World Journal of Advanced Research and Reviews
eISSN: 2581-9615 CODEN (USA): WJARAI
Cross Ref DOI: 10.30574/wjarr
Journal homepage: https://wjarr.com/

(REVIEW ARTICLE)

A definitive forecast on the regulation of generative artificial intelligence for the years 2025 to 2035

Jimmy Kinyonyi Bagonza [1,*] and Mathias Ndungu [2]

[1] *School of Computer and Information Sciences, University of the Cumberlands, Kentucky USA.*
[2] *Kogod School of Business, American University, Washington DC, USA.*

World Journal of Advanced Research and Reviews, 2026, 31(01), 523–532

Publication history: Received on 30 May 2026; revised on 06 July 2026; accepted on 08 July 2026

Article DOI: https://doi.org/10.30574/wjarr.2026.31.1.1855

Abstract

Generative artificial intelligence (AI) presents significant governance challenges due to its rapid technological advancement, emergent behavior, and the inability of existing regulatory institutions to keep pace with its development. This paper presents a qualitative, scenario-based forecast of the global regulation of generative AI between 2025 and 2035. Drawing on scenario planning, horizon scanning, and evidence from AI safety research, digital governance theory, and policy analytics, the study identifies two critical uncertainties shaping future regulatory outcomes: the level of international regulatory coordination and the pace of AI safety and interpretability advances. These uncertainties form the basis for two contrasting scenarios: Regulated Convergence, characterized by strengthened international cooperation, consolidated oversight institutions, and augmented governance supported by technical verification mechanisms; and Fragmented Acceleration, in which geopolitical competition, institutional fragmentation, and limited progress in AI safety hinder effective governance. The analysis evaluates the policy trade-offs associated with each scenario, including regulatory capacity, licensing regimes, transparency requirements, international coordination, and open-source AI development. The findings suggest that effective governance by 2035 will depend on sustained institutional investment, adaptive regulatory frameworks, technical advances in AI safety, and greater international collaboration. Without these conditions, fragmented governance is likely to remain the dominant trajectory, increasing regulatory gaps and societal risks associated with frontier AI systems.

Keywords: Generative artificial intelligence; AI governance; Scenario planning; AI regulation; Digital governance; AI safety; International regulatory coordination.

1. Introduction

Generative artificial intelligence (AI), including large language models (LLMs) such as GPT-4, Gemini, and Claude, as well as multimodal systems capable of generating text, images, audio, and code, has emerged as one of the most significant general-purpose technologies of this decade. The commercial deployment of these systems has occurred at a pace that exceeds the ability of legislative and regulatory institutions to respond, creating a governance gap with significant implications for society, the economy, and national security (Judge et al., 2024; Asensio et al., 2025). At the individual level, generative AI raises concerns relating to privacy, intellectual property, misinformation, and algorithmic bias. At the institutional level, these technologies challenge conventional regulatory oversight because their emergent behavior cannot be audited using traditional compliance methods.

The governance of generative artificial intelligence cannot be separated from electronic data governance because AI systems rely on the collection, processing, storage, and reuse of digital data throughout their lifecycle. Consequently, weaknesses in data protection, privacy, ethical oversight, and cross-border data governance become weaknesses in AI

* Corresponding author: Jimmy Kinyonyi Bagonza

governance itself. As digital technologies continue to evolve, effective AI regulation must therefore be complemented by robust data protection frameworks that safeguard individual rights while promoting responsible innovation (Asensio et al., 2025, Ndungu 2026).

Furthermore, geopolitical competition, particularly between the United States and China, has contributed to divergent national regulatory approaches, complicating prospects for coordinated global governance (Judge et al., 2024; Asensio et al., 2025).

1.1. Research Problem

Despite ongoing regulatory initiatives, including Executive Order 14110 (The White House, 2023), the European Union AI Act (European Parliament, 2024), and the NIST AI Risk Management Framework (NIST, 2023), considerable uncertainty remains regarding the institutional conditions necessary for establishing effective governance of generative AI. Existing quantitative indicators, such as market adoption and computational scaling, provide limited insight into future regulatory trajectories because legislative developments, geopolitical coordination, and advances in AI safety are shaped by political and institutional processes rather than stable historical trends. Consequently, forecasting AI regulation requires an approach capable of addressing high uncertainty and complex governance dynamics. Although several regulatory frameworks and governance proposals have been developed, there remains limited understanding of how differing levels of international coordination and advances in AI safety may shape regulatory outcomes over the next decade (Judge et al., 2024; Asensio et al., 2025).

1.2. Research Objective

This study presents a qualitative scenario-based forecast of global governance for generative AI over the period 2025–2035. Specifically, it identifies the principal drivers and uncertainties influencing future regulatory development, constructs two empirically grounded scenarios based on varying levels of international regulatory coordination and AI safety progress, and examines the governance outcomes and policy implications associated with each scenario.

1.3. Contribution of the Study

This study contributes to the growing literature on AI governance by providing an evidence-based scenario analysis of alternative regulatory futures for generative AI. Drawing on AI safety research, digital governance theory, and policy analytics, it offers a structured framework for understanding how technological progress, institutional capacity, and international cooperation may shape regulatory outcomes between 2025 and 2035. The study further identifies the policy trade-offs associated with different governance pathways and highlights institutional mechanisms that may support more effective and adaptive regulatory frameworks for generative AI (Judge et al., 2024; Hanisch et al., 2023; Asensio et al., 2025).

2. Analytical Framework and Methodology

2.1. Forecasting Focus

The central forecasting question guiding this study is: What conditions and institutional mechanisms will lead to the establishment of effective regulatory frameworks for generative AI by 2035, and what governance outcomes can be anticipated for various stakeholder groups?

The 2025–2035 time frame is analytically appropriate for several reasons. Current regulatory actions, including the Executive Order on AI (The White House, 2023), the implementation of the EU AI Act (European Parliament, 2024), and the NIST AI Risk Management Framework (NIST, 2023), are already in motion and provide essential anchoring conditions. A ten-year period is adequate to observe primary regulatory feedback effects and policy successes or failures that can foster political coalitions for further interventions without requiring forecasts that extend beyond the present expert consensus on AI capability trajectories.

2.2. Analytical Framework

This study draws upon three complementary analytical frameworks to examine future regulatory pathways for generative artificial intelligence.

The first framework is scenario planning, following Schwartz (1996) and the Global Business Network. This approach focuses on developing internally consistent alternative futures grounded in identifiable critical uncertainties. Scenario

planning provides a structured method for exploring plausible governance outcomes under conditions of high uncertainty rather than attempting deterministic prediction.

The second framework is the digital governance typology proposed by Hanisch et al. (2023), which distinguishes among analog, augmented, and automated governance modes. This framework provides the conceptual vocabulary for evaluating how regulatory architectures interact with digital technologies and for characterizing the mismatch between existing governance systems and increasingly automated AI technologies.

The third framework is derived from the AI safety literature synthesized by Judge et al. (2024). This framework grounds normative claims regarding regulatory adequacy in the technical characteristics of generative AI systems and emphasizes the challenges posed by their general-purpose nature, emergent behavior, contested safety values, and limited compatibility with conventional compliance auditing.

Together, these three frameworks provide the conceptual foundation for examining alternative governance futures and evaluating the institutional conditions under which effective regulation of generative AI may emerge.

2.3. Evidence Base and Sources

The evidence base for this analysis is drawn from peer-reviewed scholarship, regulatory documents, and policy-relevant technical literature. Judge et al. (2024) provide the primary regulatory framework by establishing the need for an adapted regulatory paradigm that reflects the emergent, black-box nature of generative AI systems. Their work identifies five principal regulatory challenges associated with AI, including its general-purpose nature, emergent behavior, contested human values, risks associated with capabilities exceeding human oversight, and limited governmental leadership in frontier model development.

Asensio et al. (2025) provide an empirical review of data technologies and policy analytics, highlighting the increasing role of machine learning, Internet of Things technologies, digital twins, and distributed ledger technologies in public policy. Their analysis identifies model explainability, cross-sector collaboration, and data accessibility as central considerations for effective AI governance and demonstrates that technological developments continue to outpace regulatory responses.

Hanisch et al. (2023) contribute a conceptual governance framework distinguishing analog, augmented, and automated governance modes. Their typology explains the structural mismatch between conventional regulatory institutions and increasingly automated technologies, providing an important lens for evaluating future governance pathways. Additional evidence is drawn from the EU AI Act, the NIST AI Risk Management Framework, the Bletchley Declaration, Executive Order 14110, and scholarship on AI governance fragmentation and technology regulation.

2.4. Scenario Development

The central forecasting question is decomposed into two critical uncertainties that serve as the scenario axes: (1) the level of international regulatory coordination, ranging from fragmentation to convergence, and (2) the pace of AI safety and interpretability research, ranging from stagnation to breakthrough. Each uncertainty is grounded in empirical evidence rather than analytical convention. These two axes define the scenario space from which two representative and empirically plausible scenarios are developed.

The first uncertainty concerns the trajectory of international regulatory coordination. While initiatives such as the Bletchley Declaration and the EU AI Act demonstrate that rhetorical convergence is achievable, binding international coordination remains uncertain because of competing national interests and institutional incentives (Büthe et al., 2022; Taeihagh et al., 2021). The second uncertainty concerns the pace of AI safety and interpretability breakthroughs. Progress sufficient to enable formal verification would substantially transform the regulatory landscape, whereas continued limitations in understanding frontier AI systems would reinforce existing governance challenges (Judge et al., 2024). The interaction between these uncertainties generates the scenario space from which the Regulated Convergence and Fragmented Acceleration scenarios are developed.

3. Key Drivers and Critical Uncertainties

3.1. Technological Drivers

The primary technological driver is the rapid scaling of foundation models. Successive generations have revealed emergent capabilities and competencies not predicted from smaller-scale performance that complicate safety testing

and regulatory assessment. As Judge et al. (2024) point out, dangerous capabilities can arise unpredictably, making exhaustive testing infeasible. Compute requirements for frontier models have doubled approximately every six to nine months. A countervailing trend is progress in mechanistic interpretability research, which aims to render neural network internal representations comprehensible to human analysts; however, this research remains nascent relative to the complexity of frontier systems. This technological dynamic directly grounds Scenario Axis 2 (the pace of safety and interpretability progress): the gap between interpretability research and frontier capability is the empirical basis for treating safety breakthroughs as genuinely uncertain rather than assured.

3.2. Economic and Market Drivers

The economics of foundation model development demonstrate strong monopolistic tendencies (Judge et al., 2024). The substantial capital required to train frontier models consolidates capability within a small group of large technology firms, primarily in the United States and China. This concentration creates structural barriers to regulatory effectiveness, as regulated entities possess technical expertise and economic influence exceeding those of potential regulators, a dynamic Asensio et al. (2025) document in data governance more broadly, where digital data sources concentrate in private platforms often inaccessible to government agencies.

At the same time, a growing open-source ecosystem, illustrated by Meta's LLaMA model and the Hugging Face platform, distributes powerful models beyond proprietary licensing regimes, complicating enforcement. The concentration of capability in a few actors while simultaneously allowing proliferation through open-source channels is a direct empirical basis for Scenario Axis 1 (international coordination): coordination is harder to achieve when the regulated entities have both the incentive and the capacity to resist binding obligations.

3.3. Political and Institutional Drivers

The ongoing geopolitical rivalry between the United States and China over AI supremacy is decisively shaping national regulatory frameworks. Both nations have articulated national AI strategies aimed at technological leadership, creating pressure against regulatory constraints perceived as competitive disadvantages. The European Union has adopted a robust regulatory approach through the AI Act (European Parliament, 2024), establishing comprehensive risk-based regulation with extraterritorial reach, a posture poised to generate regulatory spillovers analogous to the Brussels Effect observed under the GDPR.

However, international coordination mechanisms remain weak. The Bletchley Declaration of November 2023 forged rhetorical consensus on AI safety but created no binding obligations or enforcement mechanisms (UK Government et al., 2023). This institutional gap, meaningful agreement on the diagnosis but the absence of agreement on the remedy, directly anchors Axis 1. The current default is fragmentation, and convergence requires specific catalyzing conditions that are not yet present.

3.4. Critical Uncertainties

Two uncertainties dominate the scenario space, each rooted in identifiable empirical trends rather than analytical convention.

The first uncertainty is the trajectory of international regulatory coordination. The Bletchley Declaration (UK Government et al., 2023) and the European Union AI Act (European Parliament, 2024) establish that rhetorical convergence is achievable, but binding coordination remains elusive. The structural incentives documented by Buthe et al. (2022) and Taeihagh et al. (2021) explain why the degree of fragmentation or convergence by 2035 remains genuinely uncertain because the causal mechanism, a high-salience catalyzing event generating political consensus cannot be scheduled or predicted.

The second uncertainty is the pace of AI safety and interpretability breakthroughs. Judge et al. (2024) demonstrate that the black-box nature of current AI architectures is fundamentally incompatible with conventional compliance auditing. Progress sufficient to enable formal verification would substantially transform the regulatory landscape. These uncertainties are partially independent: safety breakthroughs could facilitate convergence by providing a common technical basis for standards, while geopolitical pressures could frustrate coordination even if technical tools improved. The interaction between these uncertainties generates the four-quadrant scenario space from which the two representative scenarios are developed.

4. Findings and Scenario Analysis

The interaction between the two critical uncertainties, the level of international regulatory coordination and the pace of AI safety and interpretability advances, generates two representative scenarios for the governance of generative artificial intelligence between 2025 and 2035. These scenarios are not predictions. Rather, they represent reasonable futures derived from current empirical trends, existing regulatory developments, and the analytical frameworks presented in this study. Together, they illustrate how different combinations of technological progress and institutional responses may shape the evolution of AI governance.

4.1. Scenario 1: Regulated Convergence (2025–2035)

4.1.1. Scenario Overview

In this scenario, a combination of catalyzing events, institutional learning, and technological progress drives meaningful international regulatory coordination and the gradual establishment of effective AI governance frameworks. The scenario occupies the upper-right quadrant of the uncertainty space, where moderate safety progress (Axis 2) intersects with improving international coordination (Axis 1). The empirical basis for this combination is that moderate interpretability advances provide regulators with credible shared technical standards, resolving the definitional problem identified by Judge et al. (2024), while a high-salience AI-related harm generates the political consensus that the Bletchley process failed to produce.

4.1.2. Governance Characteristics

From 2025 to 2027, the EU AI Act enters full implementation, exerting regulatory spillover pressure on firms operating within EU markets. The United States, under pressure from civil society coalitions and congressional committees, enacts a federal AI accountability framework drawing on the NIST AI Risk Management Framework (NIST, 2023), but adding mandatory disclosure, licensing, and incident reporting for frontier model developers. A bilateral US-EU working group on AI standards establishes common definitions for high-risk AI applications and shared prohibitions, including autonomous lethal systems, non-consensual biometric surveillance, and systematic deception of individuals. The evidentiary basis for this phase is the existing EU AI Act risk-classification architecture and the NIST framework's governance vocabulary, both of which provide the definitional scaffolding that negotiators would require.

Between 2027 and 2030, progress in mechanistic interpretability yields auditable formal specifications for constrained AI behaviors. Regulators in the United States, the European Union, and the United Kingdom begin requiring frontier model developers to demonstrate formal verification of red-line prohibitions as a condition of deployment licensing, an approach analogous to the FAA's airworthiness directives, which Judge et al. (2024) identify as the closest regulatory analogue. A consolidated AI safety agency is established in the United States, modeled on the lifecycle authority of the FAA and NRC, with authority spanning from pretraining data governance through deployment monitoring. An international registry of large models, proposed by Hadfield et al. (as cited in Judge et al., 2024), provides regulators with visibility into frontier development.

4.1.3. Expected Outcomes

By 2030 to 2035, the regulatory landscape has evolved into what Hanisch et al. (2023) characterize as augmented governance. Human oversight institutions are supported by automated monitoring tools, model registries, and incident-reporting systems, while retaining the ultimate authority to withdraw unsafe systems from deployment. Cross-sector collaboration among government agencies, academic researchers, and civil society, a priority Asensio et al. (2025) identify as essential for trustworthy data-driven governance, becomes institutionalized through formal advisory structures. The governance mode has shifted from analog to augmented because regulatory institutions invested in the automated compliance tools that make scaled oversight feasible.

4.2. Scenario 2: Fragmented Acceleration (2025–2035)

4.2.1. Scenario Overview

In this scenario, geopolitical competition, institutional gridlock, and the concentration of AI capability in private firms prevent the establishment of effective regulatory governance. This scenario occupies the lower-left quadrant of the uncertainty space, where safety research stagnates (Axis 2) while international coordination fragments further (Axis 1). The empirical basis is the continuation of structural conditions already observed, including the monopolistic concentration documented by Judge et al. (2024), the enforcement limitations inherent in the EU AI Act's extraterritorial

reach when confronted with non-compliant actors, and the open-source proliferation that distributes capable models beyond any licensing regime.

4.2.2. Governance Characteristics

From 2025 to 2027, the EU AI Act imposes substantial compliance burdens on European firms while larger American and Chinese developers leverage political influence to weaken enforcement provisions. Federal governance of AI in the United States remains fragmented across existing agencies, including the FTC, NHTSA, FDA, and SEC, each regulating AI applications within its respective jurisdiction and lacking the consolidated oversight authority that Judge et al. (2024) identify as essential. China's approach, characterized by government oversight of generative AI service providers but permissiveness toward state-directed applications, stands in sharp contrast to Western regulatory frameworks. The Bletchley process produces ministerial declarations but no binding treaty obligations, consistent with the structural incentive analysis of Büthe et al. (2022).

Between 2027 and 2030, the open-source AI ecosystem grows substantially, with capable models distributed through platforms that regulators struggle to monitor effectively. Jailbreaking and fine-tuning techniques that remove safety guardrails become routine, as anticipated by Judge et al. (2024). A series of AI-enabled harms, including disinformation campaigns affecting electoral integrity, AI-assisted fraud at scale, and autonomous systems causing physical harm, generates public concern but fails to produce durable legislative coalitions because the affected industries successfully argue that regulation would disadvantage national competitiveness. Regulatory responses remain sector-specific, voluntary, and uncoordinated, reflecting the broader pattern documented by Asensio et al. (2025), in which private-sector actors consistently outpace public regulatory capacity.

4.2.3. Expected Outcomes

By 2030 to 2035, the regulatory landscape remains analog in the terminology of Hanisch et al. (2023). Governance mechanisms continue to be bilateral, bureaucratic, and actor-based, while the AI systems they are intended to govern become increasingly automated and autonomous. The mismatch between governance structures and technological complexity widens rather than narrows. Asensio et al.'s (2025) warning that purely predictive, non-causal approaches can erode trust in governance is reflected in the growing use of algorithmic decision-making in consequential domains without meaningful public accountability. Furthermore, the historical lag between technological development and effective regulatory response, documented by Judge et al. (2024) as averaging several decades in the United States, persists throughout the forecast period.

5. Policy Implications

5.1. Implications under Regulated Convergence

Within the Regulated Convergence framework, the primary policy recommendations follow from the FAA and NRC analogues identified by Judge et al. (2024): consolidate oversight authority, mandate formal verification, and require independent monitoring with intervention capacity. However, each mechanism generates countervailing costs that policymakers must explicitly manage.

Licensing regimes for frontier model developers condition market access on demonstrated compliance with formal safety specifications verified by independent technical auditors. The principal benefit is a structural enforcement mechanism that self-certification cannot provide, directly addressing the information asymmetry between regulators and regulated entities. The trade-off, however, is significant. Licensing regimes favor large incumbent developers that can absorb compliance costs and may entrench the monopolistic concentration documented by Judge et al. (2024). Policymakers should therefore design tiered licensing thresholds, triggered by compute scale, deployment scope, or capability benchmarks, to avoid converting safety mandates into competitive advantages for incumbent firms.

Mandatory training data and model architecture disclosures reduce information asymmetries and enable independent audits. However, disclosure requirements create tension between transparency and intellectual property protection. Developers are likely to resist disclosures that compromise proprietary architectures, while detailed knowledge of model failure modes could be exploited by adversarial actors. Disclosure requirements should therefore be structured as confidential regulatory submissions to technically qualified auditors, analogous to pharmaceutical ingredient filings with the FDA, rather than public disclosures.

International coordination through treaty-based red-line prohibitions provides the most politically feasible form of binding agreement. The principal trade-off is jurisdictional arbitrage, where firms relocate development activities to

jurisdictions with weaker enforcement. Consequently, red-line prohibitions are most effective when paired with market access conditions that prevent systems developed outside participating treaty frameworks from being sold or deployed within participating jurisdictions.

The transition from analog to augmented governance, as described by Hanisch et al. (2023), requires sustained investment in automated compliance monitoring infrastructure. The associated trade-off is between enhanced regulatory capacity and increased surveillance risk. Regulatory design should therefore include independent oversight of the monitoring infrastructure itself to prevent abuse while maintaining public confidence in AI governance.

5.2. Implications under Fragmented Acceleration

In the Fragmented Acceleration scenario, policymakers confront the more difficult challenge of preventing catastrophic harm under conditions of institutional gridlock. The available regulatory instruments are more limited but involve significant trade-offs.

Sector-specific regulators can extend existing authorities to require safety evaluations, incident reporting, and mandatory recalls without waiting for comprehensive federal legislation. While this approach enables immediate regulatory action through agencies such as the FDA, NHTSA, and SEC, it also creates fragmented governance, jurisdictional gaps for general-purpose AI systems, inconsistent safety standards, and greater barriers to international coordination. The analysis therefore treats sector-specific regulation as a transitional rather than a permanent governance strategy.

Red-line prohibitions on the most dangerous AI applications, including autonomous lethal weapon systems, AI-enabled mass surveillance, and systematic deception of vulnerable populations, represent the most politically feasible avenue for international cooperation. However, the central trade-off lies between precision and scope. Definitions must be sufficiently precise to permit formal verification while remaining flexible enough to accommodate rapid technological change. Defining such red lines therefore remains an inherently political process constrained by technical complexity.

Open-source proliferation presents the most structurally difficult governance challenge in this scenario. Requiring open-source systems to incorporate non-removable remote off-switches and automatic regulatory self-registration could improve regulatory oversight. However, these measures would alter the nature of open-source development, potentially concentrating AI capability within proprietary firms while limiting access for researchers, civil society organizations, and public-sector agencies, particularly in lower-income countries. Policymakers should therefore carefully evaluate these distributional consequences rather than treating restrictions on open-source AI as an unqualified safety benefit.

5.3. Cross-cutting Implications for Policymakers

Across both scenarios, several strategic implications emerge regardless of which governance trajectory ultimately develops. First, governments should begin building regulatory capacity immediately through technical training, AI safety research programmes, and procurement standards for AI systems used in the public sector. Asensio et al. (2025) identify technical skills shortages as a major constraint on effective data-driven governance, and the same limitation applies to AI regulation.

Second, regulatory interventions should target the AI development pipeline before deployment. Training data governance, compute access, model architecture disclosures, and pre-deployment safety evaluations provide more effective leverage than relying solely on downstream application-specific regulation.

In addition, policymakers should recognize that effective AI governance depends on equally robust frameworks for electronic data protection, privacy, and ethical data management, as deficiencies in these areas undermine the transparency, accountability, and public trust required for responsible AI deployment (OECD, 2024).

Third, voluntary commitments by AI developers cannot substitute for legally binding requirements supported by enforcement capacity and technical verification mechanisms. Although the NIST AI Risk Management Framework provides an important foundation, its objectives cannot be consistently achieved without binding legal authority.

Finally, regulatory frameworks should be designed to remain adaptive as AI capabilities evolve. Governance mechanisms must incorporate procedures for updating regulatory requirements, integrating advances in AI safety and interpretability research, and rapidly withdrawing unsafe systems from deployment when necessary.

6. Limitations and Reflexivity

The analytical confidence of this forecast is constrained by several significant limitations, some methodological and some arising from the analyst's own situatedness within the policy discourse being examined.

6.1. Framework Constraints and Analytical Bias

The scenario framework employed here simplifies a genuinely complex, multidimensional uncertainty landscape. The two scenarios represent ideal-type endpoints; actual developments will exhibit hybrid characteristics, partial coordination, uneven interpretability progress, regional rather than global convergence, that neither scenario captures cleanly. More fundamentally, the choice of the two uncertainty axes itself reflects a prior analytical judgment about which variables matter most. Alternative framings might have prioritized the concentration of AI capability in non-democratic states, the pace of AI-enabled cyberattack capabilities, or the trajectory of AI labor displacement as the primary uncertainty drivers. These alternative framings would have produced different scenarios with different policy implications. The present framework's emphasis on regulatory coordination and technical safety progress reflects the conceptual vocabulary of the primary sources, Judge et al. (2024) and Asensio et al. (2025), and may therefore systematically underweight political-economic dynamics that those sources treat as secondary.

A related source of analytical constraint is the normative orientation of the primary sources. Judge et al. (2024), writing from the Center for Human-Compatible AI, are explicitly concerned with catastrophic AI risk and advocate for robust regulatory intervention. This orientation may lead the present analysis to underestimate the potential for self-regulatory mechanisms, market incentives, and reputational accountability to produce adequate governance outcomes without binding legislation. The possibility that the history of technology regulation provides a biased sample—that we observe regulatory success in cases where regulation was imposed, but do not observe equivalent counterfactual successes through self-regulation, should be acknowledged. Additionally, the Regulated Convergence scenario may overestimate the tractability of international coordination, given entrenched divergences in values and interests among major AI-developing nations. The analysis would benefit from systematic expert elicitation, structured interviews or Delphi surveys with specialists in AI governance spanning diverse national and disciplinary perspectives, to triangulate the scenario logic and identify blind spots.

6.2. Black Swan Events and Scenario Invalidation

A core limitation of any long-range scenario forecast is its vulnerability to events that fall outside the scenario framework entirely. Several plausible developments would invalidate not merely one scenario but the entire analytical architecture of this paper.

First, a catastrophic model failure of sufficient scale and visibility, a high-salience AI-enabled disaster that generates immediate global regulatory consensus, could collapse the timeline of both scenarios, producing binding international AI governance frameworks within years rather than a decade. The aviation analogy is instructive: the Grand Canyon mid-air collision of 1956 produced the Federal Aviation Act within two years (Judge et al., 2024). An AI-enabled event with equivalent public salience could similarly compress political timelines in ways the present framework, built on incremental institutional change, cannot capture.

Second, a geopolitical rupture of sufficient severity, a breakdown in US-China trade relations, a major AI-enabled cyberattack attributed to a state actor, or the military use of autonomous AI systems in a high-intensity conflict, could shift the entire frame of AI governance from a technical regulatory problem to a national security emergency. Under those conditions, the governance trajectory would be driven by executive and military authority rather than by the legislative and regulatory processes this analysis examines. The resulting governance architecture might be neither analog nor augmented in the Hanisch et al. (2023) typology, but something closer to a classified, security-driven control regime with minimal public accountability, an outcome entirely outside the parameter space of the present scenarios.

Third, a breakthrough in AI capabilities that produces qualitatively more powerful systems than current scaling trends project, what AI safety researchers term a discontinuous capability jump would invalidate the assumption that current governance challenges are representative of the governance challenges of 2035. Both scenarios implicitly assume that frontier AI systems in 2035 are more capable versions of current systems; a discontinuity would make the regulatory frameworks designed for today's systems inadequate for tomorrow's in ways that no present-day analysis can fully anticipate. As Judge et al. (2024) observe, the possibility of rapid recursive self-improvement that exceeds our capacity to intercede or control is precisely what motivates the urgency of establishing effective governance before such a threshold is reached.

Finally, the analysis is conducted primarily from the perspective of advanced industrial democracies and may not adequately account for governance challenges and perspectives in developing countries. Asensio et al. (2025) document that barriers to AI regulatory capacity, limited data infrastructure, skills scarcity, resource constraints, are significantly more severe in lower-income contexts, and that technology adoption in those contexts lags in ways that affect governance outcomes. A governance framework adequate to the AI development and deployment landscape of high-income democracies may be entirely inadequate to the needs of the majority of the world's population.

7. Conclusion

The regulation of generative artificial intelligence represents one of the most significant governance challenges of the coming decade. The rapid advancement of foundation models, combined with their emergent capabilities and broad societal applications, has outpaced the development of corresponding regulatory institutions. This study employed a qualitative scenario planning approach to examine plausible governance pathways for generative AI between 2025 and 2035, focusing on two critical uncertainties: the level of international regulatory coordination and the pace of AI safety and interpretability advances.

The analysis demonstrates that future regulatory outcomes will be shaped by the interaction of technological progress, institutional capacity, and geopolitical cooperation. The Regulated Convergence scenario illustrates how advances in AI safety, strengthened international collaboration, and adaptive governance mechanisms could support the development of effective regulatory systems. In contrast, the Fragmented Acceleration scenario highlights the governance risks associated with continued geopolitical competition, institutional fragmentation, and limited progress in AI safety research. These contrasting scenarios emphasize that regulatory success depends not only on technological innovation but also on the ability of governments and international institutions to coordinate policy responses under conditions of uncertainty.

The policy analysis further demonstrates that effective AI governance requires regulatory frameworks that combine technical verification mechanisms, institutional oversight, international cooperation, and sufficient flexibility to respond to rapidly evolving technologies. Although no single governance model can eliminate all risks associated with generative AI, early institutional investment, strengthened regulatory capacity, and continued advances in AI safety research can substantially improve the prospects for responsible AI development and deployment.

Ultimately, this study contributes to the growing body of literature on AI governance by providing an evidence-based scenario analysis that supports long-term strategic planning under conditions of uncertainty. As generative AI continues to evolve, policymakers will need to balance innovation with safety, accountability, and public trust while ensuring that regulatory institutions remain adaptive to future technological developments. The scenarios presented in this study provide a structured framework for anticipating these challenges and informing policy decisions over the coming decade.

Compliance with ethical standards

Disclosure of conflict of interest

The author(s) declare that there is no conflict of interest regarding the publication of this paper.

References

[1] Asensio, R., et al. (2025). *Data technologies and analytics for policy and governance: A landscape review*. *Data & Policy, 7*, e25. https://doi.org/10.1017/dap.2024.49

[2] Bresnahan, T., & Trajtenberg, M. (1995). General purpose technologies: "Engines of growth"? *Journal of Econometrics, 65*(1), 83–108. https://doi.org/10.1016/0304-4076(94)01598-T

[3] Büthe, T., Djeffal, C., Lütge, C., Maasen, S., & von Ingersleben-Seip, N. (2022). Governing AI: Attempting to herd cats? *Journal of European Public Policy, 29*(11), 1721–1752. https://doi.org/10.1080/13501763.2022.2126515

[4] European Parliament. (2024). *Regulation (EU) 2024/1689 of the European Parliament and of the Council laying down harmonised rules on artificial intelligence (Artificial Intelligence Act). Official Journal of the European Union.*

[5] Hanisch, M., Goldsby, C. M., Fabian, N. E., & Oehmichen, J. (2023). Digital governance: A conceptual framework and research agenda. *Journal of Business Research, 162*, 113777. https://doi.org/10.1016/j.jbusres.2023.113777

[6] Judge, B., Nitzberg, M., & Russell, S. (2024). When code isn't law: Rethinking regulation for artificial intelligence. *Policy and Society, 44*(1), 85–97. https://doi.org/10.1093/polsoc/puae020

[7] National Institute of Standards and Technology. (2023). *Artificial intelligence risk management framework (AI RMF 1.0) (NIST AI 100-1).* U.S. Department of Commerce. https://doi.org/10.6028/NIST.AI.100-1

[8] Ndungu, M. (2026). *E-data protection and privacy in the digital age: Risks, ethics and regulatory challenges*. *World Journal of Advanced Research and Reviews, 30*(1), 2193–2196. https://doi.org/10.30574/wjarr.2026.30.1.1081

[9] Organisation for Economic Co-operation and Development. (2024). *AI, data governance and privacy: Synergies and areas of international co-operation* (OECD Artificial Intelligence Papers No. 22). OECD Publishing. https://doi.org/10.1787/2476b1a4-en

[10] Schwartz, P. (1996). *The art of the long view: Planning for the future in an uncertain world*. Currency Doubleday.

[11] Taeihagh, A., Ramesh, M., & Howlett, M. (2021). Assessing the regulatory challenges of emerging disruptive technologies. *Regulation & Governance, 15*(4), 1009–1019. https://doi.org/10.1111/rego.12392

[12] The White House. (2023, October 30). *Executive order on the safe, secure, and trustworthy development and use of artificial intelligence (Executive Order 14110). Federal Register, 88*(210), 75191–75226.

[13] UK Government, US Government, & 26 other nations. (2023, November 1). *The Bletchley Declaration by countries attending the AI Safety Summit, 1–2 November 2023*. https://www.gov.uk/government/publications/ai-safety-summit-2023-the-bletchley-declaration