世界互联网大会 World Internet Conference

武漢大學 WUHAN UNIWERSITY 新疆大學 XINJIANG UNIVERSITY

Outcomes of World Internet Conference Think Tank Cooperation Program

Global Legislative Developments in AI Governance: Experience and Prospects

Center for Competition Law and Policy, Wuhan University

School of Law, Xinjiang University

Institute for Network Governance, Wuhan University

APRIL 2026


世界互联网大会
World Internet
Conference


武漢大學
WUHAN UNIWERSITY
新疆大學
XINJIANG UNIVERSITY

Research Team

Sun Jin

Second-Tier Professor, School of Law, Wuhan University

Director, Center for Competition Law and Policy, Wuhan University

Executive Dean, Institute for Network Governance, Wuhan University

Paziliya Yusufu

Lecturer, School of Law, Xinjiang University

Summary

The rapid development of artificial intelligence technology not only brings innovation opportunities, but also triggers complex governance challenges such as data security, algorithm bias and liability determination. In order to systematically sort out the legislative status of global artificial intelligence governance and explore future directions, this article conducts an in-depth analysis of the legislative practices of major countries and regions such as the United States, the European Union, and China, and compares the characteristics and differences of different governance models. The research adopts case analysis and comparative research methods, focusing on the institutional design of various countries in core areas such as data governance, algorithm supervision and damage liability. The study found that the legislative models of various countries are significantly differentiated: the United States focuses on innovation-friendly flexible regulation, the European Union has established a strict framework based on risk classification, and China is exploring a local path that attaches equal importance to development and security. Currently, global governance faces prominent bottlenecks such as legal lag caused by rapid technological iteration, differences in regulatory standards between countries hindering collaboration, and digital divide exacerbating social injustice. Future legislation needs to increase agility, such as introducing a regulatory sandbox mechanism; deepening international cooperation to bridge differences in rules; and paying more attention to the inclusiveness of technology to ensure that the development of artificial intelligence is in line with the common interests of mankind. The study believes that balancing technological innovation and risk prevention and control, and building an inclusive and collaborative global governance system are the keys to the healthy development of artificial intelligence.

Contents

I. Background to the Development of Artificial Intelligence and Legislative Approaches to Its Governance Worldwide

(I) Technological Evolution and Industrial Landscape: Trends in the Development of Artificial Intelligence

1. Breakthroughs in Core Technologies and the Distribution of Principal Types

Machine learning, as the cornerstone of artificial intelligence, continues to advance the capacity for identifying complex patterns within both supervised and unsupervised learning paradigms.)[1]As a significant branch thereof, deep learning — relying on multi-layer neural networks—has markedly improved the accuracy of image recognition (such as medical imaging analysis) and natural language processing; recurrent neural networks (RNN) have further enhanced the modelling of sequential data. [2]The advent of the Transformer architecture has driven a qualitative leap in natural language processing. Large-scale models, by virtue of self-attention mechanisms, are able to capture deep semantic relationships with precision, thereby laying the groundwork for the rapid emergence of generative artificial intelligence (Generative AI). [3]

From the perspective of developmental pathways, artificial intelligence exhibits a dual-track trajectory comprising both "narrow" and "general" systems. Narrow artificial intelligence (Narrow AI) is highly specialised and focused on vertical tasks within specific domains, such as medical diagnostic systems [4]and financial risk control models [5]While highly efficient in operation, such systems remain constrained by scenario-specific limitations and barriers to cross-domain transfer.[6]By contrast, the rudimentary form of artificial general intelligence (AGI), through large-scale pre-trained models, seeks to simulate cross-domain intelligence and demonstrates considerable potential for multi-task generalisation ,[7] but it still faces deficiencies in reasoning capacity as well as ethical risks.

Because of its greater maturity, narrow AI has already been deployed on a large scale in fields such as industrial automation. Although general AI is still at an early stage, it has driven the of industrial chains such as cloud computing and given rise to new business models. The diversification of technologies has also stratified the industrial ecosystem into foundational computing chips, intermediate algorithmic frameworks, and application-layer industry solutions and consumer products. While this structure accelerates the diffusion of innovation, it also gives rise to lagging standards and regulatory challenges. [8]

2. Expansion of Industrial Scale and the Global Spatial Distribution

The global artificial intelligence industry has continued to expand in scale. By 2023, the total market size had

[1] Xue, Lan, and Wang, Jingyu. "Frontier Trends in the Development of Artificial Intelligence, Governance Challenges, and Response Strategies". *Administrative Reform*, 15.8 (2024): 4–13.
[2] Chen, Lingxiang. "Legal Regulatory Pathways for Big Data Cleaning in AI Medical Devices". *Journal of Northeastern University (Social Science)*, 27.5 (2025): 117.
[3] Xie, Xiao, and Luo, Shijie. "On the Dynamic Risks of Generative Artificial Intelligence and Adaptive Governance". *Journal of Beijing University of Technology (Social Sciences Edition)*, 25.1 (2025): 112–125.
[4] World Health Organization. *Ethics and governance of artificial intelligence for health: large multi-modal models. WHO guidance.* World Health Organization, 2024.
[5] Pantanowitz, Liron, et al. "Regulatory aspects of artificial intelligence and machine learning." *Modern Pathology* 37.12 (2024): 100609.
[6] Fu, Xinhua. "Pluralistic Trends in Global AI Legislation and the Chinese Model". *SJTU Law Review*, (6) (2025): 60–73.
[7] Zhou, Hongyu, and Li, Yuyang. "Generative Artificial Intelligence Technology ChatGPT and the Modernisation of Educational Governance—Also on the Transformation of Educational Governance in the Digital Era". *Journal of East China Normal University (Educational Sciences),* 41.7 (2023): 36.
[8] Ji, Weidong, and Zhao, Zerui. "How Can Law Realise Dynamic Governance? Legislative Responses to the Uncertainty of Artificial Intelligence". *Academic Monthly*, 57.3 (2025): 98–109.

already exceeded several hundred billion US dollars ,[9] with growth driven by algorithmic optimisation, enhanced computing power, and the deep integration of AI across sectors such as medical diagnostics, financial risk control, intelligent manufacturing, and smart transportation. North America, benefiting from strong research and development capabilities and well-established capital markets, occupies a central position, leading innovation in foundational algorithms, high-performance chips, and cloud services. Europe, by contrast, places greater emphasis on the deepening of sector-specific applications and the construction of ethical frameworks, developing distinctive strengths in areas such as industrial automation, healthcare AI, and data privacy protection ,[10] although disparities remain among Member States in terms of the pace of technological transformation and industrialisation .[11]

Asia, particularly East Asia, has emerged as the fastest-growing region. China, Japan, and South Korea, propelled by government strategies and market demand, have rapidly developed their respective industrial ecosystems.[12] China, in particular, demonstrates notable strengths in applied technologies such as computer vision and natural language processing, as well as in implementation across e-commerce and smart city initiatives. However, in comparison with North America, Asia still has ground to cover in foundational theory and core hardware research. Overall, divergences across regions in terms of research focus, comparative advantages, and governance trajectories have resulted in a markedly asymmetrical pattern in the global AI industry.

3. Patterns of Future Evolution and Key Development Trends

Artificial intelligence is increasingly characterised by multi-dimensional convergence. Quantum computing is empowering machine learning to achieve foundational breakthroughs in areas such as pharmaceutical research and development, while the deep coupling of AI with biotechnology is reshaping paradigms in the life sciences. Application scenarios are expanding from isolated use cases to comprehensive domains such as the metaverse and smart cities, relying on real-time data to drive system-level intelligence upgrades, which in turn is compelling a reconfiguration of computing infrastructure towards integrated cloud – edge architectures.

At the level of industrial ecosystems, mergers and acquisitions by major technology firms are accelerating the internalisation of key technologies, while open-source communities are facilitating the democratisation of technological development. At the same time, regulatory divergences across jurisdictions are prompting adjustments in global industrial deployment, and generative AI is triggering profound restructuring of value chains in industries such as publishing.

[9] Ye, Shulan, and Li, Mengting. “Global Artificial Intelligence Governance: Progress, Dilemmas, and Prospects”. *International Studies*, 66.5 (2024): 00–118.
[10] Walter, Yoshija. “Managing the race to the moon: Global policy and governance in Artificial Intelligence regulation—A contemporary overview and an analysis of socioeconomic consequences.” *Discover Artificial Intelligence* 4.1 (2024): 14.
[11] Zhang, Xinxin, and Ji, Yu. “On the Standardisation Model in the EU Artificial Intelligence Act and Its Implications for China”. *Standard Science*, (5) (2024): 39–46.
[12] Xue, Lan, and Zhao, Jing. “International Governance of Artificial Intelligence: An Analysis Based on Technological Characteristics and Issue Attributes”. *International Economic Review,* (3) (2024): 52–69.

(II) Risk Attribution and Legal Challenges: Practical Dilemmas Arising from Artificial Intelligence

1. Risks of Data Infringement and the Protection of Personal Privacy

The application of artificial intelligence gives rise to diverse data security risks, the principal drivers of which lie in data misuse, unlawful collection, and privacy breaches. Data misuse typically manifests in unauthorised secondary use, such as the appropriation of health and physiological data for commercial analytics; unlawful collection often involves excessive background harvesting of sensitive information, including location data and contact lists; privacy breaches may arise from database compromises or the reverse engineering of individual characteristics through algorithmic inference. Such risks not only directly infringe upon individual rights and interests, but may also give rise to deeper forms of discriminatory harm.

Existing legal responses remain inadequate. Uncertainty in data ownership renders rights enforcement difficult for individuals; ambiguous boundaries of permissible use frequently result in secondary uses exceeding the scope of original consent; and remedies for infringement are often hindered by high evidentiary burdens, coupled with platforms invoking "technological neutrality" as a defence against liability.[13]

Effective governance requires the construction of a multi-layered regulatory framework. While the stringent regime under the EU General Data Protection Regulation (GDPR) offers a valuable point of reference, it may also risk constraining innovation. A more workable approach lies in the adoption of risk-based regulatory models, whereby high-risk scenarios—such as healthcare—are subject to strict controls, while data flows for foundational research are comparatively relaxed.[14] At the same time, technical safeguards should be reinforced, including differential privacy and verifiable deletion mechanisms.[15] China's algorithm filing and security assessment regimes provide an example of dynamic regulatory practice. The overarching objective is to preserve space for the responsible development of AI while safeguarding fundamental rights, which in turn necessitates the continuous adaptive refinement of legislative rules.

2. The "Black Box" Effect of Algorithms and the Tension with Fairness and Justice

Deficiencies in the transparency and explainability of AI-driven decision-making have given rise to significant concerns regarding social fairness. The "black box" effect may amplify existing biases, as evidenced in cases of gender discrimination in recruitment and discriminatory loan denials in financial services. Governance approaches vary across sectors: in healthcare, algorithmic explanation reports are often mandated, with physicians retaining a power of override; in e-commerce, systems increasingly incorporate user-managed profiling mechanisms; under the EU Artificial Intelligence Act, high-risk applications are subject to stringent documentation requirements.

[13] Fernández, José Vida. "Artificial intelligence in government: Risks and challenges of algorithmic governance in the administrative state." *Ind. J. Global Legal Stud.* 30 (2023): 65.
[14] Ebers, Martin. "Truly risk-based regulation of artificial intelligence how to implement the EU's AI Act." *European Journal of Risk Regulation* 16.2 (2025): 684-703.
[15] Lami, Bareq, et al. "The role of artificial intelligence (AI) in shaping data privacy." *International Journal of Law and Management* 68.2 (2026): 296-318.

Enhancing algorithmic auditability and accountability is therefore critical. On the technical level, this entails the development of counterfactual explanation tools; on the legal level, it calls for the establishment of tiered liability regimes. For instance, in cases of misdiagnosis by medical AI, developers may be required to demonstrate compliance with regulatory standards, while users bear responsibility for improper modifications. In parallel, independent third-party certification bodies should be instituted to conduct bias detection and assessment.

3. Attribution of Tort Liability and the Identification of Legal Subjects

The attribution of liability for harm caused by AI systems presents profound legal challenges, as traditional frameworks of liability struggle to accommodate their technological characteristics. Key points of contention centre on product liability, user liability, and developer liability. From the perspective of product liability, manufacturers of AI systems treated as ordinary products may be subject to strict liability; however, the autonomous learning capabilities of such systems often result in outcomes that deviate from original design expectations, thereby complicating causation analysis. User liability emphasises operator fault, yet the high degree of autonomy in AI systems weakens direct human control—for example, in Level 3 and above autonomous driving, where the driver assumes only a supervisory role, rendering the scope of the duty of care ambiguous. Developer liability concerns defects in algorithm design, but the "black box" nature of deep learning renders it difficult to trace errors within code, leading in some cases—such as misdiagnosis by medical AI — to a vacuum of accountability.

Representative cases illustrate these judicial dilemmas. In a 2022 German case involving a fatal accident caused by an autonomous driving system, primary liability was attributed to the manufacturer under the Product Liability Directive, yet the case exposed structural deficiencies in allocating responsibility within multi-actor systems. In disputes concerning misdiagnosis by IBM Watson for Oncology, disagreements between hospitals and developers over algorithmic responsibility revealed regulatory gaps in human–machine collaborative decision-making.

A viable solution lies in the construction of a tiered liability framework. For intelligent products with a physical embodiment (such as industrial robots), product liability rules may be strengthened, including requirements for manufacturers to maintain algorithmic decision logs. For software-based systems such as generative AI, a three-tier framework of "development–deployment–use" should be applied: developers bear responsibility for the safety of foundational models; deployers are responsible for scenario-based testing; and users must demonstrate compliant operation.[16] Drawing on the EU Artificial Intelligence Act, mandatory liability insurance for high-risk systems may be introduced to facilitate risk socialisation and distribution.

At a more fundamental level, legislative recognition of a form of "limited legal personality" for AI systems may be explored, enabling them to bear procedural responsibilities while preserving ultimate accountability with human actors.[17]

[16] Laux, Johann, Sandra Wachter, and Brent Mittelstadt. "Three pathways for standardisation and ethical disclosure by default under the European Union Artificial Intelligence Act." *Computer Law & Security Review* 53 (2024): 105957.

[17] Ruschemeier, Hannah. "AI as a challenge for legal regulation–the scope of application of the artificial intelligence act proposal." *Era Forum.* Vol. 23. No. 3. Berlin/Heidelberg: Springer Berlin Heidelberg, 2023.

(III) Governance Imperatives and Legislative Drivers: The Necessity of Regulating Artificial Intelligence

1. The Practical Need to Prevent the Alienation of Technology

Preventing the alienation of technology has become an urgent issue in AI governance. Technological systems may deviate from human-defined objectives and exhibit unpredictable behaviour, resulting in a loss of effective control over autonomous systems. For example, in financial or military contexts, iterative algorithmic processes may exceed design boundaries and depart from human value benchmarks. This phenomenon is rooted in the high degree of complexity and the "black box" nature of deep learning models, posing risks to system security and eroding human cognitive agency.

As artificial intelligence becomes deeply embedded in knowledge production and social decision-making, excessive reliance on algorithmic outputs may lead to the homogenisation of individual thinking and the erosion of critical reasoning. In fields such as healthcare, overdependence on AI-assisted systems may also result in the atrophy of professional expertise. These developments reflect the risk of technological tools, in turn, constraining and reshaping human cognition, amounting in essence to an alienation of the human–machine relationship.

In addressing such deep-seated and cumulative risks, traditional ex post liability mechanisms are of limited effectiveness. A preventive legal framework must therefore be established, intervening at the early stages of technological development. This may be achieved through institutional designs such as setting baseline requirements for algorithmic transparency, mandating the incorporation of human oversight loops, and establishing fail-safe or "kill switch" mechanisms for critical systems. Through such measures, ethical and safety requirements can be internalised into technical standards, thereby constructing institutional safeguards to ensure that the development of AI consistently serves human well-being.

2. Safeguarding the Fundamental Rights of the Public as a Baseline

The rapid development of artificial intelligence poses new challenges to fundamental rights, including the right to privacy and the right to equality. Data aggregation and analytics have begun to deconstruct traditional conceptions of privacy, while biometric technologies further intensify the risk of personal data exposure. Issues of algorithmic fairness have become increasingly prominent: automated decision-making systems in areas such as recruitment and credit evaluation may reproduce historical biases, leading to systemic discrimination against certain groups, while their opacity makes legal redress particularly difficult.
Different jurisdictions have adopted divergent regulatory approaches. The European Union has established stringent ex ante compliance obligations under the GDPR; the United States relies more heavily on ex post judicial enforcement, which has proven less effective in addressing large-scale infringements; China's Interim Measures for the Administration of Generative Artificial Intelligence Services adopts a graded and categorised regulatory framework. Each of these models faces a structural tension in that technological iteration often outpaces legislative updating.

An effective system of protection requires a departure from traditional legislative approaches. Ethical boundaries should be embedded at the early stages of

technological development — for example, prohibiting deepfake technologies that infringe upon portrait rights, or preserving human decision-making authority in medical contexts. At the same time, dynamic algorithmic evaluation mechanisms should be established, with continuous monitoring of rights-related risks through third-party auditing. Such a framework preserves space for innovation while delineating clear "red lines" in critical domains such as criminal justice and social security through the use of negative lists.

3. Institutional Safeguards for the Healthy and Orderly Development of the Industry

The rapid expansion of the artificial intelligence industry has been accompanied by market failures that threaten its long-term, sustainable development. A prominent concern is the risk of monopolisation: AI development requires vast amounts of data, highly specialised talent, and substantial computational resources. These high entry barriers, combined with network effects, tend to concentrate resources in leading firms, producing "winner-takes-all" outcomes that undermine competition, crowd out small and medium-sized enterprises, and inhibit innovation.

Another critical issue lies in increasingly complex forms of unfair competition, including data monopolies, algorithmic collusion, market foreclosure, and discriminatory pricing. More covert practices — such as the use of deepfakes for false advertising and "price discrimination based on big data" — distort market signals and harm consumer interests.

Addressing these challenges requires close coordination between industrial policy and legal regulation. Industrial policy may guide development and foster ecosystems — for example, through the establishment of funds to support foundational research and the creation of open data platforms. However, incentive-based policies alone are insufficient to resolve issues of monopoly and unfair competition, and may even exacerbate market concentration. Robust legal regulation is therefore necessary to establish clear market rules and ensure a fair competitive environment.

At the core of such regulation lies the development of a competition law framework adapted to the specific characteristics of artificial intelligence. This includes clarifying standards for assessing emerging forms of monopolistic conduct such as data dominance and algorithmic coordination; refining prohibitions and sanctions against unfair practices such as algorithmic discrimination and personalised price exploitation; and improving legal frameworks governing data ownership, circulation, and equitable access.

Legal regulation must also be forward-looking, capable of adapting to technological iteration and industrial evolution, and of maintaining a dynamic balance between encouraging innovation and preventing market distortions while safeguarding fairness. This requires a deep understanding of technological logic and its impact on market structures, the formulation of rules that are both principled and operational, and the support of effective regulatory enforcement, so as to guide the industry toward healthy and orderly development.

II. Legislative Practices in AI Governance in Major Countries and Regions

(I) United States: A Flexible Legislative Model Driven by Innovation

1. Top-Level Design: National Strategy and a Soft Law Regulatory Framework

The United States adopts a soft law approach to AI governance, relying on non-binding policy frameworks to guide industry development while emphasising regulatory flexibility. The 2019 American AI Initiative set out the goals of technological leadership and the cultivation of an innovation ecosystem, supported by industry self-regulation, technical standards (such as the non-binding risk management framework developed by the National Institute of Standards and Technology (NIST)), and inter-agency coordination mechanisms. Policy development has followed a gradual and incremental trajectory, with executive orders, white papers, and congressional hearings serving as primary channels. The 2020 Memorandum on Guidance for Regulation of Artificial Intelligence Applications requires federal agencies to conduct AI impact assessments, leaving room for flexibility across different application scenarios and reflecting a dynamic balance between technological innovation and risk control.

This framework provides diverse avenues for market actors, enabling major technology firms to participate in the co-construction of ethical norms while allowing start-ups to test innovations through regulatory sandboxes. Its underlying logic lies in activating market-driven adaptability through the coupling of industry self-discipline and responsive policymaking. At present, the federal government is exploring the introduction of more determinate guidance in high-risk areas, with a view to refining the institutional boundaries of soft governance and ensuring that regulatory effectiveness evolves in tandem with technological development.

2. Sectoral Deepening: Targeted Legislation in Key Application Scenarios

In the field of autonomous driving, the United States has adopted a tiered regulatory approach, permitting pilot programmes for low-risk applications while requiring Level 4 and above vehicles to be equipped with data recording devices. In the medical AI sector, a risk-based classification continues to apply: diagnostic algorithms must undergo clinical validation, whereas tools used to assist research and development are primarily subject to scrutiny for algorithmic bias. Data governance similarly follows a scenario-based approach: autonomous driving regulation emphasises the anonymisation of road-condition data, while the medical sector prioritises the protection of patients' biometric information.

This form of vertical legislation constructs a three-dimensional regulatory matrix of "industry–risk–data". For example, in the context of autonomous driving accidents, a "trace-back period for technical defects" has been introduced. In contrast to the European Union's horizontal and generalised framework, the United States' sector-specific rules place greater emphasis on contextual precision, providing concrete compliance guidance (such as the requirement of clinical explainability for medical algorithms). Looking ahead, the coordination and integration of rules across converging technological domains will constitute an important direction of development.

3. Evaluation of Effectiveness: Industrial Impacts of Legislative Implementation

The U.S. model of AI legislation exhibits pronounced stage-based effects. In its early phase, the flexible framework effectively stimulated industrial innovation: a relatively permissive regulatory environment contributed to nearly a threefold increase in venture capital investment between 2020 and 2023, and to the emergence of numerous breakthrough enterprises in fields such as natural language processing and computer vision. This light-touch governance approach reduced compliance costs for enterprises and accelerated technological commercialisation — for instance, algorithm-related patents in the road-testing phase of autonomous driving grew at an annual rate exceeding 25%.

As technology evolves, however, the institutional design faces new challenges of adaptability. Regulatory frameworks for foundational models remain underdeveloped, and the boundaries of lawful data use in training large language models require urgent clarification, as evidenced by multiple disputes over data rights between 2022 and 2023. In the context of systemic risk prevention, the opacity of algorithms continues to generate widespread concern in sectors such as finance and healthcare, with complaints relating to algorithmic fairness increasing sevenfold over a five-year period. At the same time, the current regulatory environment has, to some extent, contributed to increased market concentration in foundational model development, with start-ups accounting for less than 15% of market share.[18]

The existing framework also encounters coordination difficulties in balancing technological innovation and risk control. Fragmented legislative initiatives—such as the Algorithmic Accountability Act — have resulted in divergent standards across states: California tends to require annual algorithmic audits, whereas Texas allows firms to withhold key parameters on the grounds of trade secrets. This decentralised governance model has objectively increased the compliance costs of cross-state operations and diluted the overall effectiveness of risk control. In the face of the rapid rise of generative AI, the current regulatory system is being tested in areas such as deepfake governance and the attribution of liability for infringement, with over 60% of related cases suspended due to the absence of clearly established adjudicatory standards.

(II) European Union: A Comprehensive Legislative Model Grounded in Fundamental Rights

1. Institutional Linkages: The GDPR as a Precursor to AI Regulation

The core principles of the General Data Protection Regulation (GDPR) have exerted a profound influence on AI regulation. The principles of purpose limitation and data minimisation require that data processing in AI systems be confined to specific, clearly defined purposes and limited in scope. In the development of medical diagnostic AI, for example, training data must be directly relevant to clinical objectives; the collection of genetic data beyond such scope constitutes a compliance risk. Article 22 of the GDPR grants data subjects the right to refuse decisions based solely on automated processing, thereby prompting enterprises to incorporate mechanisms for human review. A notable example is a Dutch municipal authority that was held legally liable for failing to explain the decision-making logic underlying welfare distribution, and was found to have inadequately safeguarded the data subject’ s right to be informed. In substance, this provision establishes a baseline requirement for algorithmic explainability.

[18] Birkstedt, Teemu, et al. "AI governance: themes, knowledge gaps and future agendas." Internet Research 33.7 (2023): 133-167.

The large-scale training paradigm of generative AI poses new challenges to the application of the traditional principle of purpose limitation. While the use of publicly available data for model training may have a certain basis in compliance, the outputs of such models may still reconstruct personal information. In such circumstances, because the original data have been transformed into parameter weights that are difficult to trace directly, data subjects often face significant obstacles in exercising their right to erasure. This phenomenon highlights the need for further exploration in regulating the derivative value of data and reflects the necessity of expanding existing data protection frameworks to respond to emerging technological paradigms.

2. Core Analysis: The Risk-Based Classification Mechanism under the Artificial Intelligence Act

The EU Artificial Intelligence Act establishes a four-tier risk classification system, structured as a pyramid with a narrow apex and a broad base. At the top, "unacceptable risk" systems—such as those designed to manipulate human subconscious behaviour or government-run social credit scoring systems — are prohibited. Beneath this, "high-risk" applications, including those in medical devices and critical infrastructure, are subject to stringent regulatory requirements. These include the submission of technical documentation covering algorithmic logic, training data, and error evaluation, as well as ensuring that human intervention remains possible at all times. At this level, training data must also be managed to mitigate bias, and relevant technical documentation is subject to regulatory review.[19]

At the levels of "limited risk" and "minimal risk" (such as chatbots), regulatory obligations are primarily confined to transparency requirements, including disclosure obligations to users.[20] This classification mechanism determines regulatory intensity based on the consequences of application rather than the type of technology. In the context of emerging technologies such as generative AI, however, the precise delineation of these categories remains subject to further refinement through practice. By balancing compliance burdens for small and medium-sized enterprises, the Act effectively translates abstract risks into quantifiable regulatory indicators, thereby offering a valuable model for global AI governance.

3. Spillover Effects: Global Legislative Influence under the "Brussels Effect"

Through the so-called "Brussels Effect", the EU Artificial Intelligence Act exerts a significant influence on global AI governance. In order to reduce cross-border compliance costs or enhance regulatory compatibility, certain non-EU jurisdictions have chosen to align their regulatory approaches with EU standards. For instance, Brazil's draft regulatory framework draws extensively on the EU's risk classification model in defining "high-risk systems" and associated mandatory obligations. Similarly, Canada's Directive on Automated Decision-Making incorporates requirements relating to algorithmic transparency and human oversight, thereby strengthening accountability in the public sector. This diffusion of regulatory norms is largely attributable to the scale of the EU's single market and the precedential value of its relatively mature regulatory system.

[19] Van Kolfschooten, Hannah, and Janneke Van Oirschot. "The EU artificial intelligence act (2024): implications for healthcare." *Health Policy* 149 (2024): 105152.
[20] Neuwirth, Rostam J. "Prohibited artificial intelligence practices in the proposed EU artificial intelligence act (AIA)." *Computer Law & Security Review* 48 (2023): 105798.

Within the broader process of globalisation, EU rules interact dynamically with other governance approaches. The United States tends to favour industry self-regulation and sector-specific governance, relying on case-by-case enforcement by the Federal Trade Commission and targeted rules in specific domains, reflecting a distinct regulatory logic from the EU's comprehensive legislative model. China, by contrast, emphasises a coordinated approach balancing development and security—for example, by adopting a filing-based regulatory regime for generative AI that is both inclusive and prudent, thereby preserving space for domestic technological innovation. In areas such as data governance, these approaches reflect differing institutional preferences relative to EU standards.

At the same time, the EU's relatively detailed compliance requirements — such as technical documentation and continuous monitoring obligations for high-risk systems—impose higher demands on the allocation of corporate compliance resources. This constitutes an area requiring ongoing refinement in the context of cross-border application. Looking forward, the EU model will continue to interact with diverse national institutional environments and evolving technological landscapes, with the breadth and depth of its demonstrative effect to be further shaped through practice.

(III)China:Development-Oriented Legislative Exploration through Categorisation and Tiered Regulation

1. Policy Guidance: The Transition from Industrial Promotion to Normative Governance

The evolution of China's artificial intelligence policy exhibits a clear stage-based trajectory. The New Generation Artificial Intelligence Development Plan issued in 2017 established AI as a national strategy, focusing on technological breakthroughs and industrial expansion, and empowering the sector through resource allocation and platform-building. At this early stage, governance was primarily oriented toward unleashing innovation. With the deepening application of frontier technologies such as generative AI, issues relating to data compliance and algorithmic fairness have come to the forefront, and policy priorities have progressively shifted toward normative governance. The successive introduction of the Provisions on the Administration of Algorithmic Recommendation in Internet Information Services and the Interim Measures for the Administration of Generative Artificial Intelligence Services clearly reflects this transition toward a model that balances high-level security with high-quality development.

The governance principle of "inclusive and prudent regulation" has been implemented across multiple dimensions. On the one hand, inclusiveness is reflected in the reservation of space for emerging industries—for example, the introduction of a filing mechanism for generative AI to streamline compliance processes. At the same time, categorised and tiered regulation ensures differentiated oversight, thereby safeguarding innovation incentives. On the other hand, prudence is manifested in the precise control of key risks, through detailed requirements concerning security assessments, content moderation, data source verification, and the labelling of deep synthesis content, thereby reinforcing the baseline of technological ethics and market order.[21]

This policy paradigm represents a shift toward a model that accords equal weight to development and regulation. Its internal logic lies in adaptive governance

[21] Li, Yuanjun. "Foreign Practices in AI Legal Governance and China's Approach". *Journal of South-Central Minzu University (Humanities and Social Sciences),* 45.12 (2025): 190–204+212.

based on technological maturity and risk awareness: in the early stages, industrial momentum is fully released while governance experience is accumulated; as technological applications increasingly implicate public interests, legislative intervention is introduced in a timely manner to clarify behavioural boundaries. Compared with the EU's comprehensive ex ante regulatory model and the United States' reliance on industry self-regulation, China's approach places greater emphasis on dynamic risk assessment and incremental legal constraint, aiming to construct a modern governance framework that balances safety with innovation.

2. Targeted Regulation: Legislative Initiatives in Algorithms, Large Models, and Vertical Sectors

China's AI legislation is highly focused on specific technological scenarios, with algorithmic recommendation, deep synthesis, and large models emerging as key regulatory areas. The Provisions on the Administration of Algorithmic Recommendation in Internet Information Services establish a framework for the protection of user rights, requiring platforms to implement transparency mechanisms and provide users with convenient opt-out options, while imposing differentiated compliance obligations based on the nature of the algorithms. The Provisions on the Administration of Deep Synthesis in Internet Information Services clarify requirements for the labelling of generated content, establish full-process management, delineate the respective responsibilities of service providers and technical support providers, and introduce a regularised security assessment mechanism.

With respect to the governance of large models, the current regulatory approach centres on a filing system, requiring developers to submit key technical information in accordance with the law. Regulatory authorities, in turn, rely on ongoing supervision and security review, thereby forming a dynamic governance loop characterised by "filing first, supervision thereafter". In vertical sectors such as autonomous driving, regulatory development has largely proceeded through pilot legislation at the local level, which interacts constructively with national-level top-level design, jointly shaping a gradualist governance model with distinct Chinese characteristics.

3. Paradigm Summary: A Localised Model Balancing Development and Security

China has progressively developed a "dynamic balance" paradigm in AI governance, effectively coordinating technological innovation with risk prevention. This paradigm is primarily supported by the following institutional tools:

(1) Security assessment: For higher-risk AI systems, ex ante evaluations of technical reliability and social impact are conducted prior to deployment, with parameter thresholds guiding developers in mitigating potential risks;

(2) Algorithm filing: Core algorithmic mechanisms, including underlying logic, must be submitted to regulators. Compared with comprehensive disclosure regimes, this approach focuses on transparency at critical points, thereby addressing societal concerns arising from algorithmic opacity;

(3) Content labelling: Generated content must be clearly identified as such, in order to safeguard the public's right to be informed.

At its foundation lies a categorised and tiered governance system, under which differentiated regulation is applied according to the risk level of

specific application scenarios. For instance, medical diagnostic AI is subject to more stringent evaluation mechanisms, whereas industrial quality inspection systems are primarily governed through ex post monitoring. This targeted approach effectively preserves the innovative vitality of market actors.

Looking ahead, the mutual recognition and coordination of domestic technical standards, as well as deeper alignment with international rules in areas such as cross-border data flows and algorithm certification, will constitute key directions for institutional evolution. Further refinement of top-level standard-setting and sustained participation in global governance dialogues will enhance the international compatibility and cross-border coordination capacity of this governance framework.

(IV) Others: Domestic Legislative Approaches from Diverse Perspectives

1. United Kingdom: An Agile, Principles-Based Regulatory Model

The United Kingdom has adopted an agile, principles-based model for AI regulation, centred on the five core principles articulated in its AI Regulation White Paper: safety; robustness and reliability; transparency and explainability; fairness; and accountability and governance. This guiding framework provides flexible compliance references across sectors. In contrast to the EU's detailed statutory regime, this model is better suited to responding to rapid technological iteration and maintaining institutional adaptability.

In terms of implementation, the UK government relies primarily on non-binding policy instruments to advance governance, including the development of industry standards, the publication of best practice guidelines, and the promotion of self-regulatory norms.[22] At the same time, initiatives such as the "Digital Regulation Cooperation Forum" have facilitated coordination among regulators, industry, and academia, enabling the joint exploration of governance approaches aligned with technological development.

In substance, this model offers a lightweight regulatory pathway: it guides market actors through high-level principles, supplemented by dynamic guidance to address emerging challenges. It contributes to more efficient allocation of legislative resources and demonstrates distinctive practical value in balancing innovation with risk control.

2. Canada: Legislation Emphasising Algorithmic Accountability in the Public Sector

Canada's Directive on Automated Decision-Making establishes a systematic regulatory framework for algorithmic governance within the public sector. The Directive requires federal institutions to conduct a mandatory Algorithmic Impact Assessment (AIA) prior to deploying AI systems, comprehensively evaluating technical characteristics and potential risks of data bias, and to disclose the assessment results in accordance with the law.

In terms of remedies, it guarantees affected individuals the right to obtain a meaningful explanation of decision logic, and establishes an independent AI oversight

[22] Roberts, Huw, et al. "Artificial intelligence regulation in the United Kingdom: a path to good governance and global leadership?." *Internet Policy Review* 12.2 (2023): 1-31.

mechanism to handle complaints, with the authority to suspend algorithmic operations where necessary.[23] Liability is structured through a model of "dynamic accountability": during the development phase, comprehensive algorithmic logs must be retained; during the operational phase, decision-making deviations are continuously monitored, and human intervention is triggered once predefined thresholds are reached. This lifecycle-based governance approach ensures that public authorities retain ultimate control over decision-making, providing a practical model for balancing the efficiency of digital government with the protection of individual rights.

3. Japan: A Governance Approach Led by Non-Binding Guidelines

Japan's approach to AI governance is characterised by a reliance on non-binding guidelines, standing in contrast to the EU's legislative model. Its core policy instruments are grounded in industry self-regulation and social consensus, with the objective of balancing technological innovation and normative development. The policy-making process emphasises collaboration among government, academia, and industry (the so-called "government–industry–academia" model), thereby enhancing both the technical sophistication and the acceptance of regulatory frameworks.

Its "agile revision" mechanism effectively shortens the cycle for updating rules. For example, in response to risks associated with generative AI and deep synthesis technologies, provisions relating to content labelling have been introduced and updated in a timely manner.[24]

In terms of implementation, this model primarily relies on soft law instruments for guidance. Its effectiveness in high-risk domains—such as medical AI—remains to be further observed. For economies in the early stages of industrial development, however, this approach offers a flexible policy option, facilitating consensus-building in the initial phase and providing practical experience for the subsequent development of more formalised legal frameworks.

[23] Scassa, Teresa. "Administrative law and the governance of automated decision making: A critical look at Canada's directive on automated decision making." *UBCL Rev.* 54 (2021): 251.
[24] Kozuka, Souichirou. "A governance framework for the development and use of artificial intelligence: lessons from the comparison of Japanese and European initiatives." *Uniform Law Review* 24.2 (2019): 315-329.

III. The Content Dimensions and Comparative Models of Global AI Governance Legislation

(I) Legislative Core: The Institutional Characterisation of Key Elements

1. Resource Layer: Data Flows, Rights Attribution, and Security Safeguards

The core dimensions of data governance in the context of artificial intelligence encompass cross-border data flows, the allocation of data-related rights, and data security safeguards. Cross-border data flows reveal an inherent tension between sovereign regulatory control and global coordination. Certain economies, on grounds of national security and privacy protection, have adopted data localisation strategies, which, in practice, impose higher requirements on cross-border scientific research collaboration and commercial integration. In response, various jurisdictions are exploring points of regulatory convergence through bilateral agreements and regional trust frameworks.

The allocation of data-related rights presents a pressing issue requiring further clarification, particularly in high-value application scenarios such as healthcare and finance, where competing claims over data interests remain insufficiently defined. A layered rights-allocation model has been proposed as a possible approach: the original data provider retains ownership rights, the data processor is granted rights of use, and data aggregation platforms enjoy rights to economic returns derived from data utilisation. However, further refinement is still needed in delineating the precise boundaries of these rights and in developing effective mechanisms for infringement remedies.

Data security safeguards depend on the coordinated integration of technological measures, institutional arrangements, and clearly defined responsibilities. At the technical level, commonly adopted measures include end-to-end encryption, distributed storage, and dynamic access control systems. From an organisational perspective, governance frameworks emphasise hierarchical permission management and the maintenance of auditable operational logs. In addition, emergency response mechanisms specify procedures for regulatory reporting and for notifying affected individuals in the event of data breaches.

2. Technical Layer: Requirements of Algorithmic Transparency and Explainability

At the technical level, the regulation of algorithms primarily focuses on transparency and explainability. Different technical paths to transparency serve different functions. Moderate disclosure of source code may help enhance public understanding and the legitimacy of public decision-making, although in practice this must be reconciled with the protection of trade secrets. Disclosure of key model parameters (such as those required by the European Union for high-risk systems) can reveal the weighting of variables, although the cognitive barriers posed by complex models still require professional interpretation. In addition, standardised technical documentation recording design logic and data characteristics has become a routine means of governance.[25]

The applicable standards of explainability display marked contextual variation. In high-risk scenarios such as medical diagnosis and finance, there is a preference for deep explainability, requiring clear disclosure of the basis of a decision and the source of key factors; in lower-risk scenarios, such as ordinary

[25] Jia, Kai, Zhao, Jing, and Zhou, Kedi. "Global Governance of Algorithms: Theoretical Definition, Issue Framework, and Reform Paths". *Chinese Public Administration,* (06) (2022): 59–65.

recommendation systems, the focus is more on communicating the basic operating principles. As regards quantitative indicators for the appropriate depth of explanation and the harmonisation of judicial standards, these issues remain under active exploration in practice.

Algorithmic auditing is essential to ensuring the implementation of the foregoing rules. This requires auditing bodies to possess interdisciplinary expertise spanning computer science, legal ethics, and sector-specific knowledge, and the auditing process must extend across the full lifecycle, from the detection of data bias to post-deployment monitoring (with additional clinical validation required in particular sectors such as healthcare). At the same time, the provision of traceable cases, together with complaint procedures and mechanisms for human review, also constitutes a core component. As emergent characteristics of generative algorithms become increasingly prominent, ensuring that auditing standards evolve in step with cutting-edge technologies will be a major focus of future institutional development.

3. Consequence Layer: Principles and Allocation Mechanisms for the Attribution of Liability for Harm

The legal determination of liability for harm caused by artificial intelligence principally centres on three core issues: the applicable principles of attribution, the identification of responsible parties, and the mechanisms of compensation for damage.

The choice of attribution principle is the primary issue for consideration. Traditional fault-based liability requires proof of negligence or intent, but this presents evidentiary difficulties in scenarios involving autonomous algorithmic decision-making, such as sensor misjudgments in autonomous driving. Strict liability, by contrast, is grounded in the dangerousness of the activity itself (for example, diagnostic errors made by high-risk medical AI) and places greater emphasis on result-oriented responsibility. In judicial practice, the boundary between the two is generally assessed on a case-by-case basis, having regard to the risk level of the application scenario and the degree of technological controllability. In this respect, the risk classification framework under the EU Artificial Intelligence Act provides a useful point of reference.

The participation of multiple actors across the industrial chain renders the attribution of responsibility more complex. The boundaries among developers, operators, and users frequently overlap: developers may bear liability for defects in algorithmic logic or bias in training data (for example, where generative AI is alleged to infringe rights); operators primarily respond to operational risks arising after deployment (such as discriminatory pricing by recommendation algorithms); users are more commonly held liable for improper operation or technological misuse (such as the unlawful use of deep synthesis technologies). In situations involving multiple concurrent actors, detailed rules governing the apportionment of responsibility remain to be further refined, and current adjudication in fields such as autonomous driving still relies to a considerable extent on case-by-case judicial discretion.

The construction of compensation mechanisms likewise faces technical challenges. First, the assessment of loss is highly complex, as AI-related harm may be both concealed and long-term in nature (for example, latent discrimination arising from algorithmic bias), requiring comprehensive consideration of direct economic loss, personal rights violations, and consequential damage. Secondly, the setting of compensation caps must properly balance the preservation of space for industrial development with the protection of legal interests. For high-risk AI systems, reference may appropriately be made to mechanisms such as compulsory liability insurance, as used in the medical device sector. Thirdly, in light of the

difficulties of establishing causation caused by multi-source data integration or model opacity, practice is actively exploring rules that ease the burden of proof and the introduction of specialised technical appraisal procedures.

(II) Paradigm Selection: Differentiated Approaches to AI Governance

1. Centralised and Comprehensive Approach: A Unified Legislative Paradigm Covering All Sectors

The comprehensive legislative paradigm seeks to establish an overarching regulatory framework covering the entire technological lifecycle and all application scenarios. The European Union's Artificial Intelligence Act, through its four-tier risk classification system (unacceptable risk, high risk, limited risk, and minimal risk), sets mandatory compliance baselines—such as data governance requirements, documentation obligations, and human oversight — for high-risk domains including medical diagnostics. Built upon standardised assessment procedures and centralised regulatory structures, this "pyramid-shaped" framework is designed to enhance the effectiveness and enforceability of regulatory rules.

China's approach, anchored in the New Generation Artificial Intelligence Development Plan, demonstrates a pattern of coordinated evolution between policy and legislation. Its core mechanisms encompass full lifecycle governance tools, including algorithm filing and security assessment, supplemented by catalogue-based classification systems for dynamic risk identification — for example, the implementation of a tiered filing regime for generative artificial intelligence. This model seeks to address regulatory overlap and normative gaps in a systematic manner: the EU approach objectively improves compliance predictability for multinational enterprises, while China's algorithm filing system enhances regulatory penetration and oversight capability.

At the level of implementation, an inherent tension persists between the stability of comprehensive legislation and the need for adaptability in the face of rapid technological iteration. The EU's current risk classification standards still require further refinement to effectively address emerging developments such as general-purpose artificial intelligence (GPAI). While China's institutional design has incorporated a degree of flexibility, the compliance boundaries governing training data for foundational models remain to be clarified through practice. In light of the differentiated regulatory demands of vertical sectors—such as clinical validation in medical AI and real-time risk control in finance—the central challenge for this paradigm lies in maintaining regulatory stability while introducing an appropriate degree of institutional flexibility.

Table 1 Comparative Overview of Comprehensive Legislative Paradigms for Artificial Intelligence

Actors	Core Features	Risk Control Approach	Principal Advantages	Existing Challenges
European Union Artificial Intelligence Act	A unified legal framework; a four-tier risk classification system	Ex ante rule-setting; mandatory obligations for high-risk sectors	Enhances enforcement efficiency; reduces cross-border compliance costs	Lag in risk classification; ambiguity in rules governing GPAI
China's Comprehensive Governance Framework	Policy–legislation coordination; a full lifecycle regulatory network	Catalogue-based classification; tiered filing; dynamic adjustment	Addresses regulatory gaps and overlaps; strengthens regulatory penetration	Emerging issues require further refinement; limited capacity to accommodate sector-specific needs
Common Features of the Comprehensive Legislative Paradigm	System-oriented approach; unified rule structure	Comprehensive coverage; a combination of ex ante and adaptive mechanisms	Stabilises market expectations; avoids the drawbacks of fragmented legislation	Difficulty in balancing rigidity and flexibility; limited adaptability to rapid technological iteration

2. Decentralised and Sector-Specific Approach: A Fragmented Legislative Paradigm

The decentralised legislative paradigm emphasises differentiated regulation based on the risk characteristics of specific industries. In the United States, for example, machine learning – based diagnostic software in the healthcare sector is incorporated into the traditional regulatory framework for medical devices, with a focus on dynamic performance monitoring and real-time quality traceability, thereby balancing technological innovation with the protection of patient rights. In the field of financial technology, core algorithms are required to incorporate counterfactual fairness testing and mechanisms for human override, with the aim of enhancing decision-making transparency and preventing potential algorithmic bias and systemic financial risks.

In Japan, amendments to the Road Transport Vehicle Act establish a tiered liability regime for autonomous driving based on the concept of the operational design domain (ODD). At Level 3, for instance, the regulatory framework requires real-time monitoring of the driver's condition and mandates the activation of emergency braking where takeover requests are not responded to in a timely manner. Complementary requirements concerning mandatory accident data recording devices provide a statutory evidentiary basis for the objective determination of liability.

The core strength of this model lies in its high degree of sector-specific adaptability. Rooted in the technical logic of particular industries—such as biometric data processing in healthcare, real-time risk control in finance, and environmental perception in autonomous driving—it is capable of precisely defining compliance baselines tailored to technological characteristics. In practice, however, this model faces a range of regulatory coordination challenges. Multi-layered compliance requirements across sectors may increase overall compliance costs for enterprises (for example, where data from health wearables is used in insurance underwriting, triggering overlapping jurisdictions); shifting boundaries in product classification may incentivise regulatory arbitrage; and the enhancement and integration of cross-agency coordination mechanisms remains critical to improving overall governance effectiveness.

3. Dynamic Equilibrium: Inclusive Governance through the Coordination of Hard Law and Soft Law

Hybrid governance in artificial intelligence seeks to balance the interaction between binding legal rules and flexible normative instruments. The United Kingdom's approach is structured around a "principles + guidance" framework, anchored in three core dimensions: safety assurance, transparency in decision-making, and respect for fundamental rights. Regulatory authorities collaborate with the technology sector and industry stakeholders to develop sector-specific guidance, refining algorithmic standards and compliance pathways while preserving institutional flexibility.

Singapore adopts a progressive "test–learn–adjust" cycle: innovators are permitted to conduct application testing within predefined regulatory boundaries, while regulators simultaneously collect data to assess risks and iteratively refine regulatory requirements. This feedback loop effectively mitigates the inherent tension between the stability of traditional legal rules and the rapid pace of technological evolution.

Soft law instruments—such as industry standards and ethical guidelines—benefit from procedural flexibility and can respond swiftly to emerging technological scenarios, providing early-stage behavioural guidance and risk mitigation. Hard law, by contrast, focuses on

establishing baseline safeguards, clarifying the allocation of rights and responsibilities, and delivering legal certainty and enforceability. While soft law facilitates exploratory innovation, hard law consolidates the institutional foundations of governance. The interaction and complementarity between the two provide a practical framework for inclusive governance (see Table 3).[26]

Table 2 Comparative Overview of Hybrid AI Governance Practices (United Kingdom vs. Singapore)

Governance Structure	Core Mechanisms	Key Features	Challenges Addressed
Principles + Guidance	Core principles combined with sector-specific guidance; multi-stakeholder co-regulation	Preserves space for innovation; avoids rigid regulatory constraints	Rapid technological iteration
Test–Learn–Adjust Cycle	Pilot testing within predefined boundaries; iterative refinement through dynamic feedback	Continuous learning; dynamic responsiveness	Technological uncertainty

Table 3 Functional Comparison between Soft Law and Hard Law in Hybrid AI Governance

Form	Advantages	Limitations	Core Value
Industry standards, ethical guidelines, and similar instruments	Flexible; rapidly adaptable; capable of filling regulatory gaps	Limited enforceability (voluntary compliance)	Early-stage guidance; risk mitigation
Statutory provisions	Legally binding; clear allocation of responsibility; enforceable sanctions	Difficulty in adapting rapidly to emerging scenarios	Legal certainty; foundation of regulatory order

(III) Global Coordination: Communication and Convergence in Cross-Border Governance

1. The Pivotal Role of International Organisations in Harmonising Legal Rules

International organisations play a substantive coordinating role in the global governance of artificial intelligence. The OECD Principles on Artificial Intelligence, grounded in a human-centred approach, have evolved into an important reference framework for multilateral policymaking. At the level of the United Nations, efforts focus on assessing the broader implications of technological development for the Sustainable Development Goals, with an emphasis on aligning the governance priorities of different economies through multilateral dialogue mechanisms. The Group of Twenty (G20), for its part, concentrates on the structural impact of technological transformation on the global macroeconomy and labour markets, and continues to promote policy coordination and interaction among its member states.

Table 4 Comparative Overview of Three Major International Organisations in Global AI Governance

Organisation	Nature of Rules	Core Outputs	Scope of Influence
OECD	Non-binding ethical principles	"AI Principles"	38 member countries; global reference framework
United Nations AI Expert Group	Multilateral dialogue platform	Sustainability assessment reports	193 member states
G20	High-level economic coordination mechanism	Ministerial declarations and policy recommendations	G20 economies

[26] Qian, Yuzhou, Keng L. Siau, and Fiona F. Nah. "Societal impacts of artificial intelligence: Ethical, legal, and governance issues." *Societal impacts* 3 (2024): 100040.

2. Case Studies of Legislative Cooperation under Multilateral Frameworks

Within the framework of global AI governance, the cooperation mechanism of the EU – U.S. Trade and Technology Council (TTC) provides a representative example of multilateral coordination. The TTC operates through dedicated working groups, focusing on advancing mutual recognition of technical standards and coordination in risk management. Through regular dialogue, both parties engage in technical alignment on key issues such as frameworks for algorithmic transparency and classification standards for high-risk systems. In the field of biometric technologies, for instance, joint discussions have been conducted to establish baseline privacy protections, leading to the development of non-binding technical guidelines. This model of cooperation offers compliance reference points for multinational enterprises and, in practice, reduces institutional costs associated with market entry. At the same time, differences in regulatory approaches remain evident: the EU emphasises ex ante compliance review for foundational models, whereas the United States relies more heavily on industry self-regulation, reflecting distinct governance logics.

The Asia-Pacific Economic Cooperation (APEC) framework exhibits a distinctly regional approach to digital governance. Its Cross-Border Privacy Rules (CBPR) system seeks to establish trust mechanisms for data flows by coordinating data protection standards across member economies through certification processes. In the field of artificial intelligence, APEC places particular emphasis on empowering small and medium-sized enterprises, providing scenario-based compliance guidance through implementation guidelines. Compared with formal legislative approaches, APEC primarily operates through soft law mechanisms that are voluntarily adopted by member economies. For example, Singapore has developed industry self-regulatory codes on this basis, while Japan has incorporated such standards into public procurement considerations. This flexible governance model aligns with the realities of rapid technological iteration, though its effectiveness ultimately depends on the degree of coordinated implementation among member economies. Looking ahead, further development is needed in areas such as the alignment of rules relating to remedies for algorithmic fairness.

The BRICS countries are gradually building consensus in the development of ethical frameworks for artificial intelligence. Relevant frameworks adopted in 2023 articulate the principle of "inclusive growth", advocating that technological development should reflect the specific conditions of emerging markets. Participating countries exhibit converging policy preferences in the field of data governance, emphasising a balance between data security and cross-border data flows. Brazil, for instance, has advanced revisions to its personal data protection legislation with the incorporation of AI-related provisions, while South Africa has introduced social impact assessment procedures for the use of algorithms in the public sector. This model of cooperation respects national policy space and allows for differentiated compliance baselines based on varying stages of development. At the same time, in certain domains—such as security-related applications—countries retain substantial regulatory autonomy in accordance with domestic priorities.[27] Through institutional arrangements such as joint research centres, these ethical principles are increasingly being translated into substantive technical standards.

[27] Wang, Yingming, and Shi, Chao. "Ethical Considerations of Offensive Artificial Intelligence and Global Governance". *Science & Technology Review,* 43.4 (2025): 37–45.

19

Table 5 Comparative Overview of Multilateral Cooperation Practices in Global AI Governance

Entity	Focus of Cooperation	Governance Model	Key Outcomes	Institutional Features
EU–U.S. Trade and Technology Council (TTC)	Mutual recognition of technical standards; coordination in risk management; frameworks for algorithmic transparency	Divergent approaches (EU: ex ante regulatory review; U.S.: industry-led self-regulation)	Non-binding guidelines on biometric technologies	Regular dialogue through dedicated working groups
Asia-Pacific Economic Cooperation (APEC)	Cross-border privacy rules; capacity-building for small and medium-sized enterprises; implementation of AI ethics	Soft-law framework (voluntary adoption by member economies)	CBPR system; "Guidelines for the Implementation of AI Ethics Principles"	Reliance on coordination among member regulatory authorities
BRICS Countries	AI ethical guidelines; balancing data sovereignty; phased compliance mechanisms	Preservation of policy space (phased compliance approach)	2023 "AI Ethics Framework"; joint research centres	Institutional bodies facilitating the translation of principles into standards

3.Technological Barriers and the Resolution of Legal Conflicts in Legislative Coordination

In the construction of cross-border AI governance frameworks, divergences at the level of technical standards constitute a primary challenge for institutional alignment. Different jurisdictions, shaped by their respective industrial bases and technological trajectories, tend to develop distinct technical standard systems. For instance, the European Union's mandatory requirements on algorithmic transparency differ in their specific parameterisation from the more guidance-based, industry-led standards prevalent in North America. As a result, products intended for global markets face increased compliance costs.

Given that internationally recognised testing and certification mechanisms remain under development, a single AI system is often subject to multiple rounds of assessment under different regulatory regimes. This, in practice, imposes additional burdens on development cycles and may constrain innovation efficiency. Moreover, issues relating to the compatibility of core technologies — such as data interface protocols — also affect the implementation of cross-border collaborative projects.

With regard to the legal status of AI systems, different legal traditions exhibit divergent approaches. Civil law systems generally adhere to a binary structure of "subject–object" classification and remain cautious about recognising AI as possessing independent legal personality, instead attributing responsibility primarily to developers or users. By contrast, certain discussions within common law contexts explore more tailored liability frameworks for systems capable of autonomous decision-making. These conceptual differences become particularly salient in scenarios such as cross-border autonomous driving accidents, where differing legal characterisations of technological features may lead to inconsistent adjudicatory outcomes. This divergence reflects not only differences in legal tradition but also varying cultural perceptions of technological risk.

In addressing the evolution of cross-border legal conflicts, traditional conflict-of-laws doctrines are undergoing renewed examination. Some scholars advocate refining the "closest connection" principle by applying a weighted analysis of factors such as the place of development, the location of data sources, and

the place where damage occurs in order to determine the applicable law. However, the inherently distributed nature of AI technologies — characterised by cloud-based training, edge computing, and global interaction — renders the identification of such connecting factors more complex.

Proposals to establish unified conflict-of-laws rules for the digital domain also raise deeper issues concerning sovereignty coordination within international legal frameworks. In specific areas such as data privacy, judicial practice has explored the application of extraterritorial jurisdiction through doctrines such as the "market destination" principle. Although such approaches remain subject to debate, they nonetheless provide useful institutional references for resolving practical disputes. At present, however, a mature and universally accepted solution for addressing these legal conflicts has yet to emerge, and continued dialogue among jurisdictions remains essential.

IV. Evolutionary Trends and Future Vision of Global AI Governance Legislation

(I) Trend Analysis: Transformations in Legislative Philosophy

1. From Sector-Specific Regulation to Cross-Domain, Integrated Legislative Frameworks

The deep integration of artificial intelligence across sectors — such as healthcare, transportation, manufacturing, and everyday life—has led to increasing interdependence among application domains, posing challenges to traditional sector-based regulatory models. As technological elements flow across industries, fragmented legislative approaches reveal limitations in terms of both comprehensive coverage and regulatory coherence.

The development of cross-sector coordination mechanisms and the dismantling of institutional information silos have therefore become key directions for institutional evolution. The establishment of specialised coordinating bodies or the enhancement of inter-agency joint consultation mechanisms can facilitate the integration of expertise from technological development, data governance, and sectoral applications. This enables a more holistic assessment of the multidimensional impacts of emerging AI applications and supports the formulation of regulatory responses that balance innovation with risk control. For example, where a medical imaging recognition model is repurposed for environmental perception in autonomous driving, issues relating to data compliance and liability attribution must be jointly assessed by authorities responsible for healthcare, transportation, and industrial regulation to ensure the accurate application of legal rules.

The long-term objective is to construct a multi-layered governance framework covering the entire lifecycle of AI — from technological development to real-world deployment — while simultaneously addressing technological evolution, data protection, and application-specific risks. Lawmakers must integrate technological, economic, social, and ethical considerations, promoting deeper coordination across regulatory domains and building a highly adaptive governance system that enables technological advancement while ensuring its alignment with societal needs.

2. From Regional Standards to the Formation of International Consensus

Global AI governance currently exhibits a pattern of regulatory pluralism. Different economies, shaped by their respective legal traditions, industrial bases, and value orientations, have developed distinct regulatory frameworks. The European Union is characterised by stringent risk-based classification, the United States emphasises innovation efficiency and regulatory flexibility, and China seeks to balance development with security. This coexistence of regional standards reflects diverse governance priorities, but also raises compliance burdens for cross-border technological development and product deployment, thereby affecting the efficiency of global industrial coordination.[28]

International standard-setting bodies such as ISO and IEEE are playing an increasingly important role in fostering global consensus. These organisations are actively promoting alignment in foundational areas including terminology, technical specifications, testing methodologies, and ethical principles. For instance, ISO/IEC JTC 1/SC 42 focuses on core dimensions such

[28] Sun, Zhiwei. "From Technological Competition to the 'Intelligent Commons': Paradigm Shifts in Global AI Governance". *Global Review*, 18.1 (2026): 40–61+174–175.

as system trustworthiness, while IEEE advances standards related to ethically aligned design. Such efforts provide a critical pathway for bridging regional divergences and establishing a common technical language and interoperable framework.

The transition from regional standards to global consensus requires coordinated efforts across multiple dimensions. First, mutual recognition mechanisms for standards should be established, enabling certification results that meet commonly accepted benchmarks to be recognised across jurisdictions through bilateral or multilateral arrangements. Secondly, cross-border testing and certification systems should be further developed, enhancing coordination in testing environments, evaluation criteria, and data-sharing practices. Finally, greater convergence is needed in international ethical standards, particularly in areas such as human oversight, fairness, and privacy protection. The ultimate vision is to build a global regulatory network that accommodates diversity while ensuring the smooth flow of technological elements.

3. From Risk Containment to a Sustainability-Oriented Value Framework

AI legislation is undergoing a fundamental transition from a focus on baseline risk prevention to a broader value orientation centred on sustainable development. Early regulatory efforts primarily aimed to mitigate risks such as loss of control, data breaches, and algorithmic discrimination. As AI applications deepen, however, issues relating to resource efficiency, social equity, and environmental impact have become increasingly prominent. Contemporary legislative approaches are moving beyond passive risk containment toward actively guiding technological development in service of human well-being and ecological sustainability.

In practice, principles of green and low-carbon development are being incorporated into the full lifecycle management of AI systems, translated into concrete mechanisms such as energy efficiency thresholds, carbon footprint accounting, and environmental impact assessments. Requirements for large-scale systems to disclose resource consumption are intended to incentivise the development and deployment of energy-efficient algorithms and hardware.

Inclusiveness and equitable access have also emerged as central pillars of AI governance. Legal frameworks increasingly seek to ensure the fair distribution of technological benefits, for example by enhancing accessibility in public services through multilingual and barrier-free interaction, and by addressing algorithmic bias to protect vulnerable groups. Support for the development of low-cost AI tools further enables small and micro enterprises, as well as individual actors, to benefit from technological advancement.

A central objective is the integration of AI governance with the United Nations Sustainable Development Goals (SDGs). Through cross-sector coordination mechanisms — for instance, linking ethical review processes for healthcare AI with broader public health objectives — legislators can ensure that technological innovation contributes to both social prosperity and ecological balance, thereby achieving a harmonious alignment between technological potential and societal value.

(II) Path Optimisation: The Agile and Inclusive Transformation of Governance Approaches

1. Enhancing Innovation Compatibility through Regulatory Sandboxes

Regulatory sandboxes, as an experimental tool in AI governance, are designed to provide controlled real-world environments in which innovation can be tested under manageable risk conditions. Their core value lies in enabling regulators to closely observe the societal impacts and risk characteristics of emerging technologies, thereby facilitating adaptive regulatory responses.

In international practice, the United Kingdom ' s Financial Conduct Authority (FCA) pioneered the operational framework for regulatory sandboxes, which was subsequently extended by the Monetary Authority of Singapore to AI applications in the financial sector. These mechanisms typically encompass structured entry, testing, and exit processes. During the entry phase, emphasis is placed on assessing technological novelty, public interest value, and the adequacy of risk mitigation measures. The testing phase involves continuous regulatory monitoring and real-time data feedback, ensuring information symmetry between regulators and innovators. In the exit phase, regulatory decisions are made based on test outcomes: compliant projects may be authorised for market entry, while those presenting significant risks are required to undergo further refinement.

The introduction of regulatory sandboxes effectively mitigates the inherent tension between rapid technological iteration and the stability of legal rules. By providing realistic testing environments, this mechanism prevents excessive regulation from stifling innovation while enabling early-stage risk identification and mitigation. As governance experience accumulates, key directions for institutional development include enhancing the inclusiveness of sandbox participation, improving cross-agency coordination, and translating accumulated experience into more generalisable regulatory standards.

2. Building a Technology-Driven Regulatory System Integrating Technical and Institutional Tools

AI governance is increasingly shifting toward a technology-driven model, in which regulatory effectiveness is enhanced through the use of technological tools. The adoption of regulatory technology (RegTech) provides substantive support for improving regulatory oversight. For instance, the immutability of blockchain technology offers a potential solution to challenges of algorithmic transparency: by recording algorithmic operations and decision parameters in a tamper-resistant manner, it enables the creation of traceable audit trails, thereby reducing "black box" risks and providing evidentiary support for the attribution of liability.

Artificial intelligence itself is also being deployed to support the intelligent transformation of regulatory systems. Through machine learning–based analysis of operational data, regulatory systems can automatically detect anomalous patterns or emerging risks, enabling a shift from reactive enforcement to proactive risk warning. This algorithm-driven monitoring mechanism exhibits significant responsiveness, allowing regulators to identify risks at an early stage and shorten response cycles — for example, through real-time monitoring of abnormal data flows in the financial technology sector. Combined with big data analytics, regulatory processes evolve into systems of continuous learning and dynamic adjustment.

The essence of technology-enabled governance lies in constructing regulatory architectures capable of self-evolution. Such systems strengthen the foundational safeguards of safety while maintaining the innovative vitality of the industry, thereby fostering a virtuous cycle between regulation and development.

3. Advancing Ethics-Centred Legislation Based on Human–Machine Harmony

Legal regulation in the field of artificial intelligence often faces the challenge that technological development outpaces institutional supply, resulting in a lag in the application of ethical norms. Existing governance models tend to focus on ex post responses to disputes, making it difficult to prevent systemic risks at the stage of technological development. Accordingly, embedding ethical values—such as human dignity and fairness—into legal frameworks at an earlier stage has become a central focus of institutional reform.

With respect to algorithmic bias, legislation may require developers to incorporate bias detection mechanisms at the design stage, making such safeguards a condition for market access. In high-risk domains such as medical diagnostics, it is essential to ensure that human professionals retain ultimate decision-making authority, including the right of override.

The establishment of interdisciplinary ethics committees is a key institutional support for achieving these objectives. Such bodies should integrate expertise from law, ethics, and computer science, and be responsible for setting ethical evaluation benchmarks and guiding enterprises in conducting ethical impact assessments at early stages of development. In practice, this includes assessing the diversity of data samples—particularly in applications such as medical algorithms—and evaluating whether decision-making logic meets standards of professional explainability.

A dynamic ethical assessment mechanism should be implemented based on risk differentiation. Simplified procedures may apply to low-risk applications such as chatbots, while high-risk domains — including autonomous driving and financial risk control—should be subject to full lifecycle monitoring. Relevant actors should be required to submit operational data on a periodic basis for review. At the legal level, clear standards must be established to determine when technological deviation occurs — for example, where algorithmic logic diverges from widely accepted human values or produces harmful outcomes — triggering mandatory intervention. Principles such as human ultimate control should be закреnеd in binding legal provisions to ensure that technological development consistently serves the overall well-being of humanity.

(III) Vision and Outlook: Toward a Co-Governed and Shared Legal Order

1. From Competitive Safeguarding to Cooperative Governance

AI governance is transitioning from a model centred on defensive competition toward one grounded in cooperative, mutually beneficial engagement. Historically, regulatory frictions — such as restrictions on algorithm exports and data localisation requirements — have arisen from efforts to protect technological advantages. While such measures may serve national interests, they also increase global compliance costs and constrain cross-border collaboration in areas such as climate modelling and public health. Given the inherently transnational nature of AI technologies, unilateral governance frameworks

are insufficient to address complex challenges such as algorithmic bias, necessitating deeper collaboration among governments, enterprises, and research institutions. In areas such as autonomous driving standards, multi-stakeholder participation enhances the capacity to identify latent risks.[29]

In practice, the development of institutional platforms for cooperation has become a shared priority. The establishment of international AI governance funds can support broader participation in rule-making and enhance inclusiveness in global governance structures. Technology-sharing platforms can reduce duplication of research efforts; for example, open-source medical imaging models have significantly improved the global efficiency of cancer screening. At its core, this model involves the establishment of mechanisms for risk-sharing and benefit-sharing, including clear allocation of responsibilities for privacy protection and the creation of cross-border compensation funds. Achieving this transformation requires major economies to build mutual trust in frontier areas such as quantum computing regulation and to move beyond short-term competitive considerations. Ongoing multilateral dialogue mechanisms are gradually mitigating adversarial dynamics and paving the way for coordinated global governance.

2. Building a Resilient and Adaptive Global Legal Cooperation Network

In global AI governance, diverse regulatory approaches reflect differing developmental priorities but also introduce complexities in rule coordination. The cross-border flow of technological elements and the boundaries of sovereign jurisdiction are driving efforts to construct a global legal cooperation framework that is both adaptive and resilient. Such a framework must be capable of accommodating technological evolution while mediating tensions between different governance models and maintaining systemic stability.

The United Nations, together with prospective AI advisory bodies, holds significant potential as a central coordinating platform. The effective realisation of this role depends on the consolidation of international consensus. Establishing institutionalised platforms with functions such as agenda-setting, information-sharing, and policy coordination represents a key pathway forward. These platforms should regularly review national legislative developments, identify points of regulatory convergence, and mobilise expert resources for risk assessment, thereby providing substantive support for the formation of international norms.

The effectiveness of such a network also depends on enhanced cross-border enforcement cooperation. Given the global nature of AI applications, mechanisms such as joint investigations, cross-border evidence collection, and coordinated oversight in high-risk sectors are essential. These arrangements must balance enforcement efficiency with respect for digital sovereignty, while promoting mutual recognition of technical standards, certification systems, and risk assessment methodologies. Such measures not only reduce cross-border compliance costs but also foster a predictable business environment and strengthen the internal resilience of the governance system. Ultimately, this multi-layered cooperative network aims to enable consensus-building through continuous evolution while maintaining systemic stability.

3. Achieving a Dynamic Balance between Technological Empowerment and the Protection of Human Values

The evolving relationship between humans and artificial intelligence requires a refined legal framework. The principle of maintaining ultimate human control is

[29] Zhi, Zhenfeng. “Key Concerns and Future Trends in Global AI Legislation”. *Journal of Comparative Law,* (06) (2025): 17–36.

central to defining the boundaries of technology and preserving decision-making autonomy. Its institutional purpose lies in harnessing technological potential while safeguarding human dignity.

The establishment of comprehensive frameworks for assessing the societal impacts of technology is of fundamental importance. Such assessments should extend beyond considerations of technical safety or economic efficiency, incorporating ethical, social, and employment-related dimensions. Broad participation by experts, scholars, and the public enhances both the scientific validity and social legitimacy of these evaluations. The resulting findings can directly inform research directions, optimise market entry standards, and support the dynamic revision of legal rules, thereby enabling proactive risk governance.

The ultimate vision of governance is the creation of a self-adjusting legal ecosystem — one that evolves alongside technological development while remaining firmly grounded in human-centred values. Such a system should prevent the alienation of technology while promoting the expansion of human capabilities and ensuring equal access to opportunities. Particular attention to the protection of marginalised groups can help bridge the digital divide and ensure the equitable distribution of technological benefits. In this way, legal systems can achieve a sustainable dynamic balance between technological empowerment and the protection of human values.

Conclusion

Global AI governance is entering a critical phase of transition from fragmented regulation toward coordinated international cooperation. This process involves not only the alignment of legal systems but also the broader construction of an order for technological civilisation. Looking ahead, the role of the rule of law should be directed toward building a resilient and adaptive global governance network—one capable of transcending differences in development stages and cultural paradigms, and of achieving substantive consensus through diversity.

AI legislation must move beyond a purely risk-prevention paradigm toward a value-oriented framework that actively guides technological development in the service of human well-being and ecological sustainability. This transformation reflects a deeper evolution in governance philosophy: from safeguarding minimum safety thresholds to fostering an open, fair, and inclusive digital ecosystem. Within this vision, law is no longer merely a constraint, but becomes an enabling institutional framework for innovation—providing flexibility for technological advancement while ensuring ultimate human control.

The realisation of this forward-looking legal vision depends on deeper alignment across ethical integration, technology-driven governance, and collaborative regulatory frameworks. As global governance evolves from fragmented sectoral practices into a coherent legal civilisation, the benefits of technological progress can be more widely shared, narrowing the digital divide and promoting social equity. Ultimately, this process represents a profound convergence of human wisdom and the spirit of the rule of law in the digital age, ensuring that technological development remains firmly oriented toward the advancement of humanity and the realisation of a harmonious human–machine future.