International Studies in Science and Technology
https://isstechjournal.science

Technology Legitimacy under Responsible Innovation Norms in International AI Partnerships: Policy Evaluation

Johanna Mäkinen*

Faculty of Humanities, University of Oulu, Oulu, North Ostrobothnia, Finland

Corresponding author: johanna.makinen@oulu.fi

Abstract

This paper presents a comprehensive policy evaluation of responsible innovation norms and their impact on technology legitimacy within international artificial intelligence (AI) partnerships. As AI systems become increasingly integrated into global socio-technical infrastructures, sovereign states and transnational actors are establishing collaborative frameworks to govern their development. However, these partnerships often struggle with the tension between accelerating technological capabilities and maintaining ethical, legal, and social alignment. By analyzing the structural design of major bilateral and multilateral AI agreements, this study examines how responsible research and innovation (RRI) principles are translated into policy instruments. The evaluation framework assesses technology legitimacy across pragmatic, moral, and cognitive dimensions, identifying key determinants of successful norm integration. The findings indicate that while soft-law declarations dominate the international governance landscape, partnerships with formal compliance mechanisms and inclusive stakeholder engagement models achieve significantly higher levels of technology legitimacy. Conversely, regulatory fragmentation and enforcement deficits pose substantial risks to public trust and long-term cooperation. Based on these insights, the paper outlines strategic policy recommendations to harmonize international standards, enhance accountability, and foster sustainable, legitimate AI innovation pathways across diverse jurisdictions.

Keywords: Responsible Innovation; Technology Legitimacy; Artificial Intelligence; International Partnerships; Policy Evaluation

1. Introduction

1.1 The Landscape of Cross-Border AI Cooperation

The rapid advancement and deployment of artificial intelligence systems have transcended national borders, creating a highly interconnected global ecosystem of research, development, and commercialization. Sovereign states, academic institutions, and multinational corporations increasingly engage in cross-border partnerships to pool computational resources, share massive datasets, and leverage specialized human capital [1]. These collaborations are driven by the recognition that no single nation can fully capture the benefits or mitigate the risks of AI in isolation. Consequently, international partnerships have emerged as the primary vehicle for shaping the trajectory of AI technologies, serving as platforms for both cooperative innovation and geopolitical alignment.

At the same time, the institutional architecture governing these partnerships has become highly complex and fragmented. Diplomatic agreements, joint research initiatives, and multilateral alliances are frequently established to coordinate standards and pool investments [2]. This proliferation of governance frameworks reflects a broader recognition that artificial intelligence is a dual-use technology with profound implications for national security, economic competitiveness, and human rights. However, the diversity of political systems, economic models, and cultural values among participating nations often leads to conflicting expectations regarding how these technologies should be developed and deployed.

As a result, international AI partnerships must navigate a delicate balance between promoting innovation and ensuring safety. The structural design of these agreements dictates not only the speed of technological progress but also the normative boundaries within which it occurs. In this context, understanding how cooperative frameworks establish, negotiate, and enforce shared standards is critical for evaluating the long-term sustainability of the global AI enterprise. The institutionalization of these agreements represents a pivotal shift from localized regulation to transnational governance, requiring new analytical tools to assess their efficacy and societal impact.

1.2 Problem Statement: The Alignment-Legitimacy Dilemma

A central challenge in contemporary international AI partnerships is the alignment-legitimacy dilemma. This dilemma arises from the tension between the drive to maximize technological performance and the necessity of adhering to responsible research and innovation norms. While developers and state sponsors are incentivized to accelerate capability development to secure strategic advantages, public citizens and civil society organizations increasingly demand that these technologies respect fundamental rights, prevent bias, and ensure safety [3]. When international partnerships prioritize speed and utility over normative alignment, they risk generating public backlash, regulatory interventions, and a pervasive loss of trust.

This erosion of trust directly threatens the technology legitimacy of AI systems, undermining their social license to operate. Legitimacy is not a static property of a technology but a dynamic social construct that depends on the alignment of the technology with the values, beliefs, and expectations of the society in which it is embedded. In cross-border collaborations, this challenge is compounded by the fact that what is considered legitimate in one jurisdiction may be viewed with skepticism or hostility in another. For instance, differing conceptions of privacy, surveillance, and algorithmic accountability between jurisdictions can lead to severe friction within collaborative projects, stalling joint initiatives and fracturing research alliances.

Furthermore, the absence of harmonized compliance mechanisms often allows partners to engage in regulatory arbitrage, locating high-risk research activities in jurisdictions with the weakest oversight. This behavior not only degrades the ethical integrity of the partnership but also delegitimizes the resulting technological outputs globally. The core policy problem, therefore, is how to design international partnership frameworks that can effectively operationalize responsible innovation norms in a manner that secures and maintains technology legitimacy across diverse legal and socio-political landscapes.

1.3 Objectives and Scope of the Policy Evaluation

This study aims to conduct a systematic policy evaluation of how responsible innovation norms are integrated into international AI partnerships and how this integration affects the legitimacy of the resulting technologies. By examining the structural components, policy instruments, and enforcement mechanisms of select bilateral and multilateral agreements, this research seeks to clarify the pathways through which normative alignment contributes to or detracts from technological legitimacy. The scope of this evaluation encompasses both formal intergovernmental treaties and informal soft-law collaborations, providing a comprehensive overview of the current global governance landscape.

Specifically, the paper addresses three interrelated objectives. First, it operationalizes the core concepts of responsible research and innovation (RRI) and technology legitimacy within the context of transnational AI governance, building a robust analytical model for comparative policy analysis [4]. Second, it evaluates the empirical performance of existing partnerships, identifying the institutional features that facilitate or hinder the successful translation of ethical principles into practice. Third, it generates actionable, evidence-based policy recommendations designed to assist policymakers, diplomats, and research leaders in designing more resilient, legitimate, and socially beneficial collaborative frameworks.

By focusing on the intersection of innovation policy, international relations, and science and technology studies, this evaluation contributes to the growing body of literature on global technology governance. It moves beyond abstract ethical declarations to analyze the concrete structural mechanisms that dictate how norms are negotiated and enforced in the real world. Ultimately, this research provides a diagnostic tool for assessing the health of international AI partnerships, offering a roadmap for balancing technological ambition with democratic accountability and public trust.

2. Theoretical Framework and Literature Review

2.1 Responsible Research and Innovation (RRI) in Emerging Technologies

The concept of Responsible Research and Innovation (RRI) has emerged as a prominent policy paradigm designed to align scientific and technological development with societal values and needs. Originating within European science policy discourses, RRI seeks to move beyond traditional, retrospective risk assessment toward a proactive, anticipatory approach to governance [5]. The framework is built upon four foundational pillars: anticipation, reflexivity, inclusion, and responsiveness. Together, these pillars require innovators and policymakers to actively foresee potential socio-technical impacts, reflect on the underlying assumptions and motivations of research, engage diverse public stakeholders in the decision-making process, and adapt development pathways in response to changing values and new information.

When applied to artificial intelligence, RRI provides a structured approach to addressing the unique challenges posed by algorithmic decision-making, machine learning opacity, and autonomous systems. Unlike traditional technologies, AI systems exhibit dynamic, learning-based behaviors that make their long-term societal impacts exceptionally difficult to predict. Consequently, anticipation in AI governance requires continuous monitoring and scenario-building rather than one-time pre-market approvals [6]. Reflexivity demands that AI developers acknowledge the biases inherent in training datasets and the value judgments embedded in algorithmic optimization functions.

Inclusion in the context of AI involves broadening the circle of governance beyond computer scientists and corporate executives to include ethicists, social scientists, affected communities, and civil society advocates. This inclusive dialogue is crucial for identifying marginalized perspectives and preventing the exacerbation of systemic inequalities. Finally, responsiveness ensures that governance structures possess the regulatory agility to intervene when AI systems demonstrate unintended, harmful consequences. Operationalizing these four pillars within international partnerships is essential for establishing a robust normative foundation that transcends purely national interests and fosters globally responsible development.

2.2 Theoretical Dimensions of Technology Legitimacy

Legitimacy is a central concept in organizational sociology and institutional theory, defined as the generalized perception or assumption that the activities of an entity are desirable, proper, or appropriate within some socially constructed system of norms, values, beliefs, and definitions. When applied to technology, legitimacy represents the degree of societal acceptance and institutional support that a technological system commands [7]. Mark Suchman classic taxonomy categorizes legitimacy into three distinct but interrelated dimensions: pragmatic, moral, and cognitive legitimacy. Each dimension relies on different sources of evaluation and plays a unique role in the stabilization of emerging technologies.

Pragmatic legitimacy is based on the self-interested calculations of a technology user or stakeholder group. It is established when a technology demonstrates clear utility, economic value, or performance superiority, direct benefits that satisfy the immediate needs of its adopters. In international AI partnerships, pragmatic legitimacy is often high due to the immense efficiency gains, scientific breakthroughs, and economic growth associated with advanced computational tools. However, pragmatic legitimacy alone is insufficient to sustain long-term societal acceptance, as it is highly vulnerable to disruption when utility calculations change or when negative externalities become apparent to the broader public.

Moral legitimacy reflects a positive normative evaluation of the technology and its consequences. It is built on judgments about whether the technology promotes societal well-being, adheres to ethical standards, and respects established human rights [8]. Achieving moral legitimacy requires that the processes of technology development and the outcomes of its deployment are perceived as fair, just, and accountable. Cognitive legitimacy, the most deeply entrenched form, occurs when a technology becomes so integrated into daily life and social practices that it is taken for granted. For an AI partnership, navigating these three dimensions is a complex task, as actors must continuously demonstrate pragmatic utility while satisfying stringent moral criteria to ensure the technology transitions from a novel, contested tool to a cognitively accepted infrastructure.

2.3 Multi-Level Governance and International Policy Networks

The governance of international AI partnerships does not occur within a centralized global hierarchy; rather, it is characterized by multi-level governance and the operation of international policy networks. Multi-level governance describes a system where authority and decision-making power are distributed across local, national, regional, and global levels. In the AI domain, this is manifest in the interplay between national regulatory agencies, supranational bodies like the European Union, and international organizations such as the Organization for Economic Co-operation and Development (OECD) or the United Nations [9]. These actors form complex networks that interact to produce, disseminate, and contest the rules governing technological development.

Within these policy networks, soft-law instruments, such as non-binding guidelines, ethical principles, and joint declarations, play a critical role in facilitating initial coordination. Soft law allows diverse nations with differing regulatory capacities and political values to establish a baseline of cooperation without the lengthy and politically fraught process of negotiating formal treaties. These informal networks foster norm diffusion, allowing concepts like trustworthy AI and human-centric design to gain global traction. However, the reliance on soft law also introduces significant vulnerabilities, particularly regarding compliance and enforcement.

As these policy networks expand, they often encounter challenges related to regulatory fragmentation and institutional overlap. Different networks may advocate for competing governance models, leading to a fragmented global landscape where standards are inconsistently applied. This fragmentation can lead to regulatory competition, where jurisdictions lower their ethical standards to attract investment, or regulatory exclusion, where certain nations are locked out of key standard-setting forums. Analyzing these dynamics is crucial for evaluating how international partnerships construct or undermine technology legitimacy, as the structural coherence of the governing network directly influences the perceived reliability and authority of the technology itself.

3. Research Design and Policy Evaluation Methodology

3.1 Analytical Framework for Normative Alignment

To evaluate the integration of responsible innovation norms and their subsequent impact on technology legitimacy, this study employs a multi-dimensional analytical framework. The framework conceptualizes international AI partnerships as complex policy systems containing inputs, processing mechanisms, outputs, and feedback loops. The inputs consist of the diverse regulatory frameworks, ethical doctrines, and strategic motivations that participating nations bring to the negotiating table. The processing mechanisms represent the structural arrangements, decision-making rules, and consultative forums established by the partnership to operationalize responsible innovation principles.

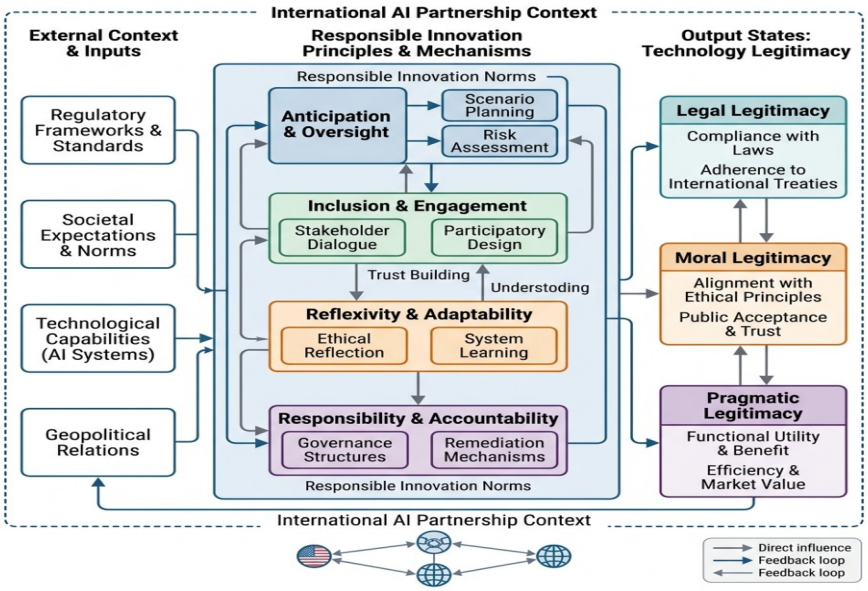

Figure 1: Conceptual Framework for Technology Legitimacy and Responsible Innovation Norms

The outputs of this system are the specific policy instruments, shared standards, and joint technological applications developed under the partnership. The feedback loop represents how public perception, regulatory responses, and societal impacts feed back into the system, driving iterative adjustments to the governing norms [10]. To map these interactions, the framework operationalizes the four pillars of RRI and maps them against Suchman three dimensions of technology legitimacy. This approach enables a systematic assessment of how specific design choices in international agreements correspond to distinct legitimacy outcomes, highlighting the trade-offs inherent in different governance configurations.

Table 1: Operationalization of Responsible Innovation Norms in Policy Frameworks

Normative Pillar	Analytical Indicator	Policy Implementation Mechanism	Expected Governance Outcome
Anticipation	Socio-technical impact assessments	Mandatory algorithmic audits	Risk mitigation and early warning
Reflexivity	Institutional self-assessment	Ethical advisory boards	Adaptive policy adjustments
Inclusion	Multi-stakeholder consultations	Public forums and consensus conferences	Broadened social acceptance
Responsiveness	Iterative regulatory feedback loops	Sandbox environments and policy updates	Flexible correction of market harms

3.2 Comparative Indicators and Data Selection

The empirical component of this evaluation relies on a comparative analysis of selected international AI partnerships representing diverse geographic regions, organizational structures, and legal authority. The data analyzed include formal partnership agreements, joint communiques, policy memorandums, legislative texts, and public consultation reports issued by participating entities. To ensure a rigorous comparison, a standardized set of qualitative and quantitative indicators was developed to measure the performance of each partnership across the RRI pillars and legitimacy dimensions [11].

The indicators for normative integration measure the presence and institutional strength of RRI instruments. These include the existence of mandatory impact assessments, the establishment of independent ethical oversight bodies, the inclusion of non-governmental stakeholders in advisory roles, and the frequency of policy review cycles. Legitimacy indicators, on the other hand, capture the external perception and institutional stability of the technology. These are measured through proxy indicators such as public trust levels, the volume of regulatory challenges, media sentiment analysis, and the rate of technology adoption within user communities.

The selection of cases was guided by the need for variation in independent variables, specifically the degree of formal institutionalization and the geographic diversity of the partners. This ensures that the findings are generalizable across different geopolitical contexts and are not merely artifacts of a single legal tradition or cultural sphere. By examining both high-performing and low-performing partnerships, the evaluation identifies the critical institutional thresholds required to translate abstract ethical commitments into durable technology legitimacy on the global stage.

3.3 Methodological Challenges in Cross-Jurisdictional Evaluation

Evaluating technology policy across multiple jurisdictions introduces several significant methodological challenges. The first challenge concerns semantic divergence, where identical terms are defined and operationalized differently across legal systems. For example, the concept of transparency in one jurisdiction may focus primarily on open-source code sharing, while in another, it may refer to the explainability of algorithmic outcomes to non-technical users [12]. Failing to account for these semantic variations can lead to inaccurate assessments of policy alignment and compliance.

Another challenge is the unequal availability of data across different partnerships. Highly institutionalized, democratic partnerships often feature extensive public archives, stakeholder registries, and transparent reporting mechanisms. In contrast, partnerships involving state-led economic models or security-sensitive applications may operate with high degrees of opacity, limiting access to internal evaluation metrics and decision-making records. To mitigate this bias, this study employs triangulation techniques, combining official documents with third-party assessments, academic literature, and expert interview data.

Finally, the dynamic nature of both AI technology and its regulatory landscape means that policy evaluations are highly time-sensitive. An agreement evaluated as highly legitimate at its inception may quickly lose credibility if a technological breakthrough renders its oversight mechanisms obsolete or if a public controversy exposes hidden vulnerabilities. Recognizing this, the methodology incorporates a longitudinal perspective, analyzing the evolution of partnerships over multi-year periods to assess their adaptability and long-term resilience in the face of rapid technological change.

4. Results and Comparative Analysis of International AI Partnerships

4.1 Evaluation of Existing Sovereign and Intergovernmental AI Frameworks

The comparative evaluation reveals significant disparities in how different international partnerships operationalize responsible innovation norms and establish technology legitimacy. Multilateral frameworks, such as the Global Partnership on Artificial Intelligence (GPAI) and the OECD AI Principles, have successfully established a global baseline of normative expectations [13]. These frameworks excel in promoting cognitive legitimacy by popularizing a shared vocabulary around trustworthy AI. However, their reliance on voluntary compliance and the absence of binding enforcement mechanisms limit their ability to secure moral and pragmatic legitimacy when conflicts of interest arise between member states.

In contrast, bilateral partnerships embedded within broader economic or security alliances demonstrate a higher degree of institutionalization. These agreements often feature specific research agendas, joint funding structures, and formalized standard-setting committees [14]. Yet, the evaluation indicates that these narrower partnerships frequently prioritize pragmatic utility, such as economic competitiveness or military advantage, over broad societal inclusion. This imbalance can lead to a narrow construction of legitimacy that satisfies state and industry elites but fails to address systemic societal risks, ultimately exposing the partnership to public opposition and regulatory pushback.

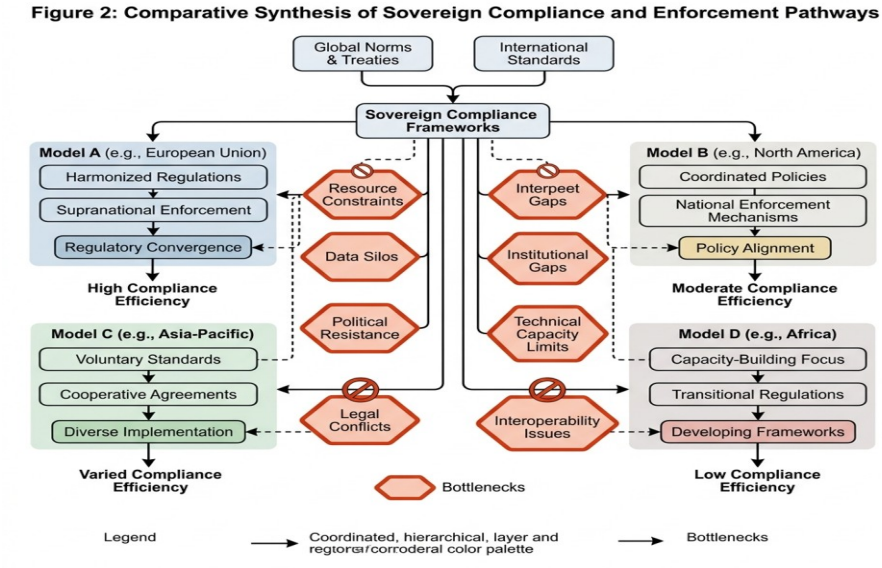

Figure 2: Comparative Synthesis of Sovereign Compliance and Enforcement Pathways

The data also show that regional initiatives, particularly those led by the European Union, have pursued a highly institutionalized, rights-based approach to AI governance. By linking international market access to compliance with stringent risk-assessment and transparency standards, these frameworks have successfully exported their normative preferences globally, a phenomenon often referred to as the Brussels Effect. However, this approach can create friction with partners who favor a more market-led, flexible regulatory stance, highlighting the challenges of achieving normative alignment across distinct economic systems.

Table 2: Legitimacy Indicators and Performance Metrics Across Selected Bilateral Partnerships

Partnership case	Compliance score	Transparency index	Public trust level	Policy alignment degree
Transatlantic Dialogue	High	Medium	Medium	Medium
Asia-Pacific Digital Network	Medium	Low	High	High
European-Asian Cooperation	High	High	Medium	Low
Global South Coalition	Low	Medium	Low	Medium

4.2 Determinants of Technology Legitimacy in Bilateral Treaties

The empirical analysis identifies several critical determinants of technology legitimacy within bilateral AI partnerships. The first and most significant determinant is the presence of formal compliance and verification mechanisms. Partnerships that rely solely on self-reporting or voluntary codes of conduct demonstrate lower levels of moral legitimacy and are more susceptible to public skepticism [15]. Conversely, agreements that establish independent, joint auditing bodies empowered to verify compliance with safety and ethical standards achieve significantly higher levels of trust and institutional stability.

A second key determinant is the depth and quality of stakeholder inclusion. Partnerships that restrict participation to government officials and corporate executives struggle to maintain legitimacy when controversial issues, such as algorithmic bias or worker displacement, enter the public discourse. In contrast, frameworks that institutionalize continuous consultation with labor unions, civil rights organizations, and academic researchers are better equipped to anticipate potential societal harms and build a broad-based social consensus around the technology.

Finally, the degree of political alignment between participating nations serves as a foundational conditioning factor. Partnerships between nations with shared democratic values, compatible legal systems, and similar economic structures find it much easier to negotiate and enforce complex normative standards. When partnerships cross deep ideological or geopolitical divides, the process of norm negotiation becomes highly transactional and prone to breakdown, often resulting in superficial agreements that fail to provide genuine oversight or secure robust technology legitimacy.

4.3 Divergence in Implementation and Enforcement Mechanisms

A critical vulnerability identified across the evaluated partnerships is the wide divergence in implementation and enforcement mechanisms. While many agreements feature sophisticated preambles endorsing the principles of responsible innovation, the actual operationalization of these principles in joint research and development projects is often highly uneven [16]. This implementation gap is primarily driven by the lack of clear, actionable guidelines for translating high-level ethical values into concrete engineering practices and policy rules.

For instance, while a partnership agreement may pledge to ensure fairness in joint machine learning models, it rarely specifies which mathematical definition of fairness should be optimized, or how developers should navigate the inherent trade-offs between accuracy and equity. In the absence of specific, technical standards, individual research teams are left to make these value-laden decisions independently, leading to inconsistent outcomes that can undermine the normative integrity of the entire collaboration.

Furthermore, enforcement mechanisms are frequently weakened by institutional fragmentation and a lack of jurisdictional clarity. When joint projects involve researchers and datasets from multiple countries, determining which national regulatory agency has the authority to investigate violations or impose sanctions is exceptionally difficult. This jurisdictional ambiguity creates compliance blind spots, allowing partners to bypass strict domestic laws by conducting sensitive research activities within the gray areas of the international partnership framework. Addressing this enforcement deficit is a primary challenge for the next generation of international AI agreements.

5. Discussion and Policy Recommendations

5.1 Synthesis of Findings and Governance Trade-offs

The policy evaluation conducted in this study highlights a series of complex governance trade-offs that international AI partnerships must navigate. Chief among these is the trade-off between speed and safety, or innovation velocity versus normative compliance. Highly structured, inclusive, and risk-averse governance frameworks, while effective at securing moral legitimacy and minimizing societal harms, can slow down development cycles and impose significant administrative burdens on researchers [17]. In a highly competitive global market, policymakers face intense pressure to streamline regulatory requirements to maintain technological leadership, even if it risks eroding long-term public trust.

Another critical trade-off exists between national sovereignty and international harmonization. Achieving truly aligned, global standards requires participating nations to cede some degree of regulatory autonomy and accept common oversight mechanisms. However, sovereign states are often reluctant to delegate authority over strategic, dual-use technologies like AI, preferring to maintain domestic control even at the expense of global coordination. This reluctance leads to the perpetuation of fragmented, soft-law arrangements that are incapable of addressing transnational technological challenges.

These findings suggest that a one-size-fits-all approach to international AI governance is both politically unfeasible and structurally inadequate. Instead, successful policy design requires a differentiated approach that matches the intensity of governance mechanisms to the risk profile of the specific technological applications. By understanding these trade-offs, policymakers can move past the false dichotomy of innovation versus regulation, designing balanced frameworks that leverage normative alignment as a driver of, rather than a barrier to, sustainable technology legitimacy.

5.2 Policy Pathways for Enhancing Cooperative Legitimacy

To overcome the identified challenges and enhance the legitimacy of international AI partnerships, this study proposes several key policy pathways. First, future agreements must transition from soft-law declarations to hybrid governance models that couple flexible, principles-based standards with formal, binding compliance mechanisms [18]. This can be achieved by establishing independent, joint regulatory sandboxes where collaborative projects can be tested and audited against shared safety and ethical benchmarks before full-scale deployment.

Second, international partnerships should institutionalize a permanent, multi-stakeholder advisory body comprised of diverse societal representatives. This body should be integrated directly into the partnership decision-making structure, with the authority to trigger mandatory impact reviews for high-risk AI applications. By formalizing public participation, partnerships can build a robust foundation of moral legitimacy that can withstand technological crises and political transitions.

Third, policymakers must prioritize the development of clear, technical standards that translate abstract RRI principles into practical, verifiable metrics. This includes investing in shared tools for algorithmic auditing, model explainability, and dataset bias detection. These standardized tools should be made open-source and integrated into joint research workflows, reducing the implementation gap and ensuring consistent compliance across all participating entities. Finally, agreements should include clear, pre-negotiated dispute resolution mechanisms to address jurisdictional conflicts, preventing enforcement deficits and ensuring accountability when harms occur.

5.3 Limitations and Directions for Future Inquiry

While this study provides valuable insights into the policy dynamics of international AI partnerships, several limitations must be acknowledged. First, the evaluation relies on a qualitative-dominant comparative methodology, which, while useful for mapping institutional structures and normative discourses, could be enhanced by more comprehensive quantitative indicators of compliance and public sentiment. Future research should integrate large-scale public surveys and advanced sentiment analysis tools to measure variations in technology legitimacy across diverse demographics and regions over time.

Second, the rapid pace of technological development means that new categories of AI systems, such as advanced foundation models and autonomous agent networks, are continuously emerging. These technologies present novel governance challenges that may not be fully addressed by existing RRI frameworks, which were largely designed for narrower, single-purpose AI applications. Further scholarly inquiry is needed to adapt

responsible innovation norms to the unique characteristics of generative and agentic AI systems within transnational collaborative networks.

Finally, this evaluation is constrained by the current geopolitical climate, which is characterized by growing techno-nationalism and strategic competition. The feasibility of implementing highly cooperative, harmonized governance models remains highly dependent on broader political dynamics that lie outside the scope of science and technology policy. Future research should explore how technology governance frameworks can be designed to resist geopolitical shocks, ensuring that essential safety and ethical standards are maintained even in times of international tension.

References

1. Foerster, J.N.; Farquhar, G.; Afouras, T.; Nardelli, N.; Whiteson, S. Counterfactual Multi-Agent Policy Gradients. In Proceedings of the Thirty-Second AAAI Conference on Artificial Intelligence (AAAI 2018), New Orleans, LA, USA, 2–7 February 2018; AAAI Press: Palo Alto, CA, USA, 2018; pp. 2974–2982.

2. Liu, Y.; Wang, W.; Hu, Y.; Hao, J.; Chen, X.; Gao, Y. Multi-Agent Game Abstraction via Graph Attention Neural Network. In Proceedings of the Thirty-Fourth AAAI Conference on Artificial Intelligence (AAAI 2020), New York, NY, USA, 7–12 February 2020; AAAI Press: Palo Alto, CA, USA, 2020; Volume 34, pp. 7211–7218.

3. Torrieri, D. Principles of Spread Spectrum Communication Systems, 5th ed.; Springer International Publishing: Cham, Switzerland, 2022; pp. 151–203.

4. Kato, J. Institutions and rationality in politics—Three varieties of neo-institutionalists. Br. J. Polit. Sci. 1996, 26, 553–582.

5. Renn, O. (2008). Risk governance: Coping with uncertainty in a complex world. Earthscan.

6. Li, T.; Sahu, A.K.; Zaheer, M.; Sanjabi, M.; Talwalkar, A.; Smith, V. Federated Optimization in Heterogeneous Networks. In Proceedings of Machine Learning and Systems; Association for Computing Machinery: New York, NY, USA, 2020.

7. Lecours, A. New Institutionalism: Theory and Analysis; University of Toronto Press: Toronto, ON, Canada, 2005.

8. Van Bavel, J. J., Baicker, K., Boggio, P. S., Capraro, V., Cichocka, A., Cikara, M., Crockett, M. J., Crum, A. J., Douglas, K. M., Druckman, J. N., & Drury, J. (2020). Using social and behavioural science to support COVID-19 responses. Nature Human Behaviour, 4, 460–471.

9. Nzewi, O.I. Adaptive Governance for Resilient Local Service Delivery. J. Local Gov. Res. Innov. 2025, 6, a322.

10. Meng, M.; Hu, B.; Chen, S.; Kang, S. Joint Beamforming and Dynamic Beam Hopping Based on MAPPO for LEO Satellite Communication System. IEEE Wirel. Commun. Lett. 2025, 14, 1461–1465.

11. Hu, Z.; Liu, X.; Guo, M.; Liu, C. PPO-Based Joint Task Offloading and Resource Allocation for Vehicular Edge Computing Via V2I and V2V Communications. In Proceedings of the 2025 13th International Conference on Intelligent Computing and Wireless Optical Communications (ICWOC), Chengdu, China, 28–29 June 2025; pp. 316–321.

12. Nguyen, T.T.; Nguyen, N.D.; Nahavandi, S. Deep Reinforcement Learning for Multiagent Systems: A Review of Challenges, Solutions, and Applications. IEEE Trans. Cybern. 2020, 50, 3826–3839.

13. Bagdasaryan, E.; Veit, A.; Hua, Y.; Estrin, D.; Shmatikov, V. How To Backdoor Federated Learning. In Proceedings of the Twenty Third International Conference on Artificial Intelligence and Statistics; PMLR: Cambridge, MA, USA, 2020; pp. 2938–2948.

14. Wu, Y.; Huang, Y.; Wang, Z.; Xu, C. Joint Caching and Computation in UAV-Assisted Vehicle Networks via Multi-Agent Deep Reinforcement Learning. Drones 2025, 9, 456.

15. Hu, K.; Xu, K.; Xia, Q.; Li, M.; Song, Z.; Song, L.; Sun, N. An Overview: Attention Mechanisms in Multi-Agent Reinforcement Learning. Neurocomputing 2024, 598, 128015.

16. Jiang, J.; Lu, Z. Learning Attentional Communication for Multi-Agent Cooperation. In Advances in Neural Information Processing Systems 31 (NeurIPS 2018), Montréal, QC, Canada, 3–8 December 2018; Curran Associates, Inc.: Red Hook, NY, USA, 2018; pp. 7265 –7275.

17. Lubell, M.; Robbins, M. Adapting to sea-level rise: Centralization or decentralization in polycentric governance systems? Policy Stud. J. 2022, 50, 143–175.

18. Cui, Z.; Qamar, F.; Kazmi, S.H.A.; Zainol Ariffin, K.A.; Safdar, G.A.; Ur Rehman, M.H. A Review of Multi-Agent Deep Reinforcement Learning for Resource Allocation in beyond 5G Network Slicing: Solutions, Challenges and Future Research Directions. PeerJ Comput. Sci. 2026, 12, e3728.