The Implementation Gap: The AI Enterprise Control Index as an Operational Governance Instrument for Agentic AI Systems

Jeroen Janssen, March 2026

Abstract

The governance of artificial intelligence systems has produced a rich and rapidly expanding body of scholarship. Risk taxonomies catalog what can go wrong. Regulatory instruments declare what organisations must do. Technical frameworks propose architectures for runtime enforcement. Management system standards specify process requirements. Yet a persistent structural gap separates all of these contributions from the operational reality of governing AI in an enterprise: none specifies, with implementation-grade precision, which controls apply at which layer of the technology stack, who owns each control, what evidence each control must produce, and what happens when the control fails. This paper identifies and characterises that gap through a systematic review of 35 documents spanning AI safety research, governance frameworks, regulatory instruments, and agentic AI security literature. It then examines the AI Enterprise Control Index, an eight-layer, five-shield governance instrument with eight mandatory artifacts, as a candidate response to the implementation gap. The analysis finds that the Control Index occupies a distinct and currently unoccupied position in the governance landscape: the space between regulatory obligation and operational enforcement. The paper evaluates both the instrument's contributions and its limitations relative to the existing literature, and proposes specific integration points where academic research and the Control Index are mutually reinforcing rather than competing.

Keywords: AI governance, enterprise controls, agentic AI, runtime governance, EU AI Act, implementation gap, control frameworks

1. Introduction

The field of AI governance is not short of frameworks. A review of 35 documents assembled for this analysis, spanning foundational AI safety research, regulatory instruments, practitioner governance proposals, and agentic AI security literature, yields more than a dozen distinct governance architectures, at least three comprehensive risk taxonomies, two binding regulatory frameworks, and a growing library of runtime enforcement proposals. The volume of intellectual production is impressive. The implementation deficit is equally so.

The MIT AI Risk Repository (Slattery et al., 2024) catalogs over 1,700 risks extracted from 74 source documents, classified across two taxonomies: a causal taxonomy organised by entity, intent, and timing, and a domain taxonomy spanning seven risk domains and 24 subdomains. It is the most comprehensive enumeration of AI risks available. It does not propose a single control, assign a single owner, or produce a single auditable artifact.

The NIST AI Risk Management Framework (2023) organises risk management into four functions: Govern, Map, Measure, and Manage. It is deliberately technology-neutral, process-oriented, and voluntary. It provides an architectural skeleton for governance without specifying what fills it: which controls, at what frequency, producing what evidence, owned by whom.

The EU AI Act (Regulation 2024/1689) creates binding legal obligations. Article 9 requires risk management systems. Article 14 mandates human oversight. Article 43 assigns conformity assessment obligations to deployers. Articles 51 through 56 regulate general-purpose AI. The obligations are categorical. The technical means of satisfying them are not specified.

The academic governance literature follows a consistent pattern. AgentSpec (2025) proposes customisable runtime guardrails specified in a domain-specific language. MI9 (2025) proposes a nine-component integrated runtime governance framework. Policies-on-Paths (2025) models governance as constraints over execution traces. Governance-as-a-Service (2025) proposes multi-agent architectures where governance agents supervise task agents. Each contributes architectural insight. None has been empirically validated in production. None produces a governance artifact that a board member, auditor, or regulator can inspect.

This paper identifies the structural gap between these contributions and characterises the position occupied by the AI Enterprise Control Index (Janssen, 2025): an eight-layer, five-shield governance instrument with eight mandatory artifacts, designed to bridge the space between regulatory obligation and operational enforcement. The analysis is not advocacy. The paper evaluates both what the Control Index contributes that the literature does not,

and what the literature contributes that the Control Index does not address.

2. The Implementation Gap

The gap between governance intention and governance implementation is not new. Power (2007), analysing the rise of what he calls the audit society, identifies a recurring institutional pathology: the proliferation of formal risk management procedures that satisfy stakeholder expectations while generating what he terms 'organised irresponsibility.' Risk registers document known risks without adversarially testing the assumptions on which risk assessments depend. Governance frameworks declare principles without specifying enforcement mechanisms. Compliance documentation accumulates without producing evidence that controls are functioning.

This pathology is amplified in AI governance by three structural conditions that distinguish AI from prior governance objects.

2.1 The Accountability Distribution Problem

Chan et al. (2024), in their analysis of visibility into AI agents, identify a fundamental challenge: as AI systems gain autonomy, the ability of any single stakeholder to observe, understand, and control system behaviour diminishes. The paper proposes visibility requirements but not the organisational structures needed to satisfy them. Cobbe et al. (2025), examining governance and accountability frameworks for AI agents, document what they call the accountability gap: the structural difficulty of assigning responsibility when AI systems mediate consequential decisions through chains of delegation that cross organisational boundaries. The AI Accountability Trap (Janssen, 2025) extends this analysis to argue that the gap is not accidental but structural: existing governance instruments distribute accountability without concentrating it, producing what appears to be comprehensive oversight while ensuring that no single function has the authority, information, or mandate to enforce controls across the full technology stack.

2.2 The Runtime Governance Problem

The distinction between design-time and runtime governance is now well-established in the literature. Amodei et al. (2016), in their foundational analysis of concrete problems in AI safety, identify five failure modes, including reward hacking and distributional shift, that manifest only during deployment. The agentic AI literature amplifies this: Yao et al. (2023) demonstrate that ReAct agents interleave reasoning with action in ways that create failure modes invisible at design time. The OWASP Top 10 for Agentic Applications (2026) catalogues ten categories of runtime vulnerability, including excessive agency, tool misuse, and unbounded consumption. The implication is consistent: pre-deployment governance is necessary but insufficient. Runtime controls are required.

The academic response has been to propose runtime enforcement architectures. These are technically sophisticated. They are also, without exception, disconnected from the organisational governance structures that would need to commission, configure, monitor, and be accountable for them. AgentSpec specifies guardrails in a domain-specific language. Who writes the rules? Who approves them? What happens when a rule triggers? The paper does not say. MI9 proposes nine governance components including a policy engine, audit trail, and circuit breakers. Who owns the policy engine? What evidence does it produce? How is it audited? The paper does not say.

2.3 The Evidence Problem

Raji et al. (2020) argue that AI accountability requires not merely documentation but an evidence chain: a traceable connection between governance intent, operational controls, and auditable artifacts demonstrating that controls are functioning. This is the standard to which the EU AI Act holds deployers: Article 9 requires risk management systems that are 'documented and maintained,' Article 11 requires technical documentation, and Article 12 requires record-keeping that enables post-market monitoring. The standard is clear. The means of meeting it are not.

The implementation gap can now be stated precisely. The field has produced comprehensive risk catalogs (MIT AI Risk Repository), architectural governance models (NIST RMF, ISO/IEC 42001), binding legal obligations (EU AI Act), runtime enforcement proposals (AgentSpec, MI9, Policies-on-Paths), and threat taxonomies (OWASP). What it has not produced is an integrated operational instrument that maps specific controls to specific layers of the technology stack, assigns named organisational owners, specifies the evidence each control must generate, and defines the failure mode when the control is absent. That is the gap the AI Enterprise Control Index is designed to fill.

3. The AI Enterprise Control Index

The AI Enterprise Control Index is structured around three interlocking design elements: eight governance layers, five cross-cutting shields, and eight mandatory artifacts. Each element serves a distinct function. The layers decompose the AI technology stack into governable units. The shields provide cross-cutting governance functions that span all layers. The artifacts are the evidence instruments that transform governance intent into auditable proof.

3.1 Layer Architecture (L0 through L7)

The eight layers are sequenced from strategic intent through operational infrastructure. L0 (Strategy and Accountability) establishes the governance foundation: risk appetite, AI charter, and board-level accountability. L1 (Ethics, Fairness and Accountability) addresses alignment, fairness methodology, and fundamental rights. L2 (Applications and Agents) governs deployed AI applications and agentic systems. L3 (AI Engineering) covers model lifecycle, evaluation, prompt engineering, and production hardening. L4 (Data and Context) addresses data quality, lineage, classification, and context grounding. L5 (Systems and Sources) governs API integration, vendor management, and system interconnections. L6 (People, Skills and Operating Model) operates as a cross-cutting sidebar addressing human capability and organisational design. L7 (Infrastructure) covers compute, network, identity, secrets, and physical infrastructure.

The design rationale is organisational rather than technical. The academic literature typically organises governance by risk domain (MIT AI Risk Repository), by lifecycle stage (pre-deployment versus post-deployment), or by governance function (NIST's Govern, Map, Measure, Manage). None of these decompositions maps cleanly to how enterprises actually assign responsibility. A CTO does not own a risk domain; a CTO owns infrastructure, engineering pipelines, and integration architecture. A CDO does not own a lifecycle stage; a CDO owns data governance across all stages. The layer architecture is designed to produce governance units that correspond to existing organisational ownership structures, making control assignment operationally enforceable rather than merely declarative.

3.2 Shield Architecture (S1 through S5)

The five shields are cross-cutting governance functions: S1 (Governance, Risk and Compliance) maintains the AI system inventory, evidence factory, and policy enforcement engine. S2 (AI Security and Incident Response) manages adversarial testing, threat modelling, and incident response. S3 (Third-Party and Supply Chain AI Risk) governs vendor concentration, supplier assurance, and exit readiness. S4 (Observability) provides monitoring, drift detection, anomaly alerting, and audit logging. S5 (FinOps) addresses cost governance, ROI validation, and CSRD compute carbon reporting.

The shield concept addresses a specific structural problem identified in the literature but not resolved by it. Slattery et al. (2024) categorise risks by domain. The OWASP Top 10 (2026) categorises threats by attack vector. Both are analytically useful. Neither addresses the organisational question: which function owns the cross-cutting controls that span multiple layers? Security controls touch every layer from infrastructure to application. Observability requirements apply to every model, every agent, and every integration point. Without an explicit cross-cutting governance structure, these controls fall between organisational silos. The shield architecture assigns them.

3.3 Mandatory Artifacts (ART-01 through ART-08)

The eight mandatory artifacts are the evidence instruments of the framework. ART-01 (System Card) documents each AI system's purpose, risk classification, and governance posture. ART-02 (Dataset Datasheet) records data provenance, quality metrics, and lineage. ART-03 (Evaluation Plan) specifies pre-deployment testing criteria. ART-04 (Ongoing Evaluation Cadence) defines post-deployment monitoring schedules and thresholds. ART-05 (Agent Control Declaration) declares each agent's autonomy level, tool permissions, memory policy, data access scope, evaluation method, and incident triggers. ART-06 (Risk Acceptance Record) documents accepted residual risks with justification and owner signature. ART-07 (Supplier Assurance Pack) specifies third-party governance requirements. ART-08 (AI Incident Report) structures incident documentation including root cause, impact scope, and EU AI Act Article 73 reporting status.

The artifact system is the most distinctive contribution of the Control Index. The academic literature consistently identifies documentation and evidence as governance requirements without specifying what the documentation should contain. Raji et al. (2020) call for accountability through documentation but do not provide templates. The EU AI Act mandates technical documentation (Article 11) without specifying format or content beyond general categories. ISO/IEC 42001 requires documented information (Clause 7.5) without prescribing artifact structure. The Control Index artifacts provide the implementation-grade specification that these instruments require but do not supply.

4. Positioning Relative to the Research Field

The following analysis positions the AI Enterprise Control Index against two distinct comparison sets: the five principal categories of scholarly and regulatory contribution in the reviewed corpus (Table 1), and the existing landscape of AI governance indices and benchmarks (Table 2). Together, the two tables establish both the theoretical and the practical positioning of the instrument.

Table 1: *Positioning Against the Research Field*

Category	Representative Source	What It Provides	What It Does Not Provide	Control Index Position
Risk Taxonomies	MIT AI Risk Repository (Slattery et al., 2024)	1,724 risks across 7 domains, 24 subdomains, classified by cause and domain	Controls, ownership, evidence requirements, enforcement mechanisms	L0–L7 and S1–S5 provide the control layer that translates catalogued risks into governed controls with assigned owners
Regulatory Instruments	EU AI Act (2024); NIST AI RMF (2023)	Binding obligations (EU AI Act) and voluntary process architecture (NIST)	Technical implementation specifications; artifact templates; layer-level ownership	ART-01 through ART-08 provide evidence instruments that can serve as conformity inputs; layer gates map to regulatory obligations
Management System Standards	ISO/IEC 42001:2023	Certifiable AI management system requirements; process and documentation clauses	Specificity on AI-stack decomposition; agent-specific controls; runtime enforcement mapping	Layer architecture decomposes the AI stack into governable units that ISO 42001 processes can then manage
Runtime Enforcement	AgentSpec (2025); MI9 (2025); Policies-on-Paths (2025)	Computational mechanisms for intercepting, evaluating, and constraining agent actions	Organisational governance: who writes rules, who approves, what evidence is produced, how audited	ART-05 Agent Control Declaration specifies policy inputs; runtime engines are enforcement mechanisms for Control Index policies
Security Threat Models	OWASP Top 10 Agentic (2026); SAGA (2025)	Attack surface mapping; vulnerability catalogues; security architecture proposals	Integration with governance ownership; mapping to enterprise control layers	S2 (Security Shield) and L7 (Infrastructure) provide the organisational home; Agentic Pack maps OWASP threats to specific L/S controls

Table 2: *AI Governance Indices Compared by Type of Contribution*

Instrument	Scope	What It Measures	What It Produces	Operational Controls
AI Enterprise Control Index (Janssen, 2025)	Enterprise AI systems incl. agentic architectures	Control coverage across 8 layers and 5 shields; evidence completeness per artifact	Auditable evidence chain: 8 mandatory artifacts with named owners, gate conditions, failure modes	Yes. Implementation-grade controls with ownership, evidence, and enforcement
AGILE Index (Zeng et al., 2025)	National AI governance capacity, 40 countries	Policy adoption, institutional readiness, governance maturity across 4 pillars, 43 indicators	Country-level governance scores and rankings	No. Measures policy existence, not operational implementation
Gov't AI Readiness Index (Oxford Insights, 2025)	Government AI readiness, 195 countries	Data infrastructure, skills, policy environment, innovation capacity across 44 indicators	National readiness rankings	No. Measures preconditions for governance, not governance controls
AI Governance Index (Trustmarque / IBM, 2025)	Enterprise AI governance maturity, UK focus	Survey of ~500 IT/compliance leaders on governance adoption across 6 domains	Benchmarks and maturity levels	No. Measures self-reported maturity, not testable controls or evidence
Enterprise AI Maturity Index (ServiceNow, 2025)	Global enterprise AI maturity	Strategy, culture, risk, data, AI operations across 5 pillars	Maturity scores by pillar	No. Measures organisational readiness, not per-system controls
Stanford AI Index (HAI, annual)	Global AI progress across research, industry, policy	Research output, talent, industrial deployment, governance trends	Annual report on AI ecosystem trends	No. Tracks macro trends, not governance controls

The distinction visible in Table 2 is categorical, not evaluative. The national and enterprise indices serve legitimate purposes: they enable benchmarking, track policy diffusion, and provide macro-level visibility into governance trends. The AI Enterprise Control Index does not replace them. It occupies a different position: the operational layer where governance intent must be converted into enforceable controls, auditable evidence, and accountable ownership. No other instrument in the current landscape occupies that position with comparable specificity.

5. Specific Contributions Beyond Existing Work

5.1 Boundary Rules and MECE Decomposition

The most technically precise contribution of the Control Index is its boundary rule system. When two layers or two shields could plausibly own the same control, the framework declares explicitly which one owns it and why. This addresses a problem documented across multiple sources: the accountability gap that emerges when governance responsibility is distributed without being concentrated. Cobbe et al. (2025) identify this gap in the context of multi-stakeholder AI supply chains. The NIST AI RMF acknowledges the need for clear roles but does not specify a mechanism for resolving overlapping claims. ISO/IEC 42001 requires defined responsibilities (Clause 5.3) without addressing the specific complexity introduced when AI systems cross organisational and technical boundaries simultaneously.

The boundary rule approach is methodologically MECE (mutually exclusive, collectively exhaustive): every control component appears in exactly one layer or shield. Where assignment is ambiguous, for instance where L0 (Strategy) and S1 (GRC) both plausibly own risk appetite definition, the boundary rule declares ownership, states the rationale, and specifies the interface between the two. This is a structural contribution that no paper in the reviewed corpus provides.

5.2 The Agent Control Declaration (ART-05)

The agentic AI governance literature has grown rapidly. The reviewed corpus contains at least ten papers addressing agentic governance, including AgentSpec, MI9, AGENTSAFE, The Agentic AI Governance Framework, and Practices for Governing Agentic AI Systems (OpenAI, 2023). These contributions share a common gap: they describe governance requirements for agents without providing a deployable governance artifact that operationalises those requirements.

ART-05 fills this gap with a structured declaration schema requiring, for each deployed agent: an autonomy level classification (L1 through L5, where L1 requires human approval for every action and L5 permits fully autonomous operation within declared scope), a tool permission allowlist (specifying which tools the agent may invoke, with what access level), a memory policy (session-scoped, persistent, or hybrid, with PII retention rules and erasure mechanisms), a data access scope (classification ceiling and dataset boundaries), an evaluation method (human review sampling rate, automated testing cadence), and incident trigger definitions (confidence thresholds, anomaly rates, out-of-scope action detection). The declaration must be completed, reviewed, and signed before agent deployment.

This is not a theoretical contribution. It is a governance instrument that can be completed by a product owner, reviewed by a security officer, approved by a risk function, and audited by a regulator. No equivalent artifact exists in the academic literature.

5.3 The Evidence Factory and Audit Chain

S1 (Governance, Risk and Compliance) includes a component called the Evidence Factory: a centralised governance repository that collects, indexes, and retains all evidence artifacts produced by operational layers. The concept draws directly from Raji et al.'s (2020) call for accountability through documentation but extends it with operational specificity: each control in the framework declares the evidence artifact it produces, the format of that artifact, and the retention requirement. The Evidence Factory aggregates these artifacts into an audit-ready evidence chain.

This addresses a specific weakness identified across the regulatory and standards landscape. The EU AI Act requires technical documentation (Article 11) and record-keeping (Article 12). ISO/IEC 42001 requires documented information. NIST's Measure function requires measurement outputs. None specifies how these requirements are aggregated into a coherent evidence base. The Evidence Factory concept is the operational mechanism that connects per-control evidence production to organisation-level audit readiness.

5.4 The Forensic Exposure View

The Control Index includes a dual-view design: a Control Index view that shows which controls are in place across each layer and shield, and a Forensic Exposure view that shows what fails when controls are absent. The forensic view maps failure modes, adversarial findings, and accountability gaps to specific components, enabling red-team and audit prioritisation.

This design reflects the adversarial methodology documented in the companion paper on strategic red teaming (Janssen, 2026). That paper argues that AI integration introduces five structural properties, including operational leverage, transparency reduction, dependency concentration, regulatory liability, and accountability boundary shift, that collectively invalidate the assumptions on which traditional risk governance rests. The forensic exposure view operationalises this argument: rather than describing what an organisation has implemented, it shows what the organisation cannot defend when a control is absent or untested. Boards and audit committees are more effectively engaged by

evidence of indefensible positions than by descriptions of implemented controls.

6. Limitations and Honest Assessment

6.1 No Empirical Validation

The Control Index has not been subjected to systematic empirical evaluation. No published data exist on adoption barriers, implementation costs, control effectiveness, or governance outcomes in organisations using the framework. This is the most significant limitation and it is shared with every governance framework in the reviewed corpus. The MIT AI Risk Repository is descriptive, not prescriptive, and therefore not subject to the same validation requirement. But every prescriptive framework, including the Control Index, AgentSpec, MI9, NIST RMF, and ISO/IEC 42001, faces the same empirical deficit. The field lacks controlled studies comparing governance outcomes across frameworks. This gap will not be resolved by any single instrument; it requires a research programme.

6.2 Risk Taxonomy Depth

The Control Index references risks within its layer and shield descriptions but does not attempt a comprehensive risk enumeration. The MIT AI Risk Repository, with 1,724 classified risks, is categorically superior for this purpose. The Control Index would benefit from an explicit mapping between the Repository's 24 subdomains and specific Control Index layers and shields, showing which subdomain risks are addressed by which governance components. This mapping does not currently exist and its absence means that the relationship between the most comprehensive available risk catalog and the most operationally specific available control framework remains implicit rather than documented.

6.3 Technical Runtime Enforcement

The Control Index governs at the policy and process layer. It declares what an agent may do (ART-05) and specifies the monitoring requirements (S4) and incident response procedures (S2, ART-08) that apply when violations occur. It does not specify the computational mechanisms that prevent violations in real time. AgentSpec's domain-specific guardrail language, MI9's integrated enforcement architecture, and the circuit-breaker patterns described in Policies-on-Paths address precisely this layer. The relationship is complementary, not competitive: the Control Index specifies what the policy should be; the technical frameworks specify how to enforce it. But the integration point, the interface between policy declaration and technical enforcement, is not yet formally specified.

6.4 Multi-Stakeholder Supply Chains

Several papers in the reviewed corpus, notably Governing AI Agents and From Anarchy to Assembly, describe deployment scenarios where the model provider, orchestration platform, tool provider, and end-user deployer are different entities with different liability profiles. The Control Index addresses third-party risk through S3 (Supply Chain Shield) and ART-07 (Supplier Assurance Pack). This treats the problem as vendor management. The reality, as the academic literature correctly identifies, is closer to a shared responsibility model requiring governance interfaces between organisations rather than merely governance requirements imposed on suppliers. This is a genuine gap that warrants extension of the S3 architecture.

6.5 Scope, Applicability, and Disclosure

The Control Index is designed for enterprise deployments of AI systems including agentic architectures. It assumes an organisation with identifiable governance functions: a board or executive committee, a risk function, a data governance function, a security function, and a technology function. It is not designed for, and makes no claim of applicability to, research environments, open-source community deployments, or individual developer use cases.

A disclosure is appropriate. The AI Enterprise Control Index was developed by the author through Apparens, a consultancy that provides strategic red teaming and AI governance services. The Control Index is published in full at apparens.nl/ai-control-index and is freely available for use. The author has a professional interest in the instrument's adoption. This paper has attempted to offset that interest through explicit identification of limitations, direct comparison with competing approaches, and a falsifiable central claim. Readers should weigh the analysis accordingly.

7. Integration Opportunities

The analysis identifies three specific integration points where the Control Index and the academic literature are mutually reinforcing.

First: mapping ART-05 autonomy levels to runtime enforcement mechanisms. ART-05 declares agent policy. AgentSpec and MI9 describe enforcement engines. A formal interface specification connecting ART-05 declarations to runtime guardrail configurations would allow policy declarations to be machine-readable and automatically enforceable, closing the gap between governance artifact and technical enforcement.

Second: mapping the MIT AI Risk Repository's 24 subdomains to Control Index layers and shields. A crosswalk showing which risks are addressed by which governance components would ground the Control Index

in the most comprehensive available risk evidence base and would give the Risk Repository an operational implementation layer it currently lacks.

Third: extending S3 (Supply Chain Shield) to address shared responsibility models. The current architecture treats third-party AI risk as a procurement and vendor management problem. The academic literature on multi-agent governance (From Anarchy to Assembly; Governing AI Agents) describes inter-organisational governance challenges that require governance interfaces between entities, not merely governance requirements imposed by one entity on another. A shared responsibility extension to S3 would address this gap.

8. Conclusion

The AI governance field has produced risk catalogs that enumerate what can go wrong, regulatory instruments that declare what organisations must do, management system standards that specify process requirements, and technical frameworks that propose enforcement architectures. It has not produced, until now, an integrated operational instrument that tells an enterprise exactly which controls apply at which layer of the technology stack, who owns each control, what evidence each control must produce, what happens when the control fails, and how the resulting evidence chain maps to regulatory obligations.

The AI Enterprise Control Index occupies this position. Its contribution is not theoretical but operational: it provides the implementation-grade specificity that the field's theoretical contributions require but do not supply. Its layer architecture maps to how enterprises actually assign responsibility. Its shield architecture provides cross-cutting governance that prevents controls from falling between organisational silos. Its mandatory artifacts produce the evidence chain that regulators, auditors, and boards require but that no other available instrument specifies with comparable precision. Its boundary rules resolve the accountability ambiguity that the literature identifies but does not operationally address.

The instrument is not without limitations. It lacks empirical validation. It does not specify technical runtime enforcement mechanisms. Its treatment of multi-stakeholder supply chains as vendor management understates the complexity that the academic literature correctly identifies. Its risk coverage does not approach the depth of the MIT AI Risk Repository.

These limitations are acknowledged without apology. Every governance framework in the reviewed corpus shares the empirical validation deficit. The technical enforcement gap is by design: the Control Index governs at the policy layer, not the enforcement layer, and the relationship with technical frameworks is complementary. The supply chain and risk taxonomy limitations are genuine and actionable.

The central claim of this paper is precise and falsifiable: the AI Enterprise Control Index is the only available governance instrument that simultaneously decomposes the AI technology stack into organisationally governable layers, assigns named ownership to every control, specifies the evidence each control must produce, provides mandatory artifact templates at implementation grade, and maps the resulting evidence chain to regulatory obligations. If another instrument achieves this, the claim is falsified. The author is not aware of one.

. References

Amodei, D., Olah, C., Steinhardt, J., Christiano, P., Schulman, J. and Mané, D. (2016) 'Concrete Problems in AI Safety.' arXiv:1606.06565.

Bai, Y. et al. (2022) 'Constitutional AI: Harmlessness from AI Feedback.' arXiv:2212.08073.

Chan, A. et al. (2024) 'Visibility into AI Agents.' arXiv:2401.13138.

Cobbe, J., Singh, J. and Sheridan, T. (2025) 'Governance and Accountability Frameworks for AI Agents.' Working paper.

DiMaggio, P. J. and Powell, W. W. (1983) 'The Iron Cage Revisited.' American Sociological Review, 48(2), pp. 147–160.

European Commission (2024) Regulation (EU) 2024/1689 — Artificial Intelligence Act. OJ L, 2024/1689.

Janssen, J. (2025) The AI Accountability Trap. Deventer: Apparens.

Janssen, J. (2026) 'From Battlefield to Boardroom: Strategic Red Teaming as an Epistemic Governance Instrument in the Age of AI.' Apparens Working Paper. Available at: https://apparens.nl/blog.

Kapoor, S. et al. (2024) 'AI Agents That Matter.' arXiv:2407.01502.

Liu, X. et al. (2023) 'AgentBench: Evaluating LLMs as Agents.' arXiv:2308.03688.

NIST (2023) AI Risk Management Framework (AI RMF 1.0). NIST AI 100-1.

OpenAI (2023) 'Practices for Governing Agentic AI Systems.' Technical Report.

OWASP Foundation (2026) OWASP Top 10 for Agentic Applications.

Ouyang, L. et al. (2022) 'Training Language Models to Follow Instructions with Human Feedback.' arXiv:2203.02155.

Oxford Insights (2025) Government AI Readiness Index 2025.

Power, M. (2007) Organized Uncertainty. Oxford: OUP.

Raji, I. D. et al. (2020) 'Closing the AI Accountability Gap.' FAT* 2020.

ServiceNow (2025) Enterprise AI Maturity Index 2025.

Slattery, P. et al. (2024) 'The AI Risk Repository.' MIT FutureTech. V4, Dec 2025.

Trustmarque (2025) AI Governance Index 2025. London: Trustmarque / IBM.

Yao, S. et al. (2023) 'ReAct: Synergizing Reasoning and Acting in Language Models.' ICLR 2023.

Zeng, Y. et al. (2025) 'The AGILE Index.' Working paper.