GOT-V1-3c-6

The Case for Formalisation

Load-Bearing Language, Institutions, Machines, and Auditable Grounding

---

Abstract

Natural language does not fail because it compresses. Compression is one of its powers. A finite observer-system could not communicate, remember, classify, coordinate, deliberate, govern, investigate, or act if every word had to restate the whole dependency structure behind its use. Language works because it carries more than it says. It condenses prior distinctions, inherited practices, social memory, technical conventions, institutional rules, categories of judgement, and action affordances into usable form. At abstract scale, a word, term, or category is a Soft Horizon: a provisional closure of wider relation into communicable, reusable form without possession of the relation as total state.

The danger begins when compressed language becomes load-bearing while the conditions that ground its use disappear from view. A term may drift across contexts while retaining authority from an earlier setting. A category may acquire procedural force after its relation to the target has weakened. A policy phrase may become a decision rule without preserving the relation it was meant to name. A scientific, legal, administrative, technical, or machine label may be justified by adjacent terms that themselves depend on the first. A word may then become not a carrier of grounded relation, but a reified Soft Horizon inside a decision path: a provisional closure treated as if it were settled warrant.

Formalisation is required where language becomes load-bearing. This does not mean natural language should be frozen, purified, replaced, or forced into a single universal convention. Ordinary language may remain adaptive, metaphorical, contested, cultural, rhetorical, and evolving. The requirement is narrower and stronger. When language is used to classify, decide, govern, enforce, allocate, investigate, automate, certify, or justify action, the operative linguistic closure must remain recoverable. The system must be able to say what relation the language is carrying, for which locus, under what availability condition, under what closure demand, through what abstraction, evidence path, comparator, residual, threshold, safeguard, and update condition.

The problem is especially acute in institutions and machine-mediated systems. Institutions convert linguistic Soft Horizons into action through roles, rules, categories, records, thresholds, procedures, permissions, incentives, archives, and authority. Machines convert language and data into action through code, model architecture, training distributions, optimisation, interfaces, metrics, classifications, and outputs. Neither substrate repairs epistemic under-specification by itself. Institutional procedure can make ungrounded language appear legitimate. Machine execution can make under-formalised language faster, more scalable, more consistent, and more opaque.

Ungrounded language creates epistemic debt: the deferred cost of acting as if the grounding burden of a compressed closure has already been paid. Such debt returns as misclassification, institutional opacity, governance failure, circular justification, machine amplification, inability to review, and failure to update. The purpose of formalisation is not to abolish uncertainty. It is to prevent uncertainty from being hidden inside authoritative language.

The governing rule is simple: load-bearing language must not carry authority beyond the conditions that ground its use.

---

1. Introduction: load-bearing language and recoverable authority

This paper concerns the formalisation of load-bearing language: the discipline required when language, as a necessary Soft Horizon of finite observer-systems, becomes authority.

The previous paper defined trust as reliance closure under residual. A finite locus cannot reopen every dependency before every act. It must rely on surfaces, signals, words, records, persons, roles, institutions, instruments, models, and machines without reconstructing the full dependency chain each time. Trust reduces audit so finite systems can continue. But where trust carriers become authority carriers, reduced audit must become recoverable. A carrier may reduce local checking; it must not carry more reliance-strength than its grounding permits.

Formalisation begins at that point.

Where reliance crosses relations produced, maintained, or governed by other distinction-conditioned loci, the trusted relation may already carry Responsibility Provenance. Trust can background that provenance without transferring it. Formalisation becomes relevant where consequential use requires enough of the carrier, claim, and responsibility path to remain recoverable. This does not make Responsibility Provenance a universal requirement of every formalised claim. It becomes material where load-bearing language makes, assigns, obscures, or relies upon an attribution of responsibility, or where a downstream decision depends on a responsibility-bearing consequential relation whose provenance matters to the claim. Such a relation may concern realised transformation, non-application, a retained or maintained condition, or material conditioning of another locus's Counterfactual Inclusion Field.

Language is one of the most important trust carriers. A term allows a finite locus to rely on prior closure-work. It carries distinction, memory, category, judgement, and practice without restating them. The word "risk" can carry a history of possible harm, evidence, probability, severity, threshold, intervention, and precaution. The word "evidence" can carry rules of admissibility, reliability, record, testimony, instrument, inference, and burden. The word "fairness" can carry ethical, legal, procedural, social, and institutional expectations. The word "safety" can carry technical standards, policy judgement, acceptable risk, and action threshold.

This compression is not the defect. A society without compressed language could not act. The problem begins when compressed language becomes operative while the conditions of its authority are no longer recoverable.

Load-bearing language is language that does consequential work inside a decision path. It converts an available or asserted relation into an operative closure that may classify a person, system, event, risk, action, harm, object, observation, model, institution, or process. It may activate authority. It may justify intervention. It may distribute resources. It may exclude. It may punish. It may certify. It may trigger an automated output. It may settle a dispute. It may guide institutional memory. It may become evidence. It may become explanation. It may become a record that later systems treat as fact.

In each case, the language may also carry reliance. A later locus may trust the term, record, category, credential, model, procedure, or institutional status as if the closure-work behind it

had already been performed adequately. A decision-maker may rely on a classification. A court may rely on a record. A public body may rely on a risk category. A platform may rely on a machine label. A regulator may rely on a certificate. A patient, citizen, worker, applicant, defendant, student, user, or community may then bear the consequence of a linguistic closure whose grounding they cannot see.

Such language cannot remain only rhetorically stable. It must be grounded enough for the relation it carries to be recovered. The claim must remain recoverable through its reference conditions. The category must remain recoverable through the relation it was meant to preserve. The evidence path supporting Epistemic Authority must remain distinguishable from any Structural Authority governing its use. The residual must remain visible. The update condition must be stated. Without such recoverability, a linguistic Soft Horizon becomes a route by which unearned closure travels through systems.

Formalisation is the discipline that makes such recoverability possible. It does not make every term rigid. It does not reduce judgement to a mechanical rule. It does not make institutions automatic. It does not make machines epistemically adequate by execution. It does not claim total closure. It asks a narrower question: when language carries authority, what conditions must remain attached so that the claim, category, decision, record, or output can be audited?

The governing rule is:

Load-bearing language must not carry authority beyond the conditions that ground its use.

Where such language carries reliance through a person, role, record, credential, institution, model, procedure, certificate, or machine output, it must also not carry reliance-strength beyond the recoverable conditions under which that carrier satisfies the relying locus's closure demand.

This rule does not apply to all speech equally. Conversation, metaphor, art, ordinary persuasion, social experimentation, conceptual invention, and evolving convention do not all require explicit formalisation. Language must remain capable of drift, play, compression, association, invention, and transformation. A term may close relation enough to communicate, coordinate, or explore without being formalised for governance. A society that tried to freeze all meaning would not become more truthful. It would become brittle and coercive.

The rule applies where language becomes operative in consequential decision paths. It applies where terms become classifications, classifications become thresholds, thresholds become actions, and actions affect loci that may not control the language by which they are judged. It applies where institutions use language to govern. It applies where machines use labels, data, prompts, rules, or learned patterns to classify and act. It applies where law, public administration, science, medicine, finance, safety systems, education, platforms, and automated governance depend on categories whose grounding is not recoverable.

The case for formalisation is therefore not a case against language.

It is a case against ungrounded authority carried by reified language.

---

2. Language failure under load

The first mistake is to treat the language problem as only a compression problem.

Compression matters. A word carries more than it states. A technical term carries prior distinctions, field practices, instruments, standards, corrections, paradigms, failures, and inherited exclusions. A legal term carries precedent, statutory context, institutional procedure, argument, interpretation, enforcement, and appeal. A medical term carries diagnostic history, measurement regimes, symptom clusters, causal models, treatment protocols, and risk thresholds. An administrative term carries policy purpose, bureaucratic category, eligibility condition, record format, and decision consequence. Each term is a Soft Horizon over a wider field of relation.

Without compression, no observer-system could act. Every claim would require the full restatement of its history. Every decision would require total reconstruction. Every institution would collapse under the burden of re-deriving the meaning of its own terms. Language is powerful because it reduces relation into repeatable form. It allows distinctions to be retained, transmitted, redeployed, contested, corrected, and coordinated. The fact that language closes relation provisionally is therefore not the error.

The danger begins when compression continues to carry authority after its grounding conditions have disappeared from the decision path. A term may remain familiar while its relation has changed. A category may remain administratively useful while its target has moved. A policy term may appear precise because it is repeated in official settings. A scientific label may remain authoritative because it descends from a once-disciplined context. A legal phrase may stabilise procedure while concealing unresolved relation. A machine label may appear neutral because it is embedded in code. In each case, a Soft Horizon begins to travel beyond the conditions that made it usable.

Compression becomes failure when the compressed term is used as if it still carried its grounding, while the system using it can no longer recover what that grounding is, what residual remained, what scope limited the original closure, or what carrier process earned reliance in the first place. The term continues to reduce audit because it still appears settled. A record is trusted because it is a record. A credential is trusted because it is credentialed. A procedural phrase is trusted because it belongs to an authorised process. A model label is trusted because it appears technical. The failure is therefore not only semantic. It is a failure of vicarious closure: the system borrows closure-work from a carrier while losing the conditions that made the borrowing legitimate.

This failure is intensified by semantic drift.

Semantic drift is not an accident or corruption of language. It is an ordinary feature of language. Terms move across contexts. They shift emphasis. They widen, narrow, specialise, popularise, politicise, institutionalise, or technicalise. A word that begins as description may become diagnosis. A diagnostic term may become identity. An identity term may become administrative category. An administrative category may become risk marker. A risk marker may become enforcement trigger. A term that once named a relation may later name a population, a procedure, a status, a funding category, a liability class, or a machine label. Soft Horizons move because language moves.

Drift is not inherently bad. It allows language to adapt to new relations. It allows communities to name emerging patterns. It allows institutions to revise outdated categories. It allows science to extend concepts. It allows law to respond to new harms. It allows technology to create new operational distinctions. A frozen language would fail the world as badly as an ungrounded one.

The problem is not drift as such. The problem is load-bearing drift without audit: the movement of a Soft Horizon into a new decision path while authority from an earlier grounding remains attached.

A drifting term can retain authority from an earlier use while functioning differently in a later one. A term may sound scientific while being used administratively. It may sound legal while being used politically. It may sound moral while being used actuarially. It may sound technical while being used rhetorically. It may sound neutral while encoding discretionary judgement. It may sound objective while resting on a hidden abstraction cut.

When semantic drift becomes load-bearing, the system must ask what the term means in the present decision path. Not what it meant historically in a general sense. Not what it means in ordinary conversation. Not what an institution would prefer it to mean. The relevant question is narrower: what relation did this term close here, for which locus, under what availability condition, under what closure demand, through which boundary, scale, evidence path, residual condition, threshold, and consequence?

Without that question, drift becomes a route of institutional manipulation. A term can move between contexts while institutional authority remains attached to it. This allows actors to exploit ambiguity without appearing to invent anything. They can say the policy already uses the term. They can say the category already exists. They can say the risk has already been identified. They can say the model already classifies the case. They can say the procedure has already been followed. But the central question remains unanswered: did the operative use of the term preserve the relation it claimed to carry?

The third failure is circular ungrounded dependency.

This occurs when a term is justified by other terms that eventually depend on the original term, or by institutional practices that assume what they are meant to establish. A decision is called reasonable because it follows the policy. The policy is called reasonable because it reflects the institution’s risk framework. The risk framework is called appropriate because it aligns with organisational standards. The standards are treated as authoritative because they are institutionalised. The institution then treats the decision as grounded because it follows the standards. At no point does any closure return to the target relation it claims to preserve.

Circular dependency is not always obvious. It often appears as a network of respectable terms. A governance decision may be justified by risk, proportionality, public interest, safeguarding, compliance, efficiency, professional judgement, evidence-led practice, and operational necessity. Each term may have legitimate uses. Each may carry some fragment of prior grounding. But if the terms are used only to support one another, and none returns to the relation being claimed, the decision path becomes semantically closed without being epistemically grounded.

The same structure appears in technical systems. A model output may be trusted because it is validated. Validation may be accepted because the benchmark is standard. The benchmark may be standard because it is widely used. It may be widely used because institutions accept it. Institutions may accept it because it produces recognised metrics. The metrics may be recognised because they correlate with prior labels. The prior labels may have been generated by the same institutional categories now being automated. The result may appear rigorous while remaining circular: trust moves from model, to benchmark, to institution, to record, to prior label, and back to model without returning to the target relation the system claims to preserve.

Circularity is especially dangerous because it can produce stability. A system can be internally coherent, procedurally repeatable, and institutionally accepted while still failing to preserve the

target relation. It may produce the same output across cases. It may maintain records. It may generate reports. It may satisfy compliance checks. It may even improve operational efficiency. But if the relation under evaluation has disappeared behind mutually supporting categories, the system has stabilised around a reified Soft Horizon rather than the world it claims to govern.

This is why formalisation must go deeper than definitional clarity. A glossary alone is not enough. A term may be defined and still ungrounded. A policy may define risk but fail to state how risk enters the decision path. A law may define harm but fail to preserve the harmed locus. A model may define success but optimise a proxy. An institution may define evidence but exclude the relevant relation through its admissibility rules. A procedure may define fairness but make the affected locus invisible.

The issue is not whether a term has a definition. The issue is whether its load-bearing use is auditable.

A load-bearing term must be recoverable at the point of use. The system must be able to say what relation the term is carrying, what alternatives it excludes, what threshold it activates, what evidence supports it, what scale it operates at, what residual remains, what discretion entered, what consequence follows, and what would force revision. If those conditions cannot be recovered, the term may remain linguistically meaningful, but it is not epistemically grounded for the decision it is being asked to carry.

This distinction matters because language often gains authority by appearing already settled. A familiar noun can make a relation appear known. A policy phrase can make a judgement appear procedural. A score can make a contested classification appear quantitative. A legal category can make a lived relation appear resolved. A machine label can make an inference appear objective. A record can make a claim appear warranted.

The appearance of settled language is one of the main ways a Soft Horizon becomes reified inside institutions.

The discipline required here is not a demand for perfect knowledge. It is a demand for visible grounding. A claim may remain uncertain. A category may remain provisional. A decision may still need to be made under residual. An institution may still need to act before final closure. A machine may still operate under probabilistic conditions. But the system must not hide the difference between grounded closure, provisional action, unresolved residual, semantic authority, and reified compression.

Language failure under load therefore has a specific structure. Compression carries hidden dependency. Drift moves terms across contexts while authority persists. Circularity lets terms support one another without returning to the relation. Institutional procedure then converts those terms into action. Machine systems may encode and scale them. The result is not merely linguistic ambiguity. It is epistemic opacity with consequence.

Formalisation begins where this opacity becomes unacceptable: where a Soft Horizon has become load-bearing and the conditions of its authority can no longer be recovered.

It does not begin by replacing natural language. It begins by asking what language is doing when it becomes operative. The question is not whether society should stop using words such as risk, harm, safety, fairness, objectivity, evidence, intelligence, knowledge, responsibility, or public interest. The question is what conditions must become explicit when such words decide, classify, allocate, enforce, automate, or justify.

The answer is not a new vocabulary alone.

It is an audit path.

3. Grounding without freezing language

The case for formalisation can be misunderstood as a demand to control language.

That would be a serious error.

Natural language must remain open. It must be able to absorb new experience, new relations, new technologies, new harms, new identities, new scientific findings, new legal conflicts, new political pressures, and new forms of social coordination. A living society cannot operate with frozen meanings. A living science cannot operate with fixed terminology immune to revision. A legal system cannot respond to novel cases if categories never adapt. A culture cannot name emerging relations if language is prohibited from generating, revising, and contesting Soft Horizons.

The aim is not to stop semantic change. The aim is to prevent semantic change from being used as ungrounded authority in load-bearing settings.

This distinction is central. A term may evolve freely in ordinary communication. It may be used poetically, politically, culturally, rhetorically, experimentally, and loosely. People may contest it. Communities may reshape it. Disciplines may specialise it. Institutions may adopt it. None of this is automatically a failure. Language is one of the means by which observer-systems discover, negotiate, and preserve relation.

But when a term enters a consequential decision path, something changes. The term is no longer only expressive. It becomes operative. It can activate a category, justify a threshold, alter a record, trigger a sanction, admit or exclude evidence, allocate resources, label a person, classify a system, determine eligibility, initiate surveillance, justify intervention, or guide machine output. At that point, the system cannot rely on the general social life of the term. It must ground the operative instance of the Soft Horizon.

Formalisation therefore does not freeze language. It preserves a recoverable time-index for the conditions under which language became consequentially operative within a decision, maintenance, or governance path. Where that use is later assessed as responsibility-bearing, the retained path supports Responsibility Provenance by preserving what relation, distinction, abstraction, evidence, residual, alternatives, threshold, authority, and materially relevant standing conditions were operative at the time.

To time-index responsibility is to make clear what the term was carrying when it became consequentially operative. A word may later change. A category may later be revised. A scientific concept may be refined. A legal standard may be interpreted differently under new cases. A machine label may be retrained. But the operative use must remain recoverable in the path where it produced, maintained, withheld, or conditioned consequence. The question is not whether the term has one meaning forever. The question is what relation this use carried here.

A public body may use a term such as vulnerability, risk, safeguarding, misconduct, public interest, extremism, efficiency, resilience, or compliance. Each term may have legitimate administrative value. Each may also be manipulated. Formalisation does not forbid the department from using the term. It requires that when the term becomes the basis for action, the department must show how it was grounded in the case.

The same is true in law. Legal systems depend on language that must remain interpretive: reasonableness, proportionality, intent, negligence, harm, fairness, duty, discrimination, best interests, public order, legitimate aim, necessity. These terms cannot be replaced by mechanical

definitions without damaging legal judgement. Yet their use must be auditable. A court, inquiry, tribunal, regulator, or review body may need to ask whether a term was applied to preserve the relation it claims to govern, or whether it functioned as a procedural shell for an ungrounded judgement.

The same is true in technical systems. A machine classifier may label content harmful, unsafe, relevant, fraudulent, high risk, eligible, anomalous, likely, or non-compliant. Such labels may be operationally necessary. But if they trigger action, their grounding must be recoverable. What training labels gave the term shape? What proxy was used? What threshold converts score into action? What false-positive and false-negative burdens are carried? What affected locus can challenge the classification? What residual remains? What update signal forces revision?

Grounding without freezing language therefore creates a middle path between two failures.

The first failure is semantic authoritarianism. This attempts to control language by imposing rigid definitions across all contexts. It treats ambiguity as a defect to be eliminated. It mistakes living language for a defective technical code. It hardens Soft Horizons everywhere, even where ordinary communication requires adaptation and contestation. It may produce surface clarity, but often at the cost of brittleness, exclusion, and inability to adapt. A society governed this way would not become more grounded; it would become more vulnerable to those who control the definitions.

The second failure is semantic opportunism. This allows institutions, authorities, or machines to exploit fluid language while avoiding accountability for its use. It treats meaning as flexible when power benefits from flexibility, and fixed when power benefits from enforcement. It allows a term to drift when challenged, but harden when applied. It allows ambiguity to survive in the institution while consequence is imposed on the affected locus.

Formalisation rejects both failures. It does not demand universal rigidity. It demands local accountability. It allows language to evolve, but requires the load-bearing instance to be recoverable. It keeps language open generally while grounding the operative closure locally.

This is why the burden of formalisation must be proportional to consequence. Not every use of a word requires audit. A conversation about risk need not become a risk framework. A metaphor about harm need not become a legal test. A public debate about fairness need not resolve into administrative categories. But when a term becomes the basis for coercion, exclusion, allocation, automated classification, institutional judgement, public authority, or irreversible action, the burden changes. The word must be grounded enough for review.

The key point is that formalisation does not replace judgement. It disciplines the conditions under which judgement becomes authoritative.[1]

A judge still judges. A doctor still diagnoses. A scientist still models. A public official still decides. A machine system may still classify. A regulator still evaluates. An institution still acts. Formalisation does not transfer or erase responsibility by converting judgement into mechanical outputs. Responsibility Provenance remains local to the consequential branch relations attributable to the relevant loci — including transformations actually performed and, where the discriminator is met, responsibility-bearing non-application, retained or maintained conditions, and materially governed field-conditioning — while any Responsibility Warrant attached to a role remains a separate allocation. It makes the claim-path visible enough that the judgement can be assessed, challenged, limited, revised, or defended without hiding behind the authority of the word.[2]

Grounding language therefore does not mean closing the future of language. It means preventing an operative closure from pretending that its future contestability, residual, and grounding conditions do not exist.

Language may remain alive.

Authority-bearing language must become answerable.

---

4. Epistemic debt

Ungrounded language creates epistemic debt.

Epistemic debt is the deferred cost of acting as if the grounding burden of a compressed closure has already been paid. A term, category, label, record, threshold, score, or procedure is used as if it carried the relation strongly enough for decision. The system proceeds. The decision is made. The record is created. The action is taken. The model is deployed. The category travels. But the grounding was not actually performed, or was performed only partially, or was performed in a context whose residual no longer travels with the closure.[3]

Debt is incurred when a system borrows closure without recording the borrowing.

A category may be used because it is administratively necessary before its relation to the target has been adequately specified. A proxy may be used because the target is hard to measure. A risk label may be used because action cannot wait. A legal standard may be applied under evidential limits. A machine classifier may be trained on inherited labels because no better labels are available. A public authority may act under threshold because delay would be harmful. These moves may be necessary. The debt is not created by finitude alone. It is created when the system lets the borrowed closure travel as if no borrowing occurred.

Formalisation does not forbid epistemic debt in every case. Finite systems often must act before full closure. The discipline of the unknown has already shown that action under residual may be justified. Trust has already shown that reduced audit may be necessary. The question is whether the residual, threshold, safeguard, and update condition remain visible. A recorded debt can be managed. An unrecorded debt becomes hidden authority.

Epistemic debt accumulates where language is useful enough to keep working while its grounding weakens.

A term may begin as a careful abstraction but later become administrative shorthand. A category may begin as a provisional tool but later become an institutional object. A record may begin as bounded registration but later become warrant. A threshold may begin as a pragmatic action point but later become proof. A model output may begin as decision support but later become decision authority. Each step reduces local work. Each step can also defer grounding.

The debt may remain invisible for a long time because the system continues to function. Ordinary cases fit the category. Records are accepted. Procedures are followed. Users comply. Outputs appear consistent. Institutions recognise one another’s terms. Machine systems reproduce the inherited labels. The very stability of the system can hide the debt. Because the language works well enough in familiar cases, the system may treat its success as evidence that the grounding burden has been paid.

But debt returns.

It may return as misclassification. A category that never preserved the target relation fails at the edge. A person, case, risk, harm, symptom, event, or system does not fit the inherited language, but the institution forces it into the category because the category is what the procedure can see. The affected locus is then misdescribed by the language that governs it.

It may return as opacity. The system can produce documents, forms, records, and policy references, but it cannot reconstruct the relation between the operative language and the decision consequence. The decision path exists administratively but not epistemically.

It may return as brittleness. A system built on under-grounded categories may function under familiar cases but fail sharply at edge cases. Because the category was never tied to its discriminator conditions, it cannot adapt intelligently. It either overextends, excludes wrongly, or demands ad hoc exceptions. The debt becomes visible when the system confronts cases its language cannot carry.

It may return as institutional defensiveness. When a decision path is under-grounded, challenge threatens not only a single decision but the institution's language structure. The institution may therefore defend the term, category, or procedure more strongly than the evidence warrants. It may treat critique as misunderstanding, non-compliance, threat, or political attack. The debt becomes governance conflict.

It may return as machine amplification. If an under-grounded institutional category is encoded into a machine system, the debt scales. What was once a local ambiguity becomes a repeatable classification architecture. What was once discretionary judgement becomes an apparently neutral output. What was once a weak record becomes training data. What was once contested language becomes automated action.

It may return as failure to update. Because the unresolved condition was not recorded, the system does not know what would count as correction. It cannot state what evidence would force revision, what threshold should be changed, what category should be retired, what model should be retrained, what decision should be reviewed, or what residual has become decisive. The system becomes unable to learn because it never preserved the condition under which learning would be required.

Epistemic debt is therefore not only an error in belief. It is a structural liability inside decision systems.

This concept matters because many institutions and machine systems are rewarded for immediate closure. They must produce decisions, outputs, reports, classifications, targets, metrics, enforcement actions, compliance results, and performance indicators. The cost of slowing down to ground a term may be visible. The cost of failing to ground it may be deferred. This creates a bias toward under-formalisation.

A department may benefit from using a broad category because it permits action. A regulator may benefit from ambiguous risk language because it preserves discretion. A platform may benefit from opaque safety labels because they enable rapid moderation. A machine-learning system may benefit from proxy metrics because they make training possible. A legal actor may benefit from indeterminate language because it can be argued flexibly. A political actor may benefit from semantic drift because it permits persuasion without accountability.

Formalisation does not eliminate these pressures.

It makes their epistemic cost visible.

The first task is to record the borrowing. If a decision is made before closure, the system should state that it is acting under residual. If a category is provisional, it should remain marked as

provisional. If a proxy is used, it should be identified as proxy. If evidence is incomplete, the missing condition should be stated. If action is justified by threshold rather than closure, the threshold should be visible. If review is required, the update trigger should be preserved. A system that records the Soft Horizon it uses is more corrigible than one that hides it.

A system that records its epistemic debt is not weak. It is more corrigible than a system that conceals it.

The deeper problem is that unrecorded epistemic debt can be mistaken for institutional strength. A confident decision path appears decisive. A clean category appears efficient. A completed form appears responsible. A numerical output appears objective. A standardised process appears fair. A repeated label appears established. The lack of visible residual can look like certainty.

Formalisation reverses that appearance. It treats the absence of recoverable grounding not as strength but as a warning sign. A decision path that cannot show its relation, locus, abstraction, evidence path, residual, threshold, safeguard, and update may still be administratively complete. It may even have produced a good outcome. But it has not carried its authority in a recoverable form.

Epistemic debt is not paid by confidence. It is not paid by repetition. It is not paid by procedure. It is not paid by machine execution. It is not paid by institutional status. It is paid only by reconstructing, preserving, or revising the grounding conditions that allow the claim to carry the strength being claimed.

This is why formalisation is best understood as memory of borrowing.

It remembers where language has borrowed closure from a prior abstraction, a category, a record, a trust carrier, a threshold, a proxy, an evidential path, or an action demand. It remembers what residual remained when the borrowing occurred. It remembers what the borrowed closure permitted and what it did not permit. It remembers what would force update.

The aim is not to eliminate every debt before action.

The aim is to prevent debt from masquerading as authority.

---

5. Institutions as language-mediated decision systems

Epistemic debt becomes most consequential when it enters institutions.

An institution is not merely a group of people. It is an organised decision structure and a durable trust carrier. It preserves roles, rules, categories, records, thresholds, procedures, permissions, incentives, archives, sanctions, and authorities across time. It can remember beyond an individual. It can decide through distributed agency. It can act through offices, departments, committees, courts, regulators, agencies, boards, platforms, schools, hospitals, professions, and administrative systems. It reduces local recalculation by allowing later actors to rely on prior classification, recorded status, authorised role, procedural pathway, or institutional memory.

That efficiency is necessary. It becomes dangerous where the institution carries reliance forward after the relation, residual, threshold, safeguard, or update condition has disappeared from view.

Institutions must therefore be treated as observer-systems in the relevant sense. They do not possess the world as total state. They encounter reality through boundary, interface, scale,

abstraction, registration, retention, redeployment, comparator, and residual. They receive applications, complaints, forms, reports, measurements, allegations, classifications, risk indicators, legal submissions, inspection records, audit findings, expert testimony, statistical summaries, operational dashboards, and machine outputs. These are not the target whole. They are institutionally available relations.

Institutional availability is not institutional warrant.

A relation may enter the institution as complaint, record, application, allegation, model output, testimony, inspection finding, or dashboard signal without being closed for the category, threshold, or action that follows. The institution must therefore preserve the difference between what became available to it, what it registered, what it classified, what it retained, and what it was entitled to treat as grounded.

The institution then classes those relations. It decides what counts as evidence, what counts as admissible, what counts as relevant, what counts as risk, what counts as harm, what counts as compliance, what counts as failure, what counts as improvement, what counts as misconduct, what counts as eligibility, what counts as urgency, what counts as public interest, and what counts as sufficient cause for action.

The institution registers. It creates a file, record, case, status, score, classification, flag, approval, refusal, warning, sanction, referral, escalation, or closure note. That registration becomes institutional memory. It may later be redeployed in another decision, another department, another legal proceeding, another model, another audit, another funding decision, another risk assessment, or another automated system. A local classification can become a durable institutional fact.

This is why institutions are language-mediated decision systems.

They do not operate only by language, but language is one of their primary carriers of action. A term becomes a category. A category becomes a threshold. A threshold becomes a decision. A decision becomes a record. A record becomes authority for later decisions. The institution thereby turns semantic structure into material consequence and can make a provisional closure durable far beyond its original scope.

This should not be read as an argument against institutions. Institutions are necessary because finite individual observers cannot carry complex social functions alone. No individual can personally remember every case, weigh every claim, inspect every system, adjudicate every dispute, maintain every standard, enforce every rule, or coordinate every public function. Institutions extend memory, stabilise procedure, distribute authority, preserve records, and make collective action possible.

That distribution also creates Responsibility Warrants. An institution may assign answerability, duty, supervision, liability, stewardship, or decision ownership so that action remains governable across distributed processes. Such allocations may be necessary, but they do not establish the Responsibility Provenance of the consequence. Formalisation must therefore keep the institutional warrant recoverable as warrant and, where responsibility attribution is claimed or relied upon, preserve enough of the local responsibility-bearing path to show which transformations, non-applications, retained or maintained conditions, or materially governed field-conditioning relations are attributable to which loci, rather than allowing role, office, container, or authority to substitute for provenance.

The problem is not that institutions formalise.

The problem is that institutions often formalise incompletely.

They may stabilise a decision structure before the relation it acts upon has been sufficiently grounded. They may preserve records without preserving residuals. They may maintain categories after semantic drift has altered their function. They may treat procedural completion as if it were epistemic adequacy. They may allow authority to attach to language whose relation to the target has become unclear.

Institutions therefore have a double character. They can reduce epistemic uncertainty by preserving records, procedures, review routes, standards, comparators, and correction mechanisms. They can also amplify epistemic error by stabilising ungrounded language across many decisions.

A single human misclassification may harm one case. An institutional misclassification may become a category. Once it becomes a category, it can reproduce. It can be embedded in forms, training, policy, metrics, machine systems, funding rules, inspection criteria, eligibility frameworks, legal thresholds, and performance indicators. It can then appear natural because the institution repeatedly acts through it. Repetition becomes familiarity. Familiarity becomes authority. Authority becomes resistance to correction.[4]

This is why load-bearing institutional language requires audit.

An institution should not be able to defend a decision merely by saying that the category existed, the form was completed, the policy was followed, the risk was recorded, the procedure was satisfied, or the model produced the output. Each of those may matter. None is sufficient by itself. The deeper question is whether the decision path preserved the relation it claimed to govern.

A governance department may say that it acted in the public interest. The audit question is not whether the phrase public interest appears in the relevant policy. The question is how the public interest was grounded in that decision path. What relation was being protected? Which public, under what horizon, at what scale, against what risk, with what evidence, through what threshold, with what excluded alternatives, and with what residual?

A regulator may say that an organisation was non-compliant. The audit question is not whether a compliance category was available. The question is what relation the category preserves, what rule or standard was invoked, what evidence path supports the finding, what discretion was exercised, what residual remains, and what consequence the classification is being asked to carry.

A court may say that a standard was satisfied. The audit question is not whether the standard has a name. The question is how the standard connected relation, evidence, burden, legal category, affected locus, residual, and remedy in the case under review.

A platform may say that content is harmful. The audit question is not whether the platform has a harm policy. The question is what harm relation was claimed, to whom, over what horizon, at what scale, through what evidence, under what threshold, with what safeguard, and with what route for challenge or update.

Institutions become dangerous when their internal availability is mistaken for the target relation itself. A case may be available to the institution only through the form by which it was submitted. A harm may be available only as a recognised category. A risk may be available only through a metric. A person may be available only as a record. A system may be available only through a dashboard. The institution then acts on what its interface can receive while treating that interface as if it had captured the relevant relation.

Formalisation must therefore preserve the institutional access path. It must ask how relation entered, what the institution could not see, what abstraction made the relation classable, what

threshold turned classification into action, what residual remained, and what update condition would force review.

The institution's authority does not remove this burden. It increases it.

Where language becomes institutional consequence, recoverability becomes a condition of legitimate authority.

---

6. Institutional language capture

Institutional language capture occurs when a term, category, record, procedure, or threshold acquires institutional force while losing recoverable relation to the target it claims to govern.

This is not merely ambiguity. It is not simply that a word has several meanings. It is the point at which institutional language can classify, allocate, exclude, authorise, sanction, certify, automate, or justify without showing the relation it preserves. The institution continues to operate. Its forms are complete. Its records are orderly. Its categories are familiar. Its decisions can be cited. But the linguistic Soft Horizon has become reified inside the decision path.

Institutional capture has several common routes.

The first is **category as target**. A category is introduced to preserve a relation under finite conditions. It cuts the field so that a person, event, claim, risk, harm, or system can be handled. The category may be necessary. But over time the category can be mistaken for the relation itself. A person becomes the risk category. A legal status becomes the lived relation. A diagnostic label becomes the whole condition. A compliance class becomes safety. An administrative status becomes identity. A model class becomes the target.

The institution then asks whether the case fits the category rather than whether the category still preserves the relation. The category becomes easier to apply than to question. It gains procedural life. It is used in forms, records, reports, dashboards, and decisions. The abstraction disappears because the category has operational force.

The correction is abstraction audit. What is the category treating as same? What is it treating as different? What does it admit? What does it exclude? What scale does it privilege? What threshold does it activate? What residual does it hide? What affected locus can challenge the cut?

The second route is **record as warrant**. A record is a retained registration. It may be valuable. It may preserve evidence, procedure, observation, testimony, classification, decision, or action history. But a record is not warrant by default. It may be incomplete, misclassified, outdated, biased, mis-scaled, poorly sourced, or created under a limited closure demand. Institutional systems often treat records as if record-presence itself supplies authority.

A recorded allegation may travel as fact. A risk flag may persist after the condition has changed. A historical status may become current identity. A decision note may be read as proof of the relation it merely registered. A machine system may treat prior records as training truth. The record becomes stronger as it travels because later users see the retained form more easily than the residual under which it was created.

The correction is record-path discipline. The institution must distinguish what was observed, what was reported, what was inferred, what was decided, what was assumed, what was unavailable, and what residual remained. A record can carry authority only to the extent that

its provenance, scope, abstraction, and update condition support the claim now being made from it.

The third route is **procedure as legitimacy**. Institutions often rely on procedure because procedure constrains discretion, creates reviewability, and supports fairness. Procedure can be a real epistemic asset. But procedure can also become laundering if completion is treated as adequacy. A form may be completed while the relevant relation is excluded by the form. A consultation may occur without affected-locus relation being preserved. An appeal may exist but be practically inaccessible. A risk assessment may be documented while its threshold remains ungrounded. A model review may be recorded while deployment residual remains hidden.

The correction is to separate procedural completion from relation-preservation. The question is not only whether the procedure was followed. The question is whether the procedure carried the relation it was supposed to carry under the closure demand being made.

The fourth route is **policy phrase as threshold**. Institutional language often turns broad terms into action triggers. Public interest, safeguarding, security, resilience, misconduct, vulnerability, risk, proportionality, compliance, safety, and efficiency can all become thresholds. They may be legitimate. But if the phrase is not grounded, it can convert flexible language into authority while preserving institutional discretion.

A policy phrase may be invoked to justify action without stating what relation it closes. The institution says the action is necessary for safety, but safety for whom, against what, over what horizon, and at what residual? It says a person is vulnerable, but vulnerable to what relation, in what context, by what evidence, and with what consequence? It says intervention is proportionate, but proportionate to what harm, what alternatives, what uncertainty, and what affected-locus burden?

The correction is threshold discipline. The institution must state what the term activates, what evidence supports activation, what alternatives were considered, what residual remains, and what would force review.

The fifth route is **residual as invisible**. Institutional outputs often prefer clean closure. A decision is approved, refused, closed, escalated, sanctioned, certified, compliant, non-compliant, eligible, ineligible, high risk, low risk, harmful, safe, reasonable, or unreasonable. Such outputs may be operationally necessary. But they can hide the unresolved remainder that should travel with them.

Residual may be present in the file but absent from the output. It may be known by a practitioner but not retained by the record. It may be mentioned in a caveat but not allowed to constrain action. It may be acknowledged in review but not attached to future use. This is residual laundering: the uncertainty is formally present but operationally erased.

The correction is residual carriage. If evidence is missing, the missing condition must be stated. If classification is provisional, it must remain marked. If action is taken under threshold, threshold must not be narrated as closure. If a record is weak, downstream reliance must be limited. If residual blocks automation, the system must not automate merely because the category exists.

The sixth route is **authority as grounding**. Institutions may treat office, expertise, legal status, professional role, certification, procedure, or historical legitimacy as if it grounded the claim. Authority may matter. It may establish competence, jurisdiction, responsibility, or procedure. But authority does not make the target relation true. It does not remove the need for evidence path, abstraction, comparator, residual, threshold, safeguard, or update.

Authority becomes capture when challenge to the relation is answered by appeal to the institution's status. The institution says, in effect: this is grounded because we are authorised to say it. That is not formalisation. It is authority replacing audit.

Institutional capture is powerful because it can appear orderly. Captured language does not always look chaotic. It may look professional, consistent, evidence-led, compliant, legal, technical, or efficient. It may produce clean dashboards, completed forms, standardised outputs, and measurable performance. But order is not grounding. Consistency is not relation-preservation. Repeatability is not warrant if what is repeated is a misclosure.

The discipline of formalisation must therefore resist two temptations.

The first is anti-institutional cynicism. This treats all institutional language as power or performance. That is too crude. Institutions can preserve relations that individuals cannot. They can stabilise memory, make reasons reviewable, expose decisions to challenge, and maintain correction paths across time.

The second is institutional deference. This treats institutional form as if it already supplies grounding. That is too weak. Institutions are finite observer-systems. They see through their categories, forms, records, procedures, and models. Their access is bounded, their abstractions are selective, and their outputs carry residual.

Formalisation does not abolish institutional language. It makes institutional language answerable.

It asks the institution to show what relation its words carry, what closure demand they satisfy, what abstraction they impose, what evidence path supports them, what comparator they survived, what residual they retain, what threshold justifies action, what safeguard limits exposure, and what update condition keeps the claim corrigible.

Institutional language becomes legitimate not because it is official, but because its authority remains recoverable.

---

7. Machines do not repair under-formalisation

Machines do not repair epistemic under-specification.

They industrialise it.

This sentence should be understood precisely. A machine system may improve many things. It may apply a rule consistently. It may process more cases than a human office. It may detect patterns human observers miss. It may preserve logs. It may compare cases. It may identify anomalies. It may enforce update schedules. It may reduce some forms of arbitrary discretion. It may reveal inconsistencies in institutional practice. It may support audit. None of these functions, by itself, grounds the Soft Horizons the machine inherits.

A machine does not become epistemically adequate merely because it is technical.

If the categories it inherits are under-grounded, if the training labels encode institutional drift, if the objective function optimises a proxy, if the threshold is not tied to consequence, if the affected locus has no challenge route, if residual cannot be carried, then machine execution may make the system more powerful without making it more warranted.[5]

The machine problem begins upstream. A model is trained, configured, evaluated, deployed, or interpreted through prior human and institutional choices. What counts as the target? What counts as a positive example? What data are available? What labels are used? What proxy is selected? What errors are tolerable? What threshold converts score into action? What explanation is required? What records are retained? What appeal is possible? What update signal matters?

Each question can import under-formalised language.

A machine may classify content as harmful. But harmful relative to what relation? Harm to whom? Over what horizon? At what scale? Under what cultural, legal, developmental, political, or contextual assumptions? What is the difference between harm, offence, disagreement, risk, illegality, misinformation, satire, distress, and abuse? If these distinctions are under-grounded, the machine may produce outputs that look decisive while preserving institutional ambiguity.

A machine may classify a person as high risk. But risk of what? To whom? Based on which evidence path? Relative to which base rate? Under what time horizon? With what false-positive burden? With what historical bias in labels? What action follows? Is the score a trigger for review, restriction, denial, surveillance, intervention, or support? If the system cannot answer, risk becomes an executable ambiguity.

A machine may assess eligibility. But eligibility may depend on categories that were never designed for automated inference. It may require judgement about need, capacity, vulnerability, income, disability, residence, intention, family relation, educational status, or compliance. If those terms drift or depend on incomplete records, the machine may reproduce administrative opacity in a more rigid form.

A machine may support governance by detecting anomalies, flagging fraud, allocating inspection resources, or ranking cases. Such uses may be valuable. But if the system treats historical institutional decisions as ground truth, it may learn the institution’s prior misclosures. If the prior record buried residuals, the model cannot recover them by training on the record. If discretion was concealed, the model may encode it as pattern. If categories were captured, the model may learn capture as regularity.

This is the danger of treating data as if it were purified observation.

Data are records. Records have provenance. Provenance has abstraction. Abstraction has residual. Residual may have been hidden by the system that generated the data. A dataset is therefore not a transparent slice of reality. It is a retained field of prior selections, omissions, categories, measures, thresholds, instruments, incentives, errors, and exclusions. A machine that learns from it inherits that field.

Technical execution can then create technical closure. A score, label, rank, prediction, recommendation, or generated statement appears complete because it is produced by a system. It may be precise. It may be repeatable. It may be fast. It may be statistically framed. It may be displayed through a polished interface. It may be endorsed by a vendor, regulator, procurement process, or institutional policy. These features can reduce audit. They do not by themselves preserve the target relation.

A machine output can become a trust carrier with unusual force. It may appear detached from human judgement while carrying human and institutional abstractions. It may appear neutral while inheriting contested categories. It may appear objective while optimising proxies. It may appear current while relying on historical records. It may appear explanatory while producing only correlation. It may appear reliable at benchmark scale while failing in deployment.

The failure is not that machines are uniquely untrustworthy. Human judgement can also be biased, opaque, inconsistent, status-driven, under-grounded, and resistant to correction. The point is not to subordinate machines to human intuition. It is to subordinate both human and machine systems to auditable relation.

The question is not whether a human or machine produced the output.

The question is whether the claim path preserved the relation it claimed to carry.

Machine systems may support formalisation where the domain has already been made recoverable enough for them to inherit through an auditable path. They may maintain audit logs, detect inconsistent category use, flag missing fields, compare current cases with prior cases, surface residuals, enforce review dates, identify drift, monitor thresholds, support retrospective analysis, and expose patterns that would otherwise remain hidden. They may help institutions remember their uncertainty rather than bury it.[6]

But machine support should come after claim-path audit, not before it. The institution must first show its categories, thresholds, evidence paths, residuals, safeguards, and update conditions. Only then can a machine system assist in applying, monitoring, comparing, or maintaining the decision path.

Where that condition is met, machines may extend discipline.

Where it is not met, machines scale opacity.

They may convert under-grounded categories into scalable outputs. They may make discretionary language appear neutral. They may treat historical misclosures as training signal. They may generate scores that institutions treat as warrant. They may increase the distance between affected loci and decision authority. They may make a weak Soft Horizon more portable, more durable, more defensible, and harder to challenge.

The rule should be strict.

Machine management may inherit audited decision domains. It must not inherit opaque institutional language, unaudited carrier trust, or historical records as if encoding were grounding.

This rule does not reject machine assistance. It makes machine assistance conditional on formalisation. A machine can help administer, monitor, compare, flag, and remember. It must not be asked to purify an under-grounded category by executing it.

Encoding is not grounding.

Automation is not audit.

Execution is not warrant.

---

8. Formalisation as preservation, not reduction

Formalisation is often feared because it appears to reduce.

It appears to reduce language to rigid definition, judgement to rule, relation to field, ambiguity to code, uncertainty to number, and living context to form. That fear is not groundless. Institutions have often used forms to control what can be seen. Technical systems have often required the world to fit their schema. Bureaucracies have often treated what cannot be

recorded as if it were irrelevant. Machine systems have often converted complex relations into labels, proxies, scores, and outputs.

But that is not the formalisation defended here.

Formalisation is preservation when it keeps the decision path answerable to the relation.

It becomes reduction when the path replaces the relation.

The distinction is decisive. A finite observer-system cannot possess the target whole. It cannot eliminate all residual. It cannot make every decision under final closure. It cannot stop language from evolving. It cannot remove judgement from institutions. It cannot guarantee that machines will preserve every relation they process. But it can preserve enough of the claim path that a load-bearing claim remains auditable.

Formalisation should preserve relation. A claim must say what it purports to carry. Risk, harm, safety, fairness, evidence, compliance, intelligence, observation, reasonableness, public interest, and machine-validated status are not self-grounding terms. Each must be tied to the relation being asserted.

Formalisation should preserve locus. A claim must say for whom or for what system the relation is being evaluated. A risk to an institution is not automatically a risk to an affected person. A benefit to a population is not automatically a benefit to each member. A distinction available to an analyst is not automatically available to the system under review. Without locus discipline, institutions define reality only from their own vantage.

Formalisation should preserve boundary and interface. A claim must show how the relevant relation entered the decision path. Was it reported, measured, observed, modelled, inferred, complained of, inspected, recorded, classified, or predicted? What was excluded by the pathway? What could not enter? Without boundary and interface discipline, the institution mistakes its access route for the target whole.

Formalisation should preserve scale and horizon. A relation may be real at one scale and misapplied at another. A short-term risk may not justify long-term classification. A local measurement may not support general policy. A statistical pattern may not settle an individual case. A historical record may not represent current condition. Without scale and horizon discipline, precision becomes overreach.

Formalisation should preserve abstraction. Every decision path cuts the world. It treats some differences as relevant and others as background. It creates classes, equivalences, thresholds, and tolerances. Formalisation should expose this cut. It should ask what the category admits, what it excludes, what it compresses, and what it cannot see. Without abstraction discipline, categories appear natural rather than constructed for a purpose under constraint.

Formalisation should preserve evidence path. A decision should distinguish what was directly registered, what was inferred, what was reported, what was modelled, what was assumed, what was unavailable, and what was excluded. Without evidence-path discipline, weak records can become strong claims.

Where the claim concerns responsibility, accountability, liability, duty, authorship, decision ownership, or causal attribution, formalisation should also preserve the responsibility path. It should distinguish Responsibility Provenance from Responsibility Warrant: which consequential branch relations are actually attributable to which loci through distinction-conditioned organisation, and what answerability or duty is instead allocated by a social, moral, legal, organisational, procedural, or institutional regime. Where application is claimed, the path should preserve enough relation to recover the relevant Counterfactual Exclusion. Where

responsibility is claimed through non-application, maintenance, retention, or cross-locus field-conditioning without application, it should preserve enough relation to recover the comparator-stable difference made to the relevant Counterfactual Inclusion Field. A warrant may determine who must act without proving who caused or conditioned the relevant branch relation; provenance may be distributed without removing the practical need for a warrant. Neither should travel as the other.

Formalisation should preserve comparator. A claim gains strength by surviving relevant alternatives, not by being the first available formulation. Formalisation should ask what weaker neighbouring claims were excluded, what live alternatives remained, and what would distinguish them. Without comparator discipline, familiar language becomes verdict.

Formalisation should preserve residual. This is the most important point. A good formalisation does not erase unresolved conditions. It carries them. If evidence is missing, the missing condition is stated. If the category is provisional, it remains marked. If action is taken before closure, the action is identified as threshold action. If uncertainty remains, it is not laundered into the output. Without residual discipline, formalisation becomes false closure.

Formalisation should preserve update. A decision path should state what would force review, reversal, narrowing, escalation, retirement, retraining, or further inquiry. Without update discipline, institutional and machine systems become locked around prior language.

These preserved conditions are not decorative. They are the difference between formalisation as domination and formalisation as accountability.

The fear that formalisation will flatten life often arises because institutions have used forms to control what can be seen. A form can become the institution's reified Soft Horizon: the boundary of what can be seen, recorded, challenged, or repaired. If the relevant relation does not fit the form, the relation disappears. This is not a failure of formalisation as such. It is a failure of formalisation without residual, affected-locus challenge, abstraction audit, and update.

A better formalisation does not say: only what fits the form exists.

It says: this is what the form could carry; this is what it could not carry; this is what remains residual; this is who is affected by the cut; this is how the cut may be challenged; this is what would force revision.

That is not reduction.

It is disciplined humility.

Formalisation also need not require numerical precision. Some relations can be measured quantitatively. Others require qualitative classification, narrative evidence, professional judgement, legal reasoning, contextual interpretation, historical reconstruction, or moral evaluation. The method should not convert all uncertainty into numbers merely to appear rigorous. A Soft Horizon is not made more disciplined merely by being quantified. False precision is another form of misclosure.

A residual may be qualitative. A weight may be comparative rather than numerical. A threshold may be legal, ethical, pragmatic, precautionary, evidential, or operational. An update trigger may be a new type of case, a contradiction, a failed prediction, an affected-locus challenge, a model drift signal, a court ruling, an audit finding, or a previously unavailable record. Formalisation requires that these conditions be explicit. It does not require that all of them be mathematical.

This point is central to law and governance. Legal systems often reason through standards rather than rigid rules. Governance decisions often require judgement under uncertainty. Public administration often operates in contexts where evidence is incomplete and action cannot wait. Formalisation should not pretend that these domains can become mechanical. It should make their judgement paths auditable.

To formalise a judgement is not to remove judgement. It is to state the relation judged, the authority invoked, the evidence used, the category applied, the discretion exercised, the residual retained, the threshold used, the safeguard imposed, and the update route preserved.

This is why formalisation must begin with the claim path rather than with a total model of the domain. It should not attempt to formalise reality as a whole, freeze language, abolish judgement, or convert uncertainty into numbers merely to appear rigorous. It should begin by preserving what a load-bearing claim must carry if its authority is to remain auditable.

A system that can state its relation, locus, abstraction, evidence path, comparator status, residual, threshold, safeguard, and update condition is not guaranteed to be correct. It is, however, more open to correction than a system that hides behind words such as objective, reasonable, compliant, safe, harmful, high risk, evidence-led, or machine-validated.

Formalisation is not the defeat of language.

It is the preservation of language where language becomes authority.

---

9. The auditable claim packet

Formalisation should not attempt to formalise reality as a whole.

It should formalise the conditions under which a load-bearing claim is allowed to carry strength.

This distinction is decisive. A finite observer-system cannot possess the target whole. It cannot eliminate all residual. It cannot make every decision under final closure. It cannot stop language from evolving. It cannot remove judgement from institutions. It cannot guarantee that machines will preserve every relation they process. It can, however, make the load-bearing claim path recoverable.

The first formal unit is therefore the auditable claim packet.

An auditable claim packet is the minimum formal container by which a consequential claim carries its relation, locus, availability condition, closure demand, reference conditions, abstraction, evidence or record path, carrier path where reliance is inherited, comparator or discriminator support, residual, scope, action threshold, safeguard, and update condition across observers, institutions, or machines.[7]

Responsibility Provenance is not an additional universal packet element. Where the claim itself concerns responsibility, accountability, liability, duty, causal attribution, or decision ownership — or where such an attribution is used to justify consequential action — the packet must recover enough of its existing locus, evidence or record, carrier, comparator, residual, scope, threshold, safeguard, update, and relevant field conditions to distinguish Responsibility Provenance from Responsibility Warrant. Where application is material to the attribution, that recovery must be sufficient to identify the relevant Counterfactual Exclusion. Where non-application, maintenance, retention, or cross-locus field-conditioning is material, it must

instead be sufficient to identify the comparator-stable difference made to the relevant Counterfactual Inclusion Field. Where no responsibility claim is being made, no separate responsibility-provenance field is required.

It is not a legal test, scientific protocol, institutional standard, machine-governance framework, or decision machine by itself. It is the disciplined answer to a practical problem: when a claim travels through a decision path, what must remain attached to it so that its authority can be audited? Where the claim operates as a Soft Horizon, the packet carries the residual, scope, threshold, safeguard, and update condition that prevent provisional closure from becoming reified authority.

A claim without such attachment can move too easily. It can leave the conversation in which it was formed. It can enter a record. It can be copied into a policy. It can become a case category. It can be represented as a score. It can be encoded into a model. It can be inherited by another department. It can be cited by a court. It can become training data. It can justify action long after the conditions of its first use have been forgotten. A Soft Horizon can travel while losing the marks that kept it soft.

The auditable claim packet slows this drift. It does not prevent a claim from travelling. It makes the travelling claim carry its conditions, so provisional closure does not become ungrounded authority merely by moving through a system.

Carrying those conditions is not yet the same as making them operative at the receiving locus. Part 2a, Paper 1c distinguishes preservation in a carrier from the receiving organisation's capacity to stabilise the relevant relation. A packet can arrive intact while its scope, residual, or threshold distinctions remain unsupported by the recipient's classing, translation, or procedural regime. A process might retain the words "provisional screening result" yet act on the result as a confirmed finding. The receiving path must therefore preserve the distinctions needed for the use being claimed, through its own organisation or available scaffolding. This does not require repeating all upstream closure-work or adding another universal packet field. It makes the existing locus, availability, abstraction, carrier, and residual conditions answerable at the point of use.

The first element is the **claim itself**. The system must state what is being asserted. This appears simple, but many institutional failures begin because the actual claim is obscured. A record may say risk, but the operative claim may be risk of physical harm, reputational harm, non-compliance, financial loss, political embarrassment, operational disruption, or future liability. A policy may say safety, but the operative claim may concern danger to a person, danger to an institution, danger to a public, or danger to the stability of a system. A machine may say likely, but the operative claim may concern probability, similarity, correlation, classification confidence, or policy threshold. The claim must be stated strongly enough to be tested.

The second element is the **relation under evaluation**. A claim is not disciplined merely because it is grammatically clear. It must name the relation it purports to preserve. Is the claim about evidence, causation, risk, identity, classification, attribution, eligibility, harm, responsibility, observation, model adequacy, explanation, procedural fairness, proportionality, machine performance, institutional competence, or action necessity? The relation determines the burden. A record that supports detection may not support explanation. A score that supports triage may not support sanction. A category that supports administrative routing may not support moral judgement.

The third element is the **locus**. The claim must state for whom, or for what system, the relation is being evaluated. A risk to an institution is not automatically a risk to an affected person. A benefit to a population is not automatically a benefit to each individual. A distinction available

to an analyst is not automatically available to the system under review. A classification useful to a regulator is not automatically adequate for the person classified. Without locus discipline, claim-strength can drift between observer, target, institution, record, model, affected subject, environment, and analyst.

The affected locus should also be visible where consequence falls on someone other than the relying locus. A department may rely on a category; a citizen may bear the error. A platform may rely on a classifier; a user may bear the restriction. A hospital may rely on a triage score; a patient may bear the risk. Formalisation must therefore ask not only who uses the claim, but who is exposed by it.

The fourth element is the **availability condition**. The system must state how the relation entered the decision path. Was it observed, measured, reported, inferred, modelled, complained of, inspected, classified, predicted, recorded, or generated? What interface made it available? What could not enter through that interface? Availability is not warrant. A complaint is not proof. A record is not truth. A model output is not target possession. A dashboard signal is not the full system. The packet must preserve how relation became available so that later use does not mistake access for closure.

The fifth element is the **closure demand**. The system must state what kind of closure the language is being asked to carry. Is the claim being used for recognition, measurement, explanation, classification, legal judgement, administrative action, scientific inference, machine output, precaution, enforcement, certification, or public justification? A relation may be closed enough for one demand and residual for another. A model may be adequate for ranking but not explanation. A legal standard may close procedure but not moral truth. A medical finding may support intervention but not final diagnosis. Without closure-demand discipline, local adequacy becomes general warrant.

The sixth element is the reference condition. The claim must state the boundary, interface, scale, horizon, discrimination regime, comparator assumptions, and, where materially relevant, the field proper to the relation being evaluated: the live formulation field where epistemic alternatives are under review, or the Counterfactual Inclusion Field where physically supported power-bearing alternatives of an IPL are being claimed. A claim may be valid at one scale and misleading at another. A risk may be short-term but not long-term. A measurement may hold within a tolerance but not outside it. A classifier may perform in one deployment context and fail in another. A category may work for aggregate monitoring and fail for individual action. Reference conditions prevent the claim from travelling beyond the conditions that made it meaningful.

The seventh element is the **abstraction or category cut**. Every load-bearing term cuts relation. It treats some differences as relevant and others as background. It creates sameness, difference, inclusion, exclusion, threshold, and tolerance. Formalisation must expose that cut. What is being treated as the same? What is being treated as different? What is inside the category? What is outside? What edge cases reveal the cut? What affected locus may be lost by the abstraction? A category that hides its cut can become target-replacement. The system stops asking whether the category preserves relation and begins acting as if category membership were enough.

The eighth element is the **evidence or record path**. The packet must distinguish direct observation, testimony, instrument output, model inference, professional judgement, historical record, proxy measure, statistical correlation, institutional file, and machine-generated classification. These are different paths with different residuals. A retained record may establish that something was registered. It does not automatically establish that the registered

relation was true, current, complete, or adequate for the present decision. The evidence path must remain visible so that weak records do not become strong claims by repetition.

The ninth element is the **carrier path** where reliance is inherited. Many consequential claims borrow closure-work from persons, roles, credentials, institutions, procedures, standards, certificates, archives, instruments, models, or machine outputs. These carriers may be legitimate. They reduce local recalculation and make complex decision systems possible. But the carrier is not the ground. A credential is not competence by itself. A certificate is not substantive adequacy by itself. A model output is not target preservation by itself. A compliance process is not warrant by itself. The packet must show what closure-work is being borrowed, from which carrier, under what limits, and with what residual.

The tenth element is **comparator or discriminator support**. A claim gains strength not merely by being asserted, recorded, repeated, or institutionalised, but by surviving relevant alternatives. What weaker neighbouring claims were excluded? What live alternatives remain? What comparator would distinguish them? What discriminator condition has been met? What candidate formulation failed? Without comparator discipline, the first available formulation becomes the verdict.

The eleventh element is **residual**. The packet must state what remains unresolved. This is not a decorative caveat. It is the condition that prevents overclaim. What evidence is missing? What abstraction may fail? What scale transition is uncertain? What affected-locus relation remains unpreserved? What comparator has not been run? What machine deployment residual remains? What cannot yet be inferred? Residual must state what it blocks. It may block final judgement but not investigation. It may block automation but not human review. It may block coercion but not support. It may block explanation but not local prediction.

The twelfth element is **scope**. The claim must state what use it permits and what use it does not permit. A risk score may support triage but not denial. A medical result may support precaution but not final attribution. A record may support inquiry but not sanction. A category may support routing but not identity. Scope prevents local closure from migrating into stronger demands.

The thirteenth element is **action threshold**. Where the claim justifies action before full closure, the packet must state why action is permitted or required. What consequence, urgency, reversibility, duty, precaution, or proportionality sets the threshold? Action threshold is not truth threshold. The fact that action is necessary does not mean the relation has been closed. The packet must keep this distinction visible.

The fourteenth element is **safeguard**. Where action proceeds under residual, safeguard constrains exposure. It may include review, appeal, human oversight, domain restriction, limited duration, independent check, notice, consent, monitoring, staged deployment, reversal condition, or withdrawal path. Safeguard is not reassurance. It is part of the warrant for acting under incomplete closure.

The fifteenth element is **update condition**. A claim that cannot be revised is not disciplined. The packet must state what would force review, narrowing, reversal, escalation, retirement, retraining, or further inquiry. The update condition may be new evidence, failed prediction, affected-locus challenge, model drift, comparator failure, audit finding, record correction, changed scale, legal change, or previously unavailable evidence. Without update, residual becomes inert and the claim hardens.

The packet is therefore not bureaucratic ornament. It is the carriage structure by which a load-bearing claim remains answerable to relation after it leaves the local setting in which it was made.[8]

This does not mean every claim requires an elaborate document. In ordinary low-stakes settings, much of the packet may remain tacit. The burden rises where consequence rises, where the claim travels, where affected loci are exposed, where automation is used, where institutional authority is invoked, where records persist, where residual is significant, where the decision is hard to reverse, or where the claim becomes a trust carrier for later systems.

The auditable claim packet is therefore the minimum discipline of formalisation.

It keeps language from travelling alone.

---

10. Proportional formalisation

Formalisation must be proportional.

A demand to formalise every use of language at full resolution would defeat the function of language. It would replace communication with procedural burden. It would mistake ordinary compression for governance failure. It would make conversation brittle, judgement timid, institutions slow, and interpretation coercive. It would also create a false sense of discipline, because a system overwhelmed by formal requirements often learns to complete forms rather than preserve relation.

The burden of formalisation should rise with consequence, mediation, opacity, contestation, reversibility, authority, and residual.

A low-consequence, local, reversible, and interpersonal use of language may require little formalisation. Ordinary trust, conversation, exploration, metaphor, internal reasoning, preliminary hypothesis, or low-stakes categorisation can often operate through tacit Soft Horizons. The correction path remains close. If misunderstanding occurs, participants can repair. If a category fails, it may be revised quickly. If an action is mistaken, the harm may be limited.

A medium-consequence use requires more visible grounding. Where a term enters a professional judgement, administrative file, scientific report, policy discussion, internal risk process, or machine-assisted recommendation, the relation, abstraction, evidence path, residual, and update condition should become more explicit. The claim may not require full external audit, but it must be recoverable enough for responsible review.

A high-consequence use requires strong formalisation. Where language classifies a person, denies a benefit, imposes sanction, triggers surveillance, guides medical intervention, supports legal judgement, authorises coercion, allocates major resources, automates action, certifies safety, or produces durable records, the claim path must carry its conditions clearly. The affected locus must be able to see, challenge, or benefit from review of the relation being claimed.

Irreversibility increases the burden. A provisional action that can be reversed may proceed under heavier residual than an action that permanently alters status, record, health, liberty, reputation, access, or opportunity. But reversibility must be assessed practically, not only formally. A record marked provisional may become durable in later systems. A temporary classification may shape future treatment. A machine-generated flag may persist in archives. A reversible-looking action may create path dependence for the affected locus.

Automation increases the burden. A human decision may be opaque, but it is often locally contestable. A machine-mediated decision may scale the same abstraction across many cases,

reproduce hidden residual, and separate affected loci from decision authority. If a claim is to be automated, the abstraction, evidence path, threshold, residual, safeguard, and update condition must be formalised before automation inherits them.

Mediation increases the burden. Where the relying locus, affected locus, decision-maker, record-maker, model-builder, and authority-holder are separated, tacit trust is insufficient. A term may travel through departments, vendors, courts, regulators, platforms, and machine systems. The more a claim travels, the more its conditions must travel with it.

Contestation increases the burden. A term that functions without dispute in ordinary practice may become load-bearing when challenged. A person, community, expert, court, affected locus, or comparator case may expose that the term does not preserve the relation it claims to carry. Formalisation should respond not by suppressing challenge, but by making the claim-path visible enough to assess whether the challenge bears on relation, abstraction, residual, threshold, or update.

Residual increases the burden. A claim with little relevant residual may travel more lightly. A claim with significant unresolved conditions must carry them. Heavy residual does not always block action, but it requires stronger marking, stronger safeguard, and clearer update. The greater the residual, the more dangerous clean closure becomes.

Authority increases the burden. The same word carries different burden when spoken in conversation, written in policy, entered in a record, invoked by a court, encoded in a model, or used to justify intervention. Authority turns language into consequence. Formalisation scales with that consequence.

Proportional formalisation also prevents two opposite failures.

The first is under-formalisation. This occurs when a system allows load-bearing language to carry authority without recoverable grounding. It produces epistemic debt, residual laundering, machine amplification, institutional defensiveness, and failure to update.

The second is over-formalisation. This occurs when a system imposes formal burden where tacit judgement, ordinary language, provisional exploration, or local repair would be more appropriate. Over-formalisation can create procedural theatre. It can encourage box-ticking. It can privilege what is easy to record over what matters. It can make the form replace the relation.[9]

The proportionality rule is therefore not a concession to convenience. It is itself part of the discipline.

Formalise enough to preserve the claim path under the strength being claimed.

Do not formalise so much that the preservation structure becomes the new misclosure.

---

11. Auditable decision-making and retrospective review

Formalisation becomes practical when it alters decision-making.

An auditable decision is not necessarily a correct decision. It is a decision whose claim path, reliance path, residual, threshold, safeguard, and update condition can be reconstructed. It allows a later observer, reviewer, court, affected locus, regulator, scientist, auditor, or machine-monitoring system to ask how the decision carried the strength it claimed at the time it was made.

A decision path should therefore preserve the movement from relation to action.

First, what relation was under review? Second, how did that relation become available? Third, what claim or category was formed? Fourth, what closure demand was being made? Fifth, what abstraction cut made the relation usable? Sixth, what evidence or record path supported the claim? Seventh, what carrier path was trusted? Eighth, what alternatives or comparators were considered? Ninth, what residual remained? Tenth, what threshold justified action? Eleventh, what safeguard constrained exposure? Twelfth, what update condition would force review?

These questions are not a mechanical checklist. They are the grammar of decision audit. A decision may answer them in different forms depending on domain. A court may answer through reasons. A hospital may answer through clinical record, differential diagnosis, consent, and follow-up. A regulator may answer through findings, standards, evidence, and enforcement threshold. A machine-mediated system may answer through data provenance, label definition, model evaluation, threshold setting, human review, monitoring, and appeal path. A scientific practice may answer through methods, controls, comparators, error estimates, and revision conditions.

The point is not sameness of format. The point is recoverability of relation.

Auditable decision-making also protects judgement. It does not replace judgement with rule-following. It makes judgement answerable. The decision-maker may still interpret, weigh, infer, prioritise, and act under uncertainty. Formalisation asks the decision-maker to preserve the structure of that judgement so it can be assessed without relying on status, confidence, or institutional authority alone.

This is especially important where discretion is necessary. Discretion does not disappear because a system formalises its claim path. In many domains, discretion is unavoidable because the relation is complex, evidence incomplete, values contested, and consequences unequal. But discretion without recoverability becomes unreviewable power. Formalised discretion remains judgement, but judgement with visible relation, category, residual, threshold, and update.

Retrospective audit is the backward-facing counterpart of auditable decision-making.

It asks how a decision was grounded at the time it was made, not merely whether the outcome later appeared good or bad. It reconstructs the relation under review, the availability condition, the evidence path, the abstraction, the live alternatives, the residual, the action threshold, the safeguard, and the update condition. It asks what the decision-maker could reasonably carry, what they failed to preserve, and what later information changed.

Retrospective audit is not blame by default. It is reconstruction. It may reveal negligence, manipulation, laundering, overclosure, under-formalisation, or bad faith. But it may also reveal a decision that was properly made under residual, even though the outcome later failed. A finite system can act responsibly and still encounter an adverse branch. Conversely, a system can produce a good outcome through an under-grounded decision path. Outcome alone does not settle decision quality.

Investigatory grounding depends on this distinction. After failure, institutions often search for the immediate cause. But a formalisation audit asks a deeper question: where did the claim path lose recoverability? Was the relation misidentified? Was the locus wrong? Was the availability condition too weak? Was the category cut hiding the affected relation? Was a record treated as warrant? Was a machine output treated as closure? Was residual known but laundered? Was action threshold confused with truth? Was safeguard symbolic? Was update impossible?

Where retrospective review asks who or what was responsible, claim-path recovery must not collapse into outcome attribution. The later result does not by itself establish earlier

Responsibility Provenance. The review should recover which distinctions, residuals, warnings, correction paths, responsibility-relevant capacities, standing conditions, and physically available alternatives were available within each locus's operative frame at the relevant time, and separately identify any Responsibility Warrant that allocated duty or answerability. Where application is claimed, the review should recover the relevant Counterfactual Exclusion; where responsibility is claimed through non-application or standing field-conditioning, it should recover the comparator-stable difference made to the relevant Counterfactual Inclusion Field. This prevents hindsight, proximity, role, container, authority, and later-discovered possibility from being mistaken for causal provenance.

Retrospective review should also preserve the live formulation field and its contraction history. A decision may appear inevitable after it has been taken because the selected formulation or action has generated records, justifications, and downstream commitments while alternatives have disappeared from view. The review must reconstruct what was live at the time and how Live Field Contraction occurred. Which formulations were available? Which were positively excluded? Which were weak but live? Which disappeared historically rather than through earned exclusion? Which comparator was missing? Which residual was known? Which affected-locus challenge was ignored? Branch-memory should preserve the material history of that contraction without treating an epistemically live formulation as proof that the corresponding causal branch was physically available to the locus. Without such branch-memory, retrospective audit becomes hindsight closure.

Affected-locus challenge is one of the most important retrospective mechanisms. Institutions often see only through their own categories. The affected locus may expose that the category did not preserve the relation it claimed to govern. A person misclassified by a record, a community harmed by a policy, a patient affected by a protocol, a user affected by a model, or a defendant affected by an evidential standard may reveal residual that the institution did not carry. Formalisation should not treat affected-locus challenge as disruption by default. It is often a route by which hidden residual becomes available.

Retrospective audit also supports future design. It turns failure into update. If a term drifted, the operative use can be narrowed. If a category collapsed, the abstraction can be revised. If a record travelled too strongly, downstream reliance can be limited. If a machine output became authority, human review and deployment limits can be changed. If a safeguard failed, its trigger can be made operational. If residual was hidden, future packets can carry it explicitly.[10]

A decision system that cannot be retrospectively audited cannot learn in a disciplined way. It may react, apologise, defend, compensate, punish, or rebrand. But unless it can reconstruct the claim path, it cannot know what must change.

Formalisation therefore serves both prospective and retrospective functions.

Prospectively, it slows overclosure before action.

Retrospectively, it makes correction possible after action.

---

12. Decision quality, outcome success, and the handoff to governance

Decision quality and outcome success must be separated.

A good outcome does not prove that the decision path was grounded. A bad outcome does not prove that the decision path was undisciplined. This distinction is essential for finite systems acting under residual.[11]

The same separation governs responsibility attribution: outcome success or failure may trigger or inform review, but it does not by itself establish which earlier transformation, non-application, retained or maintained condition, or field-conditioning relation was responsibility-bearing.

A decision may be high quality and fail. The relation may have been properly specified. The evidence path may have been the best available. The live alternatives may have been considered. The residual may have been carried. The threshold may have justified action. The safeguard may have been proportionate. The update condition may have been stated. Yet the branch selected may still produce harm, error, or failure because finite action does not control total state.

A decision may be low quality and succeed. A vague category may happen to fit. A weak record may happen to be accurate. A machine output may happen to predict the case. A policy phrase may happen to support the right action. A procedure may produce a good result despite hiding residual. Success then launders the path. The system says the outcome proves the decision was sound. But the relation was not preserved in a recoverable way.

Outcome success is evidence, but it is not closure by itself.

Outcome failure is pressure, but it is not automatic disproof.

Formalisation protects this distinction by shifting evaluation from outcome alone to claim-path quality. The question is whether the decision carried the strength it claimed under the conditions available at the time. Did it preserve relation? Did it identify locus? Did it state closure demand? Did it expose abstraction? Did it preserve evidence path? Did it retain residual? Did it act by threshold rather than false certainty? Did it impose safeguard? Did it remain updateable?

This distinction matters for law, governance, and machine-mediated decision paths because these are domains where outcome, legitimacy, authority, and public trust often become entangled.

A court may reach a result that appears acceptable while relying on language whose abstraction remains hidden. A regulator may avoid harm while using a category that cannot be defended under challenge. A public authority may produce efficiency while burying affected-locus residual. A platform may moderate harmful content while using labels that cannot be audited. A machine system may outperform a baseline while still treating historical misclosure as ground truth. The outcome may be useful. The decision path may still be under-formalised.

The reverse also matters. A court, regulator, doctor, institution, or machine-assisted process may preserve the claim path responsibly and still produce an adverse outcome. If every failure is treated as proof of bad decision quality, institutions will hide residual, overclaim certainty, and protect themselves through defensive closure. Formalisation should instead make it possible to distinguish justified action under residual from negligent overclosure.

This distinction leads directly to governance.

Once formalisation is adopted into law, governance, institutional process, certification, machine-readable schema, procurement, audit, or compliance, formalisation itself becomes load-bearing. A completed claim packet may become a defence. A formalisation template may become a bureaucratic artefact. A certificate may become a trust carrier. A method release may

become authority. A machine-readable schema may become the practical definition of the discipline. A regulator may treat compliance with formalisation as warrant. A court may treat method-use as evidence of reasonableness. A vendor may market formalisation as assurance.

The same risk returns at the level of the method.

Formalisation is introduced to prevent language from carrying authority beyond its grounding. But the word formalised can itself begin to carry authority beyond its grounding. An institution may say the decision was formalised, while residual was laundered. A vendor may say a system is audit-ready, while the target relation remains under-specified. A regulator may say the required packet was completed, while the abstraction was never challenged. A machine-readable layer may say fields were populated, while relation-preservation was absent.

The method can become a Soft Horizon.

This is not a reason to reject formalisation. It is the reason formalisation must be governed. Once the discipline becomes usable, it requires its own correction path. Releases, templates, certificates, schemas, and compliance processes may be necessary operational carriers. They reduce local recalculation and support adoption. But they must not become epistemic grounds. A packet completed without relation-preservation is not adequate. A certificate without substantive audit is not warrant. A schema that cannot carry residual should not define the method. A release should not become doctrine merely because it is official.

The handoff to the next paper is therefore precise.

Paper 6 has argued that load-bearing language requires recoverable grounding. It has shown why institutions and machines intensify the danger of under-formalisation. It has defined the auditable claim packet as the first formal unit for carrying claim-strength under residual. It has argued that the burden must be proportional and that decision quality must be separated from outcome success.

Paper 7 must now ask how formalisation itself remains corrigible once it becomes institutionalised.

The question is no longer only how claims are formalised.

The question is how the discipline of formalisation is governed without becoming another ungrounded authority.

---

13. Conclusion: formalisation under load

Formalisation is not the defeat of language.

It is the discipline required when language becomes authority.

Natural language may remain open, adaptive, contested, metaphorical, cultural, rhetorical, and alive. Language is one of the necessary Soft Horizons through which finite observer-systems communicate, classify, remember, coordinate, and act. It compresses relation into usable form. It allows finite systems to proceed without reconstructing every dependency at every moment.

The danger is not compression alone. The danger is ungrounded authority. A term may compress a prior relation and then drift into a new decision path. A category may appear stable while its grounding has disappeared. A record may be treated as warrant. A procedure may become legitimacy. A machine output may scale a hidden abstraction. An institution may

convert under-grounded language into consequence. In each case, a useful Soft Horizon becomes dangerous when it carries authority without residual, scope, comparator, safeguard, or update.

Formalisation interrupts that movement by making the claim path recoverable. It preserves relation, locus, availability condition, closure demand, boundary, interface, scale, horizon, abstraction, evidence path, carrier path, comparator, residual, threshold, safeguard, and update. It does not guarantee truth. It prevents language from carrying more authority than its grounding can support. It prevents reliance on a term, record, credential, institution, procedure, model, or machine output from carrying more strength than the trusted carrier has earned. It prevents availability, registration, procedure, machine execution, or action from being mistaken for closure.

The governing rule remains:

Load-bearing language must not carry authority beyond the conditions that ground its use.

Those conditions begin with relation, locus, availability condition, and closure demand. They remain auditable only where abstraction, evidence path, carrier path, comparator, residual, threshold, safeguard, and update continue to travel with the claim.

This does not abolish uncertainty. It makes uncertainty harder to hide.

It does not abolish judgement. It makes judgement answerable.

It does not abolish institutions. It makes institutional authority recoverable.

It does not reject machines. It prevents machine execution from being mistaken for grounding.

It does not freeze language. It makes recoverable the claim and action conditions from which later responsibility attribution can be reconstructed without inferring responsibility from outcome, role, or authority alone.

The case for formalisation is therefore narrow and strong. Ordinary language may live. Load-bearing language must answer. A word may move, change, persuade, provoke, invent, and explore. But where it classifies, governs, allocates, excludes, automates, records, certifies, or justifies action, it must carry the conditions under which its authority is earned.

Formalisation is the memory that a claim is not entitled to travel alone.

Its task is not to close the world.

Its task is to keep authority attached to its ground.

---

[1] Lipsky (1980) is a public-administration specialisation of the present claim that formalisation does not eliminate situated judgement. Street-level bureaucrats act through rules, categories, resource limits, and organisational pressures while retaining consequential discretion in how policy is applied. The present argument generalises beyond public administration and does not inherit Lipsky's institutional theory: its narrower burden is that where judgement becomes load-bearing, the relation, abstraction, evidence path, residual, threshold, safeguard, and update remain recoverable.

[2] Nissenbaum (1996) identifies distinctive accountability barriers in computerised systems, including the problem of many hands, treating bugs as inevitable, blaming the computer, and forms of ownership without corresponding liability. This is a strong comparator for the present claim that converting judgement into technical output cannot transfer or erase responsibility. The paper's Responsibility Provenance / Responsibility Warrant distinction is narrower and separately inherited from the preceding GOT architecture; Nissenbaum's accountability framework is not adopted as its source.

Endnotes

[3] Ionescu et al. (2019) use epistemic debt for accumulated technical ignorance in smart manufacturing, while Li (2026) independently uses the term for deferred verification obligations in AI-assisted research. These are prior and independent uses of the phrase. The present paper uses epistemic debt more narrowly for the deferred cost created when load-bearing claims or linguistic closures circulate as if their grounding burden had already been paid; no terminology priority or conceptual inheritance is claimed.

[4] Bowker and Star (1999) show how classification systems and standards become infrastructural, consequential, and difficult to see or contest. The present argument adds the TEM-specific burden that the operative relation, locus, abstraction, evidence path, residual, threshold, safeguard, and update remain recoverable; it does not inherit Bowker and Star's wider science-and-technology-studies framework.

[5] Selbst et al. (2019) identify abstraction traps that arise when technical fairness work is detached from the sociotechnical systems in which it operates. This is a strong domain-specific comparator for the present claim that machine execution cannot repair upstream epistemic under-specification. The present argument generalises beyond fairness and machine learning to load-bearing language across institutional and machine decision paths.

[6] Gebru et al. (2021) and Mitchell et al. (2019) provide domain-specific documentation frameworks for datasets and trained models, including provenance, composition, intended use, evaluation conditions, and limitations. They are compatible specialisations of the recoverability problem. The auditable claim packet is broader: it applies to consequential claims generally and additionally carries relation, locus, closure demand, comparator, residual, action threshold, safeguard, and update.

[7] Toulmin (1958) is a historical comparator for structured practical justification and the recoverable layout of claims and their support. The auditable claim packet does not adopt the Toulmin model: it is indexed to finite relational access and includes claim-path and reliance-path conditions, residual, threshold, safeguard, and update that perform different work.

[8] Bloomfield and Rushby's Assurance 2.0 (2022) is a particularly close methodological specialisation. Assurance cases organise claims, argument, evidence, defeaters, and residual doubts and require unresolved judgments to be made consciously and recorded. Its operative domain is assurance of safety, security, and other critical properties; the present paper proposes a general discipline for load-bearing language and consequential claim transport.

[9] Power (1999) is a critical comparator on the institutional expansion of auditing, formal scrutiny, and the production of auditable performance, including the possibility that audit practices impose their own values or generate dysfunctional consequences. The proportional-formalisation rule here is intended to block the corresponding failure in which forms, compliance, or auditability replace relation-preservation.

[10] Raji et al. (2020) develop an end-to-end framework for internal algorithmic auditing across the lifecycle of AI systems, motivated in part by the difficulty of tracing harmful outcomes back to the development choices that produced them after deployment. This is a strong domain-specific specialisation of the present paper's prospective and retrospective recoverability requirement. The auditable claim-path architecture here is broader than algorithmic auditing and does not treat lifecycle documentation or organisational audit as epistemic warrant by itself.

[11] Baron and Hershey (1988) provide empirical evidence for outcome bias in decision evaluation: otherwise comparable decisions were judged more favourably when their outcomes were favourable. The finding supports the paper's separation of decision quality from outcome success and is used as an empirical operational dependency only; no wider decision-theory framework is inherited.

References

Baron, J., & Hershey, J. C. (1988). Outcome bias in decision evaluation. Journal of Personality and Social Psychology, 54(4), 569–579. https://doi.org/10.1037/0022-3514.54.4.569

Bloomfield, R., & Rushby, J. (2022). Assessing Confidence with Assurance 2.0 (CSL Technical Report SRI-CSL-2022-02). SRI International.

Bowker, G. C., & Star, S. L. (1999). Sorting Things Out: Classification and Its Consequences. MIT Press.

Gebru, T., Morgenstern, J., Vecchione, B., Vaughan, J. W., Wallach, H., Daumé III, H., & Crawford, K. (2021). Datasheets for datasets. Communications of the ACM, 64(12), 86–92. https://doi.org/10.1145/3458723

Ionescu, T. B., Schlund, S., & Schmidbauer, C. (2019). Epistemic debt: A concept and measure of technical ignorance in smart manufacturing. In I. L. Nunes (Ed.), Advances in Human Factors and Systems Interaction (pp. 81–93). Springer. https://doi.org/10.1007/978-3-030-20040-4_8

Li, H. (2026). The calibration turn in AI-assisted research: A conceptual and methodological framework for evidence-licensed claims. arXiv:2606.31273.

Lipsky, M. (1980). Street-Level Bureaucracy: Dilemmas of the Individual in Public Services. Russell Sage Foundation.

Mitchell, M., Wu, S., Zaldivar, A., Barnes, P., Vasserman, L., Hutchinson, B., Spitzer, E., Raji, I. D., & Gebru, T. (2019). Model cards for model reporting. Proceedings of the Conference on Fairness, Accountability, and Transparency, 220–229. https://doi.org/10.1145/3287560.3287596

Nissenbaum, H. (1996). Accountability in a computerized society. Science and Engineering Ethics, 2, 25–42. https://doi.org/10.1007/BF02639315

Power, M. (1999). The Audit Society: Rituals of Verification. Oxford University Press. https://doi.org/10.1093/acprof:oso/9780198296034.001.0001

Raji, I. D., Smart, A., White, R. N., Mitchell, M., Gebru, T., Hutchinson, B., Smith-Loud, J., Theron, D., & Barnes, P. (2020). 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, 33–44. https://doi.org/10.1145/3351095.3372873

Selbst, A. D., boyd, d., Friedler, S. A., Venkatasubramanian, S., & Vertesi, J. (2019). Fairness and abstraction in sociotechnical systems. Proceedings of the Conference on Fairness, Accountability, and Transparency, 59–68. https://doi.org/10.1145/3287560.3287598

Toulmin, S. E. (1958). The Uses of Argument. Cambridge University Press.