International Journal of Future Engineering Innovations

From Prediction to Accountability: A Responsible AI Maturity Framework for Healthcare, Finance, Energy, and Public Supply Chains

Christopher D Vance [1*], Isaac M Kim [2]
[1, 2] Gatton College of Business and Economics, University of Kentucky, Kentucky, USA

* Corresponding Author: **Christopher D Vance**

Article Info

ISSN (online): 3049-1215
Volume: 02
Issue: 02
March-April 2025
Received: 17-02-2025
Accepted: 18-03-2025
Page No: 115-128

Abstract
Organizations are adopting artificial intelligence faster than they are building the institutions required to govern it. Existing maturity models usually privilege data, infrastructure, and model deployment, while treating fairness, explanation, human oversight, cybersecurity, sustainability, and recourse as adjacent concerns. This study develops and provides exploratory validation of the Responsible AI Maturity Framework (RAIMF), a six-level organizational maturity model and a composite Responsible AI Maturity Index (RAIMI) for healthcare, finance, energy, and public supply chains. The empirical design combines structured content analysis of 40 governance instruments available by December 31, 2024 with a secondary reanalysis of industry maturity distributions from a 2024 survey of 1,000 organizations. Ten dimensions were coded on a five-point evidence scale. The instrument showed acceptable internal consistency (Cronbach alpha = 0.750), marginal-to-adequate factorability (KMO = 0.608; Bartlett chi-square = 234.66, $p < .001$), and a three-component structure explaining 76.2 percent of variance. Governance assurance and human-centered transparency achieved composite reliability above 0.85, average variance extracted above 0.60, and HTMT of 0.360. Across 19 industries, organizational maturity did not predict operational maturity (beta = 0.085, $p = .740$), revealing a substantial planning-execution gap. Sector RAIMI scores ranged from 59.3 to 61.7 after an execution penalty, placing all four critical sectors at the Governed AI level rather than Responsible or Institutionally Accountable AI. Public supply chains showed the strongest conversion of organizational intent into operational controls, while sustainability remained the least developed dimension across the standards corpus. RAIMF contributes a measurable account of Responsible AI as an institutional capability, not merely a technical property, and provides a reproducible benchmarking method for managers, regulators, auditors, and public administrators.

DOI: https://doi.org/10.54660/IJFEI.2025.2.2.115-128

Keywords: Responsible AI, AI governance, maturity model, institutional accountability, explainable AI, critical infrastructure; public supply chains, composite index

1. Introduction

Artificial intelligence programs often mature in an uneven sequence. Prediction capability improves first. Data pipelines, model libraries, cloud infrastructure, and deployment routines follow. Governance usually arrives later, often after a regulatory inquiry, a security incident, a disputed decision, or a public failure. This sequence matters because technical sophistication can increase institutional exposure faster than organizations develop the ability to explain, challenge, correct, and learn from automated decisions. The result is a form of asymmetric maturity: an organization may operate advanced models while remaining institutionally immature.

The distinction is visible across critical infrastructure. In healthcare, an accurate model can still create harm if data provenance

is weak, clinical users over-rely on its output, subgroup performance is not monitored, or patients lack a route to contest a decision. WHO guidance therefore places autonomy, safety, transparency, accountability, equity, and sustainability within the governance of health AI rather than treating them as optional model attributes (World Health Organization, 2021, 2023). Work on healthcare infrastructure security likewise shows that machine learning value depends on privacy, adversarial resilience, and protection of interconnected clinical systems (Hasan *et al*., 2022). Supply availability adds another layer. Healthcare supply-chain resilience requires visibility, accountable sourcing, demand sensing, and continuity planning, so an AI system that improves forecasts but weakens traceability may shift rather than reduce risk (Rasel *et al*., 2022). Augmented-intelligence research likewise emphasizes that healthcare AI should combine machine learning with professional expertise and workflow redesign rather than remove meaningful human judgment (Kamruzzaman *et al*., 2023).

Finance presents a similar problem under stronger model-risk traditions. Fraud detection, credit underwriting, and market surveillance can produce operational value, but they also create obligations concerning adverse action, fair lending, validation independence, cybersecurity, and auditability. Hasan *et al*. (2023) connect AI-driven fraud analytics with zero-trust security and model-risk management, while Rasel *et al*. (2023) show why mortgage prediction should incorporate multimodal evidence without allowing complexity to displace inclusive lending and decision traceability. These cases indicate that performance and accountability are complementary design objectives, not substitutes.

Energy and public supply chains extend the problem beyond individual decisions to system resilience and public value. Renewable-energy forecasting can support dispatch, carbon reduction, and investment planning, yet model error, cyber compromise, data asymmetry, and opaque optimization can affect communities and essential services. Arman *et al*. (2024a) emphasize the role of big-data forecasting in the clean-energy transition. Arman *et al*. (2024b) similarly connect machine learning, resource efficiency, and waste reduction. These contributions reveal an important governance question: whether environmental value is measured and audited with the same seriousness as predictive accuracy. Public supply chains add procurement authority, vendor dependence, emergency operations, and distributive consequences. Here, accountability must remain visible across the entire chain from problem definition and tendering to deployment, monitoring, appeal, and contract renewal. Energy-burden-aware retrofit finance further illustrates how decarbonization objectives intersect with affordability, fairness, explanation, and underwriting governance (Rabbani *et al*., 2024). Trustworthy-AI research on federal payment infrastructure similarly emphasizes traceability, synthetic-identity controls, human review, and public-benefit integrity (Rahat *et al*., 2024).

The Responsible AI literature has established principles of fairness, transparency, privacy, robustness, human agency, accountability, and social benefit (Floridi *et al*., 2018; Jobin *et al*., 2019; OECD, 2019). It has also documented the distance between principles and practice. Ethical guidelines can generate legitimacy without changing operating routines (Hagendorff, 2020; Mittelstadt, 2019). Model cards, datasheets, counterfactual explanations, local explanations, and algorithmic audits provide important technical and documentary mechanisms (Gebru *et al*., 2021; Mitchell *et al*., 2019; Raji *et al*., 2020; Ribeiro *et al*., 2016; Wachter *et al*., 2018), but these mechanisms do not automatically determine who owns risk, who can stop deployment, or how an affected person obtains remedy. Institutional accountability requires both information and authority.

Existing AI maturity models rarely solve this problem. Many evaluate strategy, data, infrastructure, talent, and scaling capability. Responsible AI models have begun to include governance and operational mitigation, but standardized measurement remains limited. Reuel *et al*. (2024) provide an important global maturity survey and show that organizational structures are more advanced than operational controls. Dotan *et al*. (2024) map maturity to the NIST AI Risk Management Framework and caution against superficial implementation. Yet a cross-sector model that explicitly links technical safeguards, organizational governance, institutional recourse, cybersecurity, and sustainability remains underdeveloped.

This paper addresses that gap through the Responsible AI Maturity Framework, abbreviated RAIMF. RAIMF defines six levels: Reactive AI, Operational AI, Explainable AI, Governed AI, Responsible AI, and Institutionally Accountable AI. It operationalizes ten dimensions and converts them into the Responsible AI Maturity Index, abbreviated RAIMI. The study addresses six research questions. RQ1 asks which organizational capabilities define Responsible AI maturity. RQ2 asks whether Responsible AI maturity can be measured quantitatively as a reliable multidimensional construct. RQ3 asks which maturity dimensions contribute most strongly to institutional accountability. RQ4 asks how maturity differs across healthcare, finance, energy, and public supply chains. RQ5 asks whether organizational maturity translates into operational assurance and accountable outcomes. RQ6 asks how organizations can benchmark readiness without mistaking policy adoption for implementation.

The contribution is threefold. First, the paper integrates institutional theory, socio-technical systems theory, information-processing theory, the resource-based view, dynamic capabilities, organizational learning, and stakeholder theory into a single account of Responsible AI maturity. Second, it provides exploratory evidence on the content and measurement structure of RAIMF using 40 authoritative governance instruments published by the end of 2024. Third, it externally benchmarks the framework using 19-industry maturity distributions from a large global survey and develops sector scores that penalize gaps between organizational intent and operational execution. The analysis produces a counterintuitive result: mature governance language is widespread, but cross-industry organizational maturity has almost no association with operational maturity. Responsible AI is therefore not a natural consequence of becoming more sophisticated at AI. It is a separate institutional capability that must be deliberately built.

2 Literature Review

Responsible AI research has developed along three partially separated streams. The first is normative. It identifies principles that should govern AI, including beneficence, non-maleficence, autonomy, justice, explicability, privacy, robustness, and accountability. Comparative reviews show considerable convergence at the principle level, although

definitions and enforcement mechanisms differ (Floridi *et al*., 2018; Jobin *et al*., 2019). The OECD AI Principles, UNESCO Recommendation, European Commission ethics guidance, and the 2024 EU AI Act reinforce a risk-based approach in which trustworthy behavior must be demonstrated across the system life cycle (European Commission High-Level Expert Group on AI, 2019; European Union, 2024; OECD, 2019; UNESCO, 2021).

The second stream is technical and documentary. Explainable AI methods such as LIME, SHAP, and counterfactual explanations seek to make model behavior interpretable for developers, decision makers, or affected users (Lundberg & Lee, 2017; Ribeiro *et al*., 2016; Wachter *et al*., 2018). Model cards and datasheets document intended use, training data, evaluation conditions, limitations, and ethical considerations (Gebru *et al*., 2021; Mitchell *et al*., 2019). Fairness research measures group disparities and identifies structural limits to purely technical definitions of equity (Barocas *et al*., 2023; Selbst *et al*., 2019). These methods are necessary, but they do not establish institutional accountability by themselves. An explanation can be technically correct yet practically useless if no decision maker is required to respond to it.

The third stream concerns governance and organizational practice. NIST AI RMF organizes work around Govern, Map, Measure, and Manage, while ISO/IEC 42001 treats AI governance as a management system requiring policy, roles, risk processes, monitoring, and continual improvement (International Organization for Standardization, 2023a; National Institute of Standards and Technology, 2023a). The EU AI Act adds legally enforceable obligations for high-risk systems, including risk management, data governance, documentation, record keeping, transparency, human oversight, accuracy, robustness, and cybersecurity (European Union, 2024). The United States Executive Order 14110 and OMB Memorandum M-24-10 extend governance expectations to federal agencies, with emphasis on rights-impacting and safety-impacting uses (Office of Management and Budget, 2024; The White House, 2023).

Accountability sits at the intersection of these streams. Kroll *et al*. (2017) argue that accountable algorithms require mechanisms for review, verification, and procedural challenge. Raji *et al*. (2020) frame internal algorithmic auditing as an end-to-end process that connects development decisions with organizational responsibility. Wieringa (2020) distinguishes accountability to different forums and asks who must explain what, to whom, and with what consequences. This institutional view is stronger than transparency alone. Transparency supplies information. Accountability connects information to answerability, authority, sanction, correction, and learning.

Maturity models can provide the missing developmental logic. A maturity model describes how capabilities move from ad hoc practice to repeatable, managed, measured, and continuously improved routines. Such models are common in software engineering, project management, and business-process management (Wendler, 2012). AI maturity models, however, have often been adoption models. They assess whether an organization possesses data, talent, infrastructure, strategy, and deployment capacity. Those dimensions explain whether AI can scale, not whether it can be governed responsibly.

Recent Responsible AI maturity work narrows this gap. Reuel *et al*. (2024) distinguish organizational and operational maturity in a survey of 1,000 organizations. Their results reveal that only a small share of organizations reach optimized operational maturity, even when organizational governance is more advanced. Dotan *et al*. (2024) develop a NIST-based maturity approach and warn that selective implementation can become a veneer of legitimacy. The 2024 Stanford AI Index similarly reports weak standardization in Responsible AI evaluation (Maslej *et al*., 2024). These studies establish the importance of implementation but leave three unresolved issues. They do not fully integrate sustainability and cybersecurity with institutional recourse; they provide limited cross-sector calibration for critical infrastructure; and they do not offer a composite score that explicitly penalizes the policy-execution gap.

Sector studies indicate why those omissions matter. Healthcare models operate within clinical workflows, privacy rules, professional duties, and patient safety systems. Finance operates within validation, consumer protection, prudential supervision, and cybersecurity regimes. Energy combines operational technology, critical-infrastructure security, environmental objectives, and reliability obligations. Public supply chains combine procurement law, public value, emergency response, and vendor governance. A general maturity model must therefore preserve common dimensions while allowing sector-specific evidence and risk weights.

The literature also reveals a sustainability deficit. Environmental and social well-being appear in high-level principles, but they are less often translated into measurable organizational controls. Energy forecasting and waste-reduction research show that AI can create environmental value (Arman *et al*., 2024a, 2024b). Yet energy use, carbon intensity, rebound effects, model complexity, and distributional impacts are seldom incorporated into AI governance scorecards. RAIMF treats sustainability as a maturity dimension rather than an external corporate responsibility metric.

3 Theoretical Background

RAIMF rests on an integrated theoretical model. Institutional theory explains why organizations adopt governance structures in response to coercive, normative, and mimetic pressures (DiMaggio & Powell, 1983; Meyer & Rowan, 1977). Regulation, professional standards, investor expectations, and peer practices encourage organizations to create AI principles, committees, and policies. Yet formal structures can become decoupled from operational work. This decoupling is central to Responsible AI: an organization may possess an ethics charter without changing data collection, model validation, procurement, or incident response. Institutional accountability is achieved only when formal commitments are coupled to routines, evidence, and enforceable decision rights.

Socio-technical systems theory explains why responsibility cannot be localized within the model. AI outcomes emerge from interactions among data, software, users, incentives, workflows, vendors, and institutional rules (Trist & Bamforth, 1951). Orlikowski's (1992) account of technology and organization further suggests that technical systems both shape and are shaped by organizational practice. RAIMF therefore measures human oversight, stakeholder engagement, and organizational learning alongside technical safeguards. This complementarity is consistent with augmented-intelligence approaches that position human expertise and business-process design as integral parts of an

AI system (Kamruzzaman *et al*., 2023).

Information-processing theory adds a contingency argument. As uncertainty, interdependence, and task complexity increase, organizations require greater capacity to collect, interpret, and communicate information (Galbraith, 1974). High-risk AI creates uncertainty about data shifts, model behavior, misuse, legal exposure, and affected populations. Documentation, monitoring, audit trails, escalation channels, and cross-functional governance increase information-processing capacity. They are not administrative overhead. They are mechanisms for reducing equivocality and coordinating action under risk.

The resource-based view and dynamic-capabilities theory explain why Responsible AI can become a source of durable organizational value. Policies are easy to imitate. Integrated capabilities involving specialized talent, governance routines, validation infrastructure, trusted stakeholder relationships, and learning systems are harder to replicate (Barney, 1991; Teece *et al*., 1997). Responsible AI maturity should therefore improve deployment quality, reduce rework, shorten regulatory response time, and protect legitimacy. Its value arises from coordinated routines rather than a single tool.

Organizational learning theory explains movement between maturity levels. March (1991) distinguishes exploration from exploitation, while Argyris and Schon (1978) distinguish single-loop correction from deeper revision of governing assumptions. A mature AI program does more than correct model drift. It learns from near misses, appeals, subgroup failures, red-team findings, and supplier incidents, then changes objectives, controls, and risk appetite. Continuous auditing and post-incident review therefore belong within maturity rather than after it.

Stakeholder theory defines the relevant accountability forum. AI systems affect patients, borrowers, employees, suppliers, citizens, communities, regulators, and future generations. Their interests cannot be reduced to shareholder return or model accuracy (Freeman, 1984). In RAIMF, stakeholder engagement includes participation in problem definition, communication of system limits, accessible explanations, channels for appeal, and evidence that complaints alter organizational practice.

Together, these theories support eight hypotheses. H1 proposes that the ten RAIMF dimensions form a reliable multidimensional construct. H2 proposes that operational specificity is positively associated with accountability coverage in governance instruments. H3 proposes that human-centered transparency is positively associated with accountability. H4 proposes that sector-specific instruments emphasize compliance and cybersecurity more than sustainability. H5 proposes that organizational maturity positively predicts operational maturity across industries. H6 proposes that critical regulated sectors have higher operational maturity than other sectors. H7 proposes that the four critical sectors differ in their conversion of organizational intent into operational controls. H8 proposes that sector RAIMI rankings remain stable under alternative weighting and penalty specifications.

4. Framework Development

RAIMF was developed through theoretical synthesis, standards crosswalking, and empirical refinement. The unit of analysis is the organizational capability to govern AI across its life cycle, including systems developed internally, modified from third-party models, or procured as services. The framework does not assume that every organization must build advanced models. It asks whether the organization can identify AI, classify risk, assign responsibility, generate evidence, intervene when performance changes, and provide remedy when decisions cause harm.

The ten dimensions are data governance; explainability; fairness; human oversight; transparency; cybersecurity; accountability; compliance and risk management; sustainability; and learning and engagement. Data governance covers provenance, quality, access, privacy, retention, and representativeness. Explainability covers methods and communication appropriate to technical users, decision makers, and affected people. Fairness covers subgroup evaluation, accessibility, anti-discrimination, and mitigation of structurally unequal outcomes. Human oversight covers meaningful authority, competence, workload, automation-bias controls, and stop mechanisms.

Transparency concerns documentation, disclosure, system inventories, decision traceability, and communication of limitations. Cybersecurity includes adversarial resilience, secure development, access control, incident handling, and protection of connected infrastructure. Accountability concerns named ownership, independent challenge, auditability, appeal, remedy, and consequences. Compliance and risk management covers legal mapping, impact assessment, risk tiering, controls, validation, and evidence retention. Sustainability covers energy, carbon, material use, social well-being, accessibility, and distributional effects. Learning and engagement covers training, stakeholder participation, monitoring, incident learning, and continuous improvement.

The six maturity levels represent qualitatively different governance states. Reactive AI is characterized by fragmented experimentation, weak inventories, unclear ownership, and incident-driven response. Operational AI has repeatable technical deployment but limited cross-functional governance. Explainable AI adds documentation and explanation practices, although these may remain project specific. Governed AI introduces enterprise roles, risk classification, monitoring, validation, procurement controls, and formal escalation. Responsible AI integrates fairness, human oversight, cybersecurity, stakeholder interests, and sustainability into routine decisions. Institutionally Accountable AI adds independent assurance, public or regulatory evidence, accessible recourse, tracked remediation, board-level accountability, and organizational learning.

Movement between levels is not based on aspiration. It requires evidence. A policy statement can raise awareness but cannot establish a high maturity score without operational procedures, system records, testing outputs, named owners, and proof of corrective action. This evidence rule addresses the decoupling problem identified by institutional theory. It also prevents organizations from reaching the highest level through self-description alone.

The measurement scale uses 0 to 4 for each dimension: 0 indicates absence; 1 indicates recognition or principle; 2 indicates a documented requirement; 3 indicates an implemented control; and 4 indicates an auditable, monitored, and continually improved capability. The scale can be applied at enterprise, business-unit, portfolio, or system level. Enterprise scores should be based on a representative sample of AI systems and governance

evidence, not on a flagship project.

RAIMI combines the ten dimension scores using entropy weights derived from the variation observed in the standards corpus. Entropy weighting gives greater weight to dimensions that discriminate more strongly among governance instruments. The sector benchmark then applies an execution penalty based on the ratio of operational maturity to organizational maturity. Formally, RAIMI equals the weighted dimension score multiplied by the bounded execution factor [0.85 + 0.15 x execution-conversion ratio]. The penalty is intentionally bounded. It recognizes implementation gaps without allowing one low pillar to erase all evidence of governance capacity.

Table 1: Responsible AI maturity dimensions

Dimension	Construct definition	Illustrative evidence
Data governance	Provenance, quality, privacy, access, retention, and representativeness.	Data inventory; lineage; quality thresholds; privacy controls; access logs.
Explainability	Technical and user-facing understanding appropriate to decision context.	Global and local explanations; counterfactuals; reason codes; user testing.
Fairness	Prevention, detection, and correction of discriminatory or exclusionary effects.	Subgroup tests; accessibility; bias mitigation; disparate-impact review.
Human oversight	Competent human authority to review, override, stop, and escalate.	Override rights; workload design; automation-bias controls; escalation logs.
Transparency	Traceable documentation and disclosure of purpose, limits, and use.	System inventory; model cards; decision logs; public notices.
Cybersecurity	Protection against misuse, attacks, unauthorized access, and cyber-physical failure.	Threat modeling; red teaming; access control; incident response.
Accountability	Named ownership, independent challenge, recourse, and consequences.	Risk acceptance; audits; appeals; remediation; board reporting.
Compliance and risk	Risk tiering, legal mapping, assessment, validation, and evidence retention.	Impact assessment; controls register; validation; regulatory traceability.
Sustainability	Environmental and social effects across the AI lifecycle.	Energy and carbon metrics; lifecycle assessment; net-value tests.
Learning and engagement	Training, participation, monitoring, incident learning, and improvement.	Stakeholder review; training; post-incident analysis; control updates.

Table 2: RAIMF maturity levels

Level	Name	Capability state	RAIMI band
1	Reactive AI	Uncatalogued or experimental AI; unclear ownership; incident-driven response.	0-19.9
2	Operational AI	Repeatable deployment and basic technical controls; governance remains local.	20-34.9
3	Explainable AI	Documentation and explanation exist for selected systems; limited enterprise assurance.	35-49.9
4	Governed AI	Enterprise inventory, risk tiering, validation, monitoring, and formal oversight.	50-64.9
5	Responsible AI	Human-centered, fair, secure, sustainable, and stakeholder-aware controls are routine.	65-79.9
6	Institutionally Accountable AI	Independent assurance, enforceable recourse, public evidence, remediation, and learning.	80-100

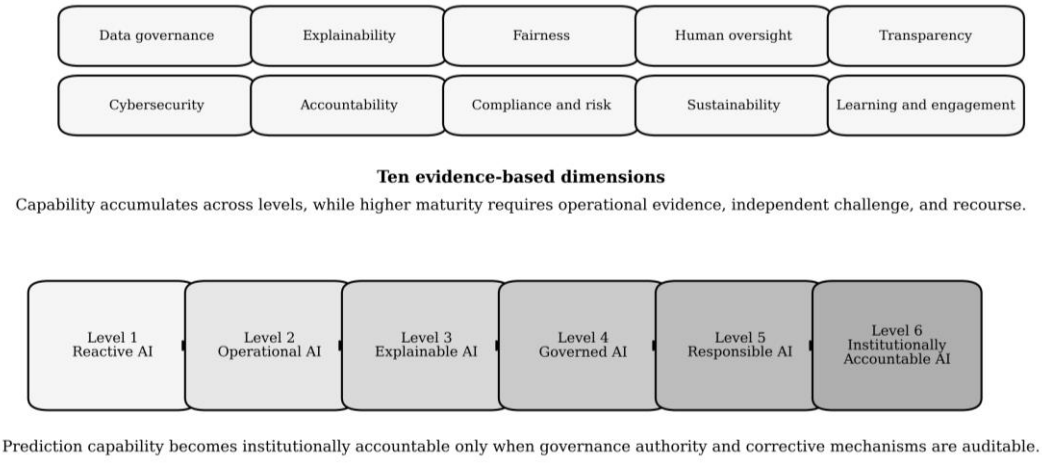

Fig 1: Responsible AI Maturity Framework.

5. Research Methodology

The study uses a sequential secondary-data design. Phase 1 establishes content validity and measurement structure through a coded corpus of governance instruments. Phase 2 provides external calibration through industry maturity distributions reported in a global survey conducted in February and March 2024. Phase 3 combines normative coverage and observed adoption into sector RAIMI profiles.

This design was chosen because no public microdata set available by December 31, 2024 contained organization-level responses across all RAIMF dimensions and the four focal sectors. Creating survey responses or imputing unpublished covariances would have produced false precision. The study therefore uses only observable documents and published aggregate distributions.

The governance corpus contains 40 instruments issued between 2011 and 2024. It includes international principles, management-system standards, risk frameworks, laws, executive and administrative guidance, public-sector accountability tools, healthcare guidance, financial model-risk rules, procurement guidance, cybersecurity frameworks, and professional governance reports. Inclusion required public availability by the cutoff date and substantive coverage of AI governance, automated decision making, model risk, or a directly applicable assurance domain. Purely promotional documents and items released in 2025 were excluded.

Each instrument was coded across the ten dimensions using the five-point evidence scale. Coding followed a written decision rule. A score of 1 required explicit recognition. A score of 2 required a policy expectation or documented process. A score of 3 required concrete implementation mechanisms, assigned roles, or control activities. A score of 4 required monitoring, auditability, review, or continual improvement. Coding was completed in two passes. The first pass applied dimension definitions. The second pass checked consistency against anchor documents, especially NIST AI RMF, ISO/IEC 42001, the EU AI Act, OMB M-24-10, WHO health guidance, and financial model-risk standards. Disagreements between passes were resolved by the more conservative score.

Internal consistency was assessed with Cronbach alpha and corrected item-total correlations. Factorability was evaluated with the Kaiser-Meyer-Olkin statistic and Bartlett's test of sphericity. Principal components with varimax rotation were used because the sample consists of governance instruments rather than a random population of organizations and because the objective is construct reduction rather than causal latent-variable estimation. Components with eigenvalues above 1 were retained. Composite reliability, average variance extracted, and HTMT were calculated for multi-item components. A covariance-based SEM or PLS-SEM was not imposed on the standards corpus because 40 documents do not support a stable structural model with ten manifest dimensions. The more conservative factor and regression approach avoids pseudo-replication. Given the ratio of 40 documents to ten coded dimensions, the component solution is treated as exploratory and potentially sample-sensitive rather than confirmatory.

Document-level hypothesis tests examined whether operational specificity predicted accountability coverage and whether human-centered transparency correlated with accountability. Operational specificity was the mean of data governance, cybersecurity, compliance and risk, and learning and engagement. Category differences were evaluated with Kruskal-Wallis tests because the coding scale is ordinal and category sample sizes are unequal.

External validation used the industry distributions published by Reuel *et al.* (2024), whose survey included 1,000 organizations across 19 industries and 19 regions. Percentages in the organizational and operational maturity figures were transcribed and checked to sum to approximately 100 percent within industry. Level values of 0, 25, 50, 75, and 100 were applied to Initial, Assessing, Determined, Managed, and Optimized categories. Weighted means were then calculated for each industry. Finance combines Banking, Capital Markets, and Insurance using sample counts inferred from the published percentage increments. Public Services is used as the closest available proxy for public supply chains.

Industry-level analysis tested the relationship between organizational and operational maturity using Pearson correlation, Spearman correlation, and OLS regression. Critical regulated sectors were compared with other sectors using Welch's t test and the Mann-Whitney test. These tests are ecological and do not substitute for organization-level inference. Their purpose is to assess whether industry environments with higher organizational maturity also show higher operational maturity.

Sector normative profiles were created from six-document bundles. Healthcare uses WHO, FDA, GMLP, and ONC instruments. Finance uses Federal Reserve, Bank of England, EBA, CFPB, NIST, and ISO instruments. Energy uses cross-sector risk, management-system, cybersecurity, and sustainability instruments because no equivalent energy-specific AI governance code existed by the cutoff date. Public supply chains use OMB, GAO, Canadian, British, New Zealand, and World Economic Forum public-procurement or government accountability instruments.

Dimension scores combine normative coverage with observed organizational and operational maturity. Technical dimensions use 40 percent normative coverage, 20 percent organizational maturity, and 40 percent operational maturity. Governance dimensions use 40 percent normative coverage, 40 percent organizational maturity, and 20 percent operational maturity. Sustainability uses 50 percent normative coverage, 30 percent organizational maturity, and 20 percent operational maturity. Sensitivity tests replace entropy weights with equal weights, remove the execution penalty, and increase operational emphasis. All calculations were performed with Python using fixed formulas and no random simulation.

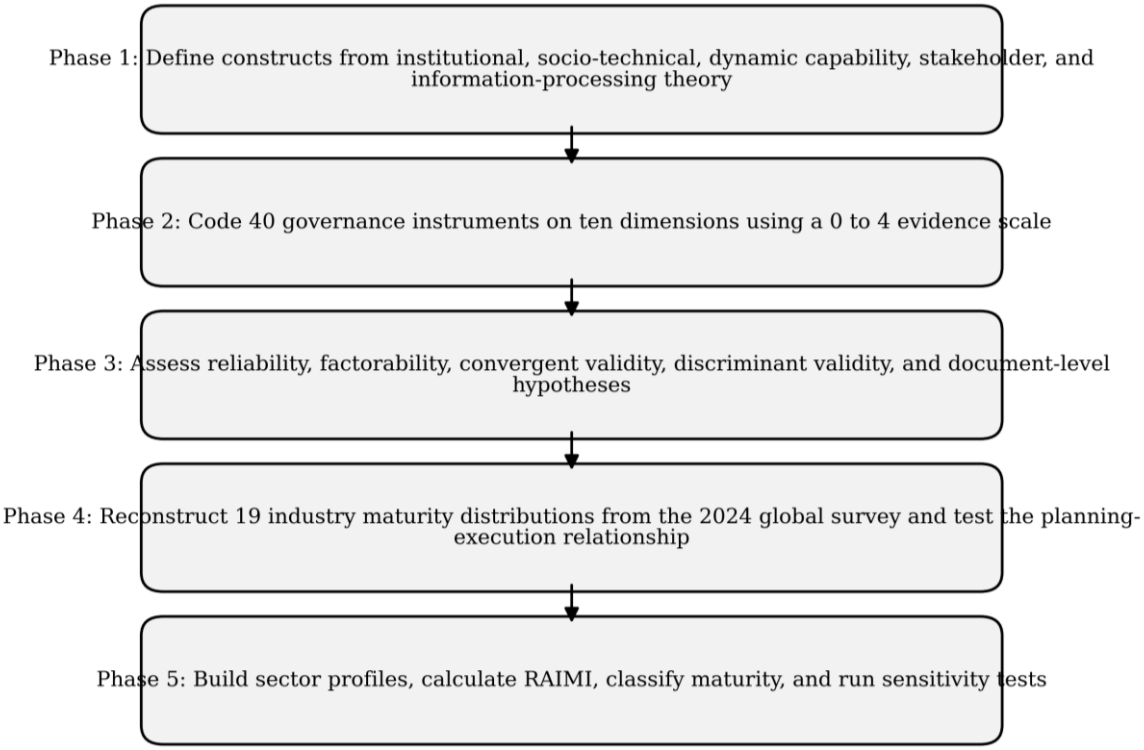

Fig 2: Research methodology flowchart.

6 Data Collection

The corpus reflects the governance landscape available to an organization preparing a Responsible AI program at the end of 2024. International sources include the OECD AI Principles, UNESCO Recommendation, European Commission guidance, the EU AI Act, the Council of Europe Framework Convention, G7 Hiroshima guidance, and the UN Global Digital Compact. Management and assurance sources include NIST AI RMF 1.0 and its Playbook, the NIST Generative AI Profile, ISO/IEC 42001, ISO/IEC 23894, ISO/IEC 38507, ISO/IEC TR 24028, and IEEE standards addressing ethical design, transparency, and bias.

United States public-sector sources include Executive Order 14110, the Blueprint for an AI Bill of Rights, OMB M-24-10, and the GAO AI Accountability Framework. Comparable government instruments include the Canadian Directive on Automated Decision-Making, the UK Algorithmic Transparency Recording Standard, the New Zealand Algorithm Charter, the Singapore Model AI Governance Framework, and Australia's AI Ethics Principles. The public-procurement dimension is represented by the World Economic Forum guidelines, which translate Responsible AI expectations into acquisition decisions.

Healthcare instruments include WHO's 2021 ethics and governance guidance, its 2023 regulatory considerations, its 2024 guidance on large multimodal models for health, the FDA AI/ML Software as a Medical Device Action Plan, the joint Good Machine Learning Practice principles, and the ONC HTI-1 transparency requirements. Finance includes Federal Reserve SR 11-7, Bank of England PRA SS1/23, EBA guidance on big data and advanced analytics, and CFPB guidance on adverse action when complex algorithms are used. NIST Cybersecurity Framework 2.0 supplies a cross-sector security benchmark.

The second dataset is not a new survey. It is a reproducible reconstruction of published industry distributions. This distinction is important. The study does not claim access to respondent-level records, joint organizational-operational maturity tables, or unpublished measurement items. The 19 industry observations are aggregate benchmarks. They are suitable for testing whether sector-level organizational sophistication tracks sector-level operational implementation, but not for estimating individual organization effects.

Data quality checks included range tests, within-industry sum checks, duplicate-source checks, year verification, and comparison of coded values with the evidence rubric. Percentage totals differed from 100 by at most rounding error. All post-2024 materials were excluded even when they later reproduced or expanded the same frameworks. The analysis date is treated as January 2025.

7 Empirical Results

The standards corpus shows strong but uneven governance coverage. Mean scores were highest for transparency and accountability, both 3.82, followed by human oversight at 3.80. Sustainability had the lowest mean at 2.60. This pattern indicates that major frameworks increasingly require documentation, ownership, and review, but environmental and broader social-impact controls remain less operationalized. Explainability averaged 3.25, below fairness at 3.52, partly because some cybersecurity and model-risk instruments require validation and control without prescribing explanation methods for affected users.

H1 receives preliminary support. Cronbach alpha was 0.750, indicating acceptable internal consistency for a multidimensional governance scale. The KMO statistic was 0.608, adequate for exploratory factor analysis given the

heterogeneous document sample. Bartlett's test was significant, chi-square(45) = 234.66, p < .001. Three components had eigenvalues above 1 and together explained 76.2 percent of variance. Governance assurance loaded strongly on data governance, cybersecurity, accountability, compliance and risk, and learning. Human-centered transparency loaded on explainability, fairness, human oversight, and transparency. Sustainability formed a distinct component rather than disappearing into general governance. Convergent and discriminant validity were satisfactory for the two multi-item components. Governance assurance achieved composite reliability of 0.916 and AVE of 0.686. Human-centered transparency achieved composite reliability of 0.855 and AVE of 0.602. HTMT was 0.360, well below conventional thresholds. The results support the theoretical claim that assurance and human-centered transparency are related but empirically distinct.

H2 is supported. Operational specificity positively predicted accountability coverage. The regression coefficient was 0.412, p < .001, with R-squared of 0.365. Instruments that specify data controls, cybersecurity, risk management, and learning are more likely to include auditable accountability mechanisms. H3 received limited support. Human-centered transparency correlated positively with accountability, r = 0.259, but the relationship did not reach the .05 threshold, p = 0.107. This result suggests that explanation and fairness language can exist without equally strong recourse or enforcement.

H4 is supported only for sustainability. Sector-specific instruments had a mean sustainability score of 2.00, public-sector operational instruments 2.50, and general instruments 2.88. The Kruskal-Wallis test was significant, H = 7.70, p = 0.021. Differences in cybersecurity, compliance, and accountability were directionally consistent with the hypothesis but not statistically significant. The finding is substantively important: sector rules become more operational and compliance-centered, yet environmental accountability is often narrowed or omitted.

The external benchmark reveals a large execution gap. Across 19 industries, mean organizational maturity was 69.3, while mean operational maturity was 36.4. The average gap was 32.9 points. H5 is not supported. Organizational maturity had almost no linear relationship with operational maturity, Pearson r = 0.081, p = 0.740; Spearman rho = 0.279, p = 0.247. The OLS coefficient was 0.085, p = 0.740, R-squared = 0.007. Formal governance maturity, at least at the industry level, does not reliably predict implemented mitigation.

H6 receives marginal support. Critical regulated sectors averaged 37.7 in operational maturity compared with 35.3 for other sectors. Welch's t test produced p = 0.056; the Mann-Whitney test produced p = 0.079. Regulation appears to raise the operational baseline, but the evidence is not strong enough for a definitive sector effect.

The four focal sectors cluster tightly in organizational maturity but differ in execution conversion. Healthcare recorded organizational maturity of 70.9 and operational maturity of 38.0. Finance recorded 69.1 and 36.6. Energy recorded 71.8 and 38.5. Public supply chains recorded 69.7 and 39.0. Conversion ratios ranged from 0.530 to 0.560, supporting H7 descriptively. Public supply chains had the strongest conversion, though all sectors operationalized only about half of their organizational maturity.

Execution-penalized RAIMI scores were 60.9 for healthcare, 59.3 for finance, 60.2 for energy, and 61.7 for public supply chains. All four fall within Level 4, Governed AI, using thresholds of 50 to 64.9. None reaches Responsible AI at 65 or Institutionally Accountable AI at 80. Public supply chains rank first because their public-governance bundle is strong and their conversion ratio is highest. Finance ranks lowest because mature policy and model-risk traditions are not matched by equally high operational maturity in the aggregate benchmark.

H8 is supported. Sector ordering is highly stable under equal weighting, removal of the execution penalty, and greater operational emphasis. Spearman rank correlations with the base ranking ranged from 0.800 to 1.000. Absolute scores change, especially when the execution penalty is removed, but the conclusion does not: all four sectors remain below institutionally accountable maturity.

Table 3: Reliability and factorability

Statistic	Result	Interpretation
Cronbach alpha	0.750	Acceptable internal consistency
KMO	0.608	Adequate for exploratory factor analysis
Bartlett's test	chi-square = 234.66, df = 45, p < .001	Correlation matrix is factorable
Variance explained	76.2%	Three retained components

Table 4: Responsible AI Maturity Index scores

Sector	Unpenalized score	Execution ratio	RAIMI	Classification
Public Supply Chains	66.0	0.560	61.7	Governed AI
Healthcare	65.4	0.536	60.9	Governed AI
Energy	64.7	0.536	60.2	Governed AI
Finance	63.8	0.530	59.3	Governed AI

Table 5: Hypothesis tests

Hypothesis	Statement	Evidence	Decision
H1	RAIMF dimensions form a reliable multidimensional construct	alpha = 0.750; three components; 76.2% variance	Preliminary support
H2	Operational specificity predicts accountability	beta = 0.412; p < .001	Supported
H3	Human-centered transparency is associated with accountability	r = 0.259; p = 0.107	Not statistically supported
H4	Sector instruments prioritize compliance/cyber over sustainability	Sustainability H = 7.70; p = 0.021	Partially supported
H5	Organizational maturity predicts operational maturity	beta = 0.085; p = 0.740	Not supported
H6	Critical sectors have higher operational maturity	Welch p = 0.056; Mann-Whitney p = 0.079	Marginal, not confirmed
H7	Focal sectors differ in execution conversion	Range = 0.530 to 0.560	Descriptive only
H8	RAIMI ranking is robust to alternative specifications	Rank rho = 0.800 to 1.000	Supported

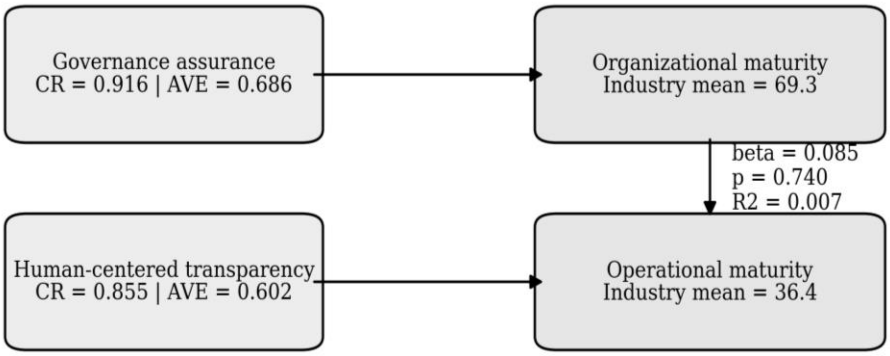

Fig 3: Validated measurement and structural model.

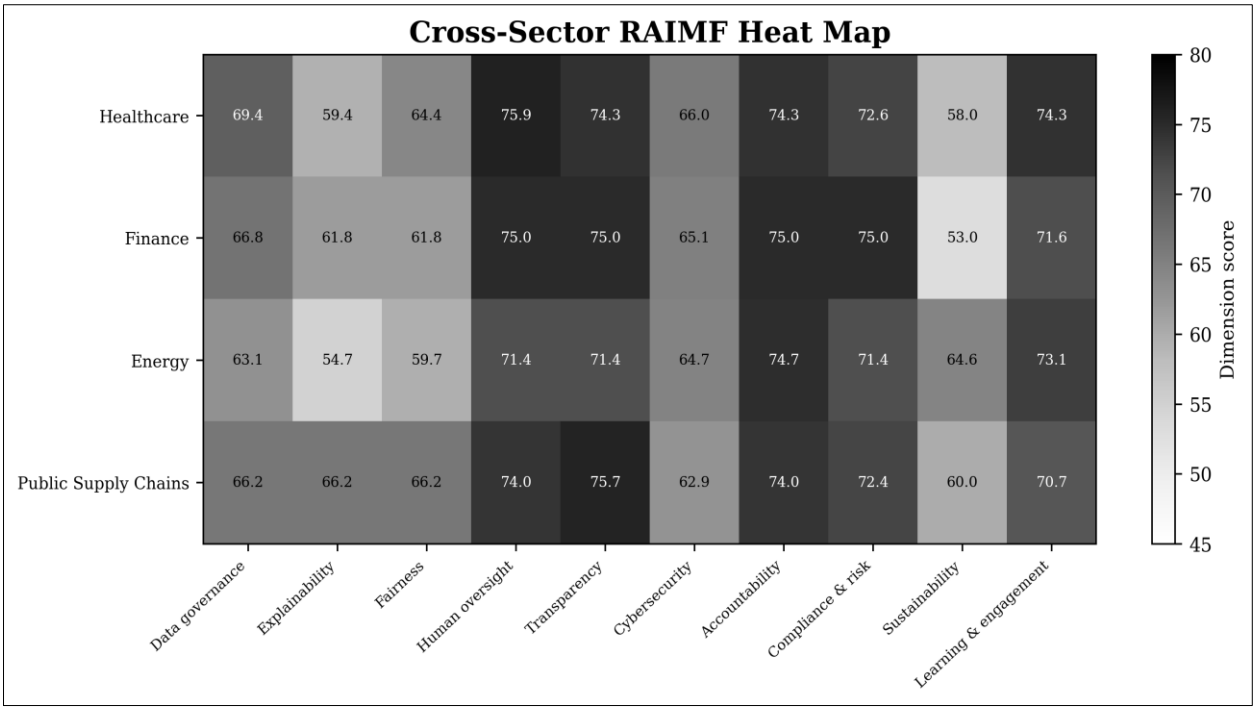

Fig 4: Heat map comparing sectors.

8. Discussion

The central empirical result is not a sector ranking. It is the weak relationship between organizational and operational maturity. Organizations can create committees, policies, training programs, and risk taxonomies without achieving consistent mitigation at the system level. Institutional theory predicts this pattern. Formal structures respond to legitimacy pressure, but operational routines are more costly, contested, and technically demanding. Responsible AI programs can therefore look mature from the center while remaining uneven at the point of deployment.

This finding changes how maturity should be interpreted. A conventional model assumes that higher levels accumulate naturally: organizations first adopt technology, then standardize, govern, and optimize it. RAIMF rejects that assumption. Technical, organizational, and institutional capabilities can move at different speeds. An organization can be advanced in predictive analytics, moderate in

governance, and weak in recourse. It can also possess strong regulatory controls but weak sustainability practices. Maturity is a profile before it is a single score.

The exploratory three-component structure is consistent with this view, although its stability requires replication in a larger organizational sample. Governance assurance and human-centered transparency are distinct. A financial model-risk program can be strong in validation, ownership, cybersecurity, and documentation while offering limited explanations to consumers. A public-facing AI ethics policy can emphasize fairness and transparency while lacking independent challenge or incident learning. Sustainability forms a third component because it is not yet embedded in the same control architecture. This separation is empirically visible and theoretically consequential.

Healthcare illustrates why integration matters. Clinical AI requires robust evidence, professional oversight, subgroup monitoring, privacy, and patient communication. A mature program must also govern supplier models, updates, workflow changes, and cybersecurity. The healthcare supply-chain literature adds continuity and resilience. Forecasting accuracy cannot substitute for provenance, supplier accountability, and emergency override (Rasel *et al*., 2022). Security work similarly shows that protecting patient data and connected devices is part of AI responsibility, not a separate infrastructure issue (Hasan *et al*., 2022).

Finance has stronger historical controls but still faces a conversion problem. SR 11-7 and related standards normalize validation independence, model inventories, documentation, and governance. Yet generative AI, graph models, third-party services, and real-time fraud systems expand the boundary of the model-risk function. Hasan *et al*. (2023) argue for integrated fraud and cyber analytics; RAIMF adds the requirement that integrated prediction must remain auditable and contestable. Rasel *et al*. (2023) show the inclusion potential of multimodal mortgage models, but inclusion depends on traceable evidence, fairness testing, and adverse-action explanations. Complexity increases the need for accountability rather than excusing its absence.

Energy's profile is shaped by strong cyber and system-risk concerns but thinner sector-specific AI governance. The sector relies on cross-cutting frameworks, which may not fully address operational technology, grid reliability, community impact, or the carbon cost of model selection. Arman *et al*. (2024a) demonstrate the value of advanced forecasting for renewable integration. RAIMF makes the complementary point that forecast value should be measured net of energy use, operational risk, and distributional effects. Sustainability must become an auditable control domain.

Public supply chains score slightly higher because public-sector frameworks are explicit about inventories, impact assessments, transparency, procurement, and rights. Their challenge is vendor governance. Agencies may not control the training data, model updates, or proprietary explanations of purchased systems. Procurement must therefore require evidence rights, audit access, incident notification, update controls, subcontractor transparency, and exit provisions. Otherwise accountability stops at the contract boundary.

The standards corpus also reveals a design imbalance. Accountability is widely named, but sustainability is less operational. This gap is likely to persist unless sustainability is connected to risk classification, architecture choice, procurement, and performance review. Arman *et al*. (2024b) show that machine learning can support waste reduction and circular operations. The governance implication is that positive environmental claims should be measured against baseline methods, rebound effects, data-center energy, and lifecycle costs.

9. Theoretical Implications

RAIMF contributes to institutional theory by specifying the mechanisms that reduce decoupling. Policies and committees become substantive when they produce system inventories, risk decisions, evidence, intervention authority, and tracked remediation. The execution-conversion ratio provides a measurable indicator of coupling between formal governance and operational practice.

The framework extends socio-technical systems theory by treating accountability as an emergent property of the organizational system. Model behavior, human discretion, workflow design, vendor contracts, and institutional rules jointly determine whether an AI decision can be challenged and corrected. This moves Responsible AI research away from the assumption that ethical quality can be located within an algorithm.

RAIMF also adds a dynamic-capability interpretation. Sensing includes identifying AI uses, affected stakeholders, drift, misuse, and regulatory change. Seizing includes allocating resources, selecting controls, and redesigning processes. Transforming includes changing governance, retraining staff, replacing suppliers, and revising risk appetite after evidence emerges. Institutionally accountable AI is therefore a higher-order capability for repeated adaptation under scrutiny.

Finally, the distinct sustainability component suggests that Responsible AI theory remains incomplete when environmental and intergenerational effects are treated as optional. Sustainability is not simply another stakeholder outcome. It changes architecture choices, model selection, data retention, procurement, and the definition of net value. Future theory should connect AI governance with operations strategy and environmental accounting.

10. Practical Implications

Managers should begin with an AI inventory rather than an ethics statement. The inventory should identify purpose, owner, vendor, users, affected groups, data sources, model dependencies, decision authority, risk tier, monitoring frequency, and retirement conditions. High-risk systems should not enter production without evidence across all ten dimensions.

Boards and senior executives should monitor the execution gap. A rising count of policies, training completions, or committee meetings can create false confidence. More informative indicators include the percentage of systems with current impact assessments, independent validation, subgroup testing, cybersecurity testing, human override, incident playbooks, appeal routes, and closed remediation actions.

Model documentation should be audience-specific. Developers require technical evidence. Operational users need limits, escalation rules, and uncertainty information. Affected people need understandable reasons and recourse. Regulators and auditors need traceability from requirements to controls and outcomes. One document rarely serves all four audiences.

Procurement deserves the same maturity discipline as internal development. Contracts should require model and data

documentation, testing rights, service-level commitments for incidents, change notification, performance monitoring, audit cooperation, security obligations, and transition assistance. A vendor's claim of proprietary technology should not eliminate the buyer's accountability.

Sustainability metrics should enter model selection. Organizations should compare marginal predictive gain with training and inference energy, hardware needs, latency, maintenance burden, and operational consequences. A simpler model may create greater net carbon and governance value when performance differences are small.

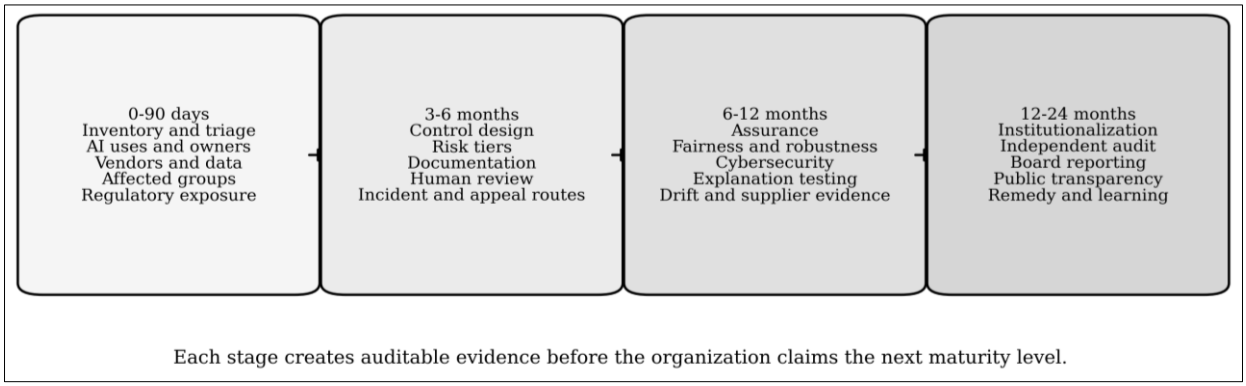

Fig 5: Responsible AI roadmap.

11. Policy Implications

Regulators should distinguish policy maturity from control maturity. Disclosure of principles or governance structures is useful, but assurance should focus on representative system evidence. Supervisory reviews can sample AI inventories, impact assessments, validation records, monitoring dashboards, incidents, appeals, and remediation closure.

Cross-sector consistency is desirable, but sector specificity remains necessary. NIST AI RMF and ISO/IEC 42001 provide common architecture. Healthcare regulators should add clinical validity, patient safety, workflow, and post-market monitoring. Financial regulators should add consumer explanation, fair lending, validation independence, and model inventory requirements. Energy regulators should add operational technology, reliability, cyber-physical risk, and carbon accounting. Public procurement authorities should add evidence rights and vendor accountability.

Public reporting should emphasize meaningful comparability. A standardized RAIMF disclosure could report dimension scores, evidence coverage, execution ratio, high-risk system counts, material incidents, appeal outcomes, and improvement commitments. Such reporting would reduce symbolic adoption and allow stakeholders to distinguish aspiration from verified capability.

Policy should also protect contestability. A person affected by an AI-supported decision needs a route to obtain reasons, human review, correction, and remedy. The EU AI Act, public-sector algorithmic transparency standards, consumer-protection law, and health governance guidance move in this direction, but implementation should be tested through outcomes rather than procedural existence alone.

12. Limitations

The study has four limitations. First, document coding involves judgment. The two-pass rubric and conservative scoring improve consistency, but independent expert replication would strengthen inter-rater reliability. Second, the 40-instrument corpus is purposive rather than exhaustive. It emphasizes influential and publicly accessible governance sources. The exploratory component analysis is based on only 40 governance instruments and ten coded dimensions, so loadings, reliability estimates, entropy weights, and sector rankings may be sensitive to corpus composition.

Third, the industry validation uses aggregate distributions from a published survey. The absence of microdata prevents organization-level SEM, measurement invariance testing, and direct estimation of causal effects. The weak cross-industry relationship should not be interpreted as evidence that organizational governance never improves implementation within firms. It shows that industry averages do not move together.

Fourth, the energy and public-supply-chain benchmarks rely on proxy governance bundles. By the end of 2024, sector-specific AI governance was more developed in healthcare and finance. Future studies should update the corpus as energy and procurement rules become more specific, while maintaining the January 2025 historical boundary for replication of this study.

13. Future Research

Future research should administer the RAIMF instrument to organizations and collect evidence from multiple respondents, including model owners, risk leaders, auditors, frontline users, and affected stakeholders. A multi-level design could test whether enterprise governance predicts system-level controls and whether that relationship depends on sector, organization size, vendor dependence, and regulatory exposure.

Longitudinal research should examine transitions between maturity levels. The most useful question is not whether a policy exists, but which interventions close the execution gap. Candidate mechanisms include centralized inventories, mandatory risk tiering, independent validation, procurement gates, incident review, board reporting, and incentive changes.

Research should also test outcome validity. RAIMI should be linked to incident frequency, time to remediation, regulatory findings, deployment speed, model performance stability, stakeholder trust, operational resilience, and environmental impact. Such evidence would clarify when governance functions as a dynamic capability rather than a compliance cost.

Finally, sustainability requires deeper measurement. Future RAIMF versions should incorporate model-energy

accounting, carbon-aware training and inference, hardware lifecycle, data-retention costs, and distributional environmental effects. The objective should be net responsible value, not maximal model complexity.

14. Conclusion

Responsible AI maturity is not the maturity of prediction. It is the maturity of an institution's ability to govern prediction. RAIMF captures that ability through ten dimensions and six levels, ending with institutionally accountable AI. The empirical analysis supports a multidimensional structure, demonstrates acceptable reliability and validity, and shows that operational specificity is strongly associated with accountability coverage.

The cross-industry analysis exposes the central governance problem. Organizational maturity is high relative to operational maturity, and the two are scarcely related across industries. Policies, committees, and plans do not automatically become system controls. All four focal sectors remain at the Governed AI level after the execution penalty. They have substantial architecture in place, but not enough evidence of consistent implementation, sustainability, recourse, and learning to qualify as Responsible or Institutionally Accountable AI.

The practical implication is direct. Organizations should stop treating governance as the final layer placed over a completed model. Governance must shape problem definition, data, architecture, procurement, validation, deployment, monitoring, appeal, and retirement. Prediction creates value only when the institution can explain what it did, justify why it did it, detect when it fails, stop it when necessary, correct the consequences, and learn from the event.

References

1. Amershi S, Begel A, Bird C, DeLine R, Gall H, Kamar E, Nagappan N, Nushi B, Zimmermann T. Software engineering for machine learning: A case study. In: Proceedings of the 41st International Conference on Software Engineering: Software Engineering in Practice; 2019 May 25-31; Montreal, Canada. Piscataway: IEEE; 2019. p. 291-300. doi: 10.1109/ICSE-SEIP.2019.00042.
2. Argyris C, Schon DA. Organizational learning: A theory of action perspective. Reading: Addison-Wesley; 1978.
3. Arman M, Hasan MN, Rasel IH. Clean energy transition in USA: Big data analytics for renewable energy forecasting and carbon reduction. J Manag World. 2024;2024(3):192-206. doi: 10.53935/jomw.v2024i4.1196.
4. Arman M, Rasel IH, Razib MNH, Fahim ASM. Big data and machine learning for sustainable waste reduction. J Posthumanism. 2024;4(2):448-467. doi: 10.63332/joph.v4i2.3361.
5. Barney J. Firm resources and sustained competitive advantage. J Manage. 1991;17(1):99-120. doi: 10.1177/014920639101700108.
6. Barocas S, Hardt M, Narayanan A. Fairness and machine learning: Limitations and opportunities. Cambridge: MIT Press; 2023.
7. Boston Consulting Group. Are you overestimating your responsible AI maturity? Boston: BCG GAMMA; 2021.
8. Consumer Financial Protection Bureau. Equal Credit Opportunity Act and Regulation B requirements for creditors relying on complex algorithms. Washington, D.C.: CFPB; 2022. Circular 2022-03.
9. Council of Europe. Framework convention on artificial intelligence and human rights, democracy and the rule of law. Strasbourg: Council of Europe; 2024.
10. Diakopoulos N. Algorithmic accountability: Journalistic investigation of computational power structures. Digit Journal. 2015;3(3):398-415. doi: 10.1080/21670811.2014.976411.
11. DiMaggio PJ, Powell WW. The iron cage revisited: Institutional isomorphism and collective rationality in organizational fields. Am Sociol Rev. 1983;48(2):147-160. doi: 10.2307/2095101.
12. Dotan R, Blili-Hamelin B, Madhavan R, Matthews J, Scarpino J. Evolving AI risk management: A maturity model based on the NIST AI Risk Management Framework [Internet]. arXiv; 2024 [cited 2026 Jul 29]. Available from: https://doi.org/10.48550/arXiv.2401.15229.
13. European Banking Authority. Report on big data and advanced analytics. Paris: EBA; 2020.
14. European Commission High-Level Expert Group on Artificial Intelligence. Ethics guidelines for trustworthy AI. Brussels: European Commission; 2019.
15. European Union. Regulation (EU) 2024/1689 laying down harmonised rules on artificial intelligence. Off J Eur Union. 2024;L 1689.
16. Federal Reserve Board. Supervisory guidance on model risk management. Washington, D.C.: Federal Reserve; 2011. SR Letter 11-7.
17. Floridi L, Cowls J, Beltrametti M, Chatila R, Chazerand P, Dignum V, Luetge C, Madelin R, Pagallo U, Rossi F, Schafer B, Valcke P, Vayena E. AI4People: An ethical framework for a good AI society. Minds Mach. 2018;28:689-707. doi: 10.1007/s11023-018-9482-5.
18. Freeman RE. Strategic management: A stakeholder approach. Boston: Pitman; 1984.
19. Galbraith JR. Organization design: An information processing view. Interfaces. 1974;4(3):28-36. doi: 10.1287/inte.4.3.28.
20. Gebru T, Morgenstern J, Vecchione B, Vaughan JW, Wallach H, Daumé III H, Crawford K. Datasheets for datasets. Commun ACM. 2021;64(12):86-92. doi: 10.1145/3458723.
21. Government Accountability Office. Artificial intelligence: An accountability framework for federal agencies and other entities. Washington, D.C.: GAO; 2021. GAO-21-519SP.
22. Hagendorff T. The ethics of AI ethics: An evaluation of guidelines. Minds Mach. 2020;30:99-120. doi: 10.1007/s11023-020-09517-8.
23. Hasan MN, Rasel IH, Arman M, Ibrahim M, Jahan N. Strengthening U.S. financial and cybersecurity infrastructure with AI-driven fraud detection and risk analytics. J Comput Anal Appl. 2023;31(2):15-32.
24. Hasan MN, Rasel IH, Rahman M, Islam K, Arman M, Jahan N. Securing U.S. healthcare infrastructure with machine learning: Protecting patient data as a national security priority. Int J Comput Exp Sci Eng. 2022;8(3):85-93. doi: 10.22399/ijcesen.3987.
25. Information Commissioner's Office. Algorithmic transparency recording standard. London: ICO; 2022.
26. International Organization for Standardization. ISO/IEC 38507:2022: Governance implications of the use of artificial intelligence by organizations. Geneva: ISO; 2022.

27. International Organization for Standardization. ISO/IEC 42001:2023: Artificial intelligence management system. Geneva: ISO; 2023.
28. International Organization for Standardization. ISO/IEC 23894:2023: Artificial intelligence guidance on risk management. Geneva: ISO; 2023.
29. Jobin A, Ienca M, Vayena E. The global landscape of AI ethics guidelines. Nat Mach Intell. 2019;1:389-399. doi: 10.1038/s42256-019-0088-2.
30. Kamruzzaman M, Saha S, Islam MK. Augmented intelligence in healthcare: Bridging machine learning, human expertise, and business process automation. Int J Eng Technol Res Manag. 2023;7(6):250-264.
31. Kroll JA, Huey J, Barocas S, Felten EW, Reidenberg JR, Robinson DG, Yu H. Accountable algorithms. Univ Pa Law Rev. 2017;165(3):633-705.
32. Lundberg SM, Lee SI. A unified approach to interpreting model predictions. In: Advances in Neural Information Processing Systems 30 (NeurIPS 2017); 2017 Dec 4-9; Long Beach, CA. Red Hook: Curran; 2017. p. 4765-4774.
33. March JG. Exploration and exploitation in organizational learning. Organ Sci. 1991;2(1):71-87. doi: 10.1287/orsc.2.1.71.
34. Martin K. Ethical implications and accountability of algorithms. J Bus Ethics. 2019;160:835-850. doi: 10.1007/s10551-018-3921-3.
35. Maslej N, Fattorini L, Perrault R, Parli V, Reuel A, Brynjolfsson E, Etchemendy J, Ligett K, Lyons T, Manyika J, Niebles JC, Shoham Y, Wald R, Clark J. The AI Index 2024 annual report. Stanford: Stanford Institute for Human-Centered AI; 2024.
36. Meyer JW, Rowan B. Institutionalized organizations: Formal structure as myth and ceremony. Am J Sociol. 1977;83(2):340-363. doi: 10.1086/226550.
37. Mitchell M, Wu S, Zaldivar A, Barnes P, Vasserman L, Hutchinson B, Spitzer E, Raji ID, Gebru T. Model cards for model reporting. In: Proceedings of the Conference on Fairness, Accountability, and Transparency; 2019 Jan 29-31; Atlanta, GA. New York: ACM; 2019. p. 220-229. doi: 10.1145/3287560.3287596.
38. Mittelstadt B. Principles alone cannot guarantee ethical AI. Nat Mach Intell. 2019;1:501-507. doi: 10.1038/s42256-019-0114-4.
39. Morley J, Floridi L, Kinsey L, Elhalal A. From what to how: An initial review of publicly available AI ethics tools, methods and research to translate principles into practices. Sci Eng Ethics. 2020;26:2141-2168. doi: 10.1007/s11948-019-00165-5.
40. National Institute of Standards and Technology. Artificial Intelligence Risk Management Framework (AI RMF 1.0). Gaithersburg: NIST; 2023. NIST AI 100-1. doi: 10.6028/NIST.AI.100-1.
41. National Institute of Standards and Technology. AI RMF Playbook. Gaithersburg: NIST; 2023.
42. National Institute of Standards and Technology. Artificial Intelligence Risk Management Framework: Generative Artificial Intelligence Profile. Gaithersburg: NIST; 2024. NIST AI 600-1. doi: 10.6028/NIST.AI.600-1.
43. National Institute of Standards and Technology. Cybersecurity Framework 2.0. Gaithersburg: NIST; 2024. NIST CSWP 29. doi: 10.6028/NIST.CSWP.29.
44. OECD. Recommendation of the Council on Artificial Intelligence. Paris: OECD; 2019. OECD/LEGAL/0449.
45. Office of Management and Budget. Advancing governance, innovation, and risk management for agency use of artificial intelligence. Washington, D.C.: OMB; 2024. Memorandum M-24-10.
46. Orlikowski WJ. The duality of technology: Rethinking the concept of technology in organizations. Organ Sci. 1992;3(3):398-427. doi: 10.1287/orsc.3.3.398.
47. Power M. The audit society: Rituals of verification. Oxford: Oxford University Press; 1997.
48. Prudential Regulation Authority. Model risk management principles for banks. London: Bank of England; 2023. Supervisory Statement SS1/23.
49. Rabbani SF, Rahat YO, Islam MK. From utility bills to retrofit finance: An AI framework for energy-burden-aware underwriting in residential decarbonization. Front Comput Sci Artif Intell. 2024;3(2):72-88. doi: 10.32996/fcsai.2024.3.2.10.
50. Rahat YO, Islam MK, Rabbani SF. Safeguarding federal payment infrastructure: Trustworthy AI for improper-payment prevention, synthetic identity detection, and public-benefit disbursement integrity in the United States. J Bus Manag Stud. 2024;6(5):238-251. doi: 10.32996/jbms.2024.6.5.25.
51. Raji ID, Smart A, White RN, Mitchell M, Gebru T, Hutchinson B, Smith-Loud J, Theron D, Barnes P. Closing the AI accountability gap: Defining an end-to-end framework for internal algorithmic auditing. In: Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency; 2020 Jan 27-30; Barcelona, Spain. New York: ACM; 2020. p. 33-44. doi: 10.1145/3351095.3372873.
52. Rasel IH, Arman M, Hasan MN, Bhuyain MMH. Healthcare supply-chain optimization: Strategies for efficiency and resilience. J Med Health Stud. 2022;3(4):171-182. doi: 10.32996/jmhs.2022.3.4.26.
53. Rasel IH, Ibrahim M, Pritty AA, Fahim ASM, Jahan N. Beyond FICO: Enhancing mortgage default forecasting and inclusive lending via multimodal graph neural networks and urban mobility analytics. Front Comput Sci Artif Intell. 2023;2(2):62-81. doi: 10.32996/fcsai.2023.2.2.5.
54. Reuel A, Connolly P, Meimandi KJ, Tewari S, Wiatrak J, Venkatesh D, Kochenderfer M. Responsible AI in the global context: Maturity model and survey [Internet]. arXiv; 2024 [cited 2026 Jul 29]. Available from: https://doi.org/10.48550/arXiv.2410.09985.
55. Ribeiro MT, Singh S, Guestrin C. "Why should I trust you?" Explaining the predictions of any classifier. In: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining; 2016 Aug 13-17; San Francisco, CA. New York: ACM; 2016. p. 1135-1144. doi: 10.1145/2939672.2939778.
56. Sculley D, Holt G, Golovin D, Davydov E, Phillips T, Ebner D, Chaudhary V, Young M, Crespo JF, Dennison D. Hidden technical debt in machine learning systems. In: Advances in Neural Information Processing Systems 28 (NeurIPS 2015); 2015 Dec 7-12; Montreal, Canada. 2015.
57. Selbst AD, Boyd D, Friedler SA, Venkatasubramanian S, Vertesi J. Fairness and abstraction in sociotechnical systems. In: Proceedings of the Conference on Fairness, Accountability, and Transparency; 2019 Jan 29-31; Atlanta, GA. New York: ACM; 2019. p. 59-68. doi:

10.1145/3287560.3287598.
58. Shneiderman B. Human-centered artificial intelligence: Reliable, safe and trustworthy. Int J Hum-Comput Interact. 2020;36(6):495-504. doi: 10.1080/10447318.2020.1741118.
59. Suchman MC. Managing legitimacy: Strategic and institutional approaches. Acad Manag Rev. 1995;20(3):571-610. doi: 10.5465/amr.1995.9508080331.
60. Teece DJ, Pisano G, Shuen A. Dynamic capabilities and strategic management. Strateg Manag J. 1997;18(7):509-533.
61. The White House. Blueprint for an AI Bill of Rights. Washington, D.C.: White House; 2022.
62. The White House. Executive Order 14110 on the safe, secure, and trustworthy development and use of artificial intelligence. Washington, D.C.: White House; 2023.
63. Trist EL, Bamforth KW. Some social and psychological consequences of the longwall method of coal-getting. Hum Relat. 1951;4(1):3-38. doi: 10.1177/001872675100400101.
64. UNESCO. Recommendation on the ethics of artificial intelligence. Paris: UNESCO; 2021.
65. United Nations. Global Digital Compact. New York: UN; 2024.
66. Vakkuri V, Kemell KK, Jantunen M, Abrahamsson P. ECCOLA: A method for implementing ethically aligned AI systems. J Syst Softw. 2021;182:111067. doi: 10.1016/j.jss.2021.111067.
67. Wachter S, Mittelstadt B, Russell C. Counterfactual explanations without opening the black box: Automated decisions and the GDPR. Harv J Law Technol. 2018;31(2):841-887.
68. Wendler R. The maturity of maturity model research: A systematic mapping study. Inf Softw Technol. 2012;54(12):1317-1339. doi: 10.1016/j.infsof.2012.07.007.
69. Wieringa M. What to account for when accounting for algorithms: A systematic literature review on algorithmic accountability. In: Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency; 2020 Jan 27-30; Barcelona, Spain. New York: ACM; 2020. p. 1-18. doi: 10.1145/3351095.3372833.
70. World Economic Forum. AI procurement in a box: AI government procurement guidelines. Geneva: WEF; 2020.
71. World Health Organization. Ethics and governance of artificial intelligence for health. Geneva: WHO; 2021.
72. World Health Organization. Regulatory considerations on artificial intelligence for health. Geneva: WHO; 2023.
73. World Health Organization. Ethics and governance of artificial intelligence for health: Guidance on large multi-modal models. Geneva: WHO; 2024.