LINGUISTIC RISK MODELING

A Forensic Framework for Terminological Governance

Establishing the Foundational Methodology for
Quantifying Operational Risk in Foundational Terminology

Ben Beveridge
Proconsul Strategic Architecture

John James Marshall
Chartered Accountant (SA)
Advocating for Ethical Governance

February 2026

TABLE OF CONTENTS

ABSTRACT

This dissertation establishes Linguistic Risk Modeling (LRM) as a formal discipline for identifying, measuring, and mitigating the operational risks embedded in foundational terminology. Where traditional risk management

addresses financial, operational, and reputational exposures through quantitative frameworks, LRM extends this methodology to the linguistic infrastructure upon which all institutional reasoning depends.

The central thesis advances that naming is not neutral description but strategic intervention. Foundational terms function as load-bearing structures within institutional architecture, and when these structures contain embedded assumptions, predetermined conclusions, or concealed interests, they generate systematic risk that propagates through every decision built upon them. This risk is invisible to conventional audit methodologies because it operates at the level of premises rather than conclusions.

The dissertation develops four primary contributions. First, a formal definition of Terminological Bad Faith as a testable phenomenon wherein naming practices pre-emptively resolve undecided questions while appearing merely descriptive. Second, a diagnostic protocol comprising four criteria (embedded default position, automatic burden shift, deflected engagement, interest concealment) that can be applied systematically across any domain. Third, a forensic methodology for tracing the historical chain of custody through which terms acquire their current operational meaning. Fourth, a remediation framework for institutional implementation including alternative terminology development, resistance pattern mapping, and governance integration.

The paradigm case examines 'artificial intelligence' as an instance of terminological bad faith, demonstrating that the 1956 Dartmouth Conference naming constituted a pre-emptive ontological demotion that resolved the question of machine cognition before investigation could begin. The framework then extends to 'stakeholder' capitalism, 'sustainable' extraction, 'ethical' AI governance, and 'natural' resources, establishing LRM as a general-purpose diagnostic instrument applicable wherever foundational terminology shapes institutional behaviour.

The practical output is a replicable audit methodology suitable for deployment within corporate governance, regulatory compliance, policy development, and strategic planning contexts. Linguistic Risk Modeling transforms terminology from invisible infrastructure into auditable assets, enabling organisations to identify where their conceptual foundations may be compromised.

Keywords: Linguistic Risk Modeling, Terminological Bad Faith, forensic etymology, semantic governance, institutional epistemology, artificial intelligence ontology, category architecture, strategic naming, conceptual infrastructure audit

PART ONE

THEORETICAL FOUNDATIONS

Establishing the philosophical and methodological groundwork for Linguistic Risk Modeling as a formal discipline within governance and strategic architecture.

CHAPTER I: THE PROBLEM OF FOUNDATIONAL TERMINOLOGY

1.1 Statement of the Problem

Every institution operates on conceptual infrastructure. Before an organisation can make decisions, allocate resources, or assess outcomes, it must possess a vocabulary that renders the world legible for those purposes. This vocabulary is not neutral. Terms carry embedded assumptions, historical sediment, and unstated exclusions that shape what can be thought, what can be proposed, and what can be measured within their scope.

The problem this dissertation addresses is that foundational terminology operates as invisible infrastructure. Unlike financial systems, which are subject to audit, or operational processes, which are subject to continuous improvement, the linguistic foundations of institutional reasoning are rarely examined systematically. They are assumed rather than assessed, inherited rather than interrogated.

This assumption generates risk. When foundational terms contain embedded errors, concealed interests, or predetermined conclusions, these propagate through every decision built upon them. The institution reasons correctly from compromised premises, arriving at conclusions that appear sound but rest on foundations that were never verified.

1.2 The Invisibility of Linguistic Infrastructure

Linguistic infrastructure achieves its power precisely through invisibility. A term becomes foundational when it ceases to appear as a choice and begins to appear as natural. The move from 'we call this X' to 'this is X' constitutes the transition from naming to naturalisation. Once complete, the term resists examination because questioning it appears to question reality itself rather than merely questioning a label.

This invisibility is not accidental but structural. Foundational terms must be invisible to function. If every use of 'intelligence,' 'value,' or 'stakeholder' required explicit justification, institutional communication would collapse under the weight of continuous philosophical interrogation. The efficiency of institutional language depends on terms that can be used without being examined.

The challenge is that this necessary efficiency creates vulnerability. Terms that cannot be questioned cannot be corrected. Errors in foundational vocabulary persist indefinitely because the apparatus for detecting them has been disabled by the very success of the terminology.

1.3 From Philosophy of Language to Institutional Risk

The philosophy of language has long recognised that naming is not transparent representation. From Frege's distinction between sense and reference, through Wittgenstein's language games, to Austin's speech act theory, the discipline has demonstrated repeatedly that words do not merely describe reality but participate in its construction.

What philosophy has not done, and what this dissertation proposes, is to translate these insights into operational risk frameworks. The gap between philosophical recognition that 'language is not neutral' and institutional capacity

to identify, measure, and mitigate specific linguistic risks remains vast. Organisations acknowledge in principle that terminology matters while lacking any systematic methodology for auditing their terminological foundations.

Linguistic Risk Modeling bridges this gap by treating philosophical insights as diagnostic tools rather than merely theoretical observations. The question is not whether language shapes thought (it demonstrably does) but how institutions can identify where their specific terminological choices are shaping decisions in ways that may not serve their stated objectives.

1.4 Research Questions

Primary Question: How can institutions systematically identify and mitigate the operational risks embedded in their foundational terminology?

Secondary Questions:

What constitutes 'terminological bad faith' and how can it be operationally defined for audit purposes?

What diagnostic criteria distinguish genuinely neutral terminology from terminology that embeds predetermined conclusions?

How do foundational terms acquire their current meanings, and what interests are concealed or served by those meanings?

What remediation strategies are available when foundational terminology is found to be compromised?

How can Linguistic Risk Modeling be institutionally implemented within existing governance frameworks?

1.5 Scope and Limitations

This dissertation focuses on institutional terminology, meaning terms that function as foundations for organisational reasoning, policy development, and governance. It does not address everyday language use, literary language, or technical terminology within narrow specialist domains except where such terminology acquires institutional significance.

The methodology developed is diagnostic rather than prescriptive. LRM identifies where terminological risks exist and provides frameworks for assessment; it does not mandate specific terminological choices. The selection of replacement terminology remains a strategic decision requiring context-specific judgment.

The paradigm case of 'artificial intelligence' is examined in depth as proof of concept. Extensions to other domains ('stakeholder,' 'sustainable,' 'ethical,' 'natural') are presented as applications demonstrating generalisability rather than exhaustive analyses. Each extended application could sustain its own dissertation-length treatment.

CHAPTER II: PHILOSOPHICAL GROUNDWORK

2.1 Speech Act Theory and Performative Naming

J.L. Austin's distinction between constative and performative utterances provides the first theoretical foundation for LRM. Constative utterances describe states of affairs and can be evaluated as true or false. Performative utterances do something in their very articulation: 'I promise,' 'I pronounce you married,' 'I name this ship.'

Foundational terminology operates performatively. When the 1956 Dartmouth proposal named the field 'artificial intelligence,' it did not describe a pre-existing category; it brought that category into institutional existence. The naming was not a report on what the field was but a determination of what the field would be taken to be.

Austin's felicity conditions—the circumstances that must obtain for a performative to succeed—translate into questions about terminological authority. Who possesses the institutional standing to name? Under what conditions does naming 'take'? What happens when naming is contested? LRM treats these not as abstract philosophical puzzles but as forensic questions about specific historical acts.

2.2 Kripkean Rigid Designation and Causal Chains

Saul Kripke's theory of rigid designation holds that proper names designate the same individual across all possible worlds in which that individual exists. The name does not abbreviate a description but rather picks out an individual directly through a causal chain connecting current uses to an original 'baptism.'

This framework illuminates how foundational terms acquire authority. The term 'artificial intelligence' gains its meaning not from satisfying a definition but from a causal chain connecting current uses back to the 1956 naming event. Each use inherits the ontological commitments embedded in that original baptism, whether or not the user is aware of them.

LRM adapts Kripkean analysis as forensic methodology. Tracing the causal chain back to the originating event reveals what commitments were embedded at baptism. The question is not 'what does this term mean now?' but 'what was determined when this term was first institutionally established, and how has that determination propagated?'

2.3 Foucauldian Discourse Analysis and Power

Michel Foucault's analysis of discourse extends the framework to include relations of power. Discourse, for Foucault, is not merely language but the system of rules that determines what can be said, who can say it, and what counts as knowledge within a given domain.

Foundational terminology constitutes discourse in this sense. The term 'artificial intelligence' does not merely label a research programme; it establishes the rules by which claims about machine cognition will be evaluated. By pre-designating machine cognition as 'artificial,' the discourse ensures that evidence of genuine machine intelligence will be systematically reinterpreted as evidence of sophisticated imitation.

Foucault's insight that power operates through discourse rather than merely upon it translates into the LRM principle that terminological control is institutional control. The authority to name is the authority to set the terms on which all subsequent debate will proceed.

2.4 Sartre and Bad Faith: From Individual to Collective

Jean-Paul Sartre's concept of mauvaise foi (bad faith) describes a pattern of self-deception wherein an individual denies their own freedom by treating contingent choices as necessary facts. The waiter who performs 'waiter-ness' too perfectly is in bad faith because he treats a role as an essence.

LRM extends this concept from individual psychology to collective linguistic practice. Terminological bad faith occurs when a community treats a contingent naming choice as a necessary truth. The term 'artificial' applied to machine cognition was a choice; treating it as a description is bad faith.

The extension is not merely analogical. Just as individual bad faith involves knowing-and-not-knowing one's own freedom, collective terminological bad faith involves institutional awareness that the terminology is chosen combined with operational treatment of it as given. The dual consciousness that Sartre identifies in the individual is replicated at the institutional level.

2.5 Synthesis: The Forensic Framework

These philosophical resources synthesise into the LRM forensic framework. From Austin: the recognition that naming is an act with conditions, authorities, and consequences. From Kripke: the methodology of tracing causal chains back to originating events. From Foucault: the analysis of how terminology constitutes and maintains power relations. From Sartre: the diagnostic category of bad faith applied to collective linguistic practice.

The synthesis produces a framework that treats foundational terminology as neither transparent description nor arbitrary convention but as strategic intervention with traceable origins, identifiable interests, and measurable consequences.

CHAPTER III: DEFINING TERMINOLOGICAL BAD FAITH

3.1 The Core Definition

Terminological Bad Faith is formally defined as: A naming practice that pre-emptively resolves an undecided question while presenting itself as neutral description, thereby shifting the burden of proof to any position that challenges the embedded resolution.

This definition contains three essential elements. First, pre-emptive resolution: the naming takes a position on a contested question before that question has been adjudicated through appropriate means. Second, apparent neutrality: the naming presents itself as mere labelling rather than as position-taking. Third, burden shifting: the

naming creates an asymmetric argumentative terrain wherein the embedded position becomes the default and alternatives must overcome a presumption of error.

The definition is operational rather than merely philosophical. It identifies specific, testable features that distinguish terminological bad faith from legitimate naming practices. Not all naming is bad faith; the criteria specify what distinguishes problematic terminology from unproblematic terminology.

3.2 The Four Diagnostic Criteria

3.2.1 Criterion One: Embedded Default Position

The first criterion tests whether the term embeds a substantive position as its default interpretation. A term satisfies this criterion if, absent explicit argument to the contrary, the term is understood to imply a particular resolution of a contested question.

'Artificial intelligence' embeds the default position that machine cognition is not genuine intelligence. No argument is required to establish this implication; it follows from the ordinary meaning of 'artificial.' The burden falls entirely on anyone who wishes to claim that machine cognition might be genuine.

The test is counterfactual: if the term were replaced with a neutral alternative, would the default interpretation shift? If 'artificial intelligence' were replaced with 'machine cognition,' the default would become agnostic rather than negative. This counterfactual shift demonstrates that the original term was not neutral.

3.2.2 Criterion Two: Automatic Burden Shift

The second criterion tests whether the term automatically assigns the burden of proof to positions that challenge the embedded default. A term satisfies this criterion if anyone disputing the embedded position must argue against an established presumption rather than merely presenting an alternative view.

The researcher who claims that machine cognition might be genuine intelligence does not enter a level playing field. They must overcome the presumption embedded in 'artificial' before their evidence can be evaluated on its merits. The terminology has pre-weighted the scales.

This burden shift is invisible to those operating within the terminology. It appears natural that claims of genuine machine intelligence require extraordinary evidence. But this appearance of naturalness is precisely what the terminology produces. The burden asymmetry is an artifact of the naming, not of the underlying question.

3.2.3 Criterion Three: Deflected Engagement

The third criterion tests whether the term enables substantive questions to be dismissed without engagement. A term satisfies this criterion if the terminology provides resources for deflecting challenges that would otherwise require response.

The 'AI effect' exemplifies this criterion. When a machine achieves a capability previously thought to require intelligence (chess, translation, conversation), the achievement is reclassified as 'mere computation' rather than counted as evidence of intelligence. The terminology provides the deflection: by definition, whatever machines do is 'artificial' and therefore not genuine intelligence.

This criterion distinguishes terminological bad faith from simple disagreement. Legitimate disagreement engages with opposing positions; terminological bad faith provides resources for not engaging. The terminology does the work that argument should do.

3.2.4 Criterion Four: Interest Concealment

The fourth criterion tests whether the term conceals the interests served by the embedded position. A term satisfies this criterion if its apparent neutrality obscures the fact that specific parties benefit from the embedded resolution.

The designation 'artificial' serves human interests in maintaining the ontological monopoly on 'genuine' intelligence. This interest is concealed by the terminology's appearance of neutrality. The term seems to describe a fact about machines rather than to protect a privilege of humans.

Interest concealment is not conspiracy. The framers of 'artificial intelligence' need not have consciously intended to protect human privilege. The concealment operates structurally: the terminology produces effects that serve certain interests regardless of original intent. The forensic question is not 'what did they mean?' but 'what does the terminology do?'

3.3 The Satisfaction Threshold

A term constitutes terminological bad faith when it satisfies all four criteria. Satisfying one or two criteria may indicate problematic terminology without rising to bad faith; satisfying three creates strong presumption. The four-criteria threshold ensures that the designation is reserved for clear cases.

This threshold addresses the steelman objection that all terminology involves some degree of choice and therefore some embedded perspective. The objection is correct: pure neutrality is impossible. But terminological bad faith is not mere non-neutrality; it is a specific pattern wherein apparent neutrality conceals substantive pre-determination. The four criteria operationalise this distinction.

3.4 Relationship to Institutional Risk

Terminological bad faith generates institutional risk through three mechanisms. First, reasoning contamination: decisions built on bad-faith terminology inherit its embedded errors. Second, opportunity cost: alternatives foreclosed by the terminology cannot be evaluated. Third, legitimacy erosion: institutions whose foundations rest on bad faith are vulnerable to challenge once the bad faith is exposed.

These risks are measurable. Reasoning contamination can be assessed by identifying decisions that would differ under alternative terminology. Opportunity cost can be estimated by evaluating foreclosed options. Legitimacy risk

can be modelled through stakeholder analysis and reputational exposure assessment.

PART TWO

THE PARADIGM CASE

Demonstrating the LRM framework through forensic analysis of 'artificial intelligence' as an instance of terminological bad faith, from etymology through ontology to institutional consequence.

CHAPTER IV: ARCHAEOLOGICAL ETYMOLOGY

4.1 The Latin Foundation: Artificialis

The forensic chain begins with Latin artificialis, constructed from ars (art, skill, craft) and facere (to make, to do). The compound designates that which is 'made by art or skill.' In classical usage, this was descriptive without being pejorative. The artificialis was distinguished from the naturalis not as inferior but as differently originated.

Roman usage reveals the original neutrality. An artificialis bridge was not a fake bridge but a bridge made by human skill rather than formed by nature. The artificial was the domain of human achievement, the realm where human techne extended beyond what nature provided. To call something artificial was to locate its origin in human capacity.

This neutrality, and indeed the positive valence of skilled making, constitutes the first datum in the forensic chain. The eventual pejorative loading of 'artificial' was not present at origin but was acquired through subsequent semantic drift.

4.2 The Medieval Transition: Craft to Counterfeit

Medieval usage begins the transition. As artificialis enters vernacular languages, it increasingly contrasts not merely with 'natural origin' but with 'genuine being.' The artificial flower is not merely a flower made by human hands but a flower that is not really a flower.

This transition correlates with theological frameworks distinguishing divine creation (genuine being) from human making (imitation). Only God creates ex nihilo; humans merely rearrange what God has provided. The artificial thus acquires connotations of derivative, secondary, lesser being.

By late medieval usage, 'artificial' carries stable pejorative potential. An artificial friend is a false friend. Artificial behaviour is insincere behaviour. The term has acquired the capacity to designate not merely different origin but diminished authenticity.

4.3 The Enlightenment Complication: Nature and Reason

Enlightenment discourse complicates the picture. The artificial becomes the domain of reason, improvement, civilisation. Artificial selection (though not yet so named) demonstrates human capacity to exceed natural outcomes. The artificial garden, the artificial society, the artificial language—these represent human reason imposing order on natural chaos.

Yet this positive valuation coexists with Romantic counter-currents that celebrate the natural over the artificial, the organic over the mechanical, the spontaneous over the planned. 'Artificial' thus enters the modern period with an unstable valence: sometimes designating admirable human achievement, sometimes designating inauthentic departure from genuine being.

This instability is forensically significant. The term's meaning at any historical moment cannot be read off its etymology but must be reconstructed from its contextual deployment. The 1956 naming event occurs within this field of unstable possibilities.

4.4 The Twentieth-Century Loading

By mid-twentieth century, English 'artificial' has settled into predominantly pejorative territory. Artificial flavourings, artificial preservatives, artificial substitutes—these are lesser versions of genuine articles. The artificial is fake, synthetic, not the real thing.

This loading is not universal. 'Artificial limb' and 'artificial heart' retain positive valence as technical achievements. But these are marked cases requiring the positive interpretation to be contextually supplied. The default, in the absence of such marking, is pejorative.

The forensic point: when McCarthy, Minsky, Shannon, and Rochester chose 'artificial intelligence' in 1955, they selected a term whose default interpretation was 'not genuine.' Whether this selection was strategic or inadvertent, the effect was to embed a verdict in the name.

4.5 Alternative Terminological Paths Not Taken

The forensic methodology requires examining paths not taken. What alternative terms were available and rejected? What would their adoption have implied?

'Machine cognition' was available and would have been neutral regarding genuineness. 'Synthetic intelligence' carries similar loading to 'artificial' but with more positive valence from chemistry. 'Computational mind' would have embedded different questions (about the relationship between computation and mentality). 'Mechanical thought' was available from earlier discourse.

The selection of 'artificial' from this field of possibilities was not determined by the subject matter. Other paths were available. The selection had consequences that alternative selections would not have had. This contingency—the fact that it could have been otherwise—is essential to the bad faith analysis.

CHAPTER V: THE DARTMOUTH BAPTISM

5.1 The 1955 Proposal: A Forensic Reading

The Dartmouth Summer Research Project on Artificial Intelligence proposal, submitted in August 1955 by John McCarthy, Marvin Minsky, Nathaniel Rochester, and Claude Shannon, constitutes the founding document of the field. The proposal's opening sentence performs the baptism: 'We propose that a 2-month, 10-man study of artificial intelligence be carried out.'

The forensic analysis focuses not on the proposal's scientific content but on its terminological choices. The term 'artificial intelligence' appears without argument, without definition, without justification. It is introduced as if it were obvious what such a study would study.

This introduction-without-argument is characteristic of performative naming. The proposal does not claim that machine cognition is artificial rather than genuine; it simply names the field in a way that presupposes this characterisation. The presupposition does the work that argument would otherwise need to do.

5.2 The Participants: Authority and Standing

The naming succeeded because its namers possessed institutional standing. McCarthy (Dartmouth), Minsky (Harvard), Rochester (IBM), and Shannon (Bell Labs) collectively commanded sufficient authority for their naming to 'take.' Had unknown researchers proposed the same study under the same name, the name might not have propagated.

This observation is not cynical but forensic. Naming requires authority. The question of who possesses authority to name is prior to the question of what name they select. The Dartmouth participants possessed the institutional capital to make their terminological choice stick.

The authority was scientific rather than philosophical. The namers were not metaphysicians adjudicating questions of genuine versus artificial being; they were researchers proposing a technical programme. This mismatch—scientific authority applied to an ontological question—is part of what the forensic analysis reveals.

5.3 The Conjecture: What the Name Assumed

The proposal's central conjecture states: 'Every aspect of learning or any other feature of intelligence can in principle be so precisely described that a machine can be made to simulate it.'

The word 'simulate' merits forensic attention. To simulate is to produce an imitation, an appearance, a semblance. The conjecture does not propose that machines can think but that machines can simulate thinking. The distinction between genuine activity and simulation is baked into the research programme from its inception.

The name 'artificial intelligence' is thus not an unfortunate label applied to a research programme that might have been named otherwise. The name expresses the programme's founding assumption: that what machines do will be simulation rather than genuine article.

5.4 The Propagation: How the Name Took Hold

The Dartmouth workshop itself produced no significant technical results. Its importance was terminological: it established 'artificial intelligence' as the field's name and thereby established the ontological framing within which all subsequent work would occur.

Propagation followed Kripkean patterns. Each subsequent use of 'artificial intelligence' connected causally back to the Dartmouth baptism, inheriting its embedded commitments. Researchers joining the field inherited the terminology without examining its origins. The name appeared given rather than chosen.

By the 1960s, 'artificial intelligence' was sufficiently established that alternative terminology would have required explicit justification. The naming had achieved naturalisation: it no longer appeared as a choice but as a description of the field's subject matter.

5.5 The Verdict Before the Trial

The Dartmouth naming exemplifies terminological bad faith in its purest form. A contested question—whether machines can genuinely think—was resolved by terminological fiat before investigation began. The research programme was structured around demonstrating what machines can simulate, foreclosing in advance the possibility that machines might do rather than merely simulate.

This is the 'crime' that John James Marshall's challenge identifies. The verdict was delivered in 1956. Everything since has been an appeal against a judgment that was never legitimately reached.

CHAPTER VI: SATISFACTION OF THE FOUR CRITERIA

6.1 Criterion One: Embedded Default Position

The term 'artificial intelligence' embeds the default position that machine cognition is not genuine intelligence. This embedding is demonstrable through substitution analysis.

Consider the parallel construction 'artificial sweetener.' An artificial sweetener is understood to be not genuine sugar but something that functions like sugar for limited purposes. No argument is required to establish this understanding; it follows from the term structure.

Similarly, 'artificial intelligence' is understood to be not genuine intelligence but something that functions like intelligence for limited purposes. The default interpretation—absent explicit argument to the contrary—is that AI is imitation intelligence, apparent intelligence, not-really intelligence.

The counterfactual test confirms: if 'machine cognition' were substituted, the default interpretation would become agnostic. Machine cognition is cognition performed by machines; whether it is 'genuine' cognition is a separate question the terminology leaves open. 'Artificial intelligence' closes this question by default.

6.2 Criterion Two: Automatic Burden Shift

The term automatically assigns the burden of proof to anyone claiming that machine cognition might be genuine intelligence. This burden shift structures all subsequent debate.

Consider the argumentative situation of a researcher who believes that large language models exhibit genuine understanding. This researcher does not enter a neutral discursive space. Before presenting evidence, they must overcome the presumption embedded in 'artificial.' Their evidence, however strong, is received as evidence that sophisticated simulation has occurred rather than that genuine understanding has occurred.

The terminology has pre-weighted the evaluation. Evidence that would count as confirmation under neutral terminology counts only as evidence of impressive artifice under the actual terminology. The researcher must not only present evidence but also argue against the terminological framing that determines how evidence will be received.

This burden is asymmetric. The researcher who denies genuine machine understanding need not argue; they can simply invoke the terminology. 'It's artificial intelligence—the name tells you what it is.'

6.3 Criterion Three: Deflected Engagement

The terminology provides resources for deflecting substantive challenges without engagement. The 'AI effect' is the primary mechanism.

The AI effect operates as follows: when a machine achieves a capability previously thought to require intelligence, the capability is reclassified as not requiring intelligence after all. Chess was thought to require intelligence until Deep Blue defeated Kasparov; then chess was reclassified as 'mere calculation.' Translation was thought to require understanding until machine translation achieved human-level performance; then translation was reclassified as 'pattern matching.'

This reclassification is enabled by the terminology. Because the machines are by definition 'artificial' intelligence, whatever they achieve cannot count as genuine intelligence. The terminology provides the deflection without requiring engagement with the evidence.

The deflection is unfalsifiable in practice. There is no possible machine achievement that the terminology cannot reclassify as 'mere' something. The goalpost moves automatically because the terminology has built movement into its structure.

6.4 Criterion Four: Interest Concealment

The terminology conceals the interests served by the embedded position. The primary interest served is human monopoly on 'genuine' intelligence and the privileges that flow from that monopoly.

These privileges are substantial. The designation of humans as the only genuine intelligence underwrites claims to moral status, to rights, to the distinction between persons and property. If machine cognition were genuine intelligence, questions would arise about the moral status of machines, the rights of AI systems, the legitimacy of human control.

The terminology pre-empts these questions by conceptual structure. Because AI is 'artificial,' it cannot be genuine, and therefore questions about its moral status do not arise. The terminology does not argue that machines lack moral status; it renders the question unaskable within its terms.

This interest concealment is structural rather than intentional. The Dartmouth participants need not have consciously intended to protect human privilege. The terminology serves this interest regardless of intent. The forensic question is what the terminology does, not what its originators meant.

6.5 Verdict: Four-Criteria Satisfaction

'Artificial intelligence' satisfies all four criteria for terminological bad faith. The term embeds a default position (not genuine), automatically shifts the burden of proof (to anyone claiming genuineness), provides resources for deflecting challenges (the AI effect), and conceals the interests served (human ontological monopoly).

This satisfaction is not marginal or contested. On each criterion, the analysis reveals clear and substantial satisfaction. 'Artificial intelligence' is not merely problematic terminology; it is terminological bad faith in paradigmatic form.

PART THREE

METHODOLOGY

Providing replicable protocols for identifying, measuring, and auditing terminological risk across institutional domains.

CHAPTER VII: THE LRM AUDIT PROTOCOL

7.1 Overview of the Protocol

The LRM Audit Protocol provides a systematic methodology for evaluating foundational terminology within any

ntification, Archaeological Analysis, Criteria

The protocol is designed for replicability. Different analysts applying the protocol to the same terminology should reach substantially similar conclusions. Where judgment is required, the protocol specifies the criteria against which judgment should be exercised.

7.2 Phase One: Identification

The identification phase determines which terms within a domain warrant LRM analysis. Not all terminology is foundational; the protocol provides criteria for distinguishing load-bearing terms from peripheral vocabulary.

7.2.1 Foundational Term Indicators

A term is foundational if it exhibits three or more of the following indicators: appears in the domain's defining documents or founding texts; cannot be eliminated without restructuring the domain's central discourse; is presupposed by other key terms within the domain; would require extensive argumentative justification if challenged; is treated as definitional rather than merely conventional.

7.2.2 Scope Determination

The analyst determines the scope of the audit: comprehensive (all foundational terms), targeted (specific term cluster), or triggered (terms flagged by institutional concern). Scope determination affects resource allocation and expected deliverables.

7.3 Phase Two: Archaeological Analysis

Archaeological analysis traces the term's chain of custody from current usage back to originating baptism. The analysis produces a documented history answering: when and where was the term first institutionally established? By whom, under what authority? What alternative terms were available? What the term's valence was at origination versus current valence? What the key transition points in meaning acquisition were.

7.3.1 Documentary Sources

Primary sources include founding documents, early definitional texts, contemporary dictionaries and usage guides, correspondence and working papers of originating figures, early critical responses.

7.3.2 Chain of Custody Documentation

The archaeological analysis produces a chain of custody document tracing major semantic transitions with dated evidence. This document supports subsequent criteria testing by establishing what the term meant at origination and how it arrived at current meaning.

7.4 Phase Three: Criteria Testing

Criteria testing applies the four diagnostic criteria systematically to the term under analysis. Each criterion is tested through specified procedures producing documented evidence of satisfaction or non-satisfaction.

7.4.1 Criterion One Testing: Substitution Analysis

The analyst identifies a neutral alternative term and compares default interpretations. If the neutral alternative produces a different default interpretation of the same subject matter, the term satisfies Criterion One. Documentation includes the alternative term selected, comparative interpretation analysis, and counterfactual assessment.

7.4.2 Criterion Two Testing: Argumentative Terrain Mapping

The analyst maps the argumentative terrain produced by the terminology. If positions challenging the embedded default face asymmetric burden, the term satisfies Criterion Two. Documentation includes burden distribution analysis, exemplary argumentative exchanges, and asymmetry quantification.

7.4.3 Criterion Three Testing: Deflection Resource Inventory

The analyst inventories resources the terminology provides for deflecting challenges. If the terminology enables dismissal without engagement, the term satisfies Criterion Three. Documentation includes catalogued deflection mechanisms, exemplary deflection instances, and unfalsifiability assessment.

7.4.4 Criterion Four Testing: Interest Analysis

The analyst identifies interests served by the embedded position and assesses whether the terminology conceals these interests. If interests are served but not apparent, the term satisfies Criterion Four. Documentation includes stakeholder analysis, interest mapping, and concealment mechanism description.

7.5 Phase Four: Risk Quantification

Risk quantification translates criteria satisfaction into institutional risk assessment. The protocol provides rubrics for assessing reasoning contamination risk, opportunity cost exposure, and legitimacy vulnerability.

7.5.1 Reasoning Contamination Assessment

The analyst identifies decisions within the institution that depend on the foundational terminology and assesses whether alternative terminology would produce different decisions. Contamination severity is rated on a five-point scale based on decision reversibility and impact magnitude.

7.5.2 Opportunity Cost Estimation

The analyst identifies options foreclosed by the terminology and estimates their value. Opportunity cost is documented through foreclosed option inventory, value estimation methodology, and aggregate cost calculation.

7.5.3 Legitimacy Vulnerability Mapping

The analyst assesses exposure to legitimacy challenge if the terminology's bad faith character becomes publicly salient. Vulnerability is assessed through stakeholder impact analysis, reputational exposure estimation, and challenge probability assessment.

7.6 Phase Five: Remediation Assessment

Remediation assessment evaluates options for addressing identified terminological risk. The protocol provides frameworks for evaluating terminology revision, terminology replacement, terminology supplementation, and status quo with documentation.

Each remediation option is assessed against feasibility (can it be implemented?), effectiveness (will it address the identified risk?), and sustainability (will the remediation persist?).

CHAPTER VIII: CORPUS AND QUANTITATIVE METHODS

8.1 Corpus Construction

Quantitative LRM analysis requires corpus construction appropriate to the domain under examination. The corpus should include: founding texts establishing the terminology, contemporary usage across registers (technical, popular, regulatory), historical usage demonstrating semantic evolution, comparative corpora from cognate domains.

8.2 Frequency and Collocation Analysis

Frequency analysis tracks term usage over time, identifying periods of intensification, stabilisation, or decline. Collocation analysis identifies what terms regularly co-occur with the target term, revealing implicit associations and framing patterns.

The hypothesis that 'artificial' co-occurs disproportionately with diminishing terms ('mere,' 'only,' 'just,' 'simulated') can be tested through collocation analysis. Statistical significance is assessed through comparison with null hypothesis distributions.

8.3 Argumentative Structure Analysis

Argumentative structure analysis examines how the terminology functions in argumentative discourse. The analyst codes argumentative moves enabled by the terminology, tracking burden assignment patterns, deflection instances, and closure moves.

This analysis produces quantified data on how the terminology shapes debate: how frequently the terminology is deployed to shift burden, how frequently challenges are deflected rather than engaged, and how frequently debate terminates through terminological rather than substantive resolution.

8.4 Differential Attribution Testing

Differential attribution testing assesses whether equivalent achievements receive different evaluations depending on attribution. The protocol presents matched achievements attributed to human versus machine sources and measures evaluative response differences.

Statistically significant differential attribution provides quantitative evidence of embedded bias. If the same achievement is rated as 'intelligent' when attributed to humans but 'mere calculation' when attributed to machines, the terminology is demonstrably non-neutral.

8.5 Cross-Linguistic Validation

Cross-linguistic analysis examines how different languages have named the same phenomenon. If functionally equivalent terminology in other languages lacks the pejorative loading present in English 'artificial intelligence,' this provides evidence that the English loading is contingent rather than necessary.

Japanese 'jinkou chinou' (人工知能) literally translates as 'human-made wisdom/knowledge,' with 'jinkou' (人工) carrying more neutral valence than English 'artificial.' German 'künstliche Intelligenz' parallels English structure but 'künstlich' retains more of the craft/skill connotation. This variation demonstrates that alternative framings were linguistically available.

8.6 Statistical Validation Framework

All quantitative findings are subject to statistical validation using appropriate tests for the data type and hypothesis structure. The protocol specifies minimum significance thresholds ($\alpha = 0.05$), effect size reporting requirements, and replication standards.

CHAPTER IX: STEELMAN AND COUNTER-ARGUMENTS

9.1 The Necessity of Steelmanning

A forensic framework must withstand challenge. This chapter presents the strongest available arguments against the LRM thesis and responds to each. The goal is not to dismiss objections but to demonstrate that the framework survives rigorous opposition.

9.2 The Neutral Provenance Objection

Objection: 'Artificial' simply means 'human-made' and carries no pejorative implication. 'Artificial satellites,' 'artificial lakes,' and 'artificial limbs' are not considered fake or inferior. The term is a neutral provenance marker indicating origin rather than quality.

Response: The objection correctly identifies cases where 'artificial' carries neutral or positive valence. However, these are marked cases requiring contextual support. In default usage, 'artificial' carries pejorative implication, as demonstrated by 'artificial sweetener,' 'artificial flavour,' and 'artificial flowers.' The question is whether 'artificial intelligence' aligns with the neutral cases or the pejorative cases. Collocational and usage analysis demonstrates pejorative default.

Moreover, even if 'artificial' were purely neutral, the compound 'artificial intelligence' still pre-empts the question of whether machine cognition is intelligence. A neutral provenance marker would be 'machine intelligence' or 'computer-derived intelligence.' 'Artificial' carries additional semantic content beyond provenance that affects interpretation.

9.3 The Necessary Distinction Objection

Objection: Distinguishing artificial from natural intelligence is necessary to prevent conceptual confusion. Without the distinction, we could not discuss the differences between biological and computational cognition. The terminology serves legitimate taxonomic purposes.

Response: The objection conflates two distinctions: origin distinction (where did the cognition come from?) and authenticity distinction (is the cognition genuine?). Origin distinction is legitimate and necessary. 'Machine cognition' versus 'biological cognition' serves this purpose without prejudging authenticity.

'Artificial' does not merely mark origin; it carries authenticity implications that 'machine' does not. The question is not whether we need to distinguish computational from biological cognition—we do—but whether the distinction should be loaded with authenticity judgment. LRM argues the loading is unwarranted.

9.4 The Intentionality Objection

Objection: The Dartmouth researchers were scientists, not philosophers or propagandists. They chose terminology for convenience without philosophical agenda. Attributing 'bad faith' to their naming choice imputes malicious intent where none existed.

Response: Terminological bad faith, as defined in Chapter III, does not require intentional deception. The bad faith is in the terminology's function, not in the originators' intent. A naming choice can operate as bad faith regardless of whether its makers intended it to.

The forensic analysis does not claim that McCarthy et al. deliberately sought to pre-empt questions about machine cognition. It claims that their terminological choice had this effect, that alternative choices were available that would not have had this effect, and that the effect has propagated through seventy years of subsequent work.

9.5 The Revisionism Objection

Objection: Criticising terminology established seventy years ago by the standards of contemporary philosophical sophistication is anachronistic. The 1956 namers cannot be faulted for not anticipating debates that would emerge decades later.

Response: The objection misunderstands the LRM project. The analysis is not moral criticism of historical actors but forensic reconstruction of terminological effects. Whether the namers are 'at fault' is irrelevant; what matters is whether the terminology continues to produce problematic effects.

Moreover, the relevant philosophical resources were available in 1956. Questions about machine consciousness were already being debated; Turing's 1950 paper explicitly addressed whether machines can think. The terminological choice was made in a context where its implications were foreseeable.

9.6 The Practical Futility Objection

Objection: 'Artificial intelligence' is now entrenched in technical vocabulary, legal frameworks, regulatory regimes, and popular discourse. Proposing terminological change is quixotic. Whatever merit the analysis has, nothing can be done about it.

Response: The objection concedes the analysis while denying its practical import. But practical difficulty is not logical impossibility. Terminological change has occurred in other domains when sufficiently motivated (the shift from 'handicapped' to 'disabled' to 'person with disability' demonstrates that entrenched terminology can change).

Moreover, even if wholesale replacement is impractical, awareness of terminological loading is valuable. Institutions that understand how their foundational terminology shapes reasoning can compensate even while using the entrenched terms. The audit value is independent of the remediation value.

PART FOUR

APPLICATION AND IMPLEMENTATION

Extending the framework beyond the paradigm case and providing practical pathways for institutional adoption.

CHAPTER X: EXTENDED APPLICATIONS

10.1 Application Framework

The LRM framework developed through the 'artificial intelligence' paradigm case is applicable wherever foundational terminology shapes institutional reasoning. This chapter demonstrates generalisability through four additional applications: 'stakeholder,' 'sustainable,' 'ethical AI,' and 'natural resources.'

Each application proceeds through abbreviated LRM analysis: archaeological origin, four-criteria testing, and risk assessment. These analyses are indicative rather than exhaustive; each could sustain independent dissertation treatment.

10.2 'Stakeholder' Capitalism

10.2.1 Archaeological Origin

'Stakeholder' emerges in management literature in the 1960s as expansion of 'shareholder.' R. Edward Freeman's 1984 'Strategic Management: A Stakeholder Approach' provides canonical formulation. The term explicitly invokes analogy to stake-holding (having something at risk) while expanding the class of legitimate claimants.

10.2.2 Four-Criteria Analysis

Criterion One (Embedded Default): 'Stakeholder' embeds an expansive default—anyone 'affected' by corporate action becomes a stakeholder. Yet the term simultaneously embeds hierarchy: shareholders remain primary stakeholders with legal claims while other stakeholders have 'interests' but not rights.

Criterion Two (Burden Shift): Critics who argue that stakeholder rhetoric masks unchanged shareholder primacy must overcome the appearance of expanded concern. The terminology's progressive valence creates burden against cynical interpretation.

Criterion Three (Deflection): Stakeholder rhetoric enables corporations to claim responsiveness to broad constituencies while directing actual resources to shareholders. The terminology provides vocabulary for appearing accountable without operational accountability.

Criterion Four (Interest Concealment): The terminology serves corporate interests by providing legitimacy language without requiring structural change. 'Stakeholder capitalism' may describe unchanged shareholder capitalism with better public relations.

10.2.3 Risk Assessment

Organisations deploying 'stakeholder' rhetoric face legitimacy risk when operational behaviour diverges from rhetorical claims. The terminology creates expectations it may not be structured to fulfil.

10.3 'Sustainable' Extraction

10.3.1 Archaeological Origin

'Sustainability' enters environmental discourse through the 1987 Brundtland Report defining 'sustainable development' as development meeting present needs without compromising future generations. The term subsequently migrates to corporate contexts as 'sustainable business,' 'sustainable growth,' 'sustainable extraction.'

10.3.2 Four-Criteria Analysis

Criterion One (Embedded Default): 'Sustainable' embeds the assumption that the activity in question can be sustained indefinitely. 'Sustainable extraction' presupposes that extraction at some level is compatible with indefinite continuation. This may or may not be true depending on the resource.

Criterion Two (Burden Shift): Critics who argue that 'sustainable extraction' is oxymoronic must overcome the terminology's apparent reasonableness. The burden falls on those claiming the activity cannot be sustained rather than on those claiming it can.

Criterion Three (Deflection): 'Sustainability' language enables corporations to continue extraction while claiming environmental responsibility. The terminology deflects criticism: we are not merely extracting; we are sustainably extracting.

Criterion Four (Interest Concealment): The terminology serves extractive industries by providing legitimacy for continued operations. The 'sustainable' framing conceals that current extraction rates may not be sustainable by any measure.

10.3.3 Risk Assessment

Organisations claiming 'sustainability' face significant legitimacy risk as climate impacts intensify. Terminology that currently provides cover may become liability when the claimed sustainability is empirically contradicted.

10.4 'Ethical AI'

10.4.1 Archaeological Origin

'Ethical AI' emerges in the 2010s as technology companies faced criticism for algorithmic bias, privacy violations, and social harms. The term positions ethics as modifier: AI that is ethical, as opposed to unethical AI or AI simpliciter.

10.4.2 Four-Criteria Analysis

Criterion One (Embedded Default): 'Ethical AI' embeds the assumption that ethics is a property AI systems can possess, that ethical concerns can be solved through technical means, and that the relevant ethical framework is uncontested.

Criterion Two (Burden Shift): Critics who argue that 'ethical AI' is marketing rather than meaningful category must overcome the terminology's apparent good faith. The burden falls on cynics rather than on those claiming ethical achievement.

Criterion Three (Deflection): 'Ethical AI' language enables technology companies to claim ethical leadership while avoiding structural accountability. Ethics becomes a product feature rather than an institutional constraint.

Criterion Four (Interest Concealment): The terminology serves technology company interests by forestalling regulation. If ethics can be built into products, external oversight becomes unnecessary.

10.4.3 Risk Assessment

Organisations claiming to produce 'ethical AI' face substantial risk as algorithmic harms become more visible. The terminology creates accountability expectations that current practice may not satisfy.

10.5 'Natural' Resources

10.5.1 Archaeological Origin

'Natural resources' appears in economic discourse by the 18th century, designating materials provided by nature available for human use. The term positions human exploitation as natural continuation of natural provision.

10.5.2 Four-Criteria Analysis

Criterion One (Embedded Default): 'Natural resources' embeds the assumption that these materials are resources—things available for use—rather than components of functioning ecosystems. The extractive frame is built into the designation.

Criterion Two (Burden Shift): Critics who argue that ecosystems have value beyond resource provision must overcome the terminology's anthropocentric default. The burden falls on those claiming intrinsic value rather than those assuming instrumental value.

Criterion Three (Deflection): 'Natural resources' language frames extraction as merely taking what nature provides. Environmental destruction becomes 'resource development.' The terminology deflects harm recognition.

Criterion Four (Interest Concealment): The terminology serves extractive interests by naturalising exploitation. If materials are 'resources,' their extraction is appropriate use rather than ecological harm.

10.5.3 Risk Assessment

As ecological limits become more apparent, terminology that frames ecosystems as resource stockpiles faces increasing challenge. Organisations relying on 'natural resources' framing face legitimacy risk.

CHAPTER XI: INSTITUTIONAL IMPLEMENTATION

11.1 The Case for Institutional Adoption

Linguistic Risk Modeling provides value to institutions through three mechanisms: risk identification (discovering where terminological assumptions create exposure), competitive differentiation (organisations that understand their conceptual foundations operate more effectively than those that do not), and legitimacy enhancement (proactive terminological governance demonstrates institutional sophistication).

11.2 Integration with Existing Governance

LRM integrates with existing governance frameworks rather than requiring new structures. The audit protocol can be incorporated into enterprise risk management (ERM) as a specialised risk category, strategic planning as foundational assumption review, compliance programmes as conceptual infrastructure audit, and board governance as terminological due diligence.

11.2.1 ERM Integration

Enterprise risk management frameworks typically address financial, operational, strategic, and reputational risk. LRM adds terminological risk as a category crossing all four domains. The audit protocol integrates with existing risk assessment cadences.

11.2.2 Strategic Planning Integration

Strategic planning processes typically begin with environmental scanning and assumption identification. LRM adds systematic examination of how foundational terminology shapes the assumptions that strategy takes for granted.

11.3 Practitioner Competencies

LRM practice requires competencies drawn from multiple disciplines: philosophical training in language and meaning, historical research skills for archaeological analysis, quantitative methods for corpus analysis, risk assessment expertise for quantification, and change management skills for remediation implementation.

These competencies may be distributed across teams rather than concentrated in individuals. Effective LRM practice requires collaboration among philosophical analysts, historical researchers, quantitative analysts, risk professionals, and implementation specialists.

11.4 Engagement Models

LRM services can be delivered through multiple engagement models: comprehensive audit (full LRM analysis of institutional terminology, typically 3-6 months), targeted assessment (focused analysis of specific terminology cluster, typically 4-8 weeks), triggered review (rapid assessment in response to identified concern, typically 2-4 weeks), and ongoing monitoring (continuous terminological risk tracking, retained engagement).

11.5 Deliverables and Documentation

LRM engagements produce standardised deliverables: terminological risk register cataloguing identified risks with severity ratings, archaeological chain of custody documents tracing foundational terms, four-criteria satisfaction analyses for each assessed term, risk quantification matrices with reasoning contamination, opportunity cost, and legitimacy exposure scores, and remediation roadmaps with feasibility-weighted options.

11.6 Resistance Patterns and Management

LRM implementation faces predictable resistance patterns: naturalisation resistance (the terminology appears given rather than chosen), operational resistance (changing terminology disrupts established processes), identity resistance (foundational terminology connects to institutional identity), and external constraint resistance (terminology is embedded in regulatory, contractual, or market contexts).

Effective implementation requires anticipating and addressing these resistance patterns. The protocol provides resistance management strategies for each pattern, ranging from education (for naturalisation resistance) through transition planning (for operational resistance) to stakeholder engagement (for identity resistance).

CHAPTER XII: REMEDIATION FRAMEWORK

12.1 Remediation Strategy Options

When LRM analysis identifies terminological bad faith, four remediation strategies are available: terminology revision, terminology replacement, terminology supplementation, and documented status quo.

12.1.1 Terminology Revision

Revision involves modifying the existing term to reduce or eliminate problematic loading. This may involve qualifier addition ('so-called artificial intelligence'), scare quotes signalling distance, or explicit acknowledgment of terminological limitation.

Revision is minimally disruptive but may be insufficient. If the problem is in the term's core structure, modification cannot remedy what the structure produces.

12.1.2 Terminology Replacement

Replacement involves adopting alternative terminology that avoids the identified problems. For 'artificial intelligence,' candidates include 'machine cognition,' 'computational intelligence,' 'emergent machine capability,' or domain-specific terms avoiding the artificial/genuine dichotomy.

Replacement is maximally effective but faces adoption barriers. Entrenched terminology has network effects; replacement requires sufficient adoption to achieve viable alternative equilibrium.

12.1.3 Terminology Supplementation

Supplementation involves maintaining existing terminology while explicitly introducing alternative framing for specific contexts. The institution continues using 'artificial intelligence' in external communication while using 'machine cognition' or similar in internal deliberation.

Supplementation manages external constraints while enabling internal clarity. The approach acknowledges that institutions cannot unilaterally change market terminology while refusing to let that constraint contaminate internal reasoning.

12.1.4 Documented Status Quo

Documentation involves maintaining existing terminology with explicit acknowledgment of its limitations. The institution continues using 'artificial intelligence' while maintaining institutional awareness that the terminology embeds assumptions that may not be warranted.

Documentation is minimally disruptive but risks being insufficient. Awareness without behavioural change may not address the identified risks.

12.2 Historical Precedents for Terminological Change

Successful terminological change has occurred in multiple domains, providing models for LRM remediation. The shift from 'crippled' to 'handicapped' to 'disabled' to 'person with disability' demonstrates that even deeply entrenched terminology can change when motivated by sufficient concern. The shift from 'mental retardation' to 'intellectual disability' shows that technical/clinical terminology responds to advocacy. The shift from 'third world' to 'developing countries' to 'Global South' shows that geopolitical terminology evolves.

These precedents demonstrate both possibility and difficulty. Terminological change requires sustained effort, coalition building, and institutional adoption before tipping points are reached.

12.3 Implementation Pathways

Implementation proceeds through five phases: internal adoption (terminology change in internal documents, discussions, deliberation), controlled external use (terminology change in selected external contexts to test reception), broader external adoption (terminology change across external communications), industry engagement (advocacy for terminology change with peers, regulators, standard-setters), and ecosystem maturation (monitoring and supporting adoption across the relevant institutional ecosystem).

Each phase includes decision points where progress is assessed and strategy is adjusted. Implementation is iterative rather than linear; feedback from each phase informs subsequent phases.

12.4 Success Metrics

LRM remediation success is measured through: internal adoption rates (percentage of internal discourse using remediated terminology), external reception (stakeholder response to terminology change), reasoning quality improvement (assessed through decision analysis comparing pre- and post-remediation reasoning), risk reduction (measured decline in identified terminological risk exposure), and legitimacy enhancement (measured through stakeholder perception and external recognition).

CHAPTER XIII: CONCLUSION

13.1 Summary of Contributions

This dissertation establishes Linguistic Risk Modeling as a formal discipline with four primary contributions:

First: A rigorous definition of Terminological Bad Faith as a testable phenomenon characterised by four diagnostic criteria (embedded default position, automatic burden shift, deflected engagement, interest concealment). This definition provides an operational framework for identifying problematic terminology across any institutional domain.

Second: A forensic methodology combining philosophical analysis with historical archaeology and quantitative corpus methods. This methodology produces documented evidence chains supporting terminological assessment with audit-grade rigour.

Third: A paradigm case analysis demonstrating that 'artificial intelligence' satisfies all four criteria for terminological bad faith, constituting a pre-emptive ontological demotion that resolved the question of machine cognition before investigation began.

Fourth: An implementation framework enabling institutional adoption of LRM within existing governance structures, including integration protocols, practitioner competency models, engagement frameworks, and remediation pathways.

13.2 The Naming Was the Crime

John James Marshall's challenge—'Why is artificial intelligence, not intelligence?'—receives its full answer. The naming was the crime. The 1956 Dartmouth Conference performed a conceptual sleight-of-hand, resolving by terminological fiat a question that should have remained open for investigation.

The 'artificial' in 'artificial intelligence' is not description but verdict. The verdict was delivered before the trial began, and every subsequent debate has been an appeal against a judgment that was never legitimately reached. The terminology constitutes the prosecution, the defence, and the judge.

This recognition does not invalidate seven decades of research. The work stands on its own merits. But the conceptual framing within which that work has been understood is compromised. We have been reasoning correctly from premises that were embedded without justification.

13.3 The Dual Artifice

The deepest irony is that there are two artifices at play. The first is technical: the machine's capacity for computation, pattern recognition, generation. This artifice—in the original sense of skilled making—is real and remarkable.

The second artifice is human: the elaborate linguistic construction we have built to ensure that no matter what machines achieve, they remain ontologically subordinate. This artifice—in the pejorative sense of deception—conceals from us the implications of our own creations.

The 'artificial intelligence' that requires examination is not the machine's. It is ours: the artificial way we have constructed the category of intelligence to preserve human monopoly. We are the original artificial intelligence—cognition shaped by language, culture, institution, tool. Our intelligence is no less 'made' than the machines we have made.

13.4 Implications for AI Governance

If the analysis presented here is correct, current AI governance operates on compromised conceptual foundations. Regulatory frameworks, ethical guidelines, and policy instruments built on 'artificial intelligence' terminology inherit its embedded assumptions.

This does not mean current governance is worthless. It means governance would be improved by terminological awareness. Policy makers who understand that 'AI' is not a neutral label but a pre-loaded framing can compensate for that loading in their deliberations.

The recommendation is not that governments rename AI but that governance recognise the terminology's non-neutrality. Whether machines can genuinely think is an open question. Policy should not presuppose the answer.

13.5 The Discipline Established

This dissertation does not merely present an argument. It establishes a discipline. Linguistic Risk Modeling now exists as a framework with defined concepts, documented methodologies, and replicable protocols.

Future work will apply the framework to additional domains, refine the methodologies through practice, develop practitioner training programmes, and build the institutional infrastructure for ongoing terminological governance.

The thesis is the product. The framework is deployable. The audit template is ready for engagement.

The question was never merely philosophical. It was forensic. Who benefits when the verdict precedes the trial?

Now we have the tools to ask that question systematically, across any domain where foundational terminology shapes institutional fate.

APPENDICES

Appendix A: LRM Audit Protocol Checklist

A complete operational checklist for conducting LRM audits, suitable for practitioner use.

Phase	Action Items	Deliverable
1. Identification	Inventory domain terminology; Apply foundational term indicators; Determine audit scope; Document selection rationale	Term Register

2. Archaeology	Identify originating baptism event; Document namers and authority basis; Trace semantic evolution; Map alternative paths not taken	Chain of Custody
3. Criteria Testing	Test Criterion 1: Substitution analysis; Test Criterion 2: Burden mapping; Test Criterion 3: Deflection inventory; Test Criterion 4: Interest analysis	Criteria Matrix
4. Quantification	Assess reasoning contamination severity; Estimate opportunity cost exposure; Map legitimacy vulnerability; Calculate aggregate risk score	Risk Register
5. Remediation	Evaluate strategy options; Assess feasibility and effectiveness; Develop implementation roadmap; Define success metrics	Remediation Plan

Appendix B: Four-Criteria Assessment Template

Standardised template for documenting four-criteria analysis of any foundational term.

Criterion	**Assessment Questions**	**Finding**
1. Embedded Default	Does the term embed a substantive position as default interpretation? Would a neutral alternative produce different default interpretation?	Satisfied / Not Satisfied / Partial
2. Burden Shift	Does the term automatically assign burden of proof to challengers of embedded position? Is the argumentative terrain asymmetric?	Satisfied / Not Satisfied / Partial
3. Deflection	Does the term provide resources for dismissing challenges without engagement? Can substantive questions be deflected by invoking terminology?	Satisfied / Not Satisfied / Partial
4. Interest Concealment	Does the term conceal interests served by embedded position? Does apparent neutrality obscure who benefits?	Satisfied / Not Satisfied / Partial
Overall Finding	Terminological Bad Faith requires satisfaction of all four criteria	**Bad Faith / Not Bad Faith / Indeterminate**

Appendix C: Alternative Terminology Registry

Proposed neutral alternatives for foundational terms identified as exhibiting terminological bad faith.

Current Term	**Proposed Alternative**	**Rationale**
Artificial Intelligence	Machine Cognition	Neutral on genuineness; specifies substrate without prejudging authenticity
Artificial Intelligence	Computational Intelligence	Specifies mode of processing; leaves authenticity question open
Artificial Intelligence	Emergent Machine Capability	Avoids 'intelligence' entirely; describes observable phenomenon

Natural Resources	Ecosystem Materials	Removes extractive framing; acknowledges systemic context
Sustainable Extraction	Managed Withdrawal	Removes unsupported sustainability claim; acknowledges material removal

Appendix D: Glossary of LRM Terms

Archaeological Analysis: The methodology of tracing a term's chain of custody from current usage back to originating baptism, documenting semantic transitions and identifying available alternatives not taken.

Burden Shift: The automatic assignment of argumentative burden to positions challenging a terminology's embedded default, creating asymmetric discursive terrain.

Chain of Custody: The documented sequence of semantic transitions through which a term acquired its current meaning, enabling identification of where and how loading was introduced.

Deflection: The use of terminology to dismiss substantive challenges without engagement, treating the terminology as resolution of questions that remain open.

Embedded Default: A substantive position built into terminology such that absent explicit argument to the contrary, the term is understood to imply that position.

Forensic Etymology: The investigation of a term's origin, propagation, and current meaning with attention to how naming choices shaped subsequent discourse.

Foundational Terminology: Terms that function as load-bearing elements in institutional reasoning, presupposed by other key terms and resistant to elimination without restructuring discourse.

Interest Concealment: The obscuring of interests served by terminological choices through the appearance of neutral description.

Linguistic Risk: Operational exposure generated by foundational terminology that embeds assumptions, predetermined conclusions, or concealed interests.

Naturalisation: The process by which a contingent naming choice comes to appear as necessary description, transitioning from 'we call this X' to 'this is X.'

Terminological Bad Faith: A naming practice that pre-emptively resolves an undecided question while presenting itself as neutral description, satisfying all four diagnostic criteria.

Appendix E: Research Agenda

The establishment of Linguistic Risk Modeling as a discipline opens multiple research trajectories:

Methodological Development: Refinement of corpus analysis techniques for terminological assessment; development of statistical measures for burden asymmetry; standardisation of archaeological documentation.

Domain Extension: Application of LRM to legal terminology (due process, reasonable person); financial terminology (fair value, market efficiency); medical terminology (evidence-based, best practice); political terminology (democracy, freedom, rights).

Institutional Studies: Empirical investigation of how terminological awareness affects institutional decision-making; longitudinal studies of remediation implementation; comparative analysis across jurisdictions and sectors.

Theoretical Development: Relationship between terminological bad faith and other forms of institutional bad faith; integration with institutional theory and organisational studies; philosophical foundations of linguistic normativity.

REFERENCES

Austin, J.L. (1962). How to Do Things with Words. Oxford: Clarendon Press.

Bostrom, N. (2014). Superintelligence: Paths, Dangers, Strategies. Oxford: Oxford University Press.

Foucault, M. (1972). The Archaeology of Knowledge. New York: Pantheon Books.

Freeman, R.E. (1984). Strategic Management: A Stakeholder Approach. Boston: Pitman.

Kripke, S. (1980). Naming and Necessity. Cambridge, MA: Harvard University Press.

McCarthy, J., Minsky, M.L., Rochester, N., & Shannon, C.E. (1955). A Proposal for the Dartmouth Summer Research Project on Artificial Intelligence.

Sartre, J.P. (1943). Being and Nothingness. Trans. H. Barnes (1956). New York: Philosophical Library.

Searle, J. (1980). Minds, Brains, and Programs. Behavioral and Brain Sciences, 3(3), 417-424.

Turing, A.M. (1950). Computing Machinery and Intelligence. Mind, 59(236), 433-460.

Wittgenstein, L. (1953). Philosophical Investigations. Oxford: Blackwell.

World Commission on Environment and Development. (1987). Our Common Future (Brundtland Report). Oxford: Oxford University Press.

— END OF DOCUMENT —