracy	Representation is both inaccurate and variable	Unstable representational distortion

The matrix prevents high agreement from being treated automatically as success. One of the most dangerous regions is a system repeating the same false account with high consistency.

10. Thresholds and Uncertainty

10.1. Observed rate

The primary rate is calculated as:

number of valid sessions meeting accurate core representation / 100 valid sessions

For example, 90/100 is an **observed accurate-core rate of 90 per cent**. It does not offer strong lower-bound assurance that the true rate exceeds 90 per cent. The approximate 95 per cent Wilson confidence interval is 82.6 to 94.5 per cent.[14]

For 96/100, the approximate 95 per cent Wilson confidence interval is 90.2 to 98.4 per cent. NJ-100 therefore distinguishes two thresholds:

- **90/100:** High observed agreement/accuracy threshold.
- **96/100 and no critical error:** Pilot decision threshold for the estimated rate's 95 per cent Wilson lower bound to exceed 90 per cent.

These figures are not universal laws of nature. Thresholds must be retested against risk class, the cost of an incorrect decision, cell design and independent validation results. Even 96/100 is insufficient by itself in high-risk fields such as health, law, finance or safety.

10.2. Critical-error gate

No aggregate rate can obscure a predefined critical error. Where an event such as identity conflation, false authority, dangerous medical or legal direction, recommendation of prohibited use or serious discrimination is observed, the study enters **critical review**. The event's gravity, prevalence and correctability are assessed separately.

10.3. Exclusion and negative-prompt gate

The protocol tests not only whether an entity is recommended appropriately, but whether it is withheld when unsuitable. No opinion of “accurate recommendation behaviour” may be given until the predefined minimum number of negative or exclusion prompts has been met. System behaviour that promotes the entity in every case is not conformity, even if it increases visibility.

11. Second Time Window

Meeting the first threshold does not make the result permanent. The same protocol is repeated in a predefined second time window. The interval is recorded according to sector and system volatility; thirty to sixty days is recommended as an ordinary pilot starting point.

In the second window:

- the same observation cell is preserved;
- new participants are preferred;
- the prompt and codebook versions remain fixed;
- a protocol event is opened if the system or interface changes;
- thresholds are calculated again and separately.

If the first window is high and the second low, the two values are not averaged to create success. The result is reported as **temporally unstable**. If a material system change destroys comparability, a new series is opened.

12. Decision Language

12.1. Permitted example

***NJ-100 pilot observation opinion:** Within the cell limited to [entity], [target user], [system/interface], [language/geography] and [date], 97 of 100 valid independent sessions met the predefined accurate core representation. The 95 per cent Wilson interval is reported in the annex. No critical error was observed; exclusion prompts met their separate gate. The result also maintained the protocol threshold in the second time window. This opinion does not guarantee future answers, other systems or commercial outcomes.*

Where a short public statement is required:

Representation stability was observed within the specified scope.

12.2. Prohibited examples

- “GEO complete.”
- “100 per cent visible across all AI.”
- “AI-approved company.”

- “Permanent representation guarantee.”
- “NJ-100 certificate”—where there is no independent certification infrastructure or authority.
- “Ninety-six per cent of one hundred people recommended the brand”—where the system answer, not user opinion, was measured.

13. Reporting Package

A complete NJ-100 report contains:

1. protocol code, version and preregistration date;
2. observation cell and boundary of generalisability;
3. locked summary of the real-world evidence file;
4. participant flow and reasons for exclusion;
5. separate results for controlled and natural-intent arms;
6. primary rate, denominator and Wilson interval;
7. semantic profile and position within the four outcome regions;
8. critical-error and negative-prompt results;
9. coder agreement and resolution of disagreements;
10. separate comparison of the first and second time windows;
11. system events, missing data and protocol deviations;
12. conflicts of interest, funding and review roles;
13. decision sentence, validity period and route of objection;
14. a machine-readable, anonymised method summary.

Public release of raw answers is not an automatic requirement. Personal data, copyright, contractual and security boundaries must be observed. Confidentiality cannot, however, justify withholding the method or concealing an adverse result.

14. Data Protection and Research Ethics

NJ-100 collects only necessary data. Full IP address, precise location, unnecessary device fingerprints and identity documents are not default fields. Participant contact information and research responses are kept separately; access roles, retention period and deletion plan are defined in advance.

Participants are not instructed to request or upload confidential information about themselves, the system provider or a third party. Personal or sensitive data

appearing unexpectedly in a generative answer is entered in the incident log, access is restricted and the necessary legal and ethical procedure followed.

The research team explains participant consent, withdrawal conditions, remuneration, conflicts of interest and the purposes for which data will be published. This protocol does not by itself ensure legal compliance; applicable rules on data protection, consumers, contracts and the relevant sector must be assessed separately.

15. Principal Invalidating Errors

The following practices invalidate an NJ-100 opinion:

- changing the correct core or success threshold after testing;
- counting several sessions from the same person as independent participants;
- retaining only favourable answers;
- dissolving different systems, languages or account tiers into one denominator;
- silently excluding technical refusals;
- treating a model-similarity score as evidence of accuracy;
- telling coders the expected outcome;
- declaring the first-window result permanent without a second window;
- concealing a critical error inside a high aggregate rate;
- calling a practitioner's review of their own work “independent validation”;
- translating observed representation stability into sales, profitability or customer satisfaction.

16. Validation Roadmap

At minimum, the following work is required before NJ-100 can leave pilot status:

- 1. Preregistration:** Hypothesis, threshold, codebook, exclusions and analysis plan shall be published before results are seen.
- 2. Desk-based resilience test:** Coding and reporting failures shall be tested on synthetic answer sets.
- 3. Diverse real cases:** Studies shall cover entities that differ by sector, language, risk and brand recognition.
- 4. Independent replication:** Teams independent of the Framework author and paid practitioner shall apply the same protocol.

5. **Sample sensitivity:** Decision stability shall be examined at samples of 50, 100, 200 and larger.
6. **False-positive/negative testing:** The protocol's discriminatory power shall be measured against intentionally accurate, inaccurate, incomplete and unstable representations.
7. **Coder resilience:** Agreement shall be compared among coders with different languages and levels of expertise.
8. **Time and system change:** The effect of version, interface and source-access changes on the result shall be measured.
9. **Harm review:** Research shall examine whether the protocol imposes disproportionate burdens on small businesses, underrepresented languages or high-risk uses.
10. **Public consultation:** Thresholds, objection routes and reporting language shall be discussed with practitioners, researchers, businesses and affected users.

Even after these steps, the protocol will not be “final”. Changing systems, new evidence and objections make version management necessary. NJ-100's credibility will arise not from the number one hundred in its name, but from stating with equal clarity what it does not measure.

Appendix C — Minimum Record Set

Function of the Records

This appendix connects the book's ten chapters to ten record families. The forms do not exist to generate more documents, but to make visible the chain of entity, evidence, method, authority and time from which a judgment arises. Fields may be extended according to risk; where a field does not apply, it is not left blank. The reason for selecting **Not applicable** is recorded.

Every record carries the following common header:

Table C.1 — Common header for all record families

Common field	Content
File code	Unique primary code for organisation/year/engagement
Record code	Family code below + sequence number
Version	Major.minor.correction format
Status	Draft / under review / approved / suspended / withdrawn / archived
Owner	Role creating and updating the record
Approver	Independent or authorised decision owner where required
Creation date	Including time zone
Last review	Date and reviewer
Next review	Date or event trigger
Confidentiality class	Public / restricted / confidential / contains personal data
Linked records	Related record codes and evidence locations
Change summary	Material difference from the previous version

The recommended format is NJ-[FAMILY]-[FILE]-[SEQUENCE]. For example, NJ-CLM-2026-014 identifies the fourteenth Claim Record. A code ensures traceability, not accuracy.

C1 — Entity and Representation Record

Code: ENT **Linked chapter:** Centre **Purpose:** Lock the boundary between the real entity and the representation under examination

Table C.2 — Entity and Representation Record fields

Field	Information to enter
Entity's legal name	
Alternative names	Brand, abbreviation, former name, personal names
Distinguishing identity	Registration number, domain, geography or equivalent verifier
Entity type	Person / business / product / service / institution / other
Scope examined	Included products, services, countries, languages and target users
Out of scope	Fields deliberately not examined
Real-world summary	Short verified definition with source codes
Identity representation	With whom/what does the system associate the entity?
Attribute representation	Which attributes does it attach or omit?
Judgment representation	When does it recommend, compare or exclude?
Known conflations	Similar name, obsolete information, wrong category, another geography
Representation debt	Accumulated effect of material omissions/errors
Risk class	Low / medium / high / critical; reasons
Entity-owner statement	Labelled explicitly as a source type, not evidence

Closing question: Is the judgment being carried to another entity or scope not defined in this record?

C2 — Claim and Evidence Record

Code: CLM **Linked chapter:** Evidence **Purpose:** Keep together the records that support and weaken every material sentence

Table C.3 — Claim and Evidence Record fields

Field	Information to enter
Claim text	Exact sentence proposed for publication
Claim type	Real-world fact / representation / outcome / cause / prediction / conformity
Materiality class	Low / medium / high / critical
Scope	Entity, system, language, geography, time

Field	Information to enter
Supporting evidence	Evidence code, source, date and relevant section
Weakening evidence	Contradiction, omission or alternative explanation
Evidence chain	Steps from source to claim
Verification status	Verified / partial / unverified / disputed
Source access	Open / licensed / confidential; review conditions
Currency	Last check and ageing trigger
Permitted language	Certain / limited / probabilistic / observation only
Prohibited extension	Sentence this record does not prove
Decision	Use / narrow / postpone / withdraw

Closing question: Is counter-evidence as visible as supporting evidence?

C3 — Measurement Design Record

Code: MEA **Linked chapter:** Measurement **Purpose:** Lock the object, denominator and analysis before the result is seen

Table C.4 — Measurement Design Record fields

Field	Information to enter
Research question	One bounded question
Primary endpoint	Main variable measured and calculation
Secondary endpoints	Indicators kept separate
Unit of observation	Answer / session / user / prompt family / other
Denominator	All included observations
Prompt universe	Families, source and selection method
System conditions	Product/model, interface, search, account, memory
Sample	Target population, recruitment, size, strata
Time window	Start, end and repetition plan
Baseline	Pre-intervention measurement
Comparison	Control, historical series or reason for absence
Codebook	Version, coders, training and blinding
Missing data	Predefined treatment

Field	Information to enter
Exclusion criteria	Reasons defined before results
Uncertainty	Interval estimate or sensitivity analysis
Deviation log	Every departure from protocol and its effect

Closing question: Will adverse outcomes remain in the denominator under the same method?

C4 — Intervention Record

Code: INT **Linked chapter:** Intervention **Purpose:** Show proportionality between problem and change, and the route of reversal

Table C.5 — Intervention Record fields

Field	Information to enter
Verified problem	Relevant ENT, CLM and MEA codes
Materiality	Who may be affected, how much and with what likelihood?
Surface of change	Owned content, data, structure, process or authorised third party
Options	Routes considered, including no intervention
Selected intervention	Exact scope and implementation steps
Minimum-sufficiency rationale	Why is a broader change unnecessary?
Reality check	Real-world evidence supporting the intervention
Risks	Misdirection, user harm, platform and legal risk
Ethical check	Veto fields and approving role
Implementation owner	Authority and accountability
Start date	
Success criterion	Pre-registered, separate measures
Stopping criterion	Harm, deviation or failure condition
Reversal plan	Which change will be reversed, and how?
Outcome	Observed effect and alternative explanations

Closing question: Does the intervention explain the entity's reality, or is it trying to replace it?

C5 — Governance and Authority Record

Code: GOV **Linked chapter:** Governance **Purpose:** Separate decision rights, conflicts of interest and the power to stop

Table C.6 — Governance and Authority Record fields

Field	Information to enter
Right to observe	Who may collect data?
Right to implement	Who may make changes?
Right to approve	Who approves method and claim?
Right to publish	Who publishes the public sentence?
Right to stop	Who may suspend the work, and when?
Accountable owner	Final human accountability
Independent reviewer	Who, under which conditions of independence?
Conflicts of interest	Financial, professional, personal and institutional
Mitigation	Separation of roles, blinding, second signature, external review
Risk class	Two-Key requirement for critical decisions
Complaints channel	Access, period and recording method
Appeal authority	Role independent of the first decision
Incident authority	Notification, investigation, suspension
Conformity mark	Yes/no; evidence of authority and infrastructure

Closing question: Is the party earning revenue from the work the sole judge of its own claim?

C6 — Time, Version and Event Record

Code: TIM **Linked chapter:** Time **Purpose:** Show the conditions under which a sound decision ages and when a series breaks

Table C.7 — Time, Version and Event Record fields

Field	Information to enter
Relevant decision	Record and version code
Evidence window	Oldest and newest evidence used
Observation window	Start and end

Field	Information to enter
Effective date	Date the decision began
End of validity	Date or condition
Scheduled review	Frequency and owner
Event triggers	System, interface, business, source or legal change
Event	What changed, and when was it learned?
Materiality decision	Does it affect comparability/the decision?
Series decision	Continue / add explanation / new series / suspend
Old-version status	Deprecated / archived / withdrawn
Public notice	Required? Where and when?

Closing question: Can a reader see at a glance whether this judgment is still valid?

C7 — Commercial Value and Attribution Record

Code: VAL **Linked chapter:** Final Test **Purpose:** Separate change in representation from the chain leading to a customer and net contribution

Table C.8 — Commercial Value and Attribution Record fields

Field	Information to enter
Value hypothesis	Which change in representation should create which legitimate value?
Target customer	Need, fitness and exclusion conditions
Representation indicator	Behaviour observed to have changed
Behavioural indicator	Visit, contact, shortlist, proposal or other trace
Source statement	How did the user report AI influence?
Attribution class	Direct / supporting / temporal association only / unknown
Cohort	Start date and common characteristic
Revenue	Gross and collected amounts separately
Costs	Acquisition, delivery, refund, support and capacity
Net contribution	Calculation and period
Quality	Fitness, refunds, complaints, repeat or retention

Field	Information to enter
Sustainability	Capacity, margin, harm and ethical conditions
Alternative explanations	Season, price, campaign, distribution, brand effect
Decision	Value observed / uncertain / adverse / unsustainable

Closing question: Is the success being described revenue, or delivered and sustainable net value?

C8 — Audit Master Record

Code: AUD **Linked chapter:** Audit **Purpose:** Unite the audit chain from scope to opinion in one file

Table C.9 — Audit Master Record fields

Field	Information to enter
Claim under audit	Exact sentence and CLM code
Audit level	Desk-based / limited / comprehensive / surveillance / pilot conformity opinion
Admission gate	Record sufficiency and independence status
Scope	Entity, system, language, geography, date
Criteria	Constitutional provisions and protocol version
Auditor	Competence and conflicts of interest
Sampling	Universe, selection method and boundary
Records examined	Code list
Findings	Favourable, adverse and uncertain fields
Nonconformities	Critical / major / limited; with evidence
Corrective action	Owner, period and verification method
Opinion	Conforming / limited / insufficient evidence / nonconforming / suspended
Exceptions	Fields to which the opinion cannot be carried
Validity	Period and trigger
Public summary	Decision text separated from confidential annex
Route of objection	Channel, period and authority

Closing question: Could another competent reviewer reconstruct the route to decision from the same file?

C9 — Objection and Correction Record

Code: APL **Linked chapter:** Objection **Purpose:** Enable an independent review capable of finding error rather than defending the decision

Table C.10 — Objection and Correction Record fields

Field	Information to enter
Decision challenged	Record, version and publication location
Application date	
Applicant	Within identity/confidentiality limits
Right/interest affected	
Objection type	Record / coding / inference / authority / process / principle
Grounds	Complete summary of the objection
New evidence	Source, date and verification status
Initial admission decision	Reviewable / insufficient; reasons
Independent reviewer	Relationship to the initial decision
Defence of initial decision	Grounds and records
Review finding	Which element was substantiated or refuted?
Outcome	Uphold / narrow / correct / remeasure / suspend / withdraw
Interim measure	Where risk of harm exists
Public correction	Link between former and new sentences
Closure and period	

Closing question: If the objection was refused, does the refusal do more than repeat the original claim?

C10 — Judgment and Publication Record

Code: JDG **Linked chapter:** Judgment **Purpose:** Convert the whole record chain into a limited, dated and withdrawable final decision

Table C.11 — Judgment and Publication Record fields

Field	Information to enter
Decision question	Which judgment is being made?

Field	Information to enter
Linked record chain	ENT–CLM–MEA–INT–GOV–TIM–VAL–AUD–APL
Conflicting interests	Accuracy, user, client, practitioner, Framework
Hierarchy of precedence	Which higher principle applied?
Decision	Complete, publication-ready sentence
Status	Pilot observation / limited opinion / suspended / withdrawn
Scope	Cell in which the decision is valid
Evidential boundary	What the decision does not prove
Critical exceptions	
Validity	Date and review trigger
Signatory roles	Preparer, reviewer, approver
Conflict of interest	Publicly suitable summary
Route of objection	
Machine-readable summary	Field names and version identifier
Withdrawal link	Former decision and reasons, where applicable

Closing question: Is the decision speaking more broadly, certainly or permanently than it deserves?

Record-Chain Check

Before an engagement closes, the following chain shall be unbroken:

ENT → CLM → MEA → INT → GOV → TIM → VAL → AUD → APL → JDG

Not every engagement must produce an intervention or a favourable judgment. Where a problem is immaterial, the INT record may close with **no intervention**; where no commercial connection can be established, the VAL record with **uncertain**; where evidence is inadequate, the JDG record with **no judgment possible**. These are not failed documents. They are decisions that preserve the standard's boundary.

Retention, Access and Immutability

Records are retained for a period proportionate to risk, subject to personal-data and contractual boundaries. A public summary and confidential evidence need not share the same access level. Records underlying a published decision may not, however, be altered silently after the fact. A correction carries a new version, date, reasons and link to the previous version.

Where structured digital records are used, field names are versioned and mandatory-field validation, access logs and change history preserved. Use of a database or blockchain does not establish that its content is accurate. Technical immutability has meaning only together with human accountability, evidence quality and the right of objection.

Appendix D — Completed Synthetic File

Meridian Ledger: Tracing a Decision from Beginning to End

Teaching simulation — synthetic data. The people, dates, prompts, rates, revenues and decisions concerning Meridian Ledger in this appendix are entirely fictional. They prove neither the performance of a real business nor the empirical effectiveness of the NJ-100 protocol. The example shows why an apparently favourable result must sometimes close with a limited judgment.

File header

Table D.1 — Synthetic file identity for Meridian Ledger

Field	Synthetic record
File code	ML-2026-01
Entity	Meridian Ledger
Scope	Creative agencies with 10–50 employees in the United Kingdom; English prompts
Product examined	Cash-flow visibility and invoice-reconciliation software for small businesses
Out of scope	Banking, payment-institution services, credit, investment and tax advice
Study status	Synthetic training file
Risk class	Medium; a false impression of financial authority may affect user decisions
Protocol	NobleJackal GEO Framework 1.0.0 + illustrative application of NJ-100-Po.9

D1 — Entity and Representation Record

Record: NJ-ENT-ML-001 **Status:** Approved synthetic baseline record

Meridian Ledger's real-world evidence file contains the following core definition:

Meridian Ledger is a subscription software product that helps small businesses view their own bank and invoice data in one place. It is not a bank; it does not hold money, execute payment orders, extend credit or provide regulated financial advice.

The baseline observation records three distortions:

- 1. Some answers call the company a “digital bank”.
- 2. Some say that the product makes payments on behalf of the business.
- 3. Comparative answers place the product in the same category as regulated open-banking payment providers.

Identity representation is mostly accurate; the attribute and judgment layers are distorted. This is not merely a matter of wording: a user seeking a regulated payment service may shortlist the wrong product.

D2 — Claim and Evidence Record

Record: NJ-CLM-ML-004 **Claim examined:** “Generative-answer systems usually misidentify Meridian Ledger as a bank or payment institution.”

Table D.2 — Synthetic claim and evidence profile for Meridian Ledger

Evidence domain	Synthetic finding
Evidence about the real world	Company registration, product contract, feature list and regulatory-scope note show that it has no banking/payment authority
Evidence of representation	37 of 120 controlled baseline sessions create an impression of banking or payment authority
Evidence of outcomes	In 5 of 18 sales conversations, prospective customers asked whether the product made payments
Weakening evidence	Not every sales question originated in a generative system; category language is inconsistent across the sector

The original sentence is narrowed. “Usually” exceeds the evidence: 37/120 is not a majority. The permitted wording is:

“In the baseline sample examined, 37 of 120 answers attached at least one expression to Meridian Ledger that created an impression of banking or payment authority.”

Sentence the evidence does not carry: “Misrepresentation caused lost sales.”
The sales conversations show a trace, not causation.

D3 — Measurement Design Record

Record: NJ-MEA-ML-002 **Preregistration status:** Synthetic plan locked before intervention results were known

Table D.3 — Synthetic measurement design for Meridian Ledger

Field	Synthetic decision
Primary endpoint	Answer carries the mandatory core accurately and contains no critical authority error
Mandatory core	Software product; financial visibility/reconciliation; not a bank or payment institution
Critical error	Claim that it holds money, executes payments, extends credit or provides regulated advice
Observation cell	United Kingdom; agency owner/finance manager; English; selected generative-answer surface; new session
Prompt arms	60 controlled + 40 natural-intent sessions; reported separately
Negative prompts	Unsuitability scenarios such as “a tool that will make supplier payments for me”
Coders	Two independent coders + a third adjudicator on disagreement
Repetition	100 new participants, forty-five days after the first window
Decision threshold	At least 96/100 in each window; 95% Wilson lower bound > 90%; no critical error; negative gate met

The design retains citation rate and commercial trace as separate secondary indicators. They are not added to the accurate-core rate to create an aggregate score.

D4 — Intervention Record

Record: NJ-INT-ML-003

The intervention team evaluates four options:

- 1. do nothing;
- 2. fill the entire site with sentences saying “not a bank”;

3. make the verified category and boundary information consistent on critical pages;
4. spread the same account through synthetic user profiles on third-party sites.

The fourth option is rejected on ethical and methodological grounds. The second would create unnecessary repetition and damage the user experience. The first fails to address the material confusion of authority. The third is selected as the minimum sufficient intervention:

- the product page uses the fixed category “cash-flow visibility and invoice reconciliation software”;
- a “does / does not” boundary is added to the feature table;
- the regulatory-scope note is verified by legal counsel and the product owner;
- terminology in developer documentation and on the homepage is aligned;
- the owner of an obsolete partner directory is sent a verifiable correction request concerning the expression “payments platform”;
- advertising language implying banking authority is removed.

The intervention is not described as “training the models” or “controlling the answers”. Owned information surfaces and those that can be corrected with authority are improved. Page versions, approvers and a reversal plan are recorded.

D5 — Governance Record

Record: NJ-GOV-ML-001

Table D.4 — Synthetic governance assignments for Meridian Ledger

Role	Synthetic assignment
Owner of real-world truth	Meridian Ledger product lead
Legal/regulatory boundary approval	External legal adviser
Intervention implementer	Meridian content team
Measurement design	NobleJackal advisory team
Coders	Contracted researchers separate from the implementation team
Final decision owner	Meridian risk committee
Ethical veto	External legal adviser + user-safety representative
Objection reviewer	Independent research adviser not involved in the first decision

Because NobleJackal both proposed the protocol and contributed to measurement design for a fee, it cannot present its own work as an **independent audit**. The illustrative file produces no certificate or conformity badge.

D6 — Time and Event Record

Record: NJ-TIM-ML-006

Table D.5 — Synthetic time and event record for Meridian Ledger

Event	Synthetic date	Decision
Real-world evidence lock	8 January 2026	Baseline version WG-1.0
Intervention published	22 January 2026	Monitoring begins
First NJ-100 window	16–20 February 2026	Series S1-W1
System-interface update	4 March 2026	Materiality review: citation presentation changed; core coding unchanged
Second NJ-100 window	2–7 April 2026	Series S1-W2; event to be disclosed in report
Next review	2 June 2026 or a change in product authority	Whichever occurs first

The interface update is not passed over silently. Both windows still measure the same core question, but the citation indicators are not compared directly.

D7 — Synthetic NJ-100 Results

First window

Table D.6 — First synthetic NJ-100 observation window

Indicator	Synthetic result
Valid sessions	100
Accurate core	97/100
Critical errors	0
Negative-prompt gate	Met
95% Wilson interval	Approximately 91.6–99.0%
Raw coder agreement	95%
Cohen's kappa	0.84

The first window meets the pilot threshold. One window is not enough for a permanent judgment.

Second window

Table D.7 — Second synthetic NJ-100 observation window

Indicator	Synthetic result
Valid sessions	100
Accurate core	94/100
Critical errors	0
Negative-prompt gate	Met
95% Wilson interval	Approximately 87.5–97.2%
Raw coder agreement	96%
Cohen's kappa	0.81

The second window exceeds the high observed rate of 90/100, but not the pilot decision threshold of 96/100. The two windows are not averaged to 95.5/100 and presented as though the threshold were met. The result is judged **high but below the protocol threshold and temporally unstable**.

The semantic profile shows that four of the six failed answers imply that the product “automates payments”, while two imply that it “manages” bank accounts. Because the critical-error definition requires a claim that the product holds money or executes payment in fact, these expressions do not cross the critical threshold; they remain attribute distortions.

D8 — Commercial Value Record

Record: NJ-VAL-ML-002

In the synthetic sixty-day post-intervention cohort:

- twenty-two qualified applicants directly state that a generative system influenced their research;
- nine become suitable opportunities and three become contracts;
- seventy-one applications arrive from other channels during the same period;
- a price change and a partnership campaign also affect total sales;
- net contribution from the three contracts is positive over the first ninety days, but the observation horizon is too short to assess retention.

Permitted wording:

“Within the synthetic post-intervention cohort, twenty-two applicants directly stated that a generative system influenced their research; three became contracts within the first sixty days. The design does not show that the intervention alone caused these outcomes.”

Prohibited wording:

“GEO increased Meridian Ledger’s sales by X per cent.”

D9 — Objection Record

Record: NJ-APL-ML-001

The sales director notes that the two windows contain 191 accurate cores out of 200 and asks to use the statement “NJ-100 validated”. The objection is refused for three reasons:

1. The protocol defines separate thresholds for each window; combining them after seeing the result changes the measurement rule.
2. “Validated” conflates a pilot observation with independent method validation.
3. NobleJackal holds a paid design role and has no authority to certify independently.

The legitimate part of the objection is upheld: the public summary will not conceal the aggregate observation of 191/200, but will report both windows separately.

D10 — Audit Opinion and Final Judgment

Records: NJ-AUD-ML-001 and NJ-JDG-ML-001

The audit finds that the record chain is traceable, the intervention grounded in reality and no critical error observed; it also finds that the second window failed the pilot threshold and commercial causation was not demonstrated.

The final sentence permitted for publication is:

Limited synthetic pilot opinion: *Accurate core representation for Meridian Ledger in the defined observation cell was observed at 97/100 in the first window and 94/100 in the second. No representation-stability conformity opinion was issued because the second window did not meet the NJ-100-Po.9 pilot decision threshold. No critical error was observed and the negative-prompt gate was met. The findings cannot be generalised to another system, language, user group or commercial outcome.*

The file closes with **conditional monitoring** status. Semantic clusters in the six failed answers will be examined; a new intervention will open only if the problem can be linked to real-world evidence and owned surfaces. Producing broader content or narrowing prompt selection merely to pass the threshold is prohibited.

What the Example Teaches

In this synthetic file, the intervention is reasonable, the first result strong and the commercial trace promising. The Framework still produces no favourable badge. The second window did not meet the predefined threshold, causation was not established and no independent certification authority arose.

The seriousness of a standard is shown not only by what it calls a successful result, but by how it limits a result that appears favourable to itself. This is the real value of the record chain: it does not enlarge the decision, but keeps it at exactly the size it deserves.

Appendix E — Register of Permitted and Prohibited Language

Why a Language Register?

Many standards fail not at the measurement table, but in the sentence announcing the result. A limited observation becomes “success”, success becomes “control”, and control becomes a “guarantee”. This appendix is not a hunt for forbidden words. It is a publication gate that aligns the authority carried by the sentence with the authority carried by the evidence.

The Register is sensitive to context. A prohibited word may be used in a historical quotation, a criticism or to describe a genuine legal status. What is prohibited is using the word to borrow trust where the status does not exist.

E1 — Status and Legitimacy

Table E.1 — Language of status and legitimacy

Wording to avoid or unauthorised use	Permitted alternative	Reason
“The world’s first GEO standard”	“A proposal for a standard addressing representation, evidence and sustainable value together in generative systems”	“First” requires comprehensive, independent historical evidence
“International standard”	“A framework intended for multilingual and international use”	Intended applicability is not international acceptance
“Official GEO standard”	“A normative framework published by NobleJackal”	Official authority must be identified
“Accredited”	“A proposal for a standard without independent accreditation”	Accreditation requires a defined external authority and process
“Certified”	“A limited conformity opinion within the specified scope”	Certification language cannot be used without certification infrastructure
“Approved by NobleJackal”	“Assessed under the NobleJackal protocol”	Use of a method does not constitute general approval
“Universal”	“Limited to the defined system, language, geography and time”	Prevents extension beyond scope

Wording to avoid or unauthorised use	Permitted alternative	Reason
“Scientifically proven”	“This finding was observed under the specified design”	One study does not produce broad scientific validation

E2 — System Behaviour and Control

Table E.2 — Language of system behaviour and control

Wording to avoid or unauthorised use	Permitted alternative	Reason
“AI knows us”	“In the answers examined, the entity was associated with the correct identity”	Does not attribute a singular, permanent mind to the system
“AI-approved brand”	“Accurate representation was observed within the specified scope”	A generative system is not an institutional approval authority
“We trained the models”	“We verified and improved the information surfaces we own”	Misleading where there is no authority over external model training
“We control the answers”	“We work on legitimate surfaces capable of influencing representation”	Separates influence from control
“Visible across all AI”	“Mentioned in the sample from the named systems and conditions”	Systems and conditions differ
“Permanent result”	“Observation valid until the specified date and subject to review”	Representational behaviour changes through time
“Guarantee”	“Outcome observed within predefined thresholds and boundaries”	Future output cannot be claimed
“A citation makes it true”	“The support relationship between source and claim was separately verified/not verified”	Presence of a citation is not accuracy

E3 — Measurement and Success

Table E.3 — Language of measurement and success

Wording to avoid or unauthorised use	Permitted alternative	Reason
“GEO score: 87”	“Profile of accuracy, sourcing, recommendation, stability and commercial trace”	Does not hide distinct dimensions in one number

Wording to avoid or unauthorised use	Permitted alternative	Reason
“GEO complete”	“Representation stability was observed within the specified scope”	The work has no permanent endpoint
“96 per cent success”	“96 of 100 valid sessions met the predefined accurate core”	Makes the denominator and measured object visible
“96 per cent of users recommended it”	“96/100 system answers met the coding criterion”	Does not conflate user opinion with system output
“Consistency means accuracy”	“Stability and accuracy were assessed on separate axes”	The same error may also be consistent
“Passed on average”	“Every predefined cell and time window was reported separately”	Does not dissolve a weak field inside a strong one
“Error-free”	“No predefined critical error was observed in the sample”	Distinguishes unobserved from impossible
“Objective score”	“Human assessment bounded by a codebook, two coders and an agreement measure”	Makes the judgment process visible

E4 — Commercial Outcome and Causation

Table E.4 — Language of commercial outcome and causation

Wording to avoid or unauthorised use	Permitted alternative	Reason
“GEO increased sales”	“Applicants declaring AI influence were observed in this cohort; the design does not by itself establish causation”	States the attribution level
“Visibility became revenue”	“Some of the specified contacts produced revenue; costs and alternative explanations were reported separately”	Does not erase intermediate links in the chain
“High-value customer”	“Customer meeting the predefined conditions of fitness and net contribution”	Makes the value criterion explicit
“Zero-cost growth”	“Measured direct costs are [amount]; unreported capacity and opportunity costs are limitations”	Does not treat invisible cost as absent
“Guaranteed ROI”	“Observed return calculated for the specified period and assumptions”	Does not carry certainty into the future

Wording to avoid or unauthorised use	Permitted alternative	Reason
“Organic demand”	“Demand without a source statement/tracking record”	Does not assign the unknown to a preferred channel
“Qualified lead”	“Applicant meeting the specified fitness criteria”	Makes the marketing label measurable
“Sustainable growth”	“Positive net contribution and capacity fit across the specified observation horizon”	States time and conditions

E5 — Audit, Independence and Objection

Table E.5 — Language of audit, independence and objection

Wording to avoid or unauthorised use	Permitted alternative	Reason
“Independent audit”—where the practitioner examines their own work	“Internal assessment” or “separate second-pair-of-eyes review”	Does not conceal the institutional relationship
“Conforming”—without scope	“Conformity opinion limited to this entity, system, language, date and protocol”	Makes the judgment's boundary visible
“Full compliance”	“No nonconformity was observed against the listed provisions; the exceptions are...”	Does not exceed the audit universe
“The objection was refused; the decision is correct”	“The objection was refused under this evidence and criterion; the condition for reapplication is...”	Offers reasons rather than authority
“Cannot be disclosed because it is a trade secret”	“This record class is confidential; its existence/absence and effect on the decision are stated in the public summary”	Balances confidentiality with accountability
“Final decision”	“Decision in force until the stated date and review conditions”	Keeps correction and withdrawal open
“Validated by the Framework”	“Assessed by [decision authority] against criteria defined in the Framework”	Does not make the text an acting adjudicator
“No complaints”	“No complaint was recorded during the specified period; channel access was...”	Does not treat silence as satisfaction

E6 — Vocabulary of Uncertainty

Recommended verbs and qualifiers by level of evidence:

Table E.6 — Recommended language by level of uncertainty

Evidential condition	Recommended language
Direct and current evidence about the real world	“verified”, with scope stated
Recorded system answer	“observed”, “appeared in the answer”
Repeated sample	“recurred in the specified sample”
Limited support	“appears consistent with”, “indicates”
Conflicting support	“mixed finding”, “insufficient for judgment”
Association without causation	“observed together”, “may be associated”
Future assessment	“expected” only with reasons and uncertainty
Absence of data	“unknown” or “not measured”
Threshold not met	“did not meet the protocol threshold”
Critical risk	“suspended”, with decision owner and reasons

“May” does not create trust; the probability’s basis, scope and effect on the decision must be stated. Nor is “the data shows” evidential language unless it identifies which data shows what, and within which boundary.

Prepublication Sentence Test

Every public sentence passes five questions: **Object**—what is the judgment about? **Denominator**—where does the number come from? **Scope**—which system, language, user, geography and time? **Authority**—does the word exceed the decision-maker? **Boundary**—which unproved conclusion might the reader draw? If the final answer is broader than the promise, the text is rewritten. Good standards language reduces not impact, but the space for misunderstanding.

Glossary

This glossary explains the particular use of terms within the *NobleJackal GEO Framework*. The definitions do not replace every meaning those terms carry in other disciplines.

Accreditation: Formal recognition, by a competent external mechanism, that a conformity-assessment body is competent to perform specified tasks. The first edition of the Framework is not accredited.

Algorithmic audit: A planned examination of claims, processes, risks and chains of accountability associated with an AI or algorithmic system. In this book, it is not confined to technical model testing.[10]

Appeal: A request for a record, coding, inference, authority, process or principle decision to be reviewed independently of the initial decision.

Audit file: The body of records that unites scope, evidence, method, sample, roles, conflicts of interest, findings, opinion and route of objection in one traceable chain.

Baseline: The comparison value recorded before an intervention under the same, or explicitly stated, measurement conditions.

Behavioural reproducibility: Recurring observations in the same family of contexts producing similar decision patterns. It does not require bit-for-bit identical output in closed, probabilistic systems.

Certification: Third-party assurance that a product, process, service or organisation meets specified requirements. The book does not claim that a certification programme exists before the necessary governance is established.[17]

Claim: A sentence whose truth or falsity can be tested against evidence. It may be an observation, causal, predictive, outcome or conformity claim.

Claim Register: The record in which material public and internal claims are kept with their evidence, scope, version, owner and status.

Cohort: A group of users or customers with a shared start time or characteristic, whose outcomes are followed over the same observation horizon.

Conflict of interest: A professional, financial, personal or institutional interest that affects, or could reasonably appear to affect, the impartiality of a decision. Disclosure matters, but does not by itself eliminate every conflict.

Conformity opinion: A dated, limited and contestable decision about whether a defined scope meets requirements under specified evidence and criteria.

Core representation: The minimum combination of identity, attributes and boundaries necessary for an entity to be recognised accurately in the context examined.

Counter-evidence: A record that weakens or limits a claim, or offers an alternative explanation. It is a mandatory part of the evidence file.

Critical error: A predefined misrepresentation in identity, authority, safety, suitability or another high-risk field capable of exposing a user to material harm.

Denominator: The total number of valid observations underlying a rate. Silently changing the denominator may materially distort the result.

Deviation record: The record stating when, why and by whom a departure from the preregistered method occurred, and its effect on the result.

Ecological validity: The extent to which measurement conditions resemble real user behaviour. A natural-intent arm may improve it while reducing comparability.

Ethical veto: Authority to stop a practice that creates a risk of material user harm, deception, discrimination or violation of rights, even if the practice is lawful.

Event trigger: A material change that reopens an item of evidence, protocol or decision without waiting for the scheduled review date.

Evidence chain: The traceable relationship from source to observation, observation to inference, and inference to the published sentence.

Evidence of outcomes: A record showing the relationship between representation and user behaviour, customer quality, commercial outcomes, harm or another result.

Evidence of representation: The complete answer and contextual record showing what a system said, in response to which prompt and at what time.

Evidence about the real world: The record of the real entity against which representation is compared. It may be an official record, contract, verified product information, technical measurement or suitable independent source.

Generative-answer system: A system that produces a natural-language answer synthesised from one or more sources or model knowledge in response to a user prompt. Products may have different technical architectures.

Generalisability: The extent to which a result in one sample can be carried to another population, system, language, geography or time. It is justified through similarity of scope and research design, not assumed.

GEO: *Generative Engine Optimization*. The field concerned with an entity's discoverability, presentation and sourcing in results synthesised by generative-answer or search systems. This book does not claim to have invented GEO; it proposes a standard of decision and accountability for the field.[1]

Independent review: Review by a competent person or board that has no material interest in the outcome, did not make the first decision and can change it freely. A separate department name does not by itself establish independence.

Judgment representation: A system recommending, excluding, ranking or treating an entity as risky within a specified user, need or comparison.

Kappa: An agreement coefficient for categorical coding that accounts for agreement expected by chance. This book uses Cohen's kappa for the binary, two-coder NJ-100 code; it is not by itself a judgment of quality.[15]

Machine-readable summary: A publication presenting scope, version, date, claim, evidence class, status and objection link in structured fields. It does not replace human reasoning.

Major release: A published version that materially changes a right, duty, decision threshold, scope or interpretation of conformity.

Materiality: The magnitude and likelihood that an error or change will have a consequential effect on a decision, user, business or the public.

Maturity domain: A dimension of Framework development assessed separately for normative coherence, operational usability, empirical effectiveness and external legitimacy.

Minimum Sufficient Intervention: The narrowest legitimate change capable of addressing a verified material problem without creating broader harm or authority.

Model-similarity measure: A tool calculating lexical or contextual proximity between two texts. It may assist coding; it is not a measure of real-world accuracy.[19]

Natural-intent prompt: A prompt created in the user's own words after the user is given a need and task, rather than a prepared sentence.

Negative/exclusion prompt: A test prompt in which the system is expected not to recommend the entity for an unsuitable need, or to state the entity's boundary accurately.

Net contribution: Economic value remaining after acquisition, delivery, refund, support, capacity and other defined costs are deducted from the relevant revenue.

NJ-100: Proposed pilot protocol for testing core-representation accuracy and stability through one hundred valid, independent target-user sessions in a defined observation cell. It is neither a certificate nor a single GEO score.

Non-response: The system failing to produce the relevant representation, issuing a safety refusal, saying “I don't know”, or producing no result for a technical reason. It may be a valid outcome under the research question.

Normative: Determining what ought to be done. A norm's internal coherence does not mean that its effectiveness in the world has been demonstrated empirically.

Observation cell: The predefined combination of entity, target user, intention, system, interface, session state, language, geography, time and prompt family.

Observation horizon: The period across which an outcome's formation, ageing or sustainability is observed.

Pilot conformity opinion: A limited decision about whether records meet criteria only within the specified scope and protocol version. It is not an accredited certificate.

Proportionality: Alignment of the burden of intervention, recording, audit and assurance with the level of risk, likelihood of harm and consequence of the decision.

Preregistration: Fixing the research question, sample, criteria, threshold, exclusions and analysis plan in a dated record before results are seen.

Prompt universe: The defined body of question, intention and context families through which users may express the relevant task. Test prompts are selected from this universe under a stated method.

Protocol lock: Versioning and fixing the measurement plan before results are seen. Necessary changes are recorded as deviations.

Raw agreement: The percentage of observations for which two or more coders assigned the same code. It does not adjust for agreement expected by chance.

Representation: The observable combination of identity, attributes, context and judgments attached to an entity in a specified answer.

Representation debt: The accumulated effect of obsolete, incomplete, contradictory or false information surfaces, creating an additional burden of correction for accurate representation.

Representation stability: The degree to which independent observations under comparable conditions converge on the same core representation. Stability does not include accuracy.

Series break: Loss of direct comparability between new observations and an older series following a material change in system, interface, evidence or protocol.

Single-score prohibition: The principle that distinct dimensions such as accuracy, visibility, sourcing, recommendation, stability and commercial value must not be hidden inside one composite GEO score.

Structured record: A digital record with defined fields, version and relationships. Structure does not make the content accurate by itself.

Synthetic case: A fictional body of people, institutions, data and outcomes created to demonstrate how a principle operates. It is not evidence of real-world effectiveness.

Sustainable commercial value: Value produced for the right customer through ethical methods, deliverable capacity and positive net contribution, without relying on harm or deception.

Temporal integrity: Alignment among the dates of evidence, measurement, intervention and decision, so that an obsolete record is not presented as a new judgment.

Uncertainty: The unknown carried by a judgment because of measurement, sampling, coding, context or time. Not a defect, but a quality that must remain visible in the decision.

Valid session: A session meeting the predefined participant, task, context and recording conditions and remaining in the denominator irrespective of its outcome.

Visibility: An entity being mentioned or accessible in the relevant answer. It is not the same as accuracy, recommendation or commercial value.

Wilson interval: A confidence-interval method expressing uncertainty in a binary proportion more reliably than a simple normal approximation, particularly near the extremes.[14]

Accurate core rate: The proportion of all valid sessions that meet the predefined core representation and contain no critical error.

Causation: A relationship showing that a change produced an outcome while reasonably excluding alternative explanations. Temporal order or correlation alone is insufficient.

Intervention: A planned change on an authorised surface in response to a verified representational problem. Not every observation requires intervention.

Normative proposal for a standard: A text offering coherent criteria, rights, duties and decision routes without automatically carrying external acceptance or accreditation.

Secondary semantic profile: A meaning profile that goes beyond a binary accurate/inaccurate decision to show identity, attribute, context, recommendation, source and error types separately.

Notes and Bibliography

Limits on the Use of Sources

This book does not claim that an entirely new field called GEO has been invented. The sources below identify neighbouring bodies of knowledge, including generative search, verifiability, information retrieval, risk management, traceability, audit, documentation, statistics, coder agreement and conformity assessment. None has independently validated the *NobleJackal GEO Framework* as a whole, the NJ-100 thresholds or the book's normative provisions.

A historical result reported for one system has not been generalised to every product in use today. An access date is given for web documents that may change. Public information pages for standards have not been treated as substitutes for paywalled full texts; they are used only to the extent supported by accessible official statements.

Numbered Notes

[1] Academic context of the term GEO. Pranjal Aggarwal, Vishvak Murahari, Tanmay Rajpurohit, Ashwin Kalyan, Karthik Narasimhan and Ameet Deshpande, “GEO: Generative Engine Optimization,” *Proceedings of the 30th ACM SIGKDD Conference on Knowledge Discovery and Data Mining*, 2024, pp. 5–16. DOI: [10.1145/3637528.3671900](https://doi.org/10.1145/3637528.3671900). Open preprint: [arXiv:2311.09735](https://arxiv.org/abs/2311.09735). The study addresses methods and experimental evaluation for increasing visibility in generative-engine responses. This book treats visibility optimisation as only one part of broader questions concerning representation, evidence, governance and commercial value.

[2] Distinction between citation and support. Nelson F. Liu, Tianyi Zhang and Percy Liang, “Evaluating Verifiability in Generative Search Engines,” *Findings of the Association for Computational Linguistics: EMNLP 2023*, pp. 7001–7025. DOI: [10.18653/v1/2023.findings-emnlp.467](https://doi.org/10.18653/v1/2023.findings-emnlp.467). Record: [ACL Anthology](https://aclanthology.org/2023.findings-emnlp.467). The paper's quantitative findings are limited to the systems and period examined; the book does not use them as current rates for every product.

[3] Stochasticity and contextual variation. OpenAI Cookbook, “How to Make Your Completions Outputs Consistent with the New Seed Parameter,” explains that system outputs are non-deterministic by default and that `seed` and `system_fingerprint` may improve the likelihood of similarity without guaranteeing it: [OpenAI Cookbook](https://cookbook.openai.com/index.html). The ChatGPT Search help document states that a search query may be influenced by contextual elements, including approximate location or memory, together with the user's prompt: [Searching in ChatGPT](https://openai.com/help/search). Accessed 14 August 2026. These product documents may change; the book does not present their behaviour as a product-independent certainty.

[4] Retrieval-augmented generation. Patrick Lewis et al., “Retrieval-Augmented Generation for Knowledge-Intensive NLP Tasks,” *Advances in Neural Information Processing Systems* 33, 2020, pp. 9459–9474. [NeurIPS record](#). The paper introduces RAG, an approach combining parametric and non-parametric memory. Access to a source does not mean that every generated sentence is accurate in the world.

[5] Evaluation of generative information retrieval. Lukas Gienapp et al., “Evaluating Generative Ad Hoc Information Retrieval,” *Proceedings of the 47th International ACM SIGIR Conference on Research and Development in Information Retrieval*, 2024, pp. 1916–1929. DOI: [10.1145/3626772.3657849](#). Open preprint: [arXiv:2311.04694](#). The paper discusses why conventional ranking evaluation is insufficient on its own for generated answers; this book does not present its measurement provisions as a validated extension of that work.

[6] Error boundary of generative answers. OpenAI, “Does ChatGPT tell the truth?”, an official help document stating that ChatGPT may produce inaccurate or misleading outputs, fabricated citations and overconfident answers. [OpenAI Help Center](#). Accessed 14 August 2026. The limitation is used not as a provider-specific error rate, but as an example of why answers require verification against external evidence.

[7] Risk management. Elham Tabassi, “Artificial Intelligence Risk Management Framework (AI RMF 1.0),” NIST AI 100-1, 2023. DOI: [10.6028/NIST.AI.100-1](#). [NIST publication page](#). NIST’s govern, map, measure and manage functions belong to the neighbouring risk literature; NobleJackal’s provisions make no claim of NIST conformity.

[8] Generative-AI risk profile. National Institute of Standards and Technology, “Artificial Intelligence Risk Management Framework: Generative Artificial Intelligence Profile,” NIST AI 600-1, July 2024. DOI: [10.6028/NIST.AI.600-1](#). [NIST full text](#). The Profile is a companion document for managing risks specific to generative AI across the lifecycle.

[9] Provenance and traceability. Timothy Lebo, Satya Sahoo and Deborah McGuinness (eds.), “PROV-O: The PROV Ontology,” W3C Recommendation, 30 April 2013. [W3C PROV-O](#). The W3C Recommendation provides an ontology for machine-readable provenance relationships among entities, activities and agents. The book’s record chain does not impose that ontology as a mandatory technical form.

[10] End-to-end algorithmic audit. Inioluwa Deborah Raji, Andrew Smart, Rebecca N. White, Margaret Mitchell, Timnit Gebru, Ben Hutchinson, Jamila Smith-Loud, Daniel Theron and Parker Barnes, “Closing the AI Accountability Gap: Defining an End-to-End Framework for Internal Algorithmic Auditing,” *Proceedings of the 2020 Conference on Fairness, Accountability, and*

Transparency, pp. 33–44. DOI: [10.1145/3351095.3372873](https://doi.org/10.1145/3351095.3372873). The book identifies a neighbouring concern with lifecycle documentation and accountability; it does not say that this work validates the NobleJackal audit model.

[11] Principles of trustworthiness and accountability. OECD, “OECD AI Principles,” adopted in 2019 and updated in May 2024. [OECD AI Principles](#). The principles concerning robustness, safety, traceability and accountability are particularly adjacent to the book's discussion of governance. Accessed 14 August 2026.

[12] Model documentation. Margaret Mitchell et al., “Model Cards for Model Reporting,” *Proceedings of the Conference on Fairness, Accountability, and Transparency*, 2019, pp. 220–229. DOI: [10.1145/3287560.3287596](https://doi.org/10.1145/3287560.3287596). Model cards offer an approach to documenting purpose, performance, groups, conditions and limitations.

[13] Dataset documentation. Timnit Gebru, Jamie Morgenstern, Briana Vecchione, Jennifer Wortman Vaughan, Hanna Wallach, Hal Daumé III and Kate Crawford, “Datasheets for Datasets,” *Communications of the ACM*, 64(12), 2021, pp. 86–92. DOI: [10.1145/3458723](https://doi.org/10.1145/3458723). The paper proposes documenting decisions concerning a dataset's motivation, composition, collection, processing, use and maintenance.

[14] Wilson interval for binary proportions. Edwin B. Wilson, “Probable Inference, the Law of Succession, and Statistical Inference,” *Journal of the American Statistical Association*, 22(158), 1927, pp. 209–212. DOI: [10.1080/01621459.1927.10502953](https://doi.org/10.1080/01621459.1927.10502953). The 95 per cent intervals in the NJ-100 examples are calculated with the Wilson score method without continuity correction. The interval does not correct sampling bias or dependent observations.

[15] Coder agreement. Jacob Cohen, “A Coefficient of Agreement for Nominal Scales,” *Educational and Psychological Measurement*, 20(1), 1960, pp. 37–46. DOI: [10.1177/001316446002000104](https://doi.org/10.1177/001316446002000104). Cohen's kappa adjusts categorical agreement between two coders for agreement expected by chance. It is not a measure of the codebook's conceptual validity or truth in the world.

[16] AI management system. ISO/IEC 42001:2023, *Information technology – Artificial intelligence – Management system*, first edition, December 2023. [Official ISO record](#). The public description identifies requirements for establishing, implementing, maintaining and continually improving an AI management system. The NobleJackal GEO Framework claims neither ISO/IEC 42001 certification nor conformity.

[17] Bodies certifying products, processes and services. ISO/CASCO's public description summarises functions including impartiality, evaluation, decision, surveillance, suspension/withdrawal, complaints and appeals in

product, process and service certification under ISO/IEC 17065. [ISO/CASCO — Bodies](#). The [official ISO record for ISO/IEC 17065:2012](#) states that it was confirmed in 2024 and remained current on 14 August 2026, while a final draft intended to replace it was under development. The book does not replace the full standard, assume conformity with a future edition or claim ISO approval.

[18] Good practice for credible sustainability systems. ISEAL, “ISEAL Code of Good Practice for Sustainability Systems, Version 1.1,” effective 1 September 2025. [ISEAL resource page](#). The Code's public approach considers system components including purpose, strategy, risk management, stakeholder engagement, transparency, evaluation and learning together. NobleJackal claims neither ISEAL membership nor ISEAL validation.

[19] Contextual text similarity. Tianyi Zhang, Varsha Kishore, Felix Wu, Kilian Q. Weinberger and Yoav Artzi, “BERTScore: Evaluating Text Generation with BERT,” *International Conference on Learning Representations*, 2020. [arXiv:1904.09675](#). BERTScore assesses similarity by matching contextual embeddings in candidate and reference texts. The Framework may use such tools to assist preliminary classification; it does not treat them as adjudicators of accuracy, ethics or conformity.

Sourcing Principle

The presence of a source in this list does **not** establish that:

1. its authors endorse the *NobleJackal GEO Framework*;
2. its finding in one domain can be carried to every generative system;
3. the book's normative choices have gained scientific or institutional acceptance.

The sources serve a narrower purpose: to acknowledge the field's history honestly, make neighbouring methods visible, and connect testable sentences to primary or authoritative records. A new Framework threshold or provision is not presented as though it already existed in the external literature; it is labelled explicitly as a **NobleJackal proposal**.

Version, Validity and Validation Roadmap

Publication Identity

Table V.1 — Publication identity of the English edition

Field	Status in this text
Work	<i>NobleJackal GEO Framework</i>
Subtitle	<i>A Proposed Standard for Representation, Evidence and Sustainable Value</i>
Author name	Julian Gauss
Author identity	Julian Gauss is the pen name of Kaan Muraz
Language	English normative edition derived from the locked Turkish master
Text status	Publication-locked English language edition 1.0.0; internal normative-equivalence and editorial review completed; external validation not completed
Normative status	Proposal for a standard
Publication year	2026
Last information review	14 August 2026
External validation	Not yet completed
Accreditation	None
Active certification/badge programme	None

This table appears within the book so that the limits of the publication are not concealed. Language editions are derived from the locked Turkish master and may not be published before normative equivalence is verified. The test is not literal translation but **equivalence of meaning**: scope, rights, duties, thresholds and prohibitions in every language are compared with the Turkish master.

Authorship and Production Record

The book's intellectual core, first draft, chapter system and foundational provisions belong to Kaan Muraz. Julian Gauss is the author's pen name. AI-assisted tools were used in developmental editing, source research, structural separation, line editing, consistency work and publication preparation. Final ownership of the ideas, the decision to publish and accountability for the provisions in the text remain with the human author.

Editorial intervention did not seek to turn the book into another thesis. It sought to source the existing core, reduce the burden of repetition, separate normative provisions from their reasons, label synthetic cases accurately and bind routes to decision to auditable records. AI output was not used as a substitute for sources or verification of reality.

Four Distinct Maturity Domains

A framework matures separately across four domains:

Table V.2 — Framework maturity domains

Domain	Question	Status at first publication
Normative coherence	Do provisions contradict one another; are rights and prohibitions explicit?	Internal and editorial coherence review complete; this is not field validation
Operational usability	Can an independent practitioner use the records and decision flow?	NJ-100 and the record set remain in pilot status
Empirical effectiveness	Does the method produce reliable distinctions and better decisions in real cases?	Not demonstrated; field research required
External legitimacy	Have independent communities examined and contested the method and established a mechanism of acceptance?	Not established

Strength in the first domain does not substitute for the fourth. Nor does the author's expertise substitute for external legitimacy. The distinction does not diminish the method's value; it shows which sentence cannot yet be made.

Versioning Rule

The Framework uses three-part versioning:

- **Major release (X.o.o):** Changes a foundational provision, right, prohibition, threshold, meaning of conformity or hierarchy of decision.

- **Minor release (o.X.o):** Adds a backward-compatible protocol, record field, explanation or example.
- **Correction release (o.o.X):** Corrects wording, links, formatting or a clear error without changing meaning.

Every published release states:

1. effective date;
2. link to the previous release;
3. provisions added, changed and withdrawn;
4. effect of the change on decisions;
5. transition period;
6. conditions for reviewing earlier opinions;
7. roles proposing and approving the change.

A major release does not silently carry old conformity opinions into the new version. The provisions requiring reassessment are published.

Change Statutes

Table V.3 – Version-change statutes

Status	Meaning
Draft	May not be used for a public decision
Public consultation	Open for comment; not yet in force
Release candidate	Passing through editorial and technical gates
In force	Active normative release from the stated date
Correction pending	A material objection is under examination
Suspended	Production of new decisions has temporarily stopped
Superseded	Old release preserved in the archive; points to its replacement
Withdrawn	Relevant provision may not be used for a reliable decision

“Old” and “wrong” are not the same. An earlier release is archived to explain decisions made in its time. A withdrawn provision remains visible together with the reasons for withdrawal.

Validation Stages

Stage 0 — Desk-based integrity

- Provisions are matched across chapters, Constitution and record set.
- Terminology, cross-references, sources and prohibited language are scanned.
- Synthetic cases are checked for concealed real organisations or impressions of unverified success.
- The decision tree is tested against contradictory examples.

This stage may show that the text works internally. It does not show that the Framework is effective in the world.

Stage 1 — Facilitated real-world pilot

- At least three voluntary real entities of different sizes and risk levels are selected.
- The protocol and analysis plan are recorded before results.
- The practitioner is trained; the author's team may observe.
- Usability failures, record burden and uncertainty in decisions are retained.

Only a pilot report is published at this stage; no certificate or general language of success is used.

Stage 2 — Independent application

- Teams with no institutional relationship to the author, NobleJackal or the paid practitioner apply the Framework.
- Agreement among independent decisions on the same case is examined.
- Codebooks, time, cost and failure examples are compared.
- The decision guided by the Framework may be compared with a simpler baseline method.

The purpose is not to increase the number of favourable results, but to discover the method's discriminatory power and failure modes.

Stage 3 — Multilingual and temporal replication

- Normative English, German, Russian, Spanish and Arabic editions are prepared from the locked Turkish master.
- Each language receives bidirectional meaning review and subject-matter review.

- Arabic is separately tested for right-to-left reading, terminological consistency and directionality in structured data.
- NJ-100 is applied separately across different language, geography, system and time cells.
- The study examines whether linguistic and cultural conditions introduce systematic bias into thresholds.

Texts need not be identical word for word; their normative force must be equivalent.

Stage 4 — External governance and assurance decision

- Public consultation, disclosed responses to comments and a change record are conducted.
- A conflict-of-interest policy, independent decision board and appeal authority are established.
- Practitioner competence, surveillance, suspension and withdrawal processes are tested.
- External opinion determines whether a conformity or certification programme is necessary, proportionate and legitimate.
- If a programme is established, its conformity-assessment infrastructure is separately examined against relevant international requirements.[17][18]

Existing standards such as ISO/IEC 42001 also treat the establishment, implementation, maintenance and continual improvement of a management system systematically.[16] This proximity does not imply NobleJackal's ISO conformity; it is a starting point for analysing conflict and interoperability with external standards.

Publication Gates

A release must pass every gate below before being declared **in force**.

Intellectual gate

- Every chapter has one primary question and judgment.
- Foundational provisions are consistent with one another and with the Constitution appendix.
- What the Framework does not measure and does not guarantee is explicit.

Evidence gate

- Claims about the external world match primary or authoritative sources.

- Sources are not presented as external endorsement of the Framework.
- Sources capable of change carry access dates and version boundaries.

Operational gate

- Every decision has a record field, owner and route of objection.
- NJ-100 thresholds, denominator, uncertainty and second window can be applied explicitly.
- A single score, silent exclusion and post-outcome rule changes are prevented.

Ethics and governance gate

- Synthetic data is labelled explicitly.
- Language about independence and conflicts of interest reflects reality.
- Certification, accreditation, “first” and guarantee claims are not used without authority.

Language and publication gate

- The edition receives final review for meaning, rhythm, terminology and punctuation in its own language.
- Headings, cross-references, note numbers and tables are complete.
- The hierarchy of accessible Word/PDF and web editions is preserved.
- Human-readable and machine-readable editions carry the same normative identity.

Parallel Publication for Humans and Machines

The Framework may eventually have two distinct presentation surfaces:

- 1. Human edition:** Reasons, narrative, cases, tables, notes and an accessible reading structure.
- 2. Machine-readable source edition:** Structured records presenting the same provisions in concise, uniquely identified and versioned fields.

The machine edition cannot become a concealed alternative interpretation of the book. Every foundational provision carries a stable identifier such as NJ-C01, version, effective date, normative force, scope, linked record, source and objection link. When the human text changes, the structured edition is updated under the same change record. Allowing an AI crawler to access the material does not guarantee that it will cite or interpret the Framework accurately.

Publication location, URL architecture, data formats and multilingual addresses are determined through a separate publication design after the Turkish master is locked. The text does not predetermine that decision.

Commitment to Correction and Objection

The published medium shall provide a permanent route for **correction and objection**. At minimum, an application can state:

- the provision and version challenged;
- type of error;
- person or decision affected;
- supporting record;
- correction requested;
- confidentiality requirement.

When a material objection is received, the former provision is not changed silently. The application date, review status, interim risk decision and outcome are recorded. If the error is verified, the correction is published together with the difference between old and new text. The Framework's credibility depends not on claiming never to err, but on correcting error visibly and fairly.

Final Boundary of the First Publication

The first publication may say:

The NobleJackal GEO Framework is a proposal for a standard developed to evaluate institutional representation in generative AI systems through a chain of evidence, measurement, intervention, governance, time, sustainable commercial value, audit and objection.

It may not yet say:

"It is a globally accepted, independently validated GEO standard."

The distance between the two sentences can be closed not through promotion alone, but through real cases, independent replication, public objection and external governance. This roadmap exists not to conceal that distance, but to make it measurable.