

Política Internacional

ISSN 2707-7330

<https://rpi.isri.cu/rpi>

✉ rpi@isri.minrex.gob.cu

Vol. VIII, No. 2 (abril-junio) | 2026

RNPS: No. 2092

<https://creativecommons.org/licenses/by-nc/4.0/>



RELACIONES INTERNACIONALES


Adaptive Political Entrepreneurship in Global Artificial Intelligence Governance: A Policy Entrepreneur Network Perspective

Emprendimiento político adaptativo en la gobernanza global de la inteligencia artificial: una perspectiva de la red de emprendedores políticos

Entrepreneuriat politique adaptatif dans la gouvernance mondiale de l'intelligence artificielle : une perspective de réseau d'entrepreneurs politiques

Empreendedorismo político adaptativo na governança global da inteligência artificial: uma perspectiva da Rede de Empreendedores Políticos

Luyao ZHANG

Ph. D. candidate in international political economy at the School of International Relations and Public Affairs of Fudan University, Shanghai, China. ✉ zhangluyao23@m.fudan.edu.cn  [0009-0005-6617-0664](https://orcid.org/0009-0005-6617-0664)

PH.D. Cuihong Cai*

Professor of international relations with the Center for American Studies at Fudan University, Shanghai, China. ✉ chcai@fudan.edu.cn  [0000-0001-8221-0244](https://orcid.org/0000-0001-8221-0244)

*Corresponding author: chcai@fudan.edu.cn

How to cite (APA, seventh edition): ZHANG, L., & Cai, C. (2026). Adaptive Political Entrepreneurship in Global Artificial Intelligence Governance: A Policy Entrepreneur Network Perspective. *Política Internacional*, VIII (Nro. 2), 192-212. <https://doi.org/10.5281/zenodo.19132609>

 <https://doi.org/10.5281/zenodo.19132609>

Este es un artículo de acceso abierto distribuido bajo los términos de la Licencia de Creative Commons Atribución-NoComercial Internacional 4.0 (CC BY-NC 4.0), que permite su uso, distribución y reproducción en cualquier medio, así como generar obras derivadas; siempre y cuando se reconozca al autor original y se haga con fines no comerciales. Las opiniones y contenidos de los documentos publicados corresponden exclusivamente a sus autores.

RECEIVED: SEPTEMBER 15, 2025

APPROVED: MARCH 12, 2026

PUBLISHED: APRIL 16, 2026

ABSTRACT Global artificial intelligence (AI) governance faces unprecedented challenges due to rapid technological advancement and geopolitical complexity. This study examines how policy entrepreneur networks operate as catalysts in global AI governance, moving beyond traditional domestic models to analyze their unique roles in international technology governance. Our study analyzes policy entrepreneur networks as catalysts, testing hypotheses on role integration, platform dependency, and adaptive design through three milestone cases: the 2019 OECD AI Principles, 2021 UNESCO Recommendation, and 2023 Bletchley Declaration. We find that effective networks must integrate technical translation, diplomatic coordination, and normative leadership. However, several limitations should be noted: first, the case selection is confined to three specific policy windows between 2018 and 2023, which may not capture all evolving governance models; second, the analysis primarily focuses on state-led initiatives within international organizations, with relatively less emphasis on independent networks driven by non-state actors such as technology corporations or civil society.

Keywords: Artificial Intelligence, Global Governance, Policy Entrepreneurs, Multiple Streams Framework, Policy Windows, Transnational Networks

RESUMEN *La gobernanza global de la inteligencia artificial (IA) se enfrenta a desafíos sin precedentes debido al rápido avance tecnológico y la complejidad geopolítica. Este estudio examina cómo las redes de emprendedores políticos actúan como catalizadores en la gobernanza global de la IA, trascendiendo los modelos nacionales tradicionales para analizar su papel único en la gobernanza tecnológica internacional. Nuestro estudio analiza las redes de emprendedores políticos como catalizadores, poniendo a prueba hipótesis sobre la integración de roles, la dependencia de plataformas y el diseño adaptativo a través de tres casos clave: los Principios de IA de la OCDE de 2019, la Recomendación de la UNESCO de 2021 y la Declaración de Bletchley de 2023. Observamos que las redes eficaces deben integrar la traducción técnica, la coordinación diplomática y el liderazgo normativo. Sin embargo, cabe señalar varias limitaciones: en primer lugar, la selección de casos se limita a tres ventanas de políticas específicas entre 2018 y 2023, lo que podría no abarcar todos los modelos de gobernanza en evolución; en segundo lugar, el análisis se centra principalmente en iniciativas estatales dentro de organizaciones internacionales, con un énfasis relativamente menor en redes independientes impulsadas por actores no estatales, como las corporaciones tecnológicas o la sociedad civil.*

Palabras clave: Inteligencia artificial, gobernanza global, emprendedores políticos, marco de múltiples corrientes, ventanas de políticas, redes transnacionales

RÉSUMÉ *La gouvernance mondiale de l'intelligence artificielle (IA) est confrontée à des défis sans précédent, dus aux progrès technologiques rapides et à la complexité géopolitique. Cette étude examine comment les réseaux d'entrepreneurs politiques agissent comme catalyseurs de la gouvernance mondiale de l'IA, en dépassant les modèles nationaux traditionnels pour analyser leurs rôles spécifiques dans la gouvernance technologique internationale. Notre étude analyse ces réseaux en tant que catalyseurs, en testant des hypothèses sur l'intégration des rôles, la dépendance à la plateforme et la conception adaptative à travers trois cas marquants : les Principes de l'OCDE sur l'IA de 2019, la Recommandation de l'UNESCO de 2021 et la*

Déclaration de Bletchley de 2023. Nous constatons que les réseaux efficaces doivent intégrer la traduction technique, la coordination diplomatique et le leadership normatif. Cependant, plusieurs limites doivent être soulignées : premièrement, la sélection des cas se limite à trois fenêtres politiques spécifiques entre 2018 et 2023, ce qui peut ne pas refléter l'ensemble des modèles de gouvernance en évolution ; deuxièmement, l'analyse se concentre principalement sur les initiatives étatiques au sein des organisations internationales, en accordant une importance relativement moindre aux réseaux indépendants pilotés par des acteurs non étatiques tels que les entreprises technologiques ou la société civile.

Mots-clés : Intelligence artificielle, gouvernance mondiale, entrepreneurs politiques, cadre à flux multiples, fenêtres d'opportunité politique, réseaux transnationaux

RESUMO *A governança global da inteligência artificial (IA) enfrenta desafios sem precedentes devido ao rápido avanço tecnológico e à complexidade geopolítica. Este estudo examina como as redes de empreendedores políticos operam como catalisadores na governança global da IA, indo além dos modelos domésticos tradicionais para analisar seus papéis únicos na governança internacional da tecnologia. Nosso estudo analisa as redes de empreendedores políticos como catalisadores, testando hipóteses sobre integração de papéis, dependência de plataforma e design adaptativo por meio de três casos marcantes: os Princípios de IA da OCDE de 2019, a Recomendação da UNESCO de 2021 e a Declaração de Bletchley de 2023. Descobrimos que as redes eficazes devem integrar tradução técnica, coordenação diplomática e liderança normativa. No entanto, várias limitações devem ser observadas: primeiro, a seleção de casos se limita a três janelas políticas específicas entre 2018 e 2023, o que pode não capturar todos os modelos de governança em evolução; segundo, a análise se concentra principalmente em iniciativas lideradas pelo Estado dentro de organizações internacionais, com relativamente menos ênfase em redes independentes impulsionadas por atores não estatais, como corporações de tecnologia ou sociedade civil.*

Palavras-chave: Inteligência Artificial, Governança Global, Empreendedores Políticos, Estrutura de Múltiplos Fluxos, Janelas Políticas, Redes Transnacionais

INTRODUCTION

As artificial intelligence technologies advance at an unprecedented pace, with developments like large language models and autonomous systems rapidly transforming industries (Chui et al., 2023), the global community faces growing pressure to establish effective governance frameworks to address the multifaceted challenges posed by these developments. However, the translation of technological imperatives into concrete policy action remains enigmatic. While traditional policy studies emphasize the role of policy entrepreneurs in facilitating policy change (Kingdon, 1995); (Mintrom & Norman, 2009), the unique characteristics of global AI

governance challenge conventional understandings of how policy entrepreneurship operates in international contexts.

Policy entrepreneurs, as conceptualized by John W. Kingdon in his Multiple Streams Framework (MSF), are individuals or organizations that invest resources in hopes of future policy returns, coupling problems, policies, and politics to open policy windows (Arnold & St. John, 2023; Mintrom & Moulton, 2019). Yet when we examine landmark achievements in global AI governance, such as the 2019 OECD AI Principles, the 2021 UNESCO Recommendation on the Ethics of AI, and the 2023 Bletchley Declaration. We found that traditional notions of policy entre-

preneurship appear insufficient to explain how these breakthroughs occurred.

This insufficiency deal to two fundamental challenges in global AI governance. First, the fragmented and competitive nature of international AI governance creates a complex multi-actor environment where no single entrepreneur can dominate agenda-setting processes (Kerry et al., 2025). Unlike traditional policy contexts where individual politicians, interest groups, or bureaucrats might drive change, global AI governance involves nation-states, international organizations, transnational expert networks, and technology corporations operating across multiple levels of authority (Cai & Zhang, 2025).

Second, AI technology presents unprecedented challenges to traditional policy entrepreneurs through what we term the Collingridge dilemma intensified (Genus & Stirling, 2018). Unlike conventional technologies that evolve predictably over decades, AI capabilities demonstrate exponential growth patterns with uncertain and emergent properties that cannot be fully anticipated (Thierer, 2018). The release of ChatGPT in November 2022, for instance, transformed public understanding of AI capabilities within months, not years, forcing policy entrepreneurs to operate under conditions of radical uncertainty. This creates a temporal mismatch between the extended timeframes required for international policy coordination and the compressed cycles of AI technological development. Policy entrepreneurs must simultaneously address current AI applications while preparing for future capabilities that may fundamentally alter the governance landscape.

Given these complexities, this study addresses a critical gap in understanding how policy entrepreneur functions in the emerging domain of global AI governance. Specifically, we ask: How do policy entrepreneurs operate differently in global AI governance compared to traditional domestic policy contexts, and how do they adapt their strategies to address the unique challenges of international AI policy coordination under conditions of technologi-

cal uncertainty? To answer this question, we adapt Kingdon's MSF to the international context, examining how policy entrepreneurs in global AI governance develop new roles, resources, and strategies that differ substantially from their domestic counterparts.

This research focuses on the years 2018 to 2023, a period where global AI governance moved from setting basic principles to managing urgent, crisis-driven safety concerns. The study explores how policy entrepreneur networks function within the international system, specifically looking at how these collective efforts differ from the individual-led models typically found in domestic policy. By applying three core hypotheses to milestone cases, we identify the specific roles and strategies these networks use to remain effective despite high technological uncertainty.

DEVELOPMENT

2.1 Adapting the Multiple Streams Framework to Global AI Governance

Kingdon's model posits that policy windows open when three streams: problems, policies, and politics converge, creating a fleeting yet powerful moment for advancing policy initiatives (Kingdon, 1995). During these moments, policy entrepreneurs actively work to facilitate the convergence of these streams, thereby enabling the opening of policy windows. While originally developed for domestic policy contexts, this framework offers substantial analytical potential for understanding how complex governance challenges, particularly those surrounding artificial intelligence, gain attention on the global stage.

The problem stream in global AI governance is characterized by what we term transnational "cascade effects" (Whyte, 2020). Unlike domestic policy problems that remain contained within national boundaries, AI-related issues exhibit inherent spillover characteristics that transcend jurisdictional limits

(Yeung & Howes, 2018). Focusing events in this domain possess unique properties: they often emerge from technological breakthroughs rather than policy failures, and their implications may not be immediately apparent to policymakers lacking technical expertise. The 2016 AlphaGo victory exemplifies how a single technological demonstration can simultaneously shift perceptions across multiple countries (Silver et al., 2016), creating synchronized windows of attention that would be impossible in traditional policy domains.

The policy stream in global AI governance operates as an extraordinarily rich and diverse “policy soup” that encompasses proposals from international organizations, academic institutions, civil society groups, technology companies, and governmental bodies across multiple governance levels. Unlike domestic policy environments where alternatives compete primarily on technical feasibility and political acceptability, global AI governance policies must also satisfy additional criteria including cross-cultural value compatibility, technological adaptability, and international legal consistency. Moreover, the interconnected nature of AI technologies creates challenges for policy development. Solutions designed to address one aspect of AI governance may have unintended consequences across multiple policy domains and jurisdictional boundaries (Taeihagh, 2021). Policy entrepreneurs must therefore develop systemic approaches that account for complex interdependencies while remaining flexible enough to adapt to rapid technological change.

The political stream in global AI governance encompasses not only traditional domestic political considerations but also international relations dynamics, geopolitical competition, and asymmetric capabilities among different countries. The strategic rivalry between major AI powers creates both opportunities for policy attention and constraints on the scope of feasible multilateral cooperation (Ding, 2024). International organizations play crucial mediating roles by providing neutral forums, but they also introduce additional layers of institutional politics and

bureaucratic procedures (Martin & Simmons, 1998). The temporal dimension of the political stream differs markedly from domestic contexts. Electoral cycles are overlaid with international summit schedules, treaty negotiation timelines, and technological development phases, creating complex and often unpredictable windows of political opportunity that policy entrepreneurs must learn to navigate.

2.2 Theoretical Hypotheses for Policy Entrepreneurs in Global AI Governance

Based on this analysis of stream adaptation, we propose three interconnected hypotheses about how policy entrepreneurs operate differently in global AI governance compared to traditional domestic contexts.

Hypothesis 1: Role Integration Imperative Policy entrepreneurs in global AI governance must simultaneously integrate three specialized roles that are typically distributed among different actors in domestic policy processes. They function as technical translators who bridge complex AI technologies with policy frameworks, diplomatic coordinators who navigate anarchic international systems to build cross-national consensus, and normative leaders who articulate shared values across diverse cultural contexts (Perry & Uuk, 2019). This role integration represents a shift from traditional policy entrepreneurship, where specialization in particular domains is typically sufficient for effectiveness.

This integration occurs because the complexity of global AI governance exceeds the capacity of any single type of traditional policy entrepreneur. Technical experts may understand the technology but lack diplomatic skills necessary for international coordination. Government officials may possess negotiating authority but lack technical knowledge for effective governance design. International bureaucrats may have institutional legitimacy but lack the agility to respond to technological developments. Success therefore requires either individual actors who master multiple competencies or closely

coordinated networks functioning as collective entrepreneurs.

Hypothesis 2: Platform Dependency Strategy Facing high international coordination costs, policy entrepreneurs in global AI governance primarily leverage existing institutional platforms rather than creating entirely new governance mechanisms. This platform dependency emerges as a rational response to transaction costs of building international consensus from scratch, particularly when addressing rapidly evolving technologies requiring agile policy responses.

Successful entrepreneurs strategically utilize the convening power, legitimacy, and established procedures of organizations like OECD, UNESCO, and the United Nations to reduce coordination costs and risks. Rather than investing resources in creating new institutions which requires extensive negotiation and may face resistance from established powers they work within existing frameworks to develop new norms, standards, and practices (Erman & Furendal, 2024). This approach enables more rapid policy development while building on established relationships and procedures that have already gained international acceptance.

However, platform dependency also creates constraints and path dependencies. Entrepreneurs must operate within existing organizational mandates, membership structures, and procedures, potentially limiting the scope and ambition of their initiatives. Platform choice influences which countries and stakeholders participate in governance processes, affecting the legitimacy and effectiveness of resulting policies.

Hypothesis 3: Adaptive Institutional Design Confronting fundamental uncertainty surrounding AI technology development and societal implications, successful policy entrepreneurs adopt flexible institutional design strategies rather than attempting to create comprehensive regulatory frameworks based on current technological capabilities. This adaptive

approach reflects understanding that traditional regulatory models assuming stable technologies and predictable implementation contexts are inadequate for governing rapidly evolving AI systems with emergent properties (Reuel & Undheim, 2024; Tan & Taihigh, 2021).

Instead of designing detailed rules based on current AI capabilities, entrepreneurs focus on establishing principles, creating monitoring mechanisms, and building institutional capacity for ongoing policy adjustment as technologies and their implications become clearer. This strategy manifests in principle-based governance frameworks that provide general guidance while allowing flexible interpretation, formal review mechanisms enabling policy evolution in response to technological developments (Radu, 2021), multi-stakeholder platforms integrating diverse expertise, and experimental features like pilot programs enabling learning rather than permanent commitment to particular approaches.

These three hypotheses are interconnected: role integration enables entrepreneurs to navigate technical and political complexities, platform dependency provides institutional vehicles for coordination, and adaptive design ensures policy resilience under uncertainty. Together, they constitute a framework for understanding how policy entrepreneurship adapts to the unique challenges of global AI governance.

3. Testing Policy Entrepreneurs' Role in Global AI Governance

The three core hypotheses proposed in the previous chapter: role integration imperative, platform dependency strategy, and adaptive institutional design require validation through specific policy practices. From the theoretical perspective of the Multiple Streams Framework (MSF), policy windows as "rare time windows that can promote policy solutions or drive policy change (Kingdon, 1995)," are often difficult to observe directly. As Birkland notes "at best we can identify key moments in policy history that seem to have turned the tide toward new

policies or new ideas (Birkland, 2020).” Therefore, milestone policies, as external manifestations of successful policy window openings, provide observable and measurable empirical phenomena.

The rapid evolution of global AI governance provides an ideal natural experimental environment for such validation. From the perspective of policy feedback theory, once milestone policies become institutionalized and enter collective policy memory, their origin stories which typically include policy window narratives become consolidated (Pierson, 1993). We chose the 2019 OECD AI Principles, the 2021 UNESCO AI Ethics Recommendation, and the 2023 Bletchley Declaration as our primary cases because they represent the most significant transitions in the field.

These three milestones were selected based on their authority, representativeness, and agenda-framing capacity. The OECD Principles established the first intergovernmental “human-centered” vision, creating a foundational normative framework for developed nations. The UNESCO Recommendation expanded this into a universal consensus of 193 member states, shifting the focus toward human rights and cultural diversity. Finally, the Bletchley Declaration responded to the breakthrough of generative AI, achieving rare coordination among major powers to focus specifically on frontier AI safety. Together, these cases represent strategic adaptations by policy entrepreneur networks across different institutional environments and technological contexts.

3.1 Case 1 2019 OECD AI Principles: Policy Entrepreneur Activities of Institutionalized Expert Networks

The first significant policy window opened in May 2019 with the adoption of the OECD AI Principles. This policy window extended from the establishment of the AIGO working group in May 2018 until the G20 adoption of related declarations in July 2019, lasting approximately two months. In terms of authority, OECD functions as an important platform

for developed country policy coordination, and principles adopted by this organization possess considerable normative influence, with the organization’s established reputation in economic and social policy coordination lending credibility to its AI governance initiatives. From a representativeness perspective, while initially covering mainly OECD member countries, the principles subsequently expanded to include Argentina, Brazil, Colombia, and other nations, demonstrating their appeal beyond the original membership base. Regarding agenda-framing capacity, these principles established a “human-centered” AI development vision at the intergovernmental level for the first time, proposing five value-oriented principles and five policy recommendations that provided a foundational normative framework for subsequent global AI governance.

The OECD Working Party on Artificial Intelligence Governance (AIGO) functioned as the primary policy entrepreneur network, exemplifying the typical characteristics of institutionalized expert coordination. According to official documentation, AIGO integrated over 70 experts from 30 countries, the EU, and stakeholder groups such as the Business and Industry Advisory Committee (BIAC) and the Trade Union Advisory Committee (TUAC), achieving balanced representation in terms of expertise, gender, and regional representation (OECD, 2019a). The network’s coordination mechanism manifested as a progressive consensus-building process: from September 2018 to February 2019, through four key meetings in Paris, MIT, Dubai, and other locations, a complete policy-making chain was established from conceptual clarification to principal establishment (OECD, 2021).

1. Role Integration Performance of AIGO

The first role of AIGO is networked implementation of technical translation capability. The AIGO network demonstrated strong technical translation capability, successfully converting complex AI technical concepts into policy-operational frameworks. The technical translation process manifested as

multi-layered conceptual conversion: first, the AI definition working group translated technical terms such as machine learning, deep learning, and neural networks into policy language as “machine-based systems that can perceive real and/or virtual environments, abstract such perceptions into models, and use model interpretations to formulate options for outcomes”(OECD, 2019c); second, technical experts and policymakers established a collaborative mechanism of concept clarification, policy translation, consensus building; finally, the resulting report became a foundational document for subsequent international standards, providing conceptual foundations for the EU AI Act, US AI initiatives, and others(Grobelnik et al., 2024).

The second role is institutionalized mechanisms of diplomatic coordination. The AIGO network achieved effective diplomatic coordination among representatives from 37 member states, demonstrating the advantages of institutionalized coordination (OECD, 2019c). The coordination mechanism manifested as a multi-round position convergence process: Phase 1 (September-December 2018) saw countries submitting position papers and identifying core points of divergence such as AI definition scope, principle binding force, and implementation responsibilities; Phase 2 (January-March 2019) promoted political position convergence through technical-level conceptual unification, with consensus on AI system definitions laying foundations for subsequent principle development; Phase 3 (April-May 2019) completed the upgrade from draft to final text through iterative revisions, reflecting a mature diplomatic coordination model of divergence management consensus consolidation text finalization. Data shows that the final five principles and five policy recommendations received unanimous approval from all OECD member states (OECD, 2019b).

The third role is discourse authority construction in normative leadership. The AIGO network played a pioneering role in constructing discourse authority in global AI governance, successfully establishing the core normative framework of “human-centric

AI.” Normative leadership manifested at three levels: first, by proposing the “human-centric AI” concept and operationalizing it into five major principles, providing value anchoring for global AI governance; then, through OECD’s authoritative endorsement, these principles acquired the status of international soft law; finally, the principles were widely cited and adapted by international organizations such as G20, UNESCO, and the EU, establishing a foundational discourse system for AI governance. Specifically, the five principles of “inclusive growth, sustainable development and well-being; human-centered values and fairness; transparency and explainability; robustness, security and safety; accountability” became standard templates for subsequent international AI governance documents.

2. Strategic Utilization of OECD Platform

The AIGO network fully utilized the existing institutional resources of the OECD platform, demonstrating the efficiency of platform dependency strategies. Resource mobilization manifested in three dimensions: first, utilizing OECD’s historical authority and professional reputation in economic policy coordination to provide legitimacy foundations for AI governance initiatives; second, activating existing expert networks, research capabilities, and multilateral coordination mechanisms, avoiding the costs of re-establishing professional networks; Third, leveraging soft law tools and advisory institutional arrangements to adapt to the uncertainty of governance needs in the emerging AI field.

The unique advantages of the OECD platform were fully leveraged by the AIGO network. The characteristic of similar values among member states significantly reduced consensus-building costs, with consistency among 37 member states in core values such as democratic institutions, market economies, and human rights protection providing a value foundation for rapid achievement of AI governance principles (OECD, 2011). The coordination efficiency provided by existing procedures and practices manifested in: standardized multilateral consultation

procedures reduced coordination costs, mature expert advisory mechanisms ensured the quality of technical input, and established peer review traditions enhanced the acceptability of principles. The international diffusion effect of OECD standards was amplified through the “OECD+ Partners” mechanism; by 2020, non-member countries including Argentina, Brazil, Colombia, Costa Rica, Malta, Peru, Romania, and Ukraine had joined the principles, demonstrating powerful demonstration effects (OECD, 2021).

The AIGO network addressed the inherent limitations of the OECD platform through innovative strategies. To address the issue of insufficient representation, a subsequent non-member state participation mechanism was established, absorbing more countries through the “OECD AI Network Observatory” (ONE AI) platform, achieving transformation from club governance to open governance. To address the issue of limited binding force, reinforcement was provided through moral authority and peer pressure mechanisms: on one hand, an annual implementation assessment mechanism was established, enhancing the effectiveness of principles through transparency and accountability; on the other hand, the influence scope was expanded through coordination and cooperation with other international organizations, such as establishing cooperative relationships with G20, ITU, IEEE, and other organizations (Cath, Wachter, Mittelstadt, Taddeo, et al., 2018). Data shows that from 2019-2023, OECD AI Principles were cited by over 40 international organizations and initiatives, becoming a benchmark framework for global AI governance.

3. Adaptive Design of Principle-based Governance Framework

The OECD AI Principles demonstrated clear adaptive institutional design characteristics, preserving policy space for rapid technological development through a principle-oriented rather than rule-oriented framework. Firstly, the five major principles used “should” rather than “must” formulations,

leaving policy space and cultural adaptation room for country implementation; Secondly, principle provisions avoided overly specific rule settings, with expressions such as “AI systems should be robust, secure and safe” providing directional guidance while maintaining openness for technological development; Finally, the institutional design choice of advisory rather than binding obligations enabled countries to implement differentially according to their own technological capabilities and governance needs.

The OECD AI Principles established systematic evolution mechanisms, reflecting the forward-looking nature of adaptive institutional design. Evolution mechanisms included regular assessment and update arrangements: a comprehensive assessment cycle every four years was established, with the 2023 assessment report demonstrating the principles’ adaptability in addressing generative AI challenges; flexible mechanisms for synchronous adjustment with technological development were established, such as updates to AI definitions in 2023-2024 in response to the rise of large language models like ChatGPT, emphasizing risk assessment and privacy protection requirements; institutionalized channels for experience learning and best practice sharing were established, with the OECD AI Policy Observatory collecting implementation experiences from various countries, forming feedback loops for policy learning. These mechanisms ensured that the principal framework could dynamically adjust with technological development.

Facing the high uncertainty of AI technological development, the OECD Principles achieved effective response through institutionalized arrangements. Response strategies manifested as multi-dimensional institutional design a dual-layer structure of principle stability and implementation flexibility was established, with core value principles remaining stable while specific implementation guidance adjusted with technological development. Furthermore, different countries were allowed to interpret and implement differentially according to their

development stages and cultural backgrounds. In addition to all these above technology-neutral principal formulations avoided lock-in to specific technological pathways. Empirical evidence shows that this design enabled the OECD Principles to effectively adapt to the technological leap from deep learning to generative AI; the 2023 updated principles maintained the core framework of the 2019 version while adding attention to emerging technology risks, demonstrating strong institutional resilience.

3.2 Case 2 2021 UNESCO AI Ethics Recommendation: Globally Inclusive Expert Networks

Following the foundational framework established by the OECD AI Principles, the second significant policy window opened in November 2021 with the adoption of the UNESCO AI Ethics Recommendation. This policy window extended from preliminary studies beginning in 2019 until formal adoption in November 2021, marking a crucial turning point in global AI governance from developed country dominance toward universal participation. In terms of authority, as a document adopted by a UN specialized agency, the UNESCO AI Ethics Recommendation possesses unquestionable international legal status and moral authority within the UN system. From a representativeness perspective, this policy's representativeness was unparalleled, representing the first global AI ethics normative instrument unanimously adopted by all 193 UN member states (UNESCO, 2019). Regarding agenda-framing capacity, the recommendation deepened and expanded AI governance from OECD's economic and social welfare perspective to more fundamental levels of human rights, dignity, cultural diversity, and environmental welfare, marking an important transition in global AI ethics governance from elite consensus to universal consensus.

The UNESCO Ad Hoc Expert Group (AHEG) functioned as the core policy entrepreneur network, exemplifying the typical characteristics of globally inclusive governance. According to official documentation, AHEG comprised 24 independent

international experts with strict balance in regional, gender, and disciplinary distribution: regional distribution covered all UNESCO regional groups, ensuring strong representation from Global South countries; gender composition achieved equal male-female configuration; disciplinary backgrounds spanned science, social sciences, economics, law, philosophy, and other fields, reflecting interdisciplinary integration characteristics. Compared to OECD's 70+ expert network, while AHEG was smaller in scale, its global representativeness and cultural diversity were significantly higher, reflecting a transition from quantitative advantage to qualitative balance.

1. Role Integration Performance of AHEG

The AHEG network faced technical translation challenges far exceeding the complexity of the OECD case, requiring the transformation of AI technology risks into ethical frameworks understandable and acceptable to different cultural backgrounds. The technical translation process embodied multiple cultural sensitivities: at the conceptual level, it needed to balance Western techno-centrism with developing countries' development needs, such as expanding the "algorithmic bias" concept into a comprehensive framework covering multi-dimensional discrimination including gender, race, religion, and culture; at the value level, it needed to coordinate ethical concepts from different civilizational traditions, such as finding balance points between individual privacy protection and collective interest priority (Cath, Wachter, Mittelstadt, & Floridi, 2018); at the practical level, it needed to consider technological capability differences among countries at different development stages, avoiding one-size-fits-all technical standards, achieving successful translation from Western technical discourse to global ethical language.

Then for mechanism design for diversified diplomatic coordination, the AHEG network demonstrated complexity management capabilities far exceeding the OECD's 37-country model in coordinating

interests among 193 member states (UNESCO, 2018, 2019). The diplomatic coordination mechanism manifested as multi-dimensional balancing strategies, it not only needed to coordinate developed countries' technology regulation needs with developing countries' development rights demands, achieving balance through differentiated responsibility principles but also needed to handle different civilizational traditions' varying understandings of AI ethics, protecting each country's cultural characteristics through cultural diversity clauses. Specific coordination mechanisms included: establishing regional representative rotation systems to ensure balanced expression of voices from all regions, setting up special working groups to handle contentious issues such as balancing regulation and innovation, and creating online consultation platforms to promote continuous dialogue. Data shows that the final recommendation received unanimous approval from 193 countries, creating a historic achievement in UN AI governance.

Finally, the AHEG network achieved a significant transition from Western-centrism to global inclusivity in normative leadership, reflecting normative innovation under multicultural contexts. Firstly, from OECD's "human-centric AI" concept to a comprehensive value framework encompassing human rights, dignity, diversity, inclusivity, and environmental sustainability; Secondly, from developed country elite consensus to universal consensus including Global South countries; Finally, from technology governance to comprehensive governance covering all social sectors including education, culture, development, and gender equality. Specific innovations included: the first explicit requirement in international AI governance documents to protect indigenous knowledge systems, reflecting deep attention to cultural diversity; establishing AI ethics impact assessment mechanisms, providing concrete implementation tools for countries; setting up the UNESCO AI Ethics Observatory, providing an institutionalized monitoring platform for global AI ethics governance (UNESCO, 2019).

2. Strategic Utilization of UNESCO Global Platform

The AHEG network fully mobilized UNESCO's unique platform advantages as a UN specialized agency, demonstrating the strategic value of global governance platforms. It not only has to obtaining authoritative status in international law through the UN framework, giving the AI Ethics Recommendation soft law binding force but also obtaining global moral authority through universal participation of 193 member states, exceeding the representativeness of any regional organization. So, the utilization strategies specifically focused on building an AI governance network by activating inter-agency consultation mechanisms and forging cooperative relationships with UNDP, ITU, WHO, and other UN agencies. These efforts were complemented by mobilizing UNESCO's established expert networks and research resources to minimize duplication, while leveraging the UN's unique convening power to ensure robust participation from developing countries. The UNESCO platform avoided direct competition with OECD through strategic differentiated positioning, achieving complementary development.

The UNESCO platform innovatively established global mobilization mechanisms, achieving unprecedented broad participation. Mobilization mechanisms manifested as multi-layered participation architectures: at the government level, official participation channels were established through countries' permanent delegations to UNESCO, ensuring effective expression of government positions; at the expert level, global expert participation was mobilized through existing professional networks such as UNESCO Chair networks and the International Council for Science; at the civil level, social mobilization was achieved through broad participation of non-governmental organizations, academic institutions, and private sectors; at the regional level, influence was expanded through cooperation with regional organizations such as the African Union, ASEAN, and the EU. The initiative introduced several key innovations, notably online participation platforms that transcended geographical and

temporal boundaries, multilingual services designed to minimize participation barriers, and comprehensive feedback mechanisms ensuring effective integration of diverse opinions.

3. Adaptive Design of Cultural Plurality

The UNESCO AI Ethics Recommendation achieved deep respect for global cultural diversity through institutionalized arrangements, reflecting innovative characteristics of adaptive institutional design. The recommendation explicitly recognized ethical values in different cultural traditions, avoiding the unidirectional orientation of Western ethical standards. In addition to this, culturally adaptive implementation guidelines were established, allowing countries to interpret differentially according to cultural backgrounds. Finally, special provisions for protecting indigenous knowledge systems were specifically established, reflecting institutionalized protection of cultural diversity. To ensure cultural sensitivity in AI ethics implementation, the framework introduced several mechanisms. Countries were required to assess cultural impacts through newly established procedures before implementing AI ethics policies. The UNESCO AI Ethics Observatory was tasked with monitoring cultural diversity through specially designed indicators, while dedicated dialogue platforms facilitated ethical exchanges across different civilizational traditions.

Facing the enormous differences among 193 member states in technological development levels and governance capabilities, the UNESCO Recommendation established flexible institutional adaptation mechanisms. Considering developing countries' limitations in AI technology capabilities, the recommendation proposed institutionalized arrangements for capacity building support, including specific measures such as technology transfer, talent development, and financial support. The refer to the implementation dimension, establishing progressive implementation pathways, allowing countries at different development stages to implement ethical requirements in phases according to their own

conditions. It also created a development-sensitive assessment system, avoiding evaluating implementation effects of countries at different development levels with uniform standards. Recognizing the resource and expertise gaps facing developing countries in AI governance, UNESCO introduced targeted innovations to address these challenges. Financial barriers were tackled through a specialized capacity building fund, while knowledge gaps were bridged via South-South cooperation mechanisms that facilitated peer-to-peer experience exchange. For nations with limited technical infrastructure, professional assistance networks were deployed to provide ongoing expert support. These combined efforts have reached more than 30 developing countries to date.

The UNESCO AI Ethics Recommendation established innovative global monitoring mechanisms, providing foundations for dynamic adjustment in adaptive institutional design. The UNESCO AI Ethics Observatory was established, achieving real-time monitoring of global AI ethics implementation through big data analysis and artificial intelligence technology; Then, a hybrid assessment model combining voluntary state reporting with multi-party evaluation was created, ensuring objectivity and comprehensiveness of monitoring results; Finally, mechanisms for civil society participation in monitoring were established, enhancing the social foundation of monitoring through participation of non-governmental organizations and academic institutions. By 2024, over 120 countries had submitted AI ethics implementation reports through the UNESCO platform, with the Observatory collecting over 10,000 AI ethics cases, providing rich practical foundations for global AI governance.

3.3 Case 3 2023 Bletchley Declaration: Crisis-driven Rapid Mobilization of Policy Entrepreneurs

The third significant policy window opened in November 2023 with the signing of the Bletchley Declaration, representing a crisis-driven policy response triggered by ChatGPT's release in November

2022. This policy window extended from the wave of expert warnings following ChatGPT's release until the successful hosting of the UK AI Safety Summit in November 2023, demonstrating unprecedented policy-making speed. In terms of authority, although the UK as a single country hosting the summit lacked the institutional authority of OECD or UNESCO, it successfully gained significant political authority by inviting 28 countries including China and the US, and choosing the historically symbolic Bletchley Park as the venue. From a representativeness perspective, signatories encompassed major AI research and development countries and different geopolitical camps, achieving rare great power coordination amid intensifying US-China competition. Regarding agenda-framing capacity, the declaration's uniqueness lay in representing the first global consensus document specifically focused on AI safety, sharply focusing the governance agenda from OECD's economic impacts and UNESCO's ethical concerns to the potential catastrophic safety risks of frontier AI, marking a direct policy response to generative AI technological breakthroughs.

The summit organization network led by the UK's Department for Science, Innovation and Technology (DSIT) exemplified the typical characteristics of crisis response policy entrepreneurs. According to official documentation, this network achieved the complete process from concept proposal to successful summit hosting within less than a year, demonstrating unprecedented policy mobilization speed. Network composition reflected cross-departmental integration characteristics: at the government level, DSIT established joint working mechanisms with the Foreign Office, Treasury, Home Office, and other departments to ensure policy coordination consistency; at the academic level, top research institutions including the Alan Turing Institute, Cambridge University, and Oxford University were rapidly mobilized to provide technical assessments and policy recommendations; at the industry level, direct dialogue channels were established with frontier AI companies such as DeepMind, OpenAI, and Anthropic to obtain first-hand technical

intelligence; at the international level, bilateral and multilateral coordination mechanisms with various countries were rapidly activated through the UK's diplomatic networks.

The UK summit organization network demonstrated acute crisis perception capabilities and rapid agenda-setting abilities. Rapid technical assessments of ChatGPT's breakthrough capabilities after its release identified potential risks and social impacts of generative AI (Hu, 2023). In addition to this, warning information from authoritative AI field experts such as Geoffrey Hinton and Yoshua Bengio was promptly captured and amplified, transforming technical concerns into policy urgency (Bengio, 2025). Moreover, media monitoring and public opinion analysis identified sharply rising social attention to AI risks. Networked operations in agenda setting manifested as: precise construction and promotion of the frontier AI safety concept through multi-channel discourse dissemination via academic papers, policy reports, and media interviews; strategic shaping of international agendas through pre-warming AI safety topics via existing multilateral mechanisms such as G7 and G20; precise positioning of summit themes focusing on frontier AI risks rather than generalized AI governance, reflecting strategic and targeted agenda setting.

1. Enhanced Role Integration in Technological Breakthrough Contexts

Facing the breakthrough development of generative AI, the UK summit organization network demonstrated powerful technical translation capabilities under high time pressure. Urgent response in technical translation manifested in three key segments: the first one is the rapid risk identification, through direct dialogue with top experts like Hinton and Bengio, quickly understanding and assessing potential risks of frontier AI systems, including capability emergence, alignment difficulties, and malicious use challenges; the second one is the precise conceptual innovation, successfully proposing and promoting the new concept of frontier AI safe-

ty, transforming complex technical risks into policy-operational governance objectives; the third one is the pragmatic policy frameworks, avoiding excessive entanglement in technical details and focusing on achievable policy options such as international cooperation mechanism building and risk monitoring network construction. Empirical evidence shows that technical content in summit documents underwent rigorous expert review, maintaining both technical accuracy and policy comprehensibility, reflecting the capability to achieve high-quality technical-policy translation within extremely short timeframes.

Then UK summit organization network achieved high-difficulty diplomatic coordination among 28 countries, particularly successfully facilitating simultaneous participation by both China and the US amid intensifying competition. Achievement mechanisms for diplomatic coordination manifested as multiple balancing strategies. It simultaneously invited major geopolitical forces including the US, China, and EU to avoid unilateralism accusations. Then, the UK summit organization network locked focus on frontier AI safety as a topic with public goods characteristics, avoiding more controversial sensitive areas such as military AI applications, data governance, and technological competition. Moreover, it also took a participation balancing strategy, including government representatives, corporate executives, and academic experts, reducing political sensitivity through multi-track diplomacy. Specific coordination mechanisms included: establishing bilateral pre-consultation procedures, conducting in-depth communication with major participants before the summit to identify key divergences and seek compromise space; setting up technology-neutral frameworks, avoiding biased statements toward specific AI technological pathways or business models; creating flexible commitment models, reducing political costs for countries through expressions of commitment to cooperation rather than specific obligations.

The UK summit organization network played innovative leadership roles in global AI governance nor-

mative reconstruction, successfully promoting strategic transformation of the governance agenda. It first established frontier AI safety as an independent international governance issue, distinct from traditional AI ethics and economic impact discussions; framework innovation, proposing a governance framework based on risk assessment-monitoring and warning-cooperative response, providing new institutional templates for international AI safety cooperation (Prime Minister's Office, 10 Downing Street et al., 2023). Then creating continuous international cooperation mechanisms through summit serialization design and establishment of the International Network of AI Safety Institutes (INAI SI). The leadership demonstrated strategic sophistication across multiple dimensions. At the symbolic level, selecting Bletchley Park, a site steeped in cryptographic and computational history, effectively constructed discourse power and generated potent symbolic capital for AI safety governance. Diplomatically, the initiative achieved remarkable inclusivity by bringing major AI powers, including China, to the table, thereby transcending the zero-sum logic of technological cold war. In terms of institutional foresight, the framework extended beyond immediate concerns about GPT-type models to accommodate governance challenges posed by potential Artificial General Intelligence (AGI) that may materialize in the future.

2. Platform Innovation Strategy of Summit Diplomacy

Facing the urgency of generative AI technological breakthroughs, the UK summit organization network identified significant limitations of existing international organization platforms, driving the necessity for platform innovation. Firstly, decision-making procedures of traditional international organizations like OECD and UNESCO typically require several years, unable to match the rapid pace of AI technological development and urgent needs for risk response; participation limitations, while OECD's membership restrictions and UNESCO's broad participation each had advantages, neither could achieve optimal configuration of including

major AI powers while maintaining decision-making efficiency; Secondly, existing platforms had formed path dependencies in AI governance, with OECD focusing on economic impacts and UNESCO emphasizing ethical values, both struggling to specifically address frontier AI safety risks. The time window from ChatGPT's release to global policy response was only 11 months, far shorter than traditional international organizations' policy-making cycles, objectively requiring more flexible and rapid coordination mechanisms. This temporal mismatch became the core driving force for platform innovation.

The UK innovatively chose bilateral summit diplomacy as a platform strategy, fully leveraging the unique advantages of this platform model. Platform advantages manifested in multiple dimensions: firstly, summit host countries possessed complete agenda-setting and rule-making authority, able to flexibly adjust meeting content and participation conditions according to technological development dynamics and political environmental changes; secondly, avoiding complex procedures and interest games within international organizations, able to achieve complete processes from issue proposal to policy output within short timeframes; finally, the uniqueness and timeliness of the summit attracted high global media attention, generating public opinion influence far exceeding traditional international organization meetings. Data shows that global media coverage of the Bletchley Summit exceeded that of concurrent OECD or UNESCO meetings by over 300%, reflecting the significant advantages of new platforms in the attention economy era.

The UK summit organization network demonstrated forward-looking design capabilities in transforming one-time diplomatic events into continuous institutional mechanisms. The institutionalization design demonstrated innovation by building temporal, structural, and adaptive dimensions into the framework. Rather than settling for a singular diplomatic moment, organizers embedded continuity through commitments to hold subsequent summits in South Korea (2024) and France (2025), establi-

shing a serialized process. Between these milestone events, the International Network of AI Safety Institutes (INAISI) would provide ongoing coordination, transforming episodic diplomacy into sustained institutional engagement. Crucially, the architecture retained flexibility for expanding participation and adjusting agendas as circumstances evolved, ensuring the platform could adapt over time. Three key functions were operationalized through specific institutional achievements. Global coordination capacity emerged as INAISI established membership spanning continents and creating a genuinely worldwide AI safety research collaboration. Industry accountability was advanced through voluntary commitment frameworks that engaged major AI companies, notably OpenAI, DeepMind, and Anthropic, in governance processes. Dynamic responsiveness was built in through annual review mechanisms designed to keep governance frameworks synchronized with the pace of technological advancement.

3. Adaptive Design under Extreme Uncertainty

Facing extreme uncertainty of frontier AI technology, the Bletchley Declaration adopted highly flexible institutional design strategies, reflecting innovative characteristics of adaptive governance. Design choice flexibility manifested in three core characteristics: commitment-oriented rather than obligation-oriented, the declaration used soft expressions such as commit to working together and dedicated to cooperation, avoiding implementation difficulties and political resistance that specific binding obligations might bring (Sheehan, 2024); principle-oriented rather than rule-oriented, focusing on establishing common understanding and cooperation frameworks rather than detailed technical standards or regulatory rules, preserving maximum policy adjustment space for rapid technological changes; process-oriented rather than outcome-oriented, emphasizing establishment of continuous dialogue and assessment mechanisms rather than one-time solutions, reflecting institutionalized adaptation to technological uncertainty. The design philosophy emphasized pragmatism over prescription at every level. Language choices reflected this

approach: framers deliberately spoke of identifying risks rather than eliminating them, acknowledging that risk assessment must be continuous and dynamic rather than definitive. This realism extended to implementation strategies, where sharing best practices replaced demands for unified standards, allowing nations to craft governance solutions responsive to their contexts. The treatment of corporate participation embodied similar flexibility, with voluntary commitments preferred over binding obligations as a means of encouraging industry self-regulation while avoiding the constraints of heavy-handed mandates.

The Bletchley Declaration established innovative dynamic monitoring and assessment mechanisms, providing institutionalized foundations for responding to rapid evolution of frontier AI technology. Monitoring and assessment mechanisms manifested as multi-layered institutional arrangements: at the technical monitoring level, real-time tracking systems for global AI capability development were established through AI safety institute networks, focusing on capability boundaries and potential risks of frontier technologies such as large language models, multimodal systems, and agents; at the risk assessment level, evidence-based risk assessment frameworks were established, providing bases for policy adjustments through regular assessment report publications; at the cooperation monitoring level, public-private cooperation information sharing mechanisms were created through combinations of voluntary corporate reporting and government regulatory information. Specific innovations included: establishing “frontier AI capability benchmarks,” providing standardized tools for assessing AI system dangerousness; setting up early warning systems, promptly alerting governments when AI capabilities experienced major breakthroughs; creating best practice databases, collecting and sharing successful experiences from various countries in AI safety governance. Empirical data shows that by 2024, the INAIS network had published the first International Scientific Report on AI Safety, providing scientific foundations for subsequent summits in South Korea and France (Center for Strategic and International Studies, 2024).

The Bletchley process embodied high institutional evolution openness, reserving sufficient adjustment space for adapting to unpredictable AI technological development. Firstly, the summit series reserved mechanisms for new member accession, with expected gradual expansion from 28 countries as AI technology globally diffused and countries’ governance needs grew; Secondly, from the first summit’s focus on “frontier AI safety” to subsequent summits potentially involving broader topics such as AI governance, AI development, and AI ethics, reflecting dynamic expansion capabilities in issue scope; Finally, avoiding institutional competition and achieving institutional complementarity through coordination and integration with existing multilateral mechanisms such as the UN, OECD, G7, and G20.

The strategic architecture prioritized adaptability across temporal, thematic, and systemic dimensions. Temporal flexibility emerged through tiered participation structures: countries could enter through observer roles and progressively deepen engagement as their capacity and commitment grew. Thematic flexibility was secured via “issue evolution roadmaps” that would guide summit agenda adjustments in response to technological breakthroughs and shifting international priorities. Systemic flexibility came through carefully designed interface standards allowing seamless coordination between the Bletchley process and other multilateral AI governance efforts. This multidimensional adaptability fundamentally distinguished the initiative from rigid crisis management tools, equipping it instead to function as sustainable long-term governance infrastructure.

CONCLUSIONS

Through systematic analysis of three milestone policy window cases: the 2019 OECD AI Principles, 2021 UNESCO AI Ethics Recommendation, and 2023 Bletchley Declaration. This research validates three core hypotheses about policy entrepreneur networks in global AI governance and reveals adaptive mechanisms of these networks under different institutional environments and technological conditions (table 1).

Table 1 Key Characteristics and Adaptive Mechanisms of Policy Entrepreneur Networks

	Network	Integrated Roles	Platform Strategy	Adaptive Design	Challenges Overcome
OECD AI Principles	AIGO (Institutionalized Experts)	Technical translation; Diplomatic coordination	Platform Dependency (Historical authority)	Principle-oriented; Evolution mechanism	High coordination costs; Technical ambiguity
UNESCO Recommendation	AHEG (Inclusive Global Experts)	Normative leadership; Cultural coordination	Strategic Differentiation (Universal legitimacy)	Cultural plurality; Developmental flexibility	North-South divide; Cultural diversity
Bletchley Declaration	DSIT (Crisis-driven Network)	Rapid risk identification; Great power mediation	Platform Innovation (Summit diplomacy)	Extreme flexibility; Process-oriented	Temporal mismatch; Geopolitical competition

The OECD case represented a relatively simple homogeneous network model: 70+ experts from 37 developed countries with similar values, possessing high consistency in technical understanding, policy preferences, and institutional traditions, making role integration and consensus-building relatively easy to achieve. The UNESCO case embodied a significantly complexified heterogeneous network model: while 24 experts were fewer in number, they represented 193 vastly different member states, requiring coordination of multiple complex factors including North-South development gaps, East-West cultural differences, and different institutional traditions, with role integration difficulty increasing exponentially. The Bletchley case demonstrated an extremely complex crisis mobilization network: integrating diverse multi-sectoral, transnational, and cross-industry actors within the extremely short timeframe of 11 months, achieving rapid coordination among 28 countries amid intensifying US-China competition.

The three cases present clear differentiation trends in platform strategies of policy entrepreneur networks, developing from simple platform dependency to strategic platform innovation. The OECD case embodied

typical platform dependency strategy: fully utilizing existing institutional authority, expert networks, and coordination mechanisms, achieving rapid AI governance breakthroughs by activating OECD's historical advantages in economic policy. The UNESCO case demonstrated strategic platform differentiation strategy: consciously differentiating positioning from the OECD model, achieving transformation from developed country club governance to global democratic governance by mobilizing the universal legitimacy, moral authority, and professional reputation of the UN system. The Bletchley case represented breakthrough platform innovation strategy: facing temporal lags and procedural constraints of existing platforms in responding to technological breakthroughs, innovatively choosing summit diplomacy models to achieve optimal combinations of efficiency, flexibility, and symbolic significance. More importantly, through serialization design and INAISI network construction, it successfully achieved transformation from one-time diplomatic events to continuous institutional mechanisms, demonstrating the institutionalization potential of platform innovation.

Finally, the three cases demonstrate an innovative progression process in adaptive institutional de-

sign, developing from technology-neutral design to culturally-sensitive design to extremely flexible design. The OECD Principles established foundational templates for adaptive design: preserving policy space for technological development through principle-oriented rather than rule-oriented frameworks; reducing country implementation resistance through soft recommendations rather than hard constraints; ensuring governance frameworks synchronized with technological development through regular assessment and update mechanisms. The UNESCO Recommendation achieved deep innovation in adaptive design: adding cultural adaptability and developmental adaptability dimensions to technological adaptability foundations. Cultural adaptability achieved localized adaptation of global ethical frameworks through mechanisms including recognizing ethical values of different civilizational traditions, establishing culturally-sensitive implementation guidelines, and protecting indigenous knowledge systems. Developmental adaptability addressed challenges posed by North-South development gaps to global governance through arrangements including capacity building support, progressive implementation pathways, and development-sensitive assessment systems. The Bletchley Declaration represented extreme innovation in adaptive design: facing extreme uncertainty of frontier AI technology, adopting institutional arrangements with maximized flexibility. This extreme adaptability manifested in three characteristics: highly open participation mechanisms, reserving space for new member accession and agenda expansion; extremely flexible commitment models, reducing political costs through voluntary commitments rather than binding obligations; forward-looking evolutionary design, reserving development space for long-term governance through summit serialization and network institutionalization. The rapid establishment of the INAIS network and publication of the first International Scientific Report on AI Safety proved the effectiveness of this extreme adaptability design.

To sum up, this research finds that successful entrepreneur networks must integrate technical trans-

lation, diplomatic coordination, and normative leadership. We propose a dynamic platform strategy where platform dependency works during stable periods, but platform innovation becomes necessary during technological breakthroughs. The core architecture of adaptive design relies on the integration of three mechanisms: stability through principle-orientation, flexibility in implementation, and evolution through dynamic assessment.

While this study focuses on state-led milestones between 2018 and 2023, the findings open several avenues for future research. As AI governance moves from defining principles to practical enforcement, further inquiry is needed into how these frameworks function across different legal and cultural settings, particularly in the Global South. Ensuring that international standards respect local developmental needs and digital sovereignty remains a key challenge for future institutional design. Additionally, there is a need to examine how private actors, such as OpenAI and Anthropic, influence global agendas through informal channels. Finally, exploring the friction between regional regulations like the EU AI Act and the interests of the broader international community will help clarify how a more inclusive and balanced governance landscape can be built.

Global AI governance is in a critical period of rapid development and profound transformation, with policy entrepreneur networks playing increasingly important roles. As AI technology continues to breakthrough and governance needs continuously evolve, policy entrepreneur networks will inevitably face new challenges and opportunities, providing broad space and rich materials for future research.

BIBLIOGRAPHIC REFERENCES

- Arnold, G., & St. John, S. (2023). Finding, distinguishing, and understanding overlooked policy entrepreneurs. *Policy Sciences*. <https://doi.org/10.1007/s11077-023-09515-4>

- Bengio, Y. (2025, June 3). Introducing LawZero. <https://yoshuabengio.org/2025/06/03/introducing-lawzero/>
- Birkland, T. A. (2020). *An introduction to the policy process: Theories, concepts, and models of public policy making* (5th ed.). Routledge.
- Cai, C., & Zhang, L. (2025). Exploring Global Artificial Intelligence Governance: A Principal-Agent Theory Perspective. *The Journal of International Studies*, 46(2), 9–35, 5.
- Cath, C., Wachter, S., Mittelstadt, B., & Floridi, L. (Eds.). (2018). *Governing artificial intelligence: Ethical, legal, and technical opportunities and challenges*. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 376(2133). <https://doi.org/10.1098/rsta.2018.0080>
- Cath, C., Wachter, S., Mittelstadt, B., Taddeo, M., & Floridi, L. (2018). Artificial Intelligence and the ‘Good Society’: The US, EU, and UK approach. *Science and Engineering Ethics*, 24(2), Article 2. <https://doi.org/10.1007/s11948-017-9901-7>
- Center for Strategic and International Studies. (2024). *The AI safety institute international network: Next steps and recommendations*. Center for Strategic and International Studies.
- Chui, M., Hazan, E., Roberts, R., Singla, A., & Smaje, K. (2023). *The economic potential of generative AI*. McKinsey Digital. <https://www.mckinsey.com/capabilities/mckinsey-digital/our-insights/the-economic-potential-of-generative-ai-the-next-productivity-frontier#introduction>
- Ding, J. (2024). *Technology and the rise of great powers: How diffusion shapes economic competition*. Princeton University Press.
- Erman, E., & Furendal, M. (2024). Artificial Intelligence and the Political Legitimacy of Global Governance. *Political Studies*, 72(2), 421–441. <https://doi.org/10.1177/00323217221126665>
- Genus, A., & Stirling, A. (2018). Collingridge and the dilemma of control: Towards responsible and accountable innovation. *Research Policy*, 47(1), 61–69.
- Grobelnik, M., Perset, K., & Russell, S. (2024). The transformation brought about by AI. ASEF. <https://asef.net/2024/05/25/marko-grobelnik-the-transformation-brought-about-by-ai>
- Hu, K. (2023). ChatGPT sets record for fastest-growing user base—Analyst note. Reuters. <https://www.reuters.com/technology/chatgpt-sets-record-fastest-growing-user-base-analyst-note-2023-02-01/>
- Kerry, C. F., Meltzer, J. P., Renda, A., & Wyckoff, A. W. (2025, February 10). *Network architecture for global AI policy*. Brookings. <https://www.brookings.edu/articles/network-architecture-for-global-ai-policy/>
- Kingdon, J. W. (1995). *Agendas, Alternatives, and Public Policies*, (2nd) HarperCollins College Publishers. New York.
- Martin, L. L., & Simmons, B. A. (1998). Theories and Empirical Studies of International Institutions. *International Organization*, 52(4), 729–757. <https://doi.org/10.1162/002081898550734>
- Mintrom, M., & Moulton, A. (2019). *Policy entrepreneurs and dynamic change*. Cambridge University Press.
- Mintrom, M., & Norman, P. (2009). Policy entrepreneurship and policy change. *Policy Studies Journal*, 37(4), 649–667. <https://doi.org/10.1111/j.1541-0072.2009.00329.x>
- OECD. (2011). *OECD guidelines for multinational enterprises*. OECD Publishing. <https://doi.org/10.1787/9789264115415-en>
- OECD. (2019a). *Mandate of the working party on artificial intelligence governance (AIGO)*. Organisation for Economic Co-operation and Development, Digital Policy Committee. <https://oecd.ai/en/network-of-experts>

- OECD. (2019b). Recommendation of the council on artificial intelligence (C(2019)3; OECD Legal Instruments). Organisation for Economic Co-operation and Development. <https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0449>
- OECD. (2019c). Scoping the OECD AI principles: Deliberations of the expert group on artificial intelligence at the OECD (AIGO) (OECD Digital Economy Papers 291). OECD Publishing. <https://doi.org/10.1787/d62f618a-en>
- OECD. (2021). State of implementation of the OECD AI principles: Insights from national AI policies (OECD Digital Economy Papers No. 311). Organisation for Economic Co-operation and Development.
- Perry, B., & Uuk, R. (2019). AI Governance and the Policy-making Process: Key Considerations for Reducing AI Risk. *Big Data and Cognitive Computing*, 3(2), Article 2. <https://doi.org/10.3390/bdcc3020026>
- Pierson, P. (1993). When effect becomes cause: Policy feedback and political change. *World Politics*, 45(4), 595–628. <https://doi.org/10.2307/2950710>
- Prime Minister's Office, 10 Downing Street, Department for Science, Innovation and Technology, & Foreign, Commonwealth & Development Office. (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/the-bletchley-declaration-by-countries-attending-the-ai-safety-summit-1-2-november-2023>
- Radu, R. (2021). Steering the governance of artificial intelligence: National strategies in perspective. *Policy and Society*, 40(2), 178–193. <https://doi.org/10.1080/14494035.2021.1929728>
- Reuel, A., & Undheim, T. A. (2024). Generative AI Needs Adaptive Governance (arXiv:2406.04554). arXiv. <https://doi.org/10.48550/arXiv.2406.04554>
- Sheehan, M. (2024). The bletchley declaration: Towards adaptive global AI governance? *AI & Society*, 39(2), 1–15
- Silver, D., Huang, A., Maddison, C. J., Guez, A., Sifre, L., van den Driessche, G., Schrittwieser, J., Antonoglou, I., Panneershelvam, V., Lanctot, M., & others. (2016). Mastering the game of Go with deep neural networks and tree search. *Nature*, 529(7587), 484–489.
- Taeihagh, A. (2021). Governance of artificial intelligence. *Policy and Society*, 40, 137–157. <https://doi.org/10.1080/14494035.2021.1928377>
- Tan, S. Y., & Taeihagh, A. (2021). Adaptive governance of autonomous vehicles: Accelerating the adoption of disruptive technologies in Singapore. *Government Information Quarterly*, 38(2), Article 2. <https://doi.org/10.1016/j.giq.2020.101546>
- Thierer, A. (2018). The Pacing Problem and the Future of Technology Regulation. The Mercatus Center. <https://www.mercatus.org/economic-insights/expert-commentary/pacing-problem-and-future-technology-regulation>
- UNESCO. (2018). Steering AI and advanced icts for knowledge societies: A ROAM perspective. UNESCO. <https://unesdoc.unesco.org/ark:/48223/pf0000265650>
- UNESCO. (2019, March). Preliminary study on the ethics of artificial intelligence. UNESCO. <https://unesdoc.unesco.org/ark:/48223/pf0000367823>
- Whyte, C. (2020). Poison, persistence, and cascade effects: AI and cyber conflict. *Strategic Studies Quarterly*, 14(4), 82–101.
- Yeung, K., & Howes, A. (2018). Government by algorithm: The role of AI in policy-making. *Policy & Internet*, 10(4), 382–396.

CONFLICT OF INTEREST

The authors declare that there are no conflicts of interest related to this article.

ACKNOWLEDGMENTS

Not applicable.

AUTHORSHIP CONTRIBUTION:

Not applicable.

Luyao Zhang: Conceptualization, Data curation, Formal analysis, Research, Methodology, Writing - original draft, Writing - revision and editing.

Cuihong Cai: Conceptualization, Data curation, Formal analysis, Acquisition of funds, Research, Methodology, Supervision, Writing - revision and editing.

FUNDING

Not applicable.

PREPRINT

Not published.

RESEARCH ETHICS STATEMENT**DATA AVAILABILITY STATEMENT**

Not applicable.

COPYRIGHT

The copyright is held by the authors, who grant Journal Política Internacional exclusive rights for first publication. The authors may enter into additional agreements for the non-exclusive distribution of the version of the work published in this journal (for example, posting in an institutional repository, on a personal website, publishing a translation, or as a book chapter), with acknowledgment that it was first published in this journal. Regarding copyright, the journal does not charge any fees for submission, processing, or publication of articles.