The 0/0 Framework: A Governance Model for Responsible AI Deployment in Generative and Synthetic Media Contexts

Francis Martinson
Department of Computer Science
North Dakota State University, USA
ORCID: 0009-0007-2235-2516

ABSTRACT

Despite substantial investment and organizational commitment, artificial intelligence initiatives continue to fail at high rates. Industry surveys consistently report failure rates between 70 and 85 percent, representing billions of dollars in lost investment and unrealized strategic value annually. These failures stem not primarily from technical limitations but from governance deficiencies: inadequate stakeholder alignment, misunderstood requirements, inappropriate risk tolerance, and insufficient iteration protocols. Building on prior work establishing risk classification taxonomies for synthetic media [1] and identifying dual-use convergence patterns in generative AI applications [2], this paper introduces the 0/0 Framework, a governance model designed to address root causes of AI project failure while enabling responsible deployment of generative AI capabilities. The framework name reflects its dual objective: achieving zero preventable harms through rigorous governance while maintaining zero tolerance for governance shortcuts that create conditions for such harms. The framework operationalizes four interconnected pillars: Value Alignment requires AI systems to embody organizational values; Risk Proportionality calibrates governance intensity to actual risk levels; Stakeholder Consideration engages affected parties meaningfully throughout system lifecycles; and Iterative Validation enables continuous adaptation based on observed outcomes. Through integration with established regulatory frameworks including the EU AI Act and NIST AI Risk Management Framework, and through a worked deployment scenario that traces the framework across a complete governance decision, the 0/0 Framework provides organizations with practical governance structures intended to reduce failure rates while supporting ethical deployment of powerful AI capabilities.

General Terms

AI Governance, Risk Management, Responsible AI

Keywords

0/0 Framework, AI Governance, Generative AI, Synthetic Media, Risk Proportionality, Value Alignment, EU AI Act, NIST AI RMF

1. INTRODUCTION

The promise of artificial intelligence has captivated organizations across every sector, driving unprecedented investment in AI capabilities. Global AI spending reached approximately 150 billion dollars in 2024 and continues to grow at double-digit annual rates. Yet this enthusiasm has not translated proportionally into realized value. Industry surveys consistently report AI project failure rates between 70 and 85 percent, with Gartner, MIT Sloan Management Review, and Boston Consulting Group all documenting similar patterns across diverse organizational contexts [3, 4, 5]. These failures represent not merely technological disappointments but substantial organizational costs: wasted investment, damaged stakeholder relationships, missed competitive opportunities, and in some cases direct harms to individuals and communities affected by poorly governed AI systems.

The causes of these failures are increasingly well understood, and they are not primarily technical. Organizations possess or can acquire technical capabilities sufficient for most AI applications. Failures occur upstream: inadequate alignment between AI capabilities and organizational values, inappropriate risk tolerance for specific applications, insufficient stakeholder engagement during development and deployment, and rigid deployment approaches that cannot adapt to emerging issues. These governance deficiencies create predictable failure modes that sophisticated technology cannot overcome.

These failure modes are amplified in generative AI contexts. The Authenticity Spectrum Framework [1] demonstrates that synthetic media outputs span from beneficial applications (Levels 1 and 2: entertainment, accessibility) to harmful misuse (Levels 4 and 5: fraud, manipulation). The Marketing-Fraud Convergence analysis [2] reveals that identical tools serve both legitimate marketing purposes and criminal fraud operations. The same large language model powering customer service automation can generate convincing phishing emails. The same avatar technology creating corporate training videos can enable executive impersonation fraud. Governance frameworks must address not only intended applications but foreseeable misuse patterns.

This paper introduces the 0/0 Framework, a governance model designed to address root causes of AI project failure while enabling responsible deployment of generative AI capabilities. The framework name reflects its dual objective: achieving zero preventable harms through rigorous governance while maintaining zero tolerance for governance shortcuts that enable such harms.

1.1 Research Objectives

This research pursues four primary objectives. First, to analyze patterns of AI project failure and identify governance-addressable root causes that existing frameworks inadequately address. Second, to develop a governance framework operationalizing responsible AI principles for practical deployment decisions. Third, to demonstrate framework alignment with existing regulatory requirements, enabling compliance integration rather than parallel governance structures. Fourth, to provide assessment tools for organizations evaluating governance maturity and identifying improvement opportunities.

1.2 Framework Nomenclature

The 0/0 Framework name reflects its dual objective through mathematical notation. The first zero represents zero preventable harms: eliminating negative outcomes that proper governance

could have prevented. The second zero represents zero tolerance for governance shortcuts, the recognition that expedient compromises in governance processes create conditions for preventable failures. The notation deliberately evokes the indeterminate form 0/0 in mathematics, signaling that governance outcomes are undefined until an organization supplies the substantive values, risk thresholds, and stakeholder commitments that resolve them.

2. BACKGROUND: AI PROJECT FAILURE

2.1 Failure Rate Evidence

Multiple independent research efforts have documented high AI project failure rates across organizational contexts. Gartner research indicates that through 2025, 85 percent of AI projects will deliver erroneous outcomes due to bias in data, algorithms, or the teams responsible for managing them [3]. MIT Sloan Management Review surveys find that only 10 percent of companies report substantial financial benefits from AI investments, despite widespread adoption efforts [4]. BCG analysis suggests that 70 percent of digital transformation initiatives fail to reach their stated goals [5]. Table 1 consolidates these findings and maps each reported failure pattern to the governance root cause it most directly reflects.

Table 1: Documented AI Failure Patterns and Governance Root Causes

Source	Reported Rate	Primary Governance Root Cause
Gartner [3]	85% erroneous outcomes	Risk miscalibration (data and algorithmic bias)
MIT Sloan [4]	90% no material benefit	Value misalignment (capability without strategic fit)
BCG [5]	70% miss stated goals	Stakeholder neglect and deployment rigidity

2.2 Root Cause Analysis

Analysis of AI project failures reveals consistent governance-related root causes. Value misalignment occurs when AI capabilities are deployed for purposes misaligned with organizational values, stakeholder expectations, or societal norms. Risk miscalibration involves inadequate assessment of potential negative consequences. Stakeholder neglect encompasses failure to engage affected parties in development and deployment decisions. Deployment rigidity describes inability to adapt deployments based on emerging evidence. Each root cause corresponds directly to one of the four pillars introduced in Section 4, and this correspondence is the organizing logic of the framework: every pillar exists to neutralize a documented failure mode rather than to express an abstract principle.

2.3 Existing Framework Limitations

Current AI governance approaches exhibit limitations. Principle-based frameworks such as IEEE Ethically Aligned Design and the Asilomar AI Principles provide valuable direction but lack operational specificity. Compliance-focused frameworks such as the EU AI Act requirements establish regulatory floors but do not guide organizations toward excellence beyond minimum compliance. Technical frameworks such as model cards and datasheets for datasets address specific artifacts but not end-to-end governance processes. The 0/0 Framework integrates these approaches into a cohesive operational model, summarized structurally in Figure 1.

3. THE 0/0 FRAMEWORK

The 0/0 Framework operationalizes responsible AI governance through four interconnected pillars: Value Alignment, Risk Proportionality, Stakeholder Consideration, and Iterative Validation. Each pillar addresses a specific failure mode, and together they provide full governance coverage. Figure 1 presents the framework architecture, showing how organizational inputs flow through the four pillars to produce governed deployment decisions, with a feedback loop returning observed outcomes to the governance process.

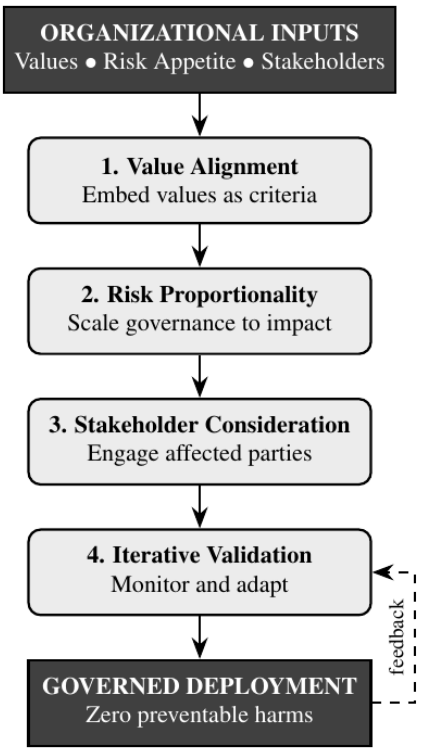

Figure 1: The 0/0 Framework architecture. Organizational inputs pass through four sequential pillars to produce governed deployment. The dashed feedback path returns observed outcomes to Iterative Validation, making governance continuous rather than one-time.

3.1 Pillar 1: Value Alignment

Value Alignment requires explicit articulation of values that AI systems should embody and verification that system behavior aligns with those values throughout development and deployment. This pillar addresses failures where technically functional systems violate organizational, stakeholder, or societal values.

Implementation requires three components. Value articulation demands that organizations explicitly state the values their AI systems should embody as operational criteria. Value operationalization translates stated values into measurable system behaviors. Alignment verification continuously assesses whether deployed systems behave consistently with stated values.

3.2 Pillar 2: Risk Proportionality

Risk Proportionality ensures that governance intensity scales appropriately with potential impact, preventing both over-governance that stifles beneficial innovation and under-governance that permits preventable harm.

Implementation requires systematic risk assessment providing evaluation of potential negative consequences across affected parties, time horizons, and reversibility dimensions. The Dual-Use Risk Assessment Framework [2] provides criteria for evaluating synthetic content technologies. Table 2 presents the Risk Proportionality Governance Matrix, which binds each risk tier to a specific governance response and an explicit escalation trigger.

Table 2: Risk Proportionality Governance Matrix

Risk Level	ASF Level	Governance	Escalation Trigger
Minimal	Level 1	Standard review	None
Limited	Level 2	Enhanced documentation	Scope change
Elevated	Level 3	Committee approval	Any incident
High	Level 4-5	Executive oversight, external audit	Pre-deployment

3.3 Pillar 3: Stakeholder Consideration

Stakeholder Consideration ensures meaningful engagement with parties affected by AI systems throughout development and deployment. Implementation requires stakeholder identification systematically identifying all parties potentially affected by AI systems, engagement mechanisms establishing appropriate channels for stakeholder input, and impact monitoring tracking actual stakeholder experiences and outcomes.

3.4 Pillar 4: Iterative Validation

Iterative Validation establishes continuous monitoring, feedback integration, and adjustment based on observed outcomes rather than one-time approval processes. Implementation requires monitoring systems providing continuous observation of system behavior and outcomes, feedback integration creating mechanisms for incorporating stakeholder reports, and adaptive response enabling rapid deployment modifications when monitoring reveals issues.

4. WORKED EXAMPLE: SYNTHETIC SPOKESPERSON DEPLOYMENT

To demonstrate the framework in operation rather than in the abstract, this section traces a single realistic deployment decision through all four pillars. The scenario concerns a mid-sized financial services firm proposing to deploy a synthetic AI spokesperson, a photorealistic avatar generating personalized video messages for customer onboarding. This case is selected because it sits squarely on the marketing-fraud convergence boundary identified in prior work [2]: the same avatar technology that improves onboarding can enable executive impersonation fraud.

4.1 Step 1: Value Alignment Assessment

The firm articulates three operative values for the deployment: truthful communication, identity transparency, and customer autonomy. Value operationalization translates these into measurable behaviors. Truthful communication requires that every synthetic video carries a visible disclosure that the spokesperson is AI-generated. Identity transparency forbids the avatar from resembling any real employee. Customer autonomy requires an opt-out to human contact in every message. The deployment fails initial alignment because the original design used a real executive's likeness, violating identity transparency, and is returned for redesign before any risk assessment proceeds.

4.2 Step 2: Risk Proportionality Classification

After redesign, the avatar is classified using Table 2. Because it generates synthetic human likeness for financial communications, it maps to Authenticity Spectrum Level 4 [1], placing it in the High risk tier. This triggers executive oversight and external audit pre-deployment, rather than the standard review that a Level 1 internal productivity tool would receive. The proportionality principle prevents two errors at once: it stops the firm from waving the avatar through under a generic software review, and it spares low-risk tools from the same heavy process.

4.3 Step 3: Stakeholder Consideration

Affected parties are enumerated: customers receiving messages, the employees whose roles the avatar partially automates, the compliance function, and the broader public exposed to potential impersonation if the model leaks. Engagement reveals a concern the technical team had not anticipated. Compliance staff note that synthetic video of financial advice may trigger disclosure obligations under financial-promotion rules, a requirement absent from the original project scope. This finding feeds directly back into Value Alignment, tightening the disclosure behavior defined in Step 1.

4.4 Step 4: Iterative Validation

Deployment proceeds with monitoring instrumented from day one. The validation plan specifies three tracked signals: customer comprehension of the AI disclosure, measured through a follow-up prompt; rate of opt-out to human contact; and any external reports of the avatar appearing outside sanctioned channels. The escalation trigger from Table 2 for the High tier specifies that any single impersonation report halts the deployment pending review. Table 3 summarizes the trace.

Table 3: Worked Example Traced Across the Four Pillars

Pillar	Action	Outcome
Value Alignment	Define 3 values; check design	Redesign forced (likeness)
Risk Proportionality	Classify via ASF	Level 4, High tier, audit
Stakeholder	Enumerate and engage	New disclosure duty found
Iterative Validation	Instrument monitoring	Halt trigger on impersonation

The example illustrates the framework's central claim in miniature. No pillar in isolation would have produced a responsible deployment. Value Alignment alone would have missed the regulatory disclosure duty; risk classification alone would have permitted the original likeness-based design; stakeholder engagement alone would have lacked an enforcement mechanism. The interlocking structure, and specifically the feedback paths between pillars, is what converts four reasonable principles into a governance process.

5. REGULATORY INTEGRATION

5.1 EU AI Act Integration

The EU AI Act establishes risk-based regulatory requirements that align naturally with 0/0 Framework Risk Proportionality [6]. The Act's risk categorization, spanning unacceptable, high-risk, limited, and minimal, maps to governance intensity calibration. High-risk AI systems under the Act require conformity assessment, post-market monitoring, and incident reporting,

requirements that 0/0 Iterative Validation operationalizes through continuous monitoring and adaptive response mechanisms.

5.2 NIST AI RMF Alignment

The NIST AI Risk Management Framework organizes AI governance around four functions: Govern, Map, Measure, and Manage [7]. The 0/0 Framework pillars map systematically, as shown in Table 4.

Table 4: Regulatory Framework Alignment

0/0 Pillar	EU AI Act	NIST RMF
Value Alignment	Art. 9, 10	Govern
Risk Proportionality	Art. 6, 7	Map, Measure
Stakeholder	Art. 13, 14	Govern
Iterative Validation	Art. 61, 62	Manage

6. IMPLEMENTATION GUIDANCE

6.1 Governance Maturity Assessment

Organizations should assess current governance maturity across each pillar before implementing the 0/0 Framework. Assessment dimensions include policy existence, process implementation, monitoring effectiveness, and adaptive capacity.

6.2 Phased Implementation

Implementation should proceed in phases appropriate to organizational context. The Foundation phase establishes basic policies and processes across all pillars, prioritizing highest-risk applications. The Enhancement phase develops more sophisticated capabilities including automated monitoring and stakeholder advisory mechanisms. The Optimization phase achieves full framework operation with continuous improvement processes.

6.3 Success Metrics

Organizations should track both process and outcome metrics. Process metrics include governance coverage, assessment completion rates, and monitoring system uptime. Outcome metrics include incident rates, time-to-detection for issues, stakeholder satisfaction, and regulatory compliance status.

7. DISCUSSION

7.1 Framework Contributions

The 0/0 Framework contributes to AI governance literature by providing an operational model that addresses documented failure patterns while enabling responsible deployment of powerful generative AI capabilities. Unlike principle-only frameworks, it provides specific implementation guidance through the four-pillar structure. Unlike compliance-only frameworks, it guides organizations toward excellence beyond minimum requirements. The worked example in Section 4 demonstrates that the pillars function as an interlocking process rather than a checklist, with findings in later pillars feeding back to refine earlier ones.

The framework's integration with prior work on synthetic media risk classification [1] and dual-use convergence [2] demonstrates applicability to the specific challenges posed by generative AI technologies.

7.2 Limitations and Future Work

Several limitations constrain the current framework. Implementation requires organizational commitment and resources that not all organizations possess. The framework provides structure but not content; organizations must still make substantive decisions about values, risk tolerance, and stakeholder prioritization. The worked example, while realistic, is illustrative rather than empirical, and does not by itself establish effectiveness.

Future work should proceed along four lines. First, empirical validation through controlled deployment studies that compare failure rates between teams using the framework and matched controls. Second, development of simplified implementation paths for resource-constrained organizations such as small enterprises and public-sector bodies. Third, automated tooling that instruments the Iterative Validation pillar, turning the monitoring signals described in Section 4 into a reusable software layer. Fourth, sector-specific extensions, beginning with financial services, where the marketing-fraud convergence is most acute, and extending to healthcare and education where the value-alignment requirements differ materially.

8. CONCLUSION

AI project failures at 70 to 85 percent rates represent not technical limitations but governance deficiencies that organizations can address through systematic frameworks. The 0/0 Framework addresses these deficiencies through four interconnected pillars: Value Alignment requires AI systems to embody organizational values; Risk Proportionality calibrates governance to actual risk levels; Stakeholder Consideration engages affected parties meaningfully; and Iterative Validation enables continuous adaptation based on observed outcomes. The worked synthetic-spokesperson deployment in Section 4 shows these pillars operating together on a single realistic decision, where no individual pillar would have sufficed and the connective feedback between them produced the responsible outcome.

By operationalizing these four pillars in integration with prior risk classification [1] and dual-use assessment [2] frameworks, organizations can deploy generative AI capabilities responsibly while avoiding governance failures that afflict most AI initiatives. As generative AI capabilities continue advancing, governance frameworks must evolve correspondingly, and the future research agenda set out above, spanning empirical validation, lightweight implementation, automated monitoring, and sector-specific extension, marks the path. The 0/0 Framework provides a foundation for responsible AI deployment that neither stifles beneficial innovation nor permits preventable harms.

9. REFERENCES

[1] F. Martinson, "The Authenticity Spectrum Framework: Classifying Deepfake and Generative AI Risks in Synthetic Media," International Journal of Computer Applications, vol. 187, no. 88, pp. 34-38, 2026. DOI: 10.5120/ijca2026926538

[2] F. Martinson, "The Marketing-Fraud Convergence: When Legitimate AI Tools Enable Financial Crime," International Journal of Computer Applications, vol. 187, no. 98, pp. 1-5, 2026. DOI: 10.5120/ijcaa6160f6a9822

[3] Gartner, "Predicts 2025: AI and the Future of Work," Gartner Research, 2024.

[4] S. Ransbotham et al., "Achieving Individual and Organizational Value with AI," MIT Sloan Management Review, 2023.

[5] Boston Consulting Group, "Most Digital Transformations Fail," BCG Henderson Institute, 2024.

[6] European Union, "Regulation (EU) 2024/1689 (AI Act)," Official Journal of the European Union, 2024.

[7] National Institute of Standards and Technology, "AI Risk Management Framework," NIST AI 100-1, 2023.

[8] L. Floridi et al., "AI4People-An Ethical Framework for a Good AI Society," Minds and Machines, vol. 28, no. 4, pp. 689-707, 2018.

[9] A. Jobin et al., "The global landscape of AI ethics guidelines," Nature Machine Intelligence, vol. 1, no. 9, pp. 389-399, 2019.

[10] B. Mittelstadt, "Principles alone cannot guarantee ethical AI," Nature Machine Intelligence, vol. 1, no. 11, pp. 501-507, 2019.

[11] T. Hagendorff, "The ethics of AI ethics: An evaluation of guidelines," Minds and Machines, vol. 30, no. 1, pp. 99-120, 2020.

[12] I. D. Raji et al., "Closing the AI Accountability Gap," FAT* '20 Proceedings, 2020.

[13] N. A. Smuha, "From a Race to AI to a Race to AI Regulation," Law, Innovation and Technology, vol. 13, no. 1, pp. 57-84, 2021.

[14] M. Whittaker et al., "AI Now Report 2018," AI Now Institute, 2018.

[15] E. Brynjolfsson and A. McAfee, "The Business of Artificial Intelligence," Harvard Business Review, 2017.