iapp | HCLTech

ARTICLE SERIES

Global AI Governance Law and Policy

AUSTRALIA, CANADA, CHINA, EU, INDIA, JAPAN, SINGAPORE, SOUTH KOREA, UAE, UK AND US

Jurisdictions worldwide are designing and implementing artificial intelligence governance laws and policies commensurate to the velocity and variety of the risks and opportunities presented by AI-powered technologies. Since the first series in 2024, the state of AI policy worldwide has evolved with many jurisdictions staking their own path.

Articles in this series, co-sponsored by HCLTech, dive into the laws, policies, broader contextual history and developments relevant to AI governance across the world. The highlighted jurisdictions have made a mark on the global conversation around AI governance and provide for a small but important snapshot of distinct approaches to AI governance in key global markets.

Each article provides a breakdown of the key sources and instruments that govern the strategic, technological and compliance landscape for AI governance in the jurisdiction through voluntary frameworks, sectoral initiatives or comprehensive legislative approaches. Special care is taken to weave together how key areas, like intellectual property or agentic AI, provide unique challenges and opportunities. Agentic AI has been a new addition to the series, as it represents the newest frontier for how organizations can realize value out of this technology. Currently, few jurisdictions have rules specific to agentic AI, but instead rely on existing legal frameworks. Read about how agentic AI is changing how organizations think about AI governance.

Contents

Published November 2025

Global AI Governance Law and Policy: Australia

By James Patto and Pamela Gupta

Australia's artificial intelligence regulatory journey has shifted from an early plan to introduce an EU-style, risk-based regime toward a more flexible, standards-led approach. What began as a move toward prescriptive guardrails and potential legislation has been seemingly overtaken by a focus on productivity, innovation and the use of existing legal frameworks. Yet this recalibration comes amid persistently low public trust in AI, creating a complex policy challenge: how to build accountability, safety and transparency without constraining the very innovation needed to realize AI's economic and social potential.

History and context

Australia's engagement with artificial intelligence builds on more than half a century of deep technology research. Australian universities and the Commonwealth Scientific and Industrial Research Organisation have long contributed to global advances in computer science, robotics, quantum computing, photonics, biotechnology and materials science. Despite strong research capability, commercialization has lagged. Limited venture capital, fragmented university-industry links and a relatively small domestic market have meant that many innovations were scaled offshore.

These structural realities have shaped Australia's broader economic profile: a nation whose prosperity rests on resource exports and advanced service sectors and a country whose "knowledge economy" has focused less on producing deep tech and more on adapting innovation to strengthen established industries. Artificial intelligence follows that pattern. Australia excels in applied domains such as mining, agriculture, health care and defense but remains a net importer of foundational AI systems and platforms.

As a result, Australia is unlikely to become a global powerhouse in AI model development. On the other hand, the country holds clear advantages in applied domains like mining automation, precision agriculture, medical diagnostics, climate science, defense and public-sector service delivery. The greatest economic benefits are expected from productivity gains and efficiency improvements rather than from AI exports.

Recognizing both opportunity and risk early, Australia was among the first countries to articulate principles for responsible AI. In November 2019, the federal government released Australia's Artificial Intelligence Ethics Principles, a voluntary framework covering fairness, transparency, privacy, accountability and human wellbeing. These principles laid the groundwork for subsequent policy, research and procurement guidance; they also signaled that AI should be pursued in line with public trust.

Institutionally, a key milestone came in 2021 with the creation of the National AI Centre under CSIRO's Data61 division to strengthen national capability and promote responsible adoption. The NAIC later moved into the Department of Industry, Science and Resources in 2024, reflecting the growing alignment between AI governance and economic strategy.

The DISR now leads AI policy as part of the broader industry and innovation portfolio. The department's posture has traditionally been risk-based, focusing on managing harms such as bias and misinformation while encouraging safe innovation. This was evident in the "Safe and Responsible AI in Australia" discussion paper published in June 2023 and its interim response that followed in January 2024, which proposed a risk-proportionate framework featuring mandatory safeguards for high-risk AI and voluntary guidance for lower-risk systems.

Australia also hosts a growing network of research and policy centres, including the Australian Institute for Machine Learning, the Responsible AI Research Centre (CSIRO, South Australian Government and University of Adelaide) and the Human Technology Institute at the University of Technology Sydney, each contributing to responsible-AI design and governance. States have also played a role, with New South Wales introducing one of the first frameworks guiding the ethical use of AI in government.

Importantly, Australia faces a pronounced trust deficit in AI adoption. According to a 2025 study by the University of Melbourne and KPMG, only 30% of Australians believe the benefits of AI outweigh its risks; just 36% of citizens trust AI systems more broadly. Approximately 78% of respondents expressed concern about negative outcomes from AI, and only 30% believe current laws and safeguards are adequate. This trust gap remains a central challenge for policymakers seeking to balance innovation with public confidence and adoption.

Approach to regulation

Australia does not have dedicated or overarching AI legislation. Instead, its regulatory approach relies on a combination of voluntary frameworks and existing non-AI specific laws. The government's position has evolved from a primarily risk-based lens toward one that increasingly seeks to harness AI's productivity and innovation benefits without stifling development.

Following the "Safe and Responsible AI in Australia" discussion paper, the government moved into a more specific phase of policy design.

In September 2024, the government released "Introducing Mandatory Guardrails for AI in High-Risk Settings," a proposals paper exploring possible ex ante obligations for high-risk AI applications. It asked stakeholders to consider what constitutes "high-risk AI," whether the proposed guardrails were fit for purpose, and how they should be implemented.

At a high level, the proposed guardrails focused on:

- Accountability and governance across the AI lifecycle.
- Privacy, data quality and data management.
- Testing, assurance and ongoing monitoring of performance and safety.
- Transparency, explainability and traceability.
- Human oversight and contestability.
- Security, integrity and record-keeping.

That same month, the government released the Voluntary AI Safety Standard to provide immediate, non-binding guidance for organizations. Closely mirroring the 10 guardrails proposed in the "Mandatory Guardrails for AI in High-Risk Settings" paper, the VAISS offered a practical preview of what future enforceable requirements might look like. These heavily drew on international benchmarks such as the EU AI Act, Canada's now defunct Artificial Intelligence and Data Act, ISO/IEC 42001, the National Institute of Standards and Technology AI Risk Management Framework and the National Institute of Standards and Technology AI Principles.

From regulation to recalibration

Based on the above, Australia appeared poised to quickly move toward a dedicated statutory regime, a lighter-touch analogue to the EU's AI Act.

By mid-2025, however, priorities shifted toward economic growth and productivity. Domestic productivity challenges and global developments have prompted Australia to reassess its posture. Within the context of the rapid rise of generative AI, early enthusiasm for EU-style regulation has given way to a more innovation-focused outlook. Moves by allies such as the U.K. to adopt approaches without prescriptive ex ante laws and the Trump administration’s explicit rejection of comprehensive AI regulation further influenced this pivot. Meanwhile, the EU AI Act's lack of global influence and mounting criticisms over its complexity and compliance burden also, no doubt, played a part.

As a result, the regulatory program was effectively paused and replaced with a broad review of existing laws and regulators. The emerging direction favours harmonisation of current frameworks over creating a new, centralized regime. The results of this review, alongside the forthcoming National AI Strategy due at the end of 2025, will determine whether further targeted reforms or coordination mechanisms are introduced. In the meantime, the National AI Centre published its Guidance for AI Adoption, a new framework replacing the VAISS, in October. While it remains too early to confirm whether this will define Australia's long-term approach, it may signal a broader shift toward standards-led rather than legislative regulation.

Public-sector governance

Progress has been slightly more tangible within the public sector where several frameworks already apply.

The Australian Government Responsible AI Policy sets minimum requirements for all Australian Public Service entities, such as mandating transparency statements and the appointment of accountability officers. The National Framework for Assurance of Artificial Intelligence in Government provides agencies with structured methods for AI assurance, testing and implementation to put the national AI Ethics Principles into practice. The Australian Government AI Technical Standard specifies design, testing and documentation requirements for AI use in government systems. The AI Data-Security Guidance was issued to address provenance, supply-chain integrity, data poisoning and model-manipulation risks.

Together, these and other instruments form a quasi-regulatory baseline that operationalizes the AI ethics principles within government practice.

Balancing innovation and protection

Australia's evolving approach seeks to balance risk management with innovation enablement. Having seemingly stepped back from a single overarching AI statute, the government appears to be intent on embedding AI oversight within existing legal and regulatory systems — a hybrid model designed for agility, coherence and international compatibility. This approach, however, faces a critical challenge highlighted above: low public trust in AI.

For AI to deliver productivity, innovation and economy-wide gains at scale, uptake is essential; uptake depends on confidence that AI will operate fairly, transparently, safely and accountably. Without that trust, citizens may resist AI-mediated decisions or services, undermining both investment and adoption. This trust gap lies at the heart of Australia's regulatory tug-of-war. Policymakers must build frameworks that reassure the public without constraining innovation, making trust both a constraint and an objective of AI regulation.

Wider regulatory environment

While Australia has paused work on a stand-alone AI law, a wide range of existing legal frameworks already apply to the development and use of AI, including those governing privacy, consumer protection and product safety, discrimination and employment, intellectual property, and online safety, together with certain sector specific laws.

Privacy

The Privacy Act 1988 (Cth) remains the primary law regulating the handling of personal information in Australia. The act is principles-based and is currently undergoing significant reform following the government's multiyear review, which commenced before the rise of generative AI. As a result, while no AI-specific amendments are currently proposed, many of the forthcoming changes would impact AI use and development. The first tranche of reforms, passed in 2024, introduced new transparency obligations around automated decision-making that will take effect in December 2026.

Australia's privacy regulator, the Office of the Australian Information Commissioner, has been proactive in interpreting the act in AI contexts and is actively regulating AI through interpretation and enforcement rather than waiting for dedicated legislation.

In October 2024, the OAIC released two companion guidance pieces clarifying how the act applies to AI. The first, Guidance on privacy and the use of commercially available AI products, is directed at organizations deploying AI tools. It emphasises due diligence when selecting vendors and outlines expectations for privacy-by-design, transparency and accountability.

The second, Guidance on privacy and developing and training generative AI models, is directed at developers and researchers. It emphasizes that even publicly available data may contain personal information and must be handled in accordance with consent, purpose-limitation and data-minimisation principles. It also cautions that AI hallucinations, or outputs inferring personal details, can themselves constitute the collection of personal information, underscoring obligations around accuracy, security and deletion of data no longer required.

The OAIC has also issued several landmark determinations relevant to AI powered facial-recognition technology, including Clearview AI in 2021, wherein scraping online images to build a FRT database breached Australian privacy law; 7-Eleven Stores in 2021; Bunnings Group in 2024; and Kmart Australia in 2025. All of these cases involved the unlawful collection of biometric information of customers by using FRT.

Additionally, the OAIC examined the use of deidentified medical imaging data from the I-MED Radiology Network shared with Annalise.ai for AI-training purposes. Following preliminary inquiries, the commissioner found that I-MED's deidentification methods and governance controls were sufficient for the data to fall outside the definition of personal information under the Privacy Act.

Consumer and product liability

Artificial intelligence-enabled products and services are governed by Australia's consumer-protection and product-liability frameworks, principally the Australian Consumer Law under the Competition and Consumer Act 2010 (Cth). The ACL is principles-based and technology-agnostic, applying broadly to emerging products and services, including those incorporating AI.

At its core, the ACL prohibits misleading or deceptive conduct, unconscionable practices and false or misleading representations. It also provides consumer guarantees requiring goods and services, including those powered by AI, to be of acceptable quality, fit for purpose and accurately described. These duties extend to AI systems that make representations, recommendations or automated decisions. A business may contravene the ACL if an AI tool exaggerates its capabilities, obscures human oversight, or produces outcomes likely to mislead consumers.

Under the ACL's product-liability provisions, suppliers and manufacturers must ensure that goods, including AI-embedded software, are safe and free from defects. Where AI contributes to physical injury, property damage or financial loss, liability may arise under negligence, statutory guarantees or the product-safety regime. In practice, this can create shared accountability across the AI supply chain, from model developers and integrators to deployers and end users.

Intellectual property

Australia's intellectual property framework presents several uncertainties for AI development and use. In 2022 and 2023, the Federal Court and Full Court confirmed that AI systems cannot be named as inventors under the Patents Act 1990 (Cth), finding that inventorship is limited to natural persons. This aligns with most common-law jurisdictions and means that patent protection for AI-generated inventions must currently be sought through a human intermediary.

In the copyright domain, governed by the Copyright Act 1968 (Cth), Australia follows a "fair-dealing" model rather than a broad "fair-use" regime. Without permission, lawful use of copyright material is limited to specific purposes such as research, criticism, news reporting and parody. None clearly extend to large-scale data scraping or model training, creating uncertainty for developers, publishers and rights holders. Questions also persist around authorship, ownership and liability for AI-generated works, particularly where human involvement is minimal or diffuse.

While these laws remain in force, their interaction with data-driven and generative AI systems is still evolving. As debates around copyright, data mining and licensing intensify, Australia's IP regime is emerging as a critical frontier for future AI regulatory reform.

Online safety and deepfakes

Australia's regulatory response to AI-generated deepfakes spans both criminal and civil regimes. Offences for creating or distributing non-consensual intimate images, including synthetic or AI-generated content, are primarily contained in the Criminal Code Act 1995 (Cth). In parallel, the Online Safety Act 2021 (Cth) empowers the regulator, the eSafety Commissioner, to order the removal of intimate images or synthetic media that depict or simulate a person without consent. Collectively, these mechanisms position Australia among the earliest jurisdictions to explicitly regulate harmful or non-consensual AI-generated material.

Discrimination and employment

Artificial intelligence-assisted decision making is also regulated under antidiscrimination and employment laws. Employers and service providers remain liable for algorithmic bias or discriminatory outcomes in hiring, promotion, credit, insurance or service delivery.

Industry-specific regulation

Certain sectors are already governed by frameworks that intersect with AI deployment.

Artificial intelligence-powered medical devices are regulated under the Therapeutic Goods Act 1989 (Cth), which establishes the framework for assessing and approving software-as-a-medical-device systems. The legislation requires such systems to demonstrate safety, efficacy and performance consistent with their intended clinical purpose.

The Australian Prudential Regulation Authority integrates AI into its prudential risk management standards, requiring financial institutions to manage operational, data and cyber risks linked to AI use.

The Security of Critical Infrastructure Act 2018 (Cth) imposes an "all-hazards" risk management framework across a wide range of key critical sectors, including risks arising from AI systems.

The Australian Securities and Investments Commission requires financial services licences holders to implement and maintain adequate risk management processes and pr ocedures, which extends to AI.

Artificial intelligence's emergence has also prompted responses from legal regulators and courts. Several professional regulators have issued guidance on responsible AI use in legal practice, emphasising duties of competence, confidentiality and supervision. In one notable case in Victoria, a practitioner lost their principal practicing certificate after submitting AI-generated material containing false citations. Australian courts have likewise begun publishing practice directions on the use of AI by litigants.

Agentic AI

While the Australian government has not specifically addressed agentic AI in any of the released policies, guidelines or laws, the broader principles for responsible and trustworthy AI are likely to be applied to the development of agentic AI as well. Notably, the 2024 amendments to the Privacy Act, which will come into effect in late 2026, have significant ramifications for automated decision-making. Covered entities must now disclose, within their privacy policies, the types of personal information used, the nature of decisions made solely by computer programs, and those where computer assistance significantly influences outcomes that could substantially affect individuals' rights or interests.

Latest developments

Australia's AI regulation journey has entered a recalibration phase rather than a standstill. While dedicated legislation remains on hold, the government and regulators are actively refining existing laws, standards, and strategies, laying the groundwork for a more integrated national approach to AI.

Review of existing laws

The government is currently reviewing whether existing laws can accommodate emergent AI risks. That review includes privacy, consumer protection, safety, antidiscrimination, product liability laws and the capacity of regulators to supervise AI-enabled systems. If these frameworks prove adequate, the original mandatory guardrails proposal may be abandoned in favour of a multi-regulator approach.

Guidance for AI Adoption

In October 2025, the NAIC released the Guidance for AI Adoption, formally updating and replacing the VAISS. The new framework provides a practical, nationally consistent blueprint for organizations seeking to responsibly govern AI. It consolidates the VAISS's 10 guardrails into six responsible AI practices that covers governance and accountability, impact assessment, risk management, transparency, testing and monitoring, and human oversight.

The first part of the Guidance for AI Adoption, "Foundations," is aimed at small and medium-sized enterprises; the second, "Implementation Practices," is intended for larger or more mature organizations. The guidance translates high-level principles into actionable steps. It is supported by a suite of resources, including an AI screening tool, policy and register templates, and reference materials outlining key definitions and risk mitigation measures.

At its core, the guidance encourages proportionate governance based on risk and fosters transparency and human accountability. While it is too early to say whether this guidance will define Australia's long-term regulatory approach, it signals a strong direction toward practical, standards-based governance as the government continues to shape its broader AI strategy.

National AI Strategy

The government has committed to producing a National AI Strategy by the end of 2025. The strategy aims to set national priorities for innovation, adoption, international alignment and trust frameworks, integrating lessons from past consultations and sectoral reviews.

Recently completed reviews

In July 2025 the Therapeutic Goods Administration published Clarifying and Strengthening the Regulation of Medical Device Software including Artificial Intelligence. The review found that Australia's risk-based, technology-neutral framework remains largely fit for purpose but needs refinement. The organization recommended updates to definitions, such as manufacturer, sponsor, supply, adaptive-AI/change-control provisions; clearer guidance for digital scribes and clinical-assist tools; and stronger evidence, monitoring and transparency requirements. Additional consultations are planned for 2025-26.

The Treasury's Final Report on AI and the Australian Consumer Law recommends targeted clarifications to clarify the ACL's application to AI systems, especially regarding definitions of goods/services, manufacturer liability and algorithmic representations. It emphasizes reinforcing the ACL's existing principles rather than introducing AI-specific laws.

Intellectual property questions

Artificial intelligence and copyright have become central to current policy debate. Key stakeholders have agreed to explore compensation frameworks for the use of copyrighted material in AI training, signaling that licensing models may feature in future reforms. At the same time, the Productivity Commission has proposed a text-and-data-mining exception to the Copyright Act 1968 (Cth) to permit certain AI-training activities. That proposal has since been rejected by the federal government, which has ruled out any copyright carve-out for AI developers. The decision marks a decisive shift in bargaining power away from technology companies potentially hoping for a legislative reprieve and towards the creative and union sectors, which are now positioned to shape the contours of future frameworks. Additionally, the government has announced that a working group will meet imminently to consider whether Australia's copyright framework needs updating.

Recommendations on next steps

Adopting OECD guidance to strengthen AI regulations and adoption

As an adherent to the OECD AI Principles and a nation striving for international compatibility in its standards-led approach, Australia can strengthen its current regulatory framework by formally adopting specific tools and policy recommendations derived from the OECD's guidance. These additions would help address the nation's persistent public trust deficit while advancing its goal of building an agile and effective governance environment.

Strengthen transparency and accountability through tools

Australia's current framework relies on existing legal systems and voluntary guidance. The OECD offers concrete tools that Australia can incorporate to enhance accountability and public transparency, moving beyond general principles.

Australia can establish a formal mechanism or participate in existing frameworks to monitor and understand AI incidents. The OECD platform provides the AI Incidents and Hazards Monitor and focuses on AI risk and accountability. Integrating a mandatory incident reporting framework, especially for high-risk systems, would allow the government to track real-world harms — a necessary step for building public trust given that 78% of Australians are concerned about negative outcomes from AI.

Australia can require organizations developing advanced AI systems to participate in a structure similar to the Hiroshima AI Process Transparency Reporting Framework. This type of reporting facilitates transparency and comparability of risk mitigation measures across the industry, directly supporting the OECD value of transparency and explainability.

Formalize definitions for interoperability

Australia's current approach involves reviewing and refining existing laws, such as consumer law and medical device regulation, rather than creating new legislation. To ensure its regulatory environment is interoperable and coherent, Australia can formally adopt the foundational OECD definitions.

Policymakers should explicitly leverage the OECD's definition of an AI system and the AI system lifecycle in their revised or new regulatory guidance. Countries, including the EU and U.S., use these definitions in their frameworks, making their adoption by Australia crucial for ensuring global interoperability. This would provide clear, internationally recognized terminology for industry and regulators, streamlining Australia's hybrid regulatory model.

Expand focus to key policy areas

While Australia's National AI Strategy is forthcoming at the end of this year, the OECD's detailed policy focus areas can guide the strategy's content, ensuring all critical aspects of AI governance are addressed.

The OECD specifically tracks AI compute and the environment. Australia can integrate policies to manage the environmental impact of large-scale AI computing, an area currently overlooked in legal review, which focuses primarily on consumer protection, privacy, and intellectual property laws.

Australia's focus on productivity gains aligns with the OECD's work on the Future of Work and the Work, Innovation, Productivity and Skills in AI program. The National AI Strategy should incorporate the OECD recommendation to build human capacity and prepare for labour market transition, ensuring the workforce is equipped for the changes AI will bring.

Given the rapid rise of generative AI, Australia should dedicate targeted policy guidance, drawing on the OECD's focus area for managing the risks and benefits of generative AI. This would complement the OAIC's existing guidance on training generative models under the Privacy Act.

Reinforce value-based principles

Australia's earlier proposals included mandatory guardrails focusing on elements like security, integrity and testing. Even with the shift toward standards-led governance, the OECD AI Principles of Corporate Governance provide a robust foundation for reinforcing core values.

Australia can explicitly structure its Guidance for AI Adoption and existing sector-specific regulations, like those governing critical infrastructure, to more strongly reflect the OECD's value of robustness, security and safety. This is critical for ensuring that AI systems are reliable and resilient, particularly in high-risk applications like medical devices and financial services.

While Australia focuses on productivity, the OECD framework promotes inclusive growth, sustainable development and well-being. Australia could embed a framework requiring AI actors to demonstrate how their systems contribute to broad societal well-being and adhere to human rights and democratic values, including fairness. This approach would help counteracting known risks, like algorithmic bias, that are currently regulated primarily through existing antidiscrimination laws.

Published September 2025

iapp | HCLTech

Global AI Governance Law and Policy: Canada

By Ashley Casovan, Carole Piovesan and Michael Pascu

Despite its population of just over 41 million, Canada has a strong track record of developing AI capabilities and talent. The country hosts numerous impactful startup accelerators, world-class researchers and universities dedicated to fostering a vibrant AI culture. Notably, it is home to several of the "godfathers of AI," including Geoffrey Hinton and Yoshua Bengio, who won the Turing Award in 2018 for their formative research on deep learning along with Yann LeCun. In October 2024, Hinton was also awarded the Nobel Prize for Physics, further cementing Canada's leadership in AI.

In 2017, Canada became the first country to launch an AI strategy, seeking to understand the implications and opportunities these powerful technologies can have on its economy and society. A cornerstone of the Pan-Canadian AI strategy is the work led by the Canadian Institute for Advanced Research. In close partnership with world-class national AI research institutes the Montreal Institute for Learning Algorithms, Vector Institute and the Alberta Machine Intelligence Institute, the vision of the AI strategy is to make Canada one of the world's most vibrant AI ecosystems.

Recognizing Canada's potential for technological advancement, the federal government, provincial governments, civil society organizations and industry have been active in seeking to create the necessary frameworks within which innovation can flourish safely and responsibly.

History and context

The federal government sets national AI standards and policies, while provinces handle localized issues like data privacy. In 2017, the federal government launched the first phase of its Pan-Canadian AI Strategy with a CAD125 million investment focusing on three pillars:

- → **Commercialization**, which involves transitioning AI research into practical applications for the private and public sectors.

- → **Standards**, which focus on developing and adopting AI standards.
- → **Talent and research**, which aim to foster academic research and enhance computing capacity for AI advancements.

In 2019, two years after launching phase one of its Pan-Canadian AI Strategy, Canada announced its Digital Charter. This charter outlines 10 principles to guide the federal government's digital and data transformation efforts, with AI playing a crucial role.

In 2022, phase two of the strategy was implemented, adding over CAD433 million to the overall budget to be utilized over the course of 10 years. The importance of AI was underscored when the Digital Charter Implementation Act was introduced to Parliament that same year. The act includes three key components: privacy reform, the establishment of a Personal Information and Data Protection Tribunal, and the introduction of a comprehensive AI and Data Act.

While concerned about the domestic implications of AI, the country also played a significant role in turning international attention and activity toward collectively working to develop AI in a responsible manner grounded in human rights. As such, Canada, along with France, was an initial driving force behind the Global Partnership on AI, a multistakeholder forum with 29 participating member nations. In 2024, the Global Partnership on AI was integrated into the Organisation for Economic Co-operation and Development's AI policy work now functioning alongside its AI Policy Observatory. Canada continues to host one of three GPAI secretariats through the International Center of Expertise in Montreal on Artificial Intelligence.

Understanding the importance of leading by example, Canada was the first country in the world to create an AI-specific legally binding instrument. With a focus on the government's use of AI, the Directive on Automated Decision-Making was launched in 2019. Designed as a risk-based policy now popularized by the likes of the EU AI Act, the DADM requires the use of a standardized algorithmic impact assessment tool to determine the risk of the system, allowing for better alignment of risk-appropriate obligations. Many of the concepts and key requirements of this policy are similar to those in related policies published today. Making the distinction between automated decision-making versus other types of AI, additional questions about the other policies may be useful. In 2023, with the same public sector scope, the government released guidelines for generative AI.

Recognizing the need to continue to build on this policy suite in light of the ever-changing nature of AI technologies, the federal government hosted a roundtable to develop an AI strategy for the public service. This strategy focuses on three main areas: building an AI-ready workforce and fostering AI growth through innovation, enabling infrastructure and engagement, and implementing tools for responsible and effective AI adoption.

In 2023, while focusing on the government's use of AI, the country brought together key industry actors to commit to a voluntary code of conduct for the safe and responsible use of generative AI. These concepts are aligned with similar international efforts like the Bletchley Declaration, a key agreement completed during the first AI Safety Summit hosted by the U.K.

To complement the existing efforts of the Pan-Canadian AI Strategy, the 2024 federal budget allocated CAD2.4 billion to advance AI with an eye on both internal use and external oversight. Of the budget, CAD2 billion is dedicated to a new AI Compute Access Fund as well as funding for a safety institute and advancement of sectoral research. This fund aims to invest in Canadian-made computing infrastructure to support AI businesses and researchers.

First minister of AI

In May 2025, Canada's new Prime Minister, Mark Carney, announced a new cabinet and appointed Evan Solomon to be the first minister responsible for AI and Digital Innovation. While portfolio-specific mandate letters have not been made public, a mandate letter to all ministers has been shared to outline priorities. Many of this government's priorities are shaped by current economic realities, including the need to foster stability and security, strengthen sovereignty and expand trade partnerships. These economic objectives are closely connected to the country's innovation agenda.

For AI in particular, the direction is significant: while Canada has long been recognized as a global hub for cutting-edge research, it has historically struggled to commercialize home-grown AI products and scale them for both domestic use and international markets. Additionally, we are starting to see similar portfolios in provinces. The province of British Columbia has appointed Minister Rick Glumac to be responsible for an Artificial Intelligence and New Technologies portfolio.

Additionally, we are starting to see similar portfolios in provinces. The province of British Columbia has appointed Minister Rick Glumac to be responsible for an Artificial Intelligence and New Technologies portfolio.

Investing in Canadian innovation

It is too early to report on the direction of this new Ministry. However, with recent funding announcements, it is clear Canada understands the economic stakes related to AI and is seeking to invest in capitalizing upon significant domestic capabilities to develop new AI technologies.

On 10 July, Solomon announced a nearly CAD100 million investment in partnership with Scale AI, a Canadian investment firm. This joint funding announcement was arranged to support 23 new projects across Canada ranging from logistics, supply chain management, health, and finance. The announcement was the first indication that Canada needs to support its AI innovators to secure future economic stability.

AI and the G7

At the beginning of 2025, Canada took on the role of the G7 Presidency. During the June Summit in Kananaskis, Alberta, the G7 leaders released a statement on AI for Prosperity. In addition to the recognition that AI would be an economic engine for all G7 nations in the future, this statement highlights the role of small and medium-sized enterprises, particularly how the SMEs will need support to adopt and develop new technologies. To assist in this goal, the creation of the G7 GovAI Grand Challenge was announced outlining that governments will work together to scale AI adoption faster. It includes a G7 AI Adoption Roadmap, which looks at public sector actions and ways that companies can adopt AI and scale their business. These efforts build upon the outcomes of the G7 Hiroshima AI Process.

Approach to regulation

In 2022, the federal government introduced Bill C-27. Aligned with global legislation trends at the time, this legislative framework proposed comprehensive oversight for AI to complement existing privacy (Part I) and consumer protection (Part II) legislation. Part III of this bill, the AI and Data Act, sought to establish a risk-based framework for regulating AI systems. While Bill C-27 made it to a second reading, the bill was ultimately not completed when the election produced a new administration in early 2025.

Similar to the EU, Canada's approach to legislating AI sought to balance protecting rights with fostering innovation. The AIDA aimed to regulate trade "by establishing common requirements, applicable across Canada, for the design, development, and use of (AI systems)" and to avoid harm by prohibiting certain conduct in relation to AI systems with a specific focus on "high-impact systems."

The AIDA did not outright ban certain AI uses, as the EU AI Act does. Instead, it classified AI systems into high-impact categories, imposing stricter risk management, transparency obligations and accountability frameworks for those who make such systems available.

At the provincial level, Québec and Ontario have taken notable steps toward regulating AI. Québec’s Law 25, a major privacy reform, includes requirements for transparency and safeguards around automated decision-making, making it one of the first provincial frameworks to directly address AI implications. In Ontario, Bill 194 passed in 2024; it focuses on strengthening cybersecurity and establishing accountability, disclosure, and oversight obligations for AI use across the public sector. In addition to legislation, some provinces have also released their own frameworks and principles to guide their use of AI, including Ontario and British Columbia.

Industry-specific regulators are also updating their guidelines and requirements. For instance, the Office of the Superintendent of Financial Institutions has released a draft update to its Model Risk Management Guideline (E-23). According to OSFI’s latest quarterly release, the finalized guideline is expected 11 Sept. If adopted in its current form, E-23 will set out enhanced expectations for how financial institutions manage model risk, explicitly extending to models that incorporate artificial intelligence and machine learning.

Additionally, several law societies across Canada, including those in Ontario, Alberta, Manitoba, Saskatchewan, and British Columbia, have released guidelines on the responsible use of AI in the legal profession.

To support these sectoral regulations, Canada is investing significant efforts in both domestic and international standards development for AI. As seen through the establishment of an AI and Data Standardization Collaborative, the federal government recognizes the role standards will play in establishing global norms and common best practices for the appropriate development and use of AI. Through the national standards body Standards Council of Canada, the federal government has played a significant role in the International Standards Organization's developments for AI. Specifically, it was one of the initial drafters of the ISO/IEC 42001 standard.

The Digital Governance Council is also a key player in setting AI standards in Canada. Through its accredited standards program, the DGC develops national guidelines for the responsible design, deployment and oversight of AI systems, helping organizations align with best practices in trust, safety and accountability.

Other guidance in AI and automated decision-making includes Health Canada's guidance document on using software as a medical device, the federal government's Guide on the use of generative AI for government institutions and the Office of the Privacy Commissioner of Canada's Principles for responsible, trustworthy, and privacy-protective generative AI technologies.

Wider regulatory environment

There are numerous enacted laws of relevance and application to various elements of the AI governance life cycle. The Personal Information Protection and Electronic Documents Act sets out important rules for how businesses use personal information. To modernize this law for the digital economy, the Consumer Privacy Protection Act was proposed as part of the Digital Charter Implementation Act, 2022. The government is also working to ensure laws governing marketplace activities stay current.

Additionally, several other frameworks apply to AI use, including:

- → The Canada Consumer Product Safety Act
- → The Food and Drugs Act
- → The Motor Vehicle Safety Act
- → The Bank Act
- → The Canadian Human Rights Act, and other provincial and territorial human rights laws
- → The Criminal Code

Data privacy and protection

The Digital Charter Implementation Act introduced the AIDA and would have overhauled the PIPEDA through the Consumer Privacy Protection Act.

Combining privacy and AI regulation makes sense because data is the key link between them. The CPPA was set to require organizations to explain any prediction, recommendation or decision made by an automated system that would have significantly impacted individuals. These explanations were expected to include the type of personal information used.

The CPPA was designed to include exceptions to consent for legitimate interest. However, it is unclear if this would have extended to the use of data to train AI systems. Under this exception, organizations would be required to identify and take reasonable measures to minimize adverse effects from using data for this purpose.

Copyright and intellectual property

The AIDA did not currently address copyright issues. Instead, it appears the government aims to tackle AI and intellectual property issues through an updated Copyright Act. In 2021, before the launch of many generative AI tools, Canada began consulting on updates to the Copyright Act. With rapid advancements in AI, especially generative AI, another public consultation was launched in December 2023.

The federal government aims to adapt the current copyright regime to address challenges posed by generative AI systems, which can produce creative content mimicking that created by humans. This raises concerns about the uncompensated use of protected works in training these AI systems, attribution and remuneration for AI-generated content, and enforcing the rights of copyright holders. Key discussion points of this consultation included text and data mining, authorship and ownership, and liability.

Consumer protection and human rights

Given the risks to human rights, including discrimination and federal, provincial and territorial human rights, laws play a crucial role in protecting individuals from AI-related harms. Redress and contestability mechanisms for discrimination, like those featured in Quebec's Law 25, are important. However, individuals affected by AI discrimination may be unaware it has occurred. In 2021, the Law Commission of Ontario, the Ontario Human Rights Commission and the Canadian Human Rights Commission launched a joint research and policy initiative to examine human rights issues in AI development, use and governance.

Regarding consumer protections, the Canada Consumer Product Safety Act and various provincial consumer protection laws address issues like misrepresentation and undue pressure while remaining technology neutral. Updates to Ontario's consumer protection legislation, Bill 142, provide insight into potential future changes. These updates include a technology-neutral approach but incorporates updates reflecting the current digital landscape. Key proposed changes include new provisions on automatic subscription renewals, unilateral contract amendments and easier mechanisms for consumers to unsubscribe from services. These amendments aim to enhance transparency and fairness in consumer transactions, especially those occurring online or through automated means.

Competition

The Competition Bureau of Canada is actively engaged in the discussion around the intersection of AI and competition. In May 2024, it published a discussion paper setting out considerations for how AI may affect competition. Key topics analyzed as a part of the paper include barriers to entry, product differentiation and market power, economies of scope and scale, network effects and competitive conduct, and consumer protection.

Agentic AI

Canada does not currently have legislation at the federal or provincial level that specifically regulates agentic AI. Given the potential broad applications of agentic AI, any potential future legislation on AI systems is expected to encompass these technologies within its scope.

At present, the government has developed the Voluntary Code of Conduct on the Responsible Development and Management of Advanced Generative AI Systems. To support public servants in their use of AI, the federal government's School of Public Service has released a course entitled "The Rise of Agentic Artificial Intelligence."

Future of AI governance in Canada

It is unclear whether the current government will retable AI legislation. Given the vigorous debate related to the AIDA and current geopolitical context, it is unlikely AI legislation would look the same as it did in 2022. However, trust continues to be a significant barrier to the public's adoption of AI in Canada. This remains a crucial factor to advancing AI adoption in any nation. This new government has provided a clear signal that AI and digital innovation are important, so it will be interesting to see how this shapes the Canadian AI governance landscape. Whether rules come in the form of additional support for industry-developed standards, top-down rules or sectoral-specific legislation developed by regulators remains to be seen.

Published December 2025

Global AI Governance Law and Policy: China

By Barbara Li

There is no doubt that China has become one of the world’s most important countries in artificial intelligence, driven by advanced innovations, robust investments and a wide range of AI applications.

China first introduced intelligent computing into its National Medium and Long-Term Technology Development Plan in 2006, laying the foundation for setting AI as a transformative technology. In 2015, the State Council of China released the national strategy of Internet Plus, identifying AI as one of the country’s strategic emerging industries. The strategy also set the goal of establishing China as a major hub for AI innovation by 2030. Following that national strategy, a comprehensive AI ecosystem has emerged. Major Chinese technology and internet companies have rapidly launched AI products and services across diverse fields.

On 27 Aug. 2025, the State Council of China issued the AI Plus Action Plan, which is widely regarded as the blueprint for the country's national AI strategy in the coming years. According to this plan, China will prioritize the use and deployment of AI in six areas: science and technology development, industrial utilisation, consumer services, public welfare, governance and security, and international collaborations. The country aims to achieve 70% AI penetration in key sectors by 2027 and 90% by 2030, with a vision of building a fully AI-powered economy and society by 2035.

Since 2021, China has introduced a series of detailed AI policies and regulations, reflecting a maturing environment that balances innovations with governance and data security. These frameworks include important AI regulations, industry standards, technical guidelines and court rulings that cover algorithms, deepfakes, generative AI, privacy, intellectual property protection, AI ethics and content labelling.

Approach to regulation

China has deliberately chosen an agile and adaptive approach to regulation, aiming to strike a balance between promoting and developing AI technologies and addressing and managing risks.

Rather than having a single and comprehensive AI law, lawmakers have initially chosen to focus on specific areas to establish the regulatory scheme. This targeted approach prioritizes high-risk or high-potential areas, such as generative AI, deep synthesis, algorithms, and AI labelling, establishing sectoral regulatory frameworks.

Although regulators across industries see AI as a key enabler for growth, the regulators in some industries have progressed faster in defining rules and governance structures. More sophisticated regulatory regimes have emerged in the financial, e-commerce, transportation, education, pharmaceutical and medical sectors.

Algorithms

The Administrative Provisions on Algorithm Recommendation for Internet Information Services, effective on 1 March 2022, marked China's earliest moves into AI regulation. These provisions apply to internet information service providers using algorithmic recommendation technologies for news feeds, blogs, short videos, chat rooms, online streaming, search results and other online services.

According to the algorithm recommendation provisions, service providers must disclose that algorithms are in use and allow users to opt out of algorithmic recommendation. The provisions underscore the essential requirements to ensure fairness and transparency; providers are prohibited from offering different prices or discriminating against users based on their personal characteristics in the context of algorithmic recommendation.

The algorithm recommendation provisions mandate service providers with public opinion attributes or social mobilization capabilities conduct risk assessments and file the algorithm with the Cyberspace Administration of China.

Failure to abide by the compliance requirements will lead to penalties, including investigations by regulators, administrative fines imposed on both the company and individuals in charge, business suspension, and, in the worst-case scenario, criminal liability.

As of October 2025, China has approved thousands of algorithm filings, reflecting a highly dynamic landscape of AI development and applications.

Deep synthesis technology

In November 2022, China released the Administrative Provisions on Deep Synthesis of Internet-based Information Services, effective beginning in January 2023. The goal was to better govern the development and adoption of deep synthesis technologies, delineating the requirements and prohibitions of deep synthesis services.

The deep synthesis provisions apply to the use of algorithms to synthetically generate or alter video, voice, text, image and other online content. Service providers are prohibited from using deep synthesis technology to produce or disseminate illegal information. They must establish the required mechanisms for user registration, algorithm review, ethics review, content monitoring, data security, personal data protection, fraud prevention, and emergency response.

Among others, these provisions mandate the appropriate labels be added to the content generated by deep synthesis technology. This requirement has been further substantiated in the AI Labelling Measures, discussed below, effective 1 Nov. 2025.

Generative AI

On 10 July 2023, the Interim Measures for Administration of Generative AI Services were issued and took effect on 15 Aug. 2023, making China the first country in the world with binding regulations for generative AI.

The generative AI measures broadly define the term to include models and technologies that can generate text, pictures, sounds, videos, codes and other content based on algorithms, models and rules. To strike a proper balance between encouraging AI innovations and addressing their security risks, the measures exclude research, development and the internal use of generative AI technologies from the compliance requirements.

However, service providers offering public-facing generative AI services are required to shoulder multiple compliance requirements, including legality and legitimacy of training data, monitoring content, upholding ethical and core social values, obtaining consent from individuals for use of personal data, protecting IP rights, maintaining transparency and accountability, preventing discrimination, and safeguarding cybersecurity and data privacy. International companies are expressly permitted to establish foreign-invested enterprises to develop and offer generative AI services in China, provided that such activities are allowed under China's foreign investment laws.

Similar to those providing algorithmic recommendation services, service providers offering generative AI services with public opinion attributes or social mobilization capabilities to their external customers must conduct security assessments and file their large language models with CAC. The mandatory filing for LLMs is required in addition to the algorithm filing under the Algorithm Recommendation Provisions.

Ethical review measures

Ethical considerations have always posed one of the central challenges in AI development for both businesses and regulators. On 8 Oct. 2023, the Ministry of Science and Technology, Ministry of Industry and Information Technology and other national governmental authorities jointly issued the Interim Measures for Ethics Review Measures, effective 1 Dec. 2023, requiring the ethical review of AI and other research and development activities in biological and medical fields.

To specifically address AI ethical complications, on 22 Aug. 2025, these governmental agencies jointly released the draft Measures for the Administration of Ethics for AI Technological Activities for public consultation.

The draft ethics measures apply to all R&D activities in China that may affect health and safety, reputation, the environment, public order and sustainability. Developers and service providers must adhere to the principles of fairness, accountability, justice, risk responsibility and respect for life and human dignity. AI projects falling within the application scope of these measures must undergo ethics review, either internally by ethics committees or externally through qualified external centers. They cannot proceed with the related AI services before the ethics review is complete.

Regulators are also preparing a detailed list of high-risk AI activities. Businesses are advised to keep a close watch on further developments.

AI labelling

The most recent legislative piece in China's AI governance bucket is the Labelling Measures for AI Generated Content, which took effect on 1 Sept. 2025. On the same date, the mandatory technical standard on AI content labelling, GB45438-2025, also became effective. Both the AI labelling measures and technical standards provide much-needed clarity and best practices for businesses to refer to when handling AI-generated content.

The labelling measures have a wide application to internet service providers that use AI to generate text, audio, video, images, virtual scenes and other content. Visible labels with AI symbols are required for chatbots, AI-written content, synthetic voices, face generation/swap and immersive scene creation or editing. Explicit labels must remain embedded with the file where AI-generated content can be downloaded, reproduced, or exported.

Other AI-generated content can use implicit labels, such as watermarks or other symbols, that are added to the data files via technical measures but are not easily perceived by users. When an implicit label is added to the metadata of the AI-generated content, the label should include key information such as content attributes, name or identifier of the service provider, and the content reference number.

Internet platforms must act as watchdogs. If they detect or suspect AI-generated content, they must alert users and may add implicit labels themselves. Non-compliance carries serious consequences, including regulatory investigations, fines, business suspensions and revocation of business permits. In severe cases, criminal liability under the Cybersecurity Law, Data Security Law and Personal Information Protection Law may be triggered.

Agentic AI

While the government has not addressed agentic AI in any specific binding law, the broader regulatory instruments that govern things like recommendation algorithms, automatic decision-making and generative AI are applicable to the development of agentic AI as well. This means that companies providing agentic AI products and services are required to conduct impact assessments, follow ethics rules, maintain content monitoring and exercise proper human oversight.

If the agentic AI products and services have public opinion attributes or social mobilization capabilities, a regulatory filing is required. Furthermore, standards and other soft-law mechanisms have been released. These include the "Technical Application Requirements for Intelligent Agents in Software Engineering – Part 1: Development of Intelligent Agents," which emphasizes technical and service capabilities of agentic models. Likewise, a recommendation purportedly led by the Chinese Academy of Information and Communications Technology, titled "ITU-T F.748.46 Requirements and Evaluation Methods of Artificial Intelligence Agents Based on Large Scale Pre-Trained Model," sets out standards for evaluating the performance of AI agents.

Wider environment and recent development

China has a robust privacy and cybersecurity legal framework that applies to AI use cases. Furthermore, China's judiciary is grappling with the issue of how copyright law applies to AI-generated works.

Privacy and cybersecurity

China's legal regime on data privacy and cybersecurity is built on the cornerstone of three national laws: CSL, DSL and PIPL. As China has not yet enacted a unified AI law, these statutes apply to AI activities. This means that, where applicable, AI developers and service providers must abide by the compliance requirements imposed by these three national laws, including without limitation, obtaining consent from data subjects before using personal data as training data, following the legal mechanisms for cross-border data transfer, and conducting impact assessments for using AI in decision-making processes, among other things.

It is important to note that on 28 Oct. 2025, China's top legislature passed major amendments to the CSL. The CSL amendments add new provisions on AI, bringing AI into China's national law for the first time. The amendments make it clear that China will support the R&D of algorithms; promote the construction of training data resources, computing power, and other AI infrastructures; and expedite rulemaking for AI ethics while firming up AI risk assessment and security governance.

The new CSL amendments will take effect on 1 Jan. 2026. Chinese regulators are anticipated to issue further detailed rules for implementation of these new amendments.

From a technical perspective, China's National Network Security Standardization Technical Committee issued the AI Governance Framework on 9 Sept. 2025, outlining principles and guidelines for governance and risk management of AI technologies.

The AI Governance Framework classifies AI risks into inherent risks and application risks. These include concerns related to LLMs and algorithms, ethical issues, bias and discrimination, contamination of training datasets, data breaches and IT vulnerabilities, criminal and illegal uses of AI, and risks within the supply chain.

The AI Governance Framework recommends adopting organizational and technical measures to address these risks. Suggested steps include ensuring transparency of AI algorithms; protecting IP rights, personal data, and privacy; enhancing AI supply chain security; implementing cybersecurity controls; classifying and grading data and prompts; ensuring traceability of AI applications; filtering and verifying AI outputs to avoid discrimination; and promoting talent development.

AI and copyright

China's courts have been the front-runners to explore how the traditional framework of copyright law applies to works generated with the assistance of AI tools. The courts are never shy about addressing critical questions, such as when AI-generated content can qualify as a "work" under the Copyright Law of the People's Republic of China and who owns the rights. In the past two years, the courts have given some ground-breaking rulings.

One landmark ruling was decided by the Beijing Internet Court in November 2023. In that case, the plaintiff used Stable Diffusion, an AI tool, to generate images from text. The court found that the plaintiff had invested meaningful human creativity by selecting prompts, adjusting parameters and selecting the final image. All these efforts, in the court's opinion, met the originality requirement under the law, and thus the AI-generated image qualified as a copyrightable work.

Similarly, in March 2025, the Changshu Court in Jiangsu province ruled in favour of copyright protection for an image generated by Midjourney and subsequently edited via Photoshop. The court ruled that the user had engaged in prompt selection and editing, resulting in sufficient originality.

However, some courts have taken a stricter line. The judges of the Zhangjiagang Court in Jiangsu province dismissed a claim for copyright protection on works generated with AI because the human author could not provide substantial evidence of creative input. The user's reliance on prompts alone, without meaningful arrangement or editing, failed the originality threshold.

These judicial developments show a remarkable trend that China's jurisprudence is adopting a balanced judicial stance along with lawmakers and regulators. On one hand, courts will grant copyright protection when human creativity is identifiable; on the other hand, the courts will closely scrutinize the human elements. If the human contribution is minimal or the output is primarily machine-driven, the courts will deny the right.

Enforcement

Regulators have stayed active in the enforcement of AI regulations. There have been multiple rounds of enforcement campaigns jointly conducted by CAC, MIIT and other governmental agencies. These efforts primarily target non-compliant activities such as failure to conduct the mandatory LLM and algorithm filings, dissemination of misinformation, violations of the PIPL when providing AI services, and insufficient organizational and technical measures to protect against cybersecurity incidents or data breaches.

With multiple new laws and regulations related to AI taking effect now or in the near future, stronger enforcement and penalties are expected in the coming months. It is crucial that businesses analyze the impact of AI regulations and carefully design and review their strategies for China's market. Prompt action to ensure compliance is equally critical.

Published October 2025

Global AI Governance Law and Policy: EU

By Isabelle Roccia, Laura Pliauškaitė, Will Simpson and Sethu S Raman

The EU has been a global first mover in adopting comprehensive digital legislation, for example, with the General Data Protection Regulation in 2016 and the EU Artificial Intelligence Act in 2024. While the GDPR arguably set a global privacy standard, creating what has been deemed the "Brussels Effect" and spurring comparable privacy laws in places like the U.K., Brazil, and many U.S. states, the AI Act is less certain to guide the global approach to regulating AI in light of the many different regulatory models emerging. Even acknowledging the heterogeneous AI governance global landscape, the AI Act remains hugely impactful.

History and context

The EU was the first jurisdiction globally to adopt a comprehensive legal regime governing the development and use of artificial intelligence. The AI Act entered into force 1 Aug. 2024 with a phased application and is the result of years of preparatory work.

The process of making this regulation a reality started in 2018 as the European Commission set out its vision for AI around three pillars: investment, socioeconomic changes and an appropriate ethical and legal framework to strengthen European values. These pillars still support the overall EU AI strategy across Europe internally and globally as Brussels ascertains its approach to shaping ethical and safe AI innovations.

The AI Act is also part of a broader and consistent approach to digital policy rulemaking by Brussels, cutting across data, infrastructure and online rights policies anchored in European values. It builds on an intricate regulatory and legislative landscape and by extension the compliance architectures that organizations have already put in place across data governance domains. In 2025 and beyond, the focus of AI Act implementation rests increasingly on its integration with the many regulatory tools with which it intersects.

Approach to regulation

The first deadline of the phased implementation of the EU AI Act kicked in 2 Feb. 2025 with the application of prohibitions on unacceptable risk AI and AI literacy requirements. On 2 Aug. 2025, obligations began for providers of general-purpose AI, for member states to appoint designated competent authorities, and the Commission launched its first review of the prohibited-AI list.

GPAI Code of Practice

The European Commission published the final version of the General-Purpose AI Code of Practice in July 2025, ahead of the formal deadline though after its earlier self-imposed deadline. The voluntary tool is meant to help providers of GPAI models demonstrate compliance with obligations in Articles 53 and 55, with a focus on transparency, copyright, and safety and security, provided they sign on to the code in its entirety.

The code took months of drafting, leveraging the expertise of over 1,000 independent representatives from industry, civil society, academia and member states. It went through several iterations to clarify and consolidate a complex table of content.

A week after it was finalized, the code was officially opened for signatures, with several companies — such as Aleph Alpha, Cohere, IBM, MistralAI and OpenAI — joining in shortly after. The Commission announced that it will narrowly focus its enforcement against signatories on the demonstration of their compliance, because adherence to the code entails increased transparency on the part of providers.

Separately, the European Commission launched in September 2025 a consultation for the development of another code of practice on AI transparency obligations, which becomes applicable on 2 Aug. 2026. The Commission will also rely on stakeholders to develop the Code of Practice, though it has not confirmed the expected date of publication.

Enforcement and regulators overiew

EU member states were tasked with designating or establishing national competent authorities — at least one market surveillance and one notifying authority per country — by 2 Aug. 2025 and identifying which one of them would act as a single point of contact. As of mid-September, about a third of member states had met the deadline.

Looking at structures that have been either formalized or announced leading up to the 2 Aug. deadline, no single governance model has emerged.

Member-state designs vary from decentralized, sector-inflected networks in some; to more centralized models in others. Communications and cybersecurity regulators often anchor coordination; DPAs are key participants but not always leads. Spain and Hungary are setting up new AI-specific bodies/authorities. Hungary will constitute a new market-surveillance authority under the act while Spain's AI Supervisory Agency is expected to act as the single point of contact.

The diversity of this pan-European architecture brings organizations and regulators alike into uncharted territory. This will have strong implications for both groups of stakeholders. Organizations should map their lead regulators and establish early engagement routes to build these relationships; agencies and regulators must find common language and interpretation of the law, and ways to cooperate and coordinate their action.

Guidelines

To facilitate the EU AI Act's implementation, the European Commission has been tasked with developing guidelines on various provisions of the act. Over the past year, it released guidance on topics such as prohibited AI practices and AI system definition.

The Commission still has a long list of deliverables to produce, including the interplay of the EU AI Act with other EU legislation. It is drafting guidelines on classifying high-risk AI systems, informed by input on both practical examples and specific high-risk AI issues, and on responsibilities along the AI value chain.

Kick-starting compliance

Launched in 2024 as a voluntary framework by the European Commission, the AI Pact was meant to help organizations kick start their compliance journey ahead of the official application of the AI Act. The pact is built on two pillars, one focused on knowledge sharing between all interested stakeholders and the other directed at providers and deployers specifically for them to share practices and demonstrate voluntary commitments to prepare for the early implementation of the AI Act. Since its launch, the AI Office hosted a series of AI Pact webinars to explore topics such as the architecture of the AI Act, AI literacy obligations, and the GPAI Code of Practice.

In addition, the EU AI Office launched an AI Act Service Desk as a central initiative to help stakeholders navigate the AI Act's requirements. As part of the Service Desk, the AI Act Single Information Platform provides the AI Act Explorer, an online tool to navigate the AI Act legal text; a compliance checker to assist in evaluating whether AI systems and general-purpose AI models meet the requirements set by the AI Act; and a portal to national resources.

European AI Strategy

In April 2025 the European Commission launched the AI Continent Action Plan to transform Europe into a leading AI continent. This initiative defines the EU's current approach to AI and includes several actions to boost trustworthy and human-centric AI development and, as a result, enhance the EU's innovation and competitiveness while ensuring that democratic values and fundamental rights are safeguarded. These actions fall under five strategic areas: computing infrastructure, data, skills, development of algorithms and adoption, and simplify rules.

Computing infrastructure

The Commission is scaling computational power through AI Factories and planned AI Gigafactories — EuroHPC-linked supercomputers, data centers and talent pipelines open to startups, SMEs, research and public users. At least 13 AI Factories are slated to be operational by 2026, with up to five AI gigafactories to follow, all offering access to European users across industry, research and the public sector.

A proposed EU Cloud and AI Development Act aims to triple cloud capacity over 5–7 years and may add security and data-localization requirements for critical workloads. The proposal for this is currently expected in March 2026.

Data

A European Data Union Strategy would simplify rules, expand access via Common European Data Spaces and establish Data Labs within AI Factories; a Commission Communication is due this year

Skills

The EU will develop and attract AI talent through an AI Skills Academy, European Digital Innovation Hubs and links to (giga) factories and research programs.

Development of algorithms and adoption

The Apply AI Strategy published on 8 October by the Commission aims to accelerate AI development, adoption and use, across the EU's strategic and public sectors, such as health care, pharmaceuticals, manufacturing, construction and defense. It promotes European AI solutions and encourages organizations to adopt an "AI first" policy.

Simplify rules

Streamlining documentation, record keeping and incident/information-sharing obligations across digital policies are being considered to ease AI Act implementation without weakening core safeguards.

Funding and Investment

The EU will need to pool enough funding to support these ambitions. It launched the InvestAI Initiative at the beginning of this year to mobilize 200 billion euros of investment in AI. The funding is planned to come from different sources, including existing EU funding initiatives, such as Horizon Europe, Digital Europe, but also private investment.

Agentic AI

Agentic AI, autonomous, adaptive systems, raise novel risks but fall squarely within the AI Act's technology-neutral, risk-based approach.

High-risk agentic uses such as employment, education, and law enforcement must meet Chapter III obligations — risk management, data governance, documentation, quality management, post-market monitoring — and ensure effective human oversight (Article 14).

Unacceptable-risk behaviors such as manipulation and exploitation of vulnerabilities are prohibited under Article 5. Agentic AI that is also GPAI entails additional provider obligations, e.g., technical documentation, and cooperation with authorities.

Because agentic systems can self-update, risk levels may evolve — underscoring continuous monitoring and in-life change control.

Many agentic uses will also trigger the GDPR, including automated decision-making limits and core data-protection principles, and, for connected devices, the Data Act's access and interoperability rules.

The bottom line is that many EU digital laws are technology-neutral and designed to adapt to future technological innovation. Although agentic AI may not be mentioned by name within a law, the scope of the law may still capture this new application of machine learning.

Wider regulatory environment

The AI Act and broader European AI initiatives fold into a comprehensive policy agenda that cuts across digital responsibility domains. Several policy areas could be subject to legislative changes, which would have implications for the AI legislative framework. Among others, two would have big impacts.

AI liability remains a question mark on the Commission's agenda. After proposing a dedicated Directive in 2022, it was withdrawn early this year due to lack of consensus on the general direction to take. It remains to be seen whether and how this topic will be addressed at the EU level.

The EU copyright framework is also in question, prompting a discussion on its necessary review and update to reflect technological developments, particularly in AI.

Future of AI governance in EU

The EU AI Act's implementation requires the development of standards that will translate the act's legal requirements into technical requirements. In April, CEN-CENELEC, an organization tasked with standard development, noted possible delays in the delivery of standards.

Over the summer, public debate has mounted with some stakeholders calling to postpone the AI Act's remaining implementation deadlines. European Commission Executive Vice-President Henna Virkkunen expressed in June the possibility of postponing "some parts of the AI Act in the coming months," though firmly brushed off any equivocating on the implementation itself.

This pressure comes at a time when AI is caught in trade and political friction between Brussels and Washington. The complex global geopolitics surrounding the trans-Atlantic relationship at this moment have also prompted the EU to reaffirm both its sovereignty to design its own rules as well as its ambitions to promote transparent, accountable, and human-centric AI.

Published July 2025

iapp | HCLTech

Global AI Governance Law and Policy: India

By Rahul Matthan and Sanah Javed

From a single computer in Kolkata's Indian Statistical Institute in the 1950s to powering the world's digital infrastructure, India's technical journey has been tremendously transformative. In the 1990s, the government focused on the 'information technology' sector by encouraging the export of software services. This resulted in India eventually emerging as a global IT hub first as an outsourcing destination and eventually as the birthplace of digital public infrastructure that offers a new form of governance using population-scale digital architecture. Today, India has the world's largest digital identity system, the biggest digital payments system by volume and a population that is, for the most part, digital by default.

This dramatic digital transformation has naturally fuelled the adoption of AI. According to the Stanford Artificial Intelligence Index Report 2025, India ranks second in the list of countries with the highest AI skill penetration from 2015 to 2024 and is among the top 10 countries in the world that received the most private investment in AI from 2013 to 2024.

India's AI policy has two different facets. One set of policy initiatives focuses on promoting AI adoption amid rapid advances in generative AI. The government has pledged to invest USD1.25 billion in AI development, and, to that end, launched the IndiaAI Mission. Additional initiatives promote the integration of AI across a range of use cases and sectors. The other set of policy initiatives relates to the governance of AI and the risks that it could pose. While India currently does not have standalone AI legislation, existing intellectual property, data protection, cybersecurity and content regulations are being adapted to apply to AI.

Historically, India's digital space was exclusively regulated by the Information Technology Act, 2000, an omnibus law that addresses online contracts, data protection, cybercrimes and digital harms, such as phishing and identity theft. The IT Act has undergone numerous amendments to respond to new threats in the digital space. Originally intended to regulate computers and electronic records, it has been used to regulate a range of digital products and services due to the expansive definition of computer resources. AI systems and models will likely fall within the purview of its legislative scope.

Various subordinate legislations were enacted under the IT Act, including the IT (Reasonable Security Practices and Procedures and Sensitive Personal Data or Information) Rules, 2011, which governed how entities should process personal data. Until they were replaced by the Digital Personal Data Protection Act, 2023, these rules served as the primary data protection regulation. Other laws, such as the Bharatiya Nyaya (Second) Sanhita, 2023, and the Indian Copyright Act, 1957, broadly extend to the digital space and also apply to AI.

Approach

In 2018, India's National Institution for Transforming India, the government's apex policy think tank, released the National Strategy on Artificial Intelligence. The policy took an "AI for all" approach and aimed to address challenges of accessibility, affordability and skilled expertise.

The NSAI set out four key areas. One part concentrated on research to boost core and applied research in the AI field. Another focused on reskilling the current workforce to facilitate large-scale employment generation through AI; this initiative called for realigning the education sector to harness the potential AI. A third priority promoted investments in AI and product development that would enhance AI adoption. Finally, there was a focus to manage concerns around ethics, privacy and security, such as through the use of privacy preserving technologies.

In 2021, the IndiaAI Mission was launched to develop a comprehensive ecosystem to foster AI innovation by democratising access to compute, enhancing data quality, developing indigenous AI capabilities, attracting top AI talent and promoting ethical AI. The comprehensive national-level program houses initiatives like the IndiaAI Compute Capacity, which intends to scale AI computing infrastructure by deploying over 18,000 graphics processing units through strategic public-private partnerships.

The Ministry of Electronics and Information Technology indicated it was looking to replace the IT Act with new legislation in 2023, tentatively titled the Digital India Act. Since the Indian government is unlikely to introduce a standalone legislation on AI, it is likely the Digital India Act will also apply to potential AI risks.

The MeitY issued an AI advisory addressed to all intermediaries and platforms on 15 March 2024. It required them to ensure that their use of AI models, large language models and generative AI do not make it possible for users to share unlawful content on the platform. Additionally, they were required to ensure the use of AI technology does not result in any bias or discrimination or threaten the integrity of the electoral process.

All platforms and intermediaries were obliged to test their AI models; if they were unreliable in any way, intermediaries and platforms had to appropriately label them as such. Users had to be informed through terms of service or user agreements that their user account could be terminated if they were dealing with unlawful information.

The MeitY has recently constituted a subcommittee to examine whether the IT Act, the Bharatiya Nagarik Suraksha Sanhita, 2023, various content laws, the Information Technology (Intermediary Guidelines and Digital Media Ethics Code) Rules, 2021, cyber security laws, DPDPA and India's intellectual property regulations are sufficient to regulate AI. Since AI spans multiple sectors, the view was that a fragmented regulatory approach may result in inefficiencies. To address this concern, the committee recommended an integrated government approach, calling on the MeitY and principal scientific adviser, who advises the prime minister on matters of science and technology, to establish an empowered mechanism to coordinate on AI governance.

While there is no official guidance on how existing intellectual property laws apply to AI, the NSAI notes that the IP regime must enable AI developers to innovate. The policy, therefore, has recommended the creation of a task force comprised of the Ministry of Corporate Affairs and the Department for Promotion of Industry and Internal Trade to consider the issue.

The Consumer Protection Act, 2019, offers broad protections to consumers that could extend to AI harms. The government recently issued the "Guidelines for Prevention and Regulation of Dark Patterns", prohibiting the use of dark patterns on the grounds that they amount to deceptive advertising.

Sector-specific AI regulations

Various sectoral regulators are looking to regulate AI within the specific domains they supervise. Some key initiatives are listed below:

Financial sector

The Reserve Bank of India developed the Framework for Responsible and Ethical Enablement of Artificial Intelligence committee to study AI adoption by financial institutions and review AI regulation in the context of the global financial sector. The committee is yet to release its report.

The Securities and Exchange Board of India now requires that all registered mutual fund offerings using AI or machine learning applications and systems file a report with the board on a quarterly basis. These reports will detail how the AI/machine learning project is implemented, the safeguards put in place to prevent abnormal behavior of the AI/machine learning system, and if there are key controls in the AI/machine learning system in accordance with SEBI's cyber security control requirements.

Health care sector

The Indian Council of Medical Research released the Ethical Guidelines for Application of AI in Biomedical Research and Healthcare in 2023. Some of these guidelines include ensuring a "Human in the Loop" model, minimizing risk by implementing safety standards, securing data privacy and protection, defining accountability and liability for AI actions and promoting accessibility, equity and inclusiveness.

Under the Telemedicine Practice Guidelines, 2020, any technology platform based on AI or machine learning is prohibited from counselling patients or prescribing medication. However, the technology could be used to assist and support a registered medical practitioner in carrying out patient evaluation, diagnosis or management.

Telecommunication sector

The Telecommunication Engineering Centre, a technical wing of the Department of Telecommunication, has recognized that AI enables real-time decision making and will significantly influence upcoming technologies, such as satellite broadband, drone communication and the metaverse. In this context, TEC published a report on AI system fairness and invited stakeholder input on developing new standards to assess and rate the robustness of AI systems in telecom networks and digital infrastructure.

Defense

The Artificial Intelligence in Defence report sets out a risk-based assessment framework to integrate AI applications into defense operations. In this report, the Indian Department of Defence recognizes the wide array of AI applications that are critical to the defense sector.

Cybersecurity

The Indian Computer Emergency Response Team is the national nodal agency for responding to cybersecurity incidents. It has previously released an advisory on the security implications of using AI language-based applications.

Foundation models

India has developed indigenous foundation AI models, both large and small language models. India's first indigenous large language model, Sarvam-1, is trained on datasets in other languages apart from English to perform multilingual tasks; BharatGen is a government-funded AI-based LLM for Indian languages.

According to the MeitY subcommittee, multiple stakeholders are involved in the lifecycle of a foundation model, such as data principals, data providers, AI developers including model builders and AI deployers including app builders and distributors. Accordingly, it is critical that the distribution of responsibilities between different players is clear.

Agentic AI

While the Indian government has not specifically addressed agentic AI in any of the released policies and guidelines, the larger principles on responsible and trustworthy AI are likely to be used to apply to the development of agentic AI as well. Various tech companies in India have deployed AI agents to optimize their logistics services. Infosys Limited, an Indian multinational technology company, has developed generative AI agents for client applications.

Enforcement

Presently, AI-related enforcement is carried out under existing legal frameworks. For instance, the BNSS outlines offenses related to cybercrimes, creation and dissemination of deepfakes and AI-generated misinformation, impersonation-based cheating and privacy violations — particularly when deepfakes exploit a person's image. These are common risks that may arise from the use and deployment of AI systems and models.

The privacy rules and DPDPA address any violations of an individual's privacy rights. Accordingly, any AI system that uses user data would need to ensure compliance with the requirements under these laws. Similarly, since AI systems and models are trained on large datasets that may contain copyrighted material or other intellectual property, they must comply with the ICA and other intellectual property laws.

Latest developments

The Digital India Act has not been released for public consultation and it is unlikely that any AI law will materialize this year.

An Inter-Ministerial AI Coordination Committee/Governance Group is likely to be set up to develop a common roadmap and government approach to regulate AI. Some of the key goals of this committee would aim to strengthen existing laws to minimize AI-related risks and harm, harmonize existing efforts and initiatives around common technologies, provide legal clarity on the development and use of AI and create a policy environment that enables the use of AI for beneficial use-cases.

In parallel, various developments are underway to address sector-specific considerations and challenges posed by AI, such as the RBI's FREE-AI committee report, TEC's standards on AI, and policy frameworks released by the healthcare and defense sectors. The government is also deliberating on how existing legal frameworks, such as copyright law, pose challenges to the advancement of AI.

Published November 2025

iapp | HCLTech

Global AI Governance Law and Policy: Japan

By Takaya Terakawa and Jorge Sanz

Japan's approach to an artificial intelligence policy is both unique and strategically nuanced. While firmly committed to the shared democratic values upheld by jurisdictions like the U.S. and Europe, Japan has pursued mutual respect and interoperability grounded in a recognition of diverse global values.

While the policy framework might initially suggest prudence, its substance is decidedly proactive and firm. It reflects a deliberate, strategic orientation toward becoming "the most AI-friendly country in the world" — a compelling vision this article will explore in depth, tracing its historical development.

Japan has consistently positioned itself as a global leader in promoting the practical application of AI. This initiative formally began in 2016 with the government's introduction of "Society 5.0" within its Fifth Science and Technology Basic Plan. Society 5.0 envisions a future that achieves concurrent economic advancement and the resolution of critical social challenges through a deep integration of cyber and physical spaces.

At its core, the concept aims to create a human-centered society where big data, leveraged by AI, delivers tailored information to individuals. Robotics also play a key role by addressing pressing issues like labor shortages, regional depopulation, and demographic aging. Ultimately, this vision seeks to foster a society where people of all generations can live with mutual respect and dignity.

Building upon the Society 5.0 vision, the government released the "Social Principles of Human-Centric AI" in 2019. These principles articulate three foundational values: dignity — respecting human worth; diversity and inclusion — enabling all individuals to pursue happiness; and sustainability — ensuring long-term viability and well-being. In alignment with these core values, the government subsequently issued a series of supportive guidance documents, including the "AI R&D Guidelines," the "AI Utilization Guidelines," and the "Governance Guidelines for the Implementation of AI Principles."

A core component of Japan's AI governance framework is agile governance, which emphasizes continuous, rapid iteration using the plan-do-check-act cycle to ensure adaptive and responsive policy development. This concept was further elaborated in the 2022 report, "Agile Governance Update: How Governments, Businesses and Civil Society Can Create a Better World By Reimagining Governance." Demonstrating this adaptive spirit, in 2024, Japan consolidated its previously dispersed guidelines into a comprehensive document, the "AI Business Operator Guidelines," which was subsequently revised in 2025 to reflect evolving technological and societal needs.

Approach to regulation

Japan's approach emphasizes agility and pluralistic interoperability. In May 2025, Japan enacted the Act on Promotion of Research and Development, and Utilization of AI-related Technology, also referred to as the AI Promotion Act, which came into full effect in September. Unlike the EU's AI Act, which emphasizes regulatory oversight, the AI Promotion Act is designed primarily to support and accelerate the development and deployment of AI technologies. It represents a proactive and facilitative approach to AI governance.

The act is underpinned by four fundamental principles. First, it aims to enhance AI research and development capabilities and strengthen international competitiveness. Second, the act promotes comprehensive and strategic efforts by all stakeholders across the entire AI lifecycle. Third, it enables transparency in the development and use of AI while implementing necessary measures to mitigate associated risks. Finally, the act commits to taking a leading role in international cooperation on AI governance and standards.

Consistent with Japan's broader commitment to agile governance, the AI Promotion Act establishes a framework for implementing a continuous plan-do-check-act cycle to realize these principles.

Plan

The Artificial Intelligence Strategy Headquarters, chaired by the prime minister, is responsible for formulating the AI Basic Plan that outlines national strategies and priorities.

Do

Implementation is carried out by a wide range of stakeholders, including national and local governments, research institutions, private enterprises, and the general public. Private sector actors are expected to cooperate with government initiatives, with their responsibilities defined as "best-effort obligations" rather than legally binding mandates. However, violations of existing laws — such as the Act on the Protection of Personal Information, Copyright Act, or other sector-specific regulations — remain subject to legal penalties. Under Japan's copyright law, the act of searching for and collecting content, like news articles from websites, without permission for use in AI responses could be judged as copyright infringement. This is based on the principle that such usage may "unfairly harm the interests of the copyright holder." This legal risk was made clear in 2025 when newspaper companies filed related lawsuits, such as the Yomiuri Shimbun lawsuit against Perplexity. Thus, it would be inaccurate to claim that Japan's AI governance framework lacks enforceable provisions.

Check

The national government is tasked with monitoring domestic and international trends in AI research and applications. It also investigates cases where AI technologies are developed or used for illicit purposes or in inappropriate ways that may infringe on citizens' rights. Based on these analyses, the government formulates countermeasures and provides guidance, advice, and information to relevant parties. The AI Safety Institute and the G7 Hiroshima AI Process reporting framework are expected to play a key role in these promising initiatives whose effects will be seen over the next decade.

Act

This phase involves revising laws and guidelines, as well as updating the AI Basic Plan to reflect new insights and developments.

The AI Promotion Act is thus positioned as a foundational element in Japan's broader effort to build the societal infrastructure necessary for realizing the Society 5.0 vision — a human-centered, technologically advanced society that harmonizes innovation with inclusivity and sustainability.

Other laws, regulations and guidelines

Among the various legal reforms accompanying Japan's AI advancement, the most significant changes have occurred in the realm of copyright law. The 2018 amendment to the Copyright Act introduced a groundbreaking provision (Article 30-4) that permits the use of copyrighted works for purposes of information analysis, including AI development and training, without requiring prior authorization from rights holders, provided the use is not intended to replicate the work's expressive content.

Demonstrating the limits of this safe harbor provision, the government formally requested that OpenAI refrain from copyright infringement after its Sora 2 text-to-video model generated the likenesses of copyrighted anime and video game characters. Furthermore, in the generation and utilization phase of AI, the standard rules of copyright infringement apply, meaning that AI-generated content is judged under the same criteria as human-created works. Relatedly, government agencies have launched GENIAC, an initiative to raise the level of foundation model development capabilities in Japan.

In June 2025, the government adopted a Basic Policy on the Ideal Data Utilization System, signaling a shift toward more flexible use of personal data. The policy outlines a proposed amendment to the APPI, aiming to allow the sharing of personal data, including sensitive data, with third parties for purposes, such as statistical analysis and AI development, without requiring individual consent. This move is intended to facilitate innovation while maintaining appropriate safeguards.

As previously noted, support for businesses implementing AI governance includes the publication of the AI Business Operator Guidelines, which consolidate key principles and best practices. In addition, the government has released a "Contract Checklist for AI Use and Development," designed to facilitate the responsible and effective deployment of AI by providing practical guidance on drafting and reviewing contracts related to AI technologies.

Turning to the public sector, the government has taken proactive steps to institutionalize AI governance and promote its responsible use within administrative operations. In this context, the Guidelines for the Procurement and Utilization of Generative AI for the Evolution and Innovation of Administration were established. These guidelines define the government's approach to AI promotion, governance, procurement, and utilization, laying the foundation for a comprehensive framework that enables the active and responsible adoption of AI technologies across government functions.

On the public sector front, the government has taken proactive steps to institutionalize AI governance and released the Guidelines for the Procurement and Utilization of Generative AI in Public Administration. These guidelines establish a framework for the responsible adoption of generative AI in government operations, focusing on risk management, transparency, and ethical use.

A notable innovation in this space is the development of Gennai platform, a government-backed generative AI system designed to enhance administrative efficiency and service delivery. Generative AI has already begun to be deployed within various ministries and agencies, supporting administrative tasks and enhancing operational efficiency. This marks a significant step toward the digital transformation of public services and reflects Japan's commitment to leveraging AI not only in the private sector but also as a strategic asset in governance.

The Hiroshima AI Process

The Hiroshima AI Process, launched in 2023 under Japan's G7 presidency, marked an important milestone for global AI governance coordination. A central component of this initiative is the Hiroshima AI Process' Comprehensive Policy Framework, which was developed to promote the safe, secure, and trustworthy deployment of advanced AI systems. This international framework includes guiding principles and a code of conduct aimed at fostering responsible AI innovation across borders.

Building on this, the HAIP introduced a reporting framework in 2024 to enhance transparency and accountability in AI system development. Operationalizing under the auspices of the Organisation for Economic Co-operation and Development starting in 2025, the framework allows AI developers to voluntarily respond to a standardized set of questions about their practices. These responses are made publicly available, enabling companies to benchmark their efforts against peers and build essential trust with users and stakeholders.

By encouraging voluntary participation and public disclosure, the HAIP not only promotes global best practices but also strengthens international collaboration toward a shared vision of trustworthy AI.

Agentic AI

While the government has not specifically addressed agentic AI in any of the released policies, guidelines or laws, the broader principles on responsible and trustworthy AI are likely to be used to apply to the development of agentic AI.

Conclusion

Japan's approach to AI policy carves out a distinct path. It does not pursue AI supremacy like the policy proposed by the U.S. government in 2025, nor does it aim for the broad, high-risk regulation adopted by the EU AI Act and related relevant legislation. Instead, Japan's policy is rooted in the Society 5.0 vision: creating an environment that ensures the safety and security of its citizens and realizes a society where diverse happiness can be achieved for every individual.

Under this vision, the use of intrusive AI is naturally unacceptable as it threatens public safety and security. Similarly, unfair bias hinders the realization of individual happiness and is therefore also prohibited. Undesirable AI applications will be addressed using Japan's existing laws and regulations.

In summary, while Japan's approach to AI regulation differs from that of Western countries, its objective remains clear and consistent: to establish a robust framework that safeguards individual rights and freedoms while charting a path toward a desirable future society. Rather than imposing rigid constraints, Japan's legal and policy architecture seek to enable innovation through trust, transparency, and inclusive governance.

Published July 2025

Global AI Governance Law and Policy: Singapore

By Darren Grayson Chng

Singapore has solidified its position as a global leader in AI governance, balancing innovation with ethical considerations through a flexible, principles-based approach. Tortoise Media's Global AI Index continues to rank Singapore in third place, below the U.S. and China, in respect of level of investment, innovation and implementation of AI. The publication describes Singapore as "Asia's most dynamic AI hub after China" and having "made big advancements on absolute metrics, especially on AI research and development."

History and context

AI governance in Singapore began with three initiatives in 2018. First, an Advisory Council on the Ethical Use of AI and Data was appointed. This led to the publication of a discussion paper on the responsible development and deployment of AI. Finally, the government supported the launch of a research program on the governance of AI and data use.

In 2019, the Singapore government published its first National AI Strategy, outlining plans to drive AI innovation and adoption across the economy to generate significant social and economic impact. The Smart Nation and Digital Government Group, which operates within the prime minister's office and is administered by the Ministry of Digital Development and Information, designed and oversaw this strategy. The NAIS represented a whole-of-government approach to advance and oversee AI development and governance.

An updated strategy, NAIS 2.0, was launched in December 2023 to address recent challenges and uplift Singapore's economic and social potential over the next three to five years. NAIS 2.0 seeks to achieve two goals: develop areas of excellence in AI that advance innovation and maximize economic impact, and empower individuals, businesses and communities to use AI with confidence, discernment and trust.

This updated strategy emphasizes that the government will support experimentation and innovation while ensuring AI is developed and used responsibly and lawfully within existing safeguards.

Action 13 is a crucial element of NAIS 2.0; it requires the government to regularly review and update its frameworks to consider updates to broader standards and laws to support effective use of AI and reflect emerging principles, concerns, and technological advancements. This demonstrates Singapore's style of governance — agile and adaptive.

The following key events illustrate Singapore's AI governance journey over the years:

2018

The Advisory Council on the Ethical Use of AI and Data was appointed; a discussion paper on the responsible development and deployment of AI was published; and a research program on the governance of AI and data use was launched.

2019

NAIS was published and the Model AI Governance Framework was launched. The framework aimed to provide private sector organizations with readily implementable guidance on key ethical and governance issues when deploying AI solutions.

2020

The updated Model Framework was released, containing additional considerations and refining the original Framework for greater relevance and usability; the World Economic Forum's Implementation and Self-Assessment Guide for Organizations was published, which aimed to help organizations assess the alignment of their AI governance practices with the Model Framework; and a Compendium of Use Cases was released, which illustrated how organizations implemented accountable AI governance practices and aligned AI governance practices with the Model Framework.

2022

The AI Verify testing framework was launched. Positioned as the world's first AI testing toolkit, it enables companies to demonstrate accountability and responsible AI practices through testing and validation.

2023

NAIS 2.0 was released. The Infocomm Media Development Authority launched the AI Verify Foundation to support the development and use of AI Verify. The inaugural Singapore Conference on AI explored potential challenges that could limit society's capacity to leverage AI to benefit people and communities.

2024

AI Verify and IMDA launched the Model AI Governance Framework for Generative AI and extended the 2020 Model Framework to cover nine trust dimensions from accountability to "AI for Public Good." The IMDA and Rwanda's Ministry of Information Communication Technology and Innovation launched the AI Playbook for Small States.

Project Moonshot, one of the world's first large language model evaluation toolkits, was rolled out. It is an open-source tool that brings together benchmarking and red teaming within a single platform.

The GenAI Sandbox launched, a tool that enables local small and medium-sized enterprises greater access to generative AI.

The inaugural Singapore AI Safety Red Teaming Challenge was held with eight other Asia Pacific countries.

Approach to regulation

Rather than rush into legislation, Singapore has deliberately taken an incremental approach to AI governance. The government actively promotes AI adoption as a driver of national growth by publishing flexible frameworks that are regularly reviewed and updated. The national administration adopts a light-touch approach grounded in clear principles for responsible development and use and encourages voluntary compliance.

Singapore sees AI governance as a shared responsibility between government, industrial and research institutions. The government acts as an enabler and facilitator. Industry is empowered to apply governance principles using a risk-based approach.

The AI Verify framework maps to standards around the world, like ISO/IEC 42001:2023, the U.S. National Institute of Standards and Technology AI Risk Management Framework, and Hiroshima Process International Code of Conduct — this shows that companies operating in Singapore can reuse their compliance efforts globally and, according to ISO/IEC JTC1 / SC42 Artificial Intelligence Chairman Wael William Diab, "demonstrates Singapore’s strong support in advancing global harmonisation in a practical way."

Risk-based comprehensive legislation and policy

Singapore has not enacted a comprehensive AI-specific law but has developed governance frameworks that organizations can voluntarily adopt instead. There are also sector-specific regulations that apply to AI systems in particular industries and numerous other laws of relevance and application to various elements of the AI governance lifecycle.

The Model Framework for Generative AI emphasizes risk-based assessments and management throughout the AI lifecycle, from development to monitoring. Organizations are encouraged to tailor the governance measures to address the specific risks they encounter.

Sector-specific legislation

→ Financial services

The Monetary Authority of Singapore, the nation's central bank and integrated financial regulator, was the first sectoral authority to implement AI governance regulation. In 2018, the FEAT principles — a set of principles focused on fairness, ethics, accountability and transparency — was created by the MAS and financial industry to support responsible AI use.

To operationalize these principles, the MAS worked with industry partners to create the Veritas framework in 2019. Veritas enables financial institutions to evaluate their AI and data-analytic-driven solutions against the FEAT principles. In 2023, the framework was updated and the Veritas Toolkit version 2.0 was released.

As a part of Singapore's NAIS, the FEAT principles and Veritas aim to foster a progressive and trusted environment for AI adoption in the financial sector.

The announcement of Project Mindforge was another significant development in 2023. The project was a collaboration between the MAS and key partners from the banking, insurance and capital market sectors with two objectives: to develop a clear and concise framework for the responsible use of generative AI in finance, and to drive innovation to solve common industry-wide challenges and enhance risk management using generative AI.

In early 2024, a whitepaper detailing the generative AI risk framework was published, identifying seven risk dimensions in the areas of accountability and governance, monitoring and stability, transparency and explainability, fairness and bias, legal and regulatory, ethics and impact, and cyber and data security.

After reviewing how banks manage AI model risk, the MAS published an information paper that outlined best practices. The MAS encouraged all financial institutions to reference these practices when developing and deploying AI.

→ Health care

The health care sector does not have any laws related to AI. The Ministry of Health published the AI in Healthcare Guidelines in 2021 to enhance patient safety and foster trust in AI technologies by sharing best practices with AI developers and deployers. These guidelines were co-developed with the Health Sciences Authority and Synapxe; it complements the HSA's Regulatory Guidelines for Software Medical Devices.

→ Legal

In 2024, the Supreme Court released a Guide on the Use of Generative AI Tools by Court Users, setting out general principles and guidance in relation to the use of generative AI tools in legal proceedings.

The Ministry of Law announced in 2025 that it was developing guidelines to help legal professionals be "smart buyers and users of generative AI tools."

Foundation or general-purpose models

Singapore does not have laws that specifically regulate foundation models or general-purpose AI models. Instead, the national approach is based on voluntary guidelines, such as the Model Framework for Generative AI, and technical toolkits that apply to a broad range of AI systems, including foundation and generative AI models.

Agentic AI

Singapore does not have laws that specifically regulate agentic AI and doesn't have a legal definition for the term. Agentic AI would fall under the scope of Singapore's voluntary guidelines and technical toolkits.

In 2025, the Government Technology Agency of Singapore's AI Practice Group published the Agentic AI Primer to provide a clear framework for how AI agents could autonomously pursue objectives and execute tasks within various domains, especially in the public sector.

Enforcement

There aren't any AI-specific laws or AI enforcement agencies in Singapore. Enforcement is limited to existing laws, such as those governing data protection, cybersecurity, copyright and online safety. These laws are enforced by the pertinent regulators or authorities.

Wider regulatory environment

Numerous legal frameworks are applicable to various elements of the AI governance lifecycle.

Data protection

The Personal Data Protection Act governs the collection, use, disclosure and care of personal data. It provides a baseline standard for the protection of personal data and complements sector-specific regulatory provisions, such as those found in the Banking Act and Insurance Act.

In 2024, the Personal Data Protection Commission issued the Advisory Guidelines on Use of Personal Data in AI Recommendation and Decision Systems. The guidelines clarify how the PDPC will interpret the PDPA. While not legally binding, organizations regard the guidelines as standards that should be followed.

Also in 2024, the PDPC launched a Proposed Guide on Synthetic Data Generation to help organizations understand synthetic data generation techniques and potential use cases, particularly for AI.

Cybersecurity

In 2024, the Cyber Security Agency of Singapore released guidelines and a Companion Guide on Security AI Systems to help system owners secure AI throughout its lifecycle. The CSA emphasized that the Companion Guide was not prescriptive, but a community-driven resource that curated practical measures, security controls and best practices from industry and academic partners to complement the guideline document.

Copyright

The Copyright Act protects original expressions of ideas in tangible form. Like the U.S. and unlike the U.K., the work must also have been created by an identifiable human author.

Copyright infringement occurs when a copyright owner's rights are violated, which can occur when copyrighted work is copied, distributed, performed or displayed without the copyright owner's permission. There are two exceptions under Singapore's copyright law: the computational data analysis exception per Sections 243-244 and the fair use exception per Sections 190-191.

The Intellectual Property Office of Singapore has confirmed the computational data analysis exception permits the use of copyrighted material for sentiment analysis, text and data mining, and machine learning, subject to specified conditions. The fair-use exception permits the use of non-substantial parts of copyrighted work for non-profit, educational purposes.

Willful copyright infringement is a criminal offense when it is significant and carried out for commercial gain.

Online safety

In 2022, the Online Safety (Miscellaneous Amendments) Act was passed to enhance online safety for users. Providers of "online communication services," i.e., electronic services that allow Singapore users to access or communicate content via the internet with significant reach or impact, must comply with the Codes of Practice issued by IMDA. So far, this new rule only covers social media services; IMDA has issued one Code of Practice for online safety. The amendments also empower IMDA to issue directives to address specified "egregious content" accessible to users in Singapore.

The Online Criminal Harms Act, passed in 2023, enables the government to more effectively address online activities of a criminal nature.

Established by the Agency for Science, Technology and Research in April 2023, the Centre for Advanced Technologies in Online Safety conducts research on technological capabilities to prevent digital harm. The organization focuses on detecting deepfakes and misinformation, applying watermarks, tracing the origin of digital content, and helping vulnerable groups verify online information.

Anti discrimination

In 2024, the Ministry of Manpower clarified that regardless of the technological tools used to aid employment decisions, all employers in Singapore must comply with the Tripartite Guidelines on Fair Employment Practices. The guidelines outline principles of fair employment practices, such as hiring on the basis of merit — including skills, experience or ability to perform the job — regardless of age, race, gender, religion, disability or marital status and family responsibilities.

The ministry said that if certain AI use results in discriminatory employment practices, workers or job applicants could approach the Tripartite Alliance for Fair and Progressive Employment Practices for assistance. As of 13 Nov. 2024, the alliance had not received any complaints of discrimination arising from the use of AI tools.

International cooperation on AI

As part of Singapore's goal to establish itself as an international partner for AI innovation and governance, the government wants to continue building international networks with key partner countries and leading AI companies.

So far, half of Singapore's digital economy agreements contain AI modules that promote the adoption of ethical AI governance frameworks and, where appropriate, the alignment of governance and regulatory frameworks. The intention is to establish a shared set of AI governance and ethics principles with international partners so organizational compliance is more straightforward and easily achievable.

The U.S. National Institute of Standards and Technology and IMDA published a crosswalk in October 2023, mapping the NIST's AI Risk Management Framework 1.0 to AI Verify. IMDA said this joint effort signaled both parties' common goal of balancing AI innovation and maximizing the benefits of AI technology while mitigating technology risks.

In November 2023, Singapore's prime minister participated in the AI Safety Summit. The country joined 27 others in signing the Bletchley Declaration, agreeing to work together to prevent "catastrophic harm, either deliberate or unintentional," which may arise from AI computer models and engines.

Singapore released the Association of Southeast Asian Nations Guide on AI Governance and Ethics during the fourth ASEAN Digital Ministers' meeting in February 2024.

In May 2024, Singapore — through the AI Verify Foundation — published a Memorandum of Intent in collaboration with MLCommons, a leading AI engineering consortium recognized by the U.S. NIST. The safety initiative developed a common set of benchmarks, tools and testing approaches for generative AI models.

While Singapore is not a member of the Organisation for Economic Co-operation and Development, representatives were invited to take part in its Expert Group on AI. Singapore is a founding member of the Global Partnership on AI, an international initiative to promote responsible use of AI that respects human rights and democratic values.

Latest developments

Despite being only halfway through 2025, Singapore has introduced a significant number of projects and programs.

At the AI Action Summit in February, Singapore's Minister for Digital Development and Information, Josephine Teo, announced three new AI governance initiatives to enhance the safety of AI for both Singapore and the global community.

First, IMDA and the AI Verify Foundation launched the Global AI Assurance Pilot, which helps codify emerging norms and best practices around the technical testing of generative AI applications.

Following a landmark collaboration between Singapore and Japan, the International Network of AI Safety Institutes released a joint testing report that evaluated the safety of LLMs across diverse linguistic and cultural environments. The primary objective was to assess whether the safeguards built into LLMs, such as protections against generating harmful, illegal, or biased content, were effective in non-English settings.

As co-lead of the testing and evaluation track under the AISI network, Singapore brought together global linguistic and technical experts from the AISI network to conduct tests across ten languages: Cantonese, English, Farsi, French, Japanese, Kiswahili, Korean, Malay, Mandarin Chinese and Telugu. The experts also conducted tests across five harm categories: violent crime, non-violent crime, IP, privacy and jailbreaking.

Following the 2024 multicultural and multilingual AI safety red teaming exercise, the IMDA published the Singapore AI Safety Red Teaming Challenge Evaluation Report 2025, which sets out a consistent methodology for testing AI safeguards across diverse languages and cultures.

In April 2025, Singapore hosted the SCAI: International Scientific Exchange on AI Safety, gathering more than 100 minds from academia, industry, and government to set priorities for shaping reliable, secure, and safe AI. The follow-up report, "The Singapore Consensus on Global AI Safety Research Priorities," was published in May 2025; the report aims to advance impactful research and development efforts to rapidly develop safety and evaluation mechanisms, and foster a trusted ecosystem where AI is harnessed for the public good.

Also in May, the IMDA launched an upgraded version of its LLM Multimodal Empathetic Reasoning and Learning in One Network, as well as the MERaLION Consortium, a collaborative platform that brings together local and global industry players, research and development institutions and leading technology companies, to develop practical AI applications such as multilingual customer support, health and emotional insight detection and agentic decision-making systems.

The IMDA launched a four-week public consultation in May on the "Starter Kit for Safety Testing of LLM-Based Applications." The kit is a set of voluntary guidelines on how to think about and conduct testing for common risks in apps. It sets out a structured approach to pre-deployment testing from app output to app components and contains recommended tests for the four key risks commonly encountered in apps today — hallucination, undesirable content, data disclosure and vulnerability to adversarial prompt.

Published August 2025

Global AI Governance Law and Policy: South Korea

By Sam Jungyun Choi, Hwan Kyoung Ko, Matt Younghoon Mok, Hyun Wo Kim, Sunghee Chae, D. Yoon Chae and Hyewon Chin

South Korea is a global powerhouse in several industries, including IT, semiconductors and batteries, enabling the country to emerge as a key player in artificial intelligence. The adoption of South Korea's Act on the Development of Artificial Intelligence and Establishment of Trust has been a watershed moment in the development of the nation's artificial intelligence policy. The AI Basic Act is scheduled to come into effect on 22 January 2026 and will be the world's second comprehensive AI law after the EU AI Act.

Many of the technical details of the obligations under the act is delegated to the enforcement decrees, which are currently being prepared by the Ministry of Science and Information and Communication Technology. There have also been other regulatory developments relating to AI, including the Personal Information Protection Act, Copyright Act, and Monopoly Regulation and Fair Trade Act.

The Lee Jae Myung presidential administration, which came into power on 4 June 2025, views AI as a key driver of South Korea's economic growth. President Lee announced his vision to position South Korea among the world's top three players by creating an AI-related industrial innovation ecosystem, building the world's most advanced AI infrastructure, introducing legislation and governance systems, and developing AI talent. Many experts anticipate South Korea's AI industry to undergo significant transformations, fueled by such technological growth and recent regulatory development.

Approach to regulation

The South Korean government has pursued policies to help promote the AI industry by recognizing the autonomy of the private sector and avoiding onerous regulations. At the forefront of the government's efforts is the AI Basic Act, which is aimed at "protect[ing] the rights and dignity of the people, improve[ing] their quality of life and strengthen[ing] national competitiveness by stipulating the basic matters necessary for the safe development of artificial intelligence and the establishment of trust."

Regulation under the AI Basic Act

Under the AI Basic Act, AI is defined as an electronic implementation of human intellectual abilities, such as learning, reasoning, perception, decision making and language comprehension. An AI system is defined as a system powered by AI capable of making predictions, suggestions and decisions that affect real and virtual environments for a given goal with various levels of autonomy and adaptability. The AI Basic Act defines AI business operators as corporations, organizations, individuals and government bodies conducting AI-related businesses; business operators fall into one of two categories: AI development business operators who develop and offer AI or AI utilization business operators who provide AI products or services powered by AI developed by the former. While not an exact match, an AI development business operator would constitute a provider under the EU AI Act, whereas an AI utilization business operator would be considered a deployer under the same framework. The AI Basic Act's obligations currently equally apply to both types of AI business operators, although it remains to be seen if forthcoming enforcement decrees or associated guidelines will specify any differentiated obligations for each type of business operator.

The AI Basic Act adopts a risk-based and comprehensive regulatory framework. The act imposes different obligations on AI business operators depending on the type of AI being provided. For example, operators of high-impact AI — defined as systems that significantly affect or poses risks to human life, physical safety, or fundamental rights and are used in critical domains such as nuclear power, energy, traffic control, recruitment, and loan evaluation — are required to:

- → Assess whether their AI qualifies as "high-impact AI" before deployment. The operator may optionally confirm the assessment through the Minister of the MSIT.
- → Inform the users in advance that their products are powered by high-impact AI.
- → Implement a comprehensive framework of safety and reliability measures for their high-impact AI.

Additionally, AI business operators are encouraged to conduct impact assessments to evaluate the potential effects of their high-impact AI on people's fundamental rights.

Generative and high-performance AI

The AI Basic Act does not mention foundation or general-purpose models. Instead, it sets out obligations for generative AI and high-performance AI — i.e., AI systems that either surpass a predetermined threshold of cumulative compute used during training.

The AI Basic Act defines generative AI as an AI system that generates text, sound, images, videos or other outputs by mimicking the structure and features of the input data. Operators are required to notify users in advance that their products are being powered by generative AI. Operators must also clearly label products or services that have been created by generative AI. When virtual outputs may be mistaken as real — often referred to as deepfakes — operators are required to provide clear notifications or labels indicating the potential for misinterpretation.

Operators of high-performance AI systems are required to identify, assess and mitigate risks throughout the AI's lifecycle. Operators must also establish a risk management system for monitoring and responding to AI-related safety incidents.

Agentic AI

There are currently no government regulations or policies specifically targeting agentic AI. However, as its use becomes more prevalent, the question of whether it qualifies as "high-impact AI" under the AI Basic Act is expected to be addressed. There have also been discussions about the need for trust-based governance and multi-layered safeguards for agentic AI.

Enforcement

The minister of the MSIT may launch a fact-finding investigation if it receives a report/complaint or otherwise suspects that any of the following requirements under the AI Basic Act have been violated: compliance with the safety and reliability standards for high-impact AI; labeling requirements for generative AI outputs, as well as notification or labeling requirements for deepfake outputs; and implementation of safety measures and reporting of compliance results for high-performance AI.

Upon confirmation of a violation, the minister may issue an order directing the offending party to suspend or correct the non-compliant action. Failure to comply with an MSIT order can result in fines of up to KRW30 million, or approximately USD21,700.

Wider regulatory environment

Data protection

South Korea's data protection authority, the Personal Information Protection Commission is spearheading policy reform to ensure the Personal Information Protection Act is aligned with the realities of the AI era.

The central goal is to facilitate safe and responsible use of personal information for AI development through measures such as encouraging the use of pseudonymized data for AI training and promoting frameworks to make pseudonymization more accessible. Another measure encourages proposing amendments to the PIPA to permit the processing of lawfully collected personal information for secondary purposes — including AI development — under certain safeguards. Additionally, the PIPC published "Guidelines on Publicly Available Personal Information for AI Development and Services" in July 2024.

At the same time, from a privacy protection perspective, the PIPC issued guidelines in December 2024 in the form of the "AI Privacy Risk Management Model for Safe Use of AI and Data" to help data controllers identify and mitigate privacy risks associated with the development and deployment of AI technologies.

Copyright and intellectual property

South Korea's copyright framework is still evolving in response to the unique challenges posed by AI development. In particular, there is ongoing legislative discussion regarding whether text and data mining for AI training qualifies as fair use under the Copyright Act. At present, there is no binding precedent or regulatory guidance clarifying this issue.

The Korea Copyright Commission is working on clarifying authorship and copyright recognition for AI-generated and AI-assisted works and developing standards for the use of copyrighted materials in AI training datasets. The KCC is also promoting the enactment of the Right of Publicity Act to regulate the commercial use of individuals' names, likenesses and voices, particularly in the context of deepfakes and AI-generated content.

Antitrust policy

While enforcement activity in the AI space remains nascent, the Korea Fair Trade Commission is closely monitoring potential competition issues in AI markets. Following a market survey involving more than 50 major domestic and global AI firms, the KFTC published a policy report, "Generative AI and Competition" in December 2024. The report identified key antitrust concerns, including barriers to entry in AI markets, the structures of vertical and horizontal competition, and concerns over data and infrastructure monopolization.

The KFTC is currently reviewing potential regulatory reforms based on the report and plans to conduct further studies on anti-competitive practices related to data collection and usage in AI systems.

Consumer protection

To protect consumers from misleading claims about AI capabilities, also known as AI washing, the KFTC is collaborating with the Korea Consumer Agency to monitor such practices. Where the KCA identifies deceptive advertising in violation of the Act on Fair Labeling and Advertising, it may issue corrective recommendations.

While the recommendations do not carry binding legal force, noncompliance may prompt the KFTC to initiate formal investigations, which can lead to administrative fines. In addition, false or exaggerated advertising may result in criminal penalties of up to two years' imprisonment or a fine of up to KRW150 million, approximately USD108,500; consumers may also pursue civil remedies under applicable law.

Health care sector

The Ministry of Food and Drug Safety has taken a proactive stance in regulating AI-powered medical devices. In January 2025, the KFDA released the world's first Guideline for Approval and Examination of Generative AI Medical Devices, setting the approval and evaluation standards for such devices. It also co-published the Guiding Principles for conducting Clinical Trial for Machine Learning-enabled Medical Devices in collaboration with Singapore's Health Sciences Authority.

Meanwhile, AI-based medical devices are expected to be classified as high-impact AI once the AI Basic Act comes into force in January 2026, thereby becoming subject to stricter compliance obligations.

Financial sector

AI is increasingly being utilized in South Korea's financial sector for credit assessment, fraud detection, customer service, and investment and risk management. Although there is currently no AI-specific legislation in place within the financial sector, the Financial Services Commission and the Financial Supervisory Service have issued a series of guidance documents — including the AI Guidelines for the Financial Sector, the Guidelines for AI Development and Utilization and AI Security Guidelines — to address the evolving financial landscape shaped by AI adoption.

These guidelines promote responsible AI implementation by emphasizing core principles such as fairness, transparency, and ethics, as well as requirements for data validation and internal controls. They also provide direction on high-risk applications, including AI-based credit scoring, which may fall under the category of high-impact AI in the AI Basic Act.

Latest developments and next steps

AI has emerged as a national priority under the current presidential administration. The government has ambitions to build a comprehensive national support system for AI research and development talent cultivation and an infrastructure to foster AI as a strategic industry.

Recent developments in AI policy include a proposed three-year grace period for regulatory enforcement under the AI Basic Act and plans to strengthen the National Artificial Intelligence Council. The proposed Special Act on Fostering and Supporting the Artificial Intelligence Industry outlines comprehensive support measures for AI-related enterprises. Additionally, recent legislative changes have broadened permissible use of personal information, including raw data. These changes apply when pseudonymization alone is insufficient to achieve research objectives and introduce sector-specific standards for AI and data processing.

South Korea's AI regulatory regime remains growth-oriented but is rapidly evolving. Policymakers are striving to strike a balance between fostering innovation and safeguarding public interests. Given the fast pace of legislative developments and the wide range of affected sectors, stakeholders should closely monitor South Korea's evolving AI regulatory trajectory.

Published December 2025

Global AI Governance Law and Policy: UAE

By Masha Ooijevaar, Ben Gibson, Eve Brady, Hannah Torpey, Swapnil Govind Dambe and Varun Kharbanda

The United Arab Emirates has steadily built a national framework for artificial intelligence over the past decade with a focus on integrating AI into government services, economic planning, and infrastructure. In 2017, the UAE launched its first AI strategy and appointed a Minister of State for Artificial Intelligence, Digital Economy, and Remote Work Applications Office, initiating a government-led effort to explore how AI could be applied across public services and national development.

Since then, the UAE has expanded its efforts through the UAE National Strategy for Artificial Intelligence 2031, which outlines goals for integrating AI into sectors such as health care, education, and transportation. It is supported by initiatives that build technical capacity, attract investment, and promote responsible innovation.

The UAE's level of investment — both internally and via strategic bilateral agreements with global players and governments, such as the U.S.-UAE AI Acceleration Partnership — demonstrates the government's commitment to AI. It positions AI not merely as a computing resource but as a critical national asset.

While there is currently no dedicated AI legislation in the UAE, other than one provision of the Dubai International Financial Center Data Protection Law discussed below, the government has introduced various mechanisms to manage the development and use of AI. These include ethical guidelines, sector-specific standards, and institutional bodies such as the Artificial Intelligence and Advanced Technology Council. In parallel, the UAE has intensified its engagement in cross-border AI standards harmonization efforts, actively contributing to international forums such as ISO/IEC JTC 1/SC 42 Artificial Intelligence and cooperating with organizations like the Organisation for Economic Co-operation and Development and UNESCO on AI ethics and governance frameworks. This reflects the nation's strategic intent to align domestic AI governance with evolving global norms and reinforce trust in responsible AI adoption.

The UAE's approach to AI governance remains policy-driven with legal oversight primarily addressed through broader frameworks such as data protection and cybersecurity. Governance continues to evolve through a combination of strategic initiatives, institutional development, and non-binding guidance while formal regulatory structures specific to AI are still taking shape.

From a governance perspective, the UAE's approach represents a hybrid regulatory model — one that blends state-led strategic direction with decentralized implementation. This model is increasingly viewed as a "sandbox state" approach, where policy frameworks precede and inform eventual statutory codification. This contrasts with the EU’s prescriptive AI Act and the U.S.'s sectoral self-regulation, underscoring the UAE's pragmatic and innovation-friendly stance.

History and context

The UAE's engagement with AI began in 2017 when the federal government launched the UAE Strategy for Artificial Intelligence and appointed the world's first minister of state for AI. This marked a shift in national policy toward embedding AI into public administration and economic planning. The strategy was framed as part of the broader UAE Centennial 2071 vision, which aims to position the country as a global leader in innovation and advanced technology. It also built upon the UAE's smart government and city initiatives, particularly the Smart Dubai 2021 Strategy, which laid the foundational digital infrastructure and governance frameworks that later enabled large-scale AI implementation across public services and urban ecosystems.

Early moves to establish a minister of AI pre-empted most OECD jurisdictions and positioned the country among the first nations to institutionalize AI governance at the cabinet level. This foresight reflects an understanding that AI policy cannot be confined to digital transformation portfolios alone and must be integrated with national competitiveness, education, and industrial policy.

The country's initial strategy focused on improving government performance, reducing costs, and enhancing service delivery through AI. It identified priority sectors, including transportation, health care, education, energy and space, and introduced mechanisms such as the UAE Council for Artificial Intelligence and Blockchain to oversee implementation. The strategy also called for the development of AI capabilities within government entities, including the appointment of a chief AI officer across ministries and federal bodies.

Over time, the UAE has introduced AI readiness and maturity assessment frameworks for federal and local entities, requiring ministries and agencies to track their AI adoption performance using defined key performance indicators aligned with the Artificial Intelligence Strategy 2031. This systematic approach ensures measurable progress toward embedding AI into decision-making, operations, and citizen services.

The UAE expanded its ambitions through the Artificial Intelligence Strategy 2031, which set out eight strategic objectives. These include building a reputation as an AI destination, developing a fertile ecosystem for AI innovation, integrating AI into customer-facing services, and ensuring strong governance and ethical oversight. The strategy also emphasized the importance of attracting talent, supporting research, and creating the infrastructure needed to support AI deployment. These objectives are directly aligned with the UAE's broader national competitiveness strategy and Digital Economy Vision, reinforcing how AI acts as a key lever for gross domestic product diversification beyond hydrocarbons and for advancing national productivity and sustainability goals.

To support these goals, the UAE launched several initiatives, including the establishment of the National Program for Artificial Intelligence, the Mohamed bin Zayed University of Artificial Intelligence, and partnerships with international technology firms and academic institutions. The MBZUAI has played a pivotal role in nurturing indigenous AI talent and advancing applied research in frontier areas, such as large language models, climate-tech AI, and sustainable computing, strengthening the UAE's sovereign capability and positioning it as a regional thought leader in responsible AI innovation. These efforts have been supplemented by targeted investments in AI infrastructure, such as data centers and cloud platforms, and the development of open-source AI models like Falcon 3.

In 2024, the UAE strengthened its institutional framework by establishing the Artificial Intelligence and Advanced Technology Council in Abu Dhabi. The council is tasked with developing policies and strategies related to AI research, infrastructure, and investment. It plays a coordinating role across government and industry and is intended to support Abu Dhabi's positioning as a regional hub for advanced technology. The council also reflects the UAE's growing engagement in international AI governance, including participation in multilateral forums and global standard-setting initiatives.

A notable policy theme emerging through these frameworks is the UAE's emphasis on "human-centric AI," ensuring that technological advancement remains aligned with ethical, transparent, and inclusive development. This approach resonates with global responsible AI principles, embedding fairness, accountability, privacy, and security at the core of the nation's AI transformation journey.

Although the UAE has not enacted a dedicated AI law, it has issued ethical guidelines and charters to guide responsible development. The UAE Charter for the Development and Use of AI, released in 2024, sets out twelve principles covering human oversight, data privacy, transparency, and fairness. They inform future regulatory developments and guide both public and private sector actors.

Legal oversight of AI is currently limited and primarily addressed through broader regulatory frameworks that intersect with AI-related activities. Federal Decree-Law No. 45 of 2021 Regarding the Protection of Personal Data includes provisions relevant to automated processing. The Dubai International Financial Centre has amended its Data Protection Law to address AI-related transparency, governance and accountability. Although these measures form part of broader legal frameworks, they do not constitute standalone, comprehensive AI legislation like the EU AI Act.

From a comparative law standpoint, the UAE's incremental legislative layering — where AI governance is diffused through data, cyber, and consumer protection regimes — illustrates a functional distributed regulatory model. Such a model may risk regulatory fragmentation, especially in cross-sectoral AI use cases, such as generative AI in finance or health care, that fall between jurisdictional mandates.

The country's approach to AI governance has therefore been shaped more by policy frameworks and institutional developments than by legal regulation. It reflects a model focused on enabling innovation while gradually building the structures needed to address emerging risks and regulatory challenges.

Regulatory approach

At present, the UAE does not have dedicated legislation exclusively governing AI. Oversight falls across various government agencies and bodies, such as the Artificial Intelligence Office of the Ministry of Health and Prevention for the health care sector, the Telecommunications and Digital Government Regulatory Authority, and the Ministry of State for Artificial Intelligence, Digital Economy and Remote Work Applications Office, also known as the Ministry of AI. Additionally, the UAE Data Office is empowered to investigate data breaches linked to AI systems; the UAE Council for Artificial Intelligence has been established as a specialized committee to strengthen the governance and coordination of AI initiatives across government entities.

Furthermore, as AI becomes more embedded in decision-making and automation, the concept of AI liability is emerging, raising important questions about how existing tort, contract, and product safety laws may be interpreted to address harms or errors arising from autonomous systems. The country's evolving regulatory approach will likely continue to refine these accountability mechanisms, balancing innovation with the protection of individual and societal interests.

In practice, many UAE organizations navigate a mosaic of overlapping obligations — data privacy under the PDPL, content regulation under media laws, and cyber risk under Federal Law No. 34 of 2021. For corporate compliance officers, this necessitates an AI compliance-by-design mindset, embedding algorithmic audit trails, explainability logs, and ethical impact assessments from procurement to deployment. Such internal governance measures are likely to form the baseline for any forthcoming federal AI law.

Personal Data Protection Law

The UAE Personal Data Protection Law is the country's first comprehensive federal data protection framework, enacted in 2022 to regulate the collection, processing, and storage of personal data across private sectors. It is modelled on the EU General Data Protection Regulation, incorporating core principles such as lawfulness, fairness, transparency, purpose limitation, data minimization and accountability.

Although the PDPL does not contain standalone provisions dedicated to AI, it has implications for AI systems, particularly those involving automated decision-making and personal data processing. Article 18 of the PDPL addresses automated processing, granting data subjects the right to object to decisions made solely through automated means and to request human intervention. The law also includes broader data subject rights, such as access, correction, and erasure, that apply to AI-driven systems handling personal data. For example, Article 15 allows individuals to request the correction or deletion of inaccurate or unlawfully processed data, which may arise in the context of machine learning models trained on personal datasets.

Additionally, the PDPL requires organizations to conduct an assessment of the impact of personal data protection in cases where processing is likely to result in a high risk to the privacy and confidentiality of individuals. This obligation is particularly relevant to AI systems that involve profiling, large-scale processing of sensitive data, or decisions with legal or similarly significant effects.

At the time of publication, the PDPL is in force but not yet enforceable. Enforcement will begin following the expiry of a six-month grace period granted to organizations once the implementing regulations are published. These regulations, which are expected to provide further clarity on compliance obligations, have not yet been released. Additionally, the UAE Data Office, designated as the federal data protection regulator, is not yet fully operational.

Consumer Protection Law

Federal Law No. 15 of 2020 on Consumer Protection, as amended along with its implementing regulations, restricts the use of consumer data collected by businesses to only the execution of the relevant transaction. Any additional use requires informed consumer consent. This includes use of the data for profiling, analysis, pattern recognition and predictive purposes, all areas where AI may play a role.

Cybercrime Law

Federal Law No. 34 of 2021 on Countering Rumours and Cybercrime includes various broad provisions that define offenses relating to the use of computing systems and information networks. These offenses include invasion of privacy and unauthorized access to records and information. An article within the law criminalizes processing activities such as acquisition, disclosure and modification of personal data without authorization. To the extent that AI models are trained on data from various sources, it is important to consider the provenance of such data and the legal bases on which it is legitimately shared and processed to avoid inadvertent exposure under the Cybercrime Law.

From a policy-risk perspective, the Cybercrime Law acts as an implicit AI accountability statute, given its broad language around data misuse, defamation, and misinformation. The absence of intent thresholds in several provisions means that developers and deployers of generative AI tools face strict exposure, even where harm is unintentional.

The Cybercrime Law also criminalizes the online publication of media content that infringes the country's media laws, spreads "fake news" including the use of bots to disseminate inaccurate information, harms the interests of the state or attacks foreign states, and contains misleading advertising. The use of generative AI in creative processes or for amplifying messages can pose serious criminal liability risks if not carefully managed.

Defamatory content is also a criminal offense as well as a civil offense; the truth of a statement is not necessarily a defense to a defamation allegation. The test is not whether the statement is true, but rather whether it exposes the target to contempt and ridicule and whether it was reasonable to make the statement. On that basis, any generative AI tools that create images or depictions of real or identifiable characters or describe any real or identifiable persons risk committing defamation if not moderated.

Financial Free Zones

The UAE's legal system is composed of federal laws that apply nationwide; emirate-level legislation, such as in Dubai and Abu Dhabi; and regulations issued by various free zones. Among these are the DIFC and the Abu Dhabi Global Market, two financial free zones with independent legal frameworks including their own data protection regimes. These jurisdictions are excluded from the scope of the PDPL and operate under separate supervisory authorities.

The DIFC has introduced targeted provisions within its data protection framework to address the use of AI and autonomous systems. Regulation 10 of the DIFC Data Protection Regulations governs the use of autonomous and semi-autonomous systems in personal data processing, including AI, and sets out specific obligations for organizations deploying such technologies.

Regulation 10 establishes a set of general design principles for AI systems, requiring that they be developed and operated in accordance with ethical, fair, transparent, secure, and accountable standards. Systems must be able to process personal data only for purposes that are either human-defined or human-approved or for purposes defined by the system itself, strictly within human-defined constraints. These principles are intended to mitigate bias, ensure explainability, protect data confidentiality, and establish clear lines of responsibility.

The DIFC's Regulation 10 is one of the first subnational AI governance instruments globally, pre-dating most G20 jurisdictions. Its principles echo the OECD's 2019 AI Recommendations and the ISO/IEC 42001 standard under development, suggesting the DIFC's intent to align with international regulatory interoperability frameworks.

The regulation also imposes detailed notice and transparency requirements. Where autonomous systems are used in applications or services that process personal data, deployers and operators must provide clear and explicit notice to users at the point of initial use. This notice must describe the nature of the system, including whether it operates independently of human direction, and explain its impact on individual rights.

Entities engaging in high-risk processing must appoint an autonomous systems officer, with responsibilities comparable to those of a data protection officer, and ensure that their system complies with any audit and certification requirements established by the DIFC commissioner of data protection.

In support of these obligations, a certification framework has been developed to assist organizations in demonstrating compliance. This framework is intended to align with international standards and provide a basis for future audit and certification requirements that may be established by the DIFC commissioner of data protection.

Health care sector

Emirate-level authorities have taken steps to develop AI-specific policies in the health care sector. The Dubai Health Authority issued a policy framework for the use of AI in health care services, outlining requirements for transparency, accountability, and patient safety in AI-enabled systems used across Dubai's health infrastructure. Similarly, the Abu Dhabi Department of Health published its Policy on the Use of Artificial Intelligence in the Healthcare Sector in 2018. The policy applies to all licensed health care providers, insurers, researchers, and pharmaceutical manufacturers operating in Abu Dhabi and sets out principles for responsible AI adoption, including risk management, data governance, and patient safety.

Importantly, both frameworks anticipate the evolution of AI oversight mechanisms, including potential certification, validation, and audit requirements for AI-based medical devices and software-as-a-medical-device solutions. These forward-looking provisions align with global regulatory trends seen under frameworks such as the U.S. Food and Drug Administration's AI/machine learning-based medical device guidance and the EU Medical Device Regulation, underscoring the UAE's intent to harmonize its health care AI governance with international best practices.

The health care sector shows how the UAE is turning AI policy into practical regulation. Abu Dhabi and Dubai increasingly align AI approval with existing medical-device and data-exchange processes, creating an implicit layer of oversight without new legislation. This integration, comparable in spirit to the EU's Medical Device Regulation, allows innovation to progress while maintaining patient-safety assurance and regulatory accountability.

The UAE Charter for the Development and Use of Artificial Intelligence

The UAE Charter for the Development and Use of Artificial Intelligence, issued in June 2024 by the minister of AI, serves as a cornerstone for responsible AI governance in the country. It aligns closely with the UAE Strategy for Artificial Intelligence 2031, emphasizing human well-being, safety, privacy, and transparency in the design and deployment of AI systems. The charter articulates twelve guiding principles to ensure ethical and inclusive AI implementation grounded in robust governance, accountability, and compliance with both international and domestic laws.

While the charter is not a binding legal instrument, it plays a strategic role as the ethical foundation for all sectoral AI regulatory initiatives in the UAE. Much like the EU AI Act's horizontal risk-based model, it provides a unifying ethical and governance framework that ensures coherence and consistency across diverse, highly regulated sectors, including banking, finance, public administration, media, mobility, and health care. This approach reinforces the UAE's commitment to responsible, human-centric AI while enabling innovation across the national digital ecosystem.

The charter's soft-law nature is equally significant. By remaining flexible rather than prescriptive, it enables adaptive regulatory evolution without stifling innovation — a concept increasingly endorsed in the Gulf Cooperation Council's emerging AI legal culture.

Smart Dubai AI Ethics Principles

In 2019, the Smart Dubai Office, now part of Digital Dubai, released a set of AI Ethics Principles and Guidelines to support the responsible development and deployment of AI across the emirate. These principles are intended to guide public and private sector organizations in ensuring that AI systems are designed and used in ways that promote fairness, accountability, transparency, and human benefit. The framework includes commitments to mitigate bias, ensure explainability, and uphold data security and individual rights. The guidelines emphasize that individuals should be able to challenge significant automated decisions and that accountability for AI outcomes must rest with human actors, not the systems themselves. The initiative also introduced a self-assessment tool to help organizations evaluate the ethical performance of their AI systems.

Whitepaper on the Responsible Metaverse Self-Governance Framework

The UAE's AI office, working with the Dubai Department of Economy and Tourism, has released a whitepaper on the Responsible Metaverse Self-Governance Framework, which lays out nine self-regulatory principles to help shape the metaverse's ethical and responsible growth. While it does not directly regulate AI, the whitepaper treats AI as a key building block of the metaverse and weaves AI governance considerations into the broader framework.

Agentic AI

While there is no single comprehensive agentic AI law, the existing framework — particularly Regulation 10 of the DIFC Data Protection Regulations, the UAE Charter for the Development and Use of Artificial Intelligence and the establishment of the minister of AI — reflects the UAE's emerging approach to regulating autonomous AI systems. Sector-specific policies such as the emirate health authorities' policies on AI in health care and Law No. 9 of 2023 Regulating the Operation of Autonomous Vehicles in the Emirate of Dubai appear to reinforce this evolving framework, requiring AI systems, including those with autonomous or agentic capabilities, to operate within legal and ethical boundaries.

Importantly, the UAE is beginning to differentiate between autonomous and agentic AI in its policy discussions. Autonomous AI refers to self-operating systems functioning within predefined parameters, while agentic

AI denotes systems capable of adaptive decision-making, learning, and acting on inferred intent. To manage the risks and governance challenges posed by such advanced models, the country is increasingly leveraging regulatory sandboxes, such as the Abu Dhabi Global Market's digital sandbox, to enable controlled experimentation and supervised validation of agentic AI applications in sectors like finance, logistics, and mobility.

Latest developments

The UAE's approach to AI governance is evolving from foundational frameworks to more tangible regulatory mechanisms. Since launching the UAE Charter for the Development and Use of Artificial Intelligence, the country established the world's first AI-enabled Regulatory Intelligence Office within the Cabinet in April 2025. The office connects federal and local laws with judicial rulings, executive procedures, and public services through a centralized AI system. This system monitors the real-world impact of laws and suggests updates based on large-scale data analysis. Officials describe this as a shift toward AI-driven regulation, aimed at accelerating legislative processes and improving responsiveness. While ambitious, experts emphasize the need for human oversight and safeguards against bias and reliability risks.

This innovation positions the UAE at the frontier of AI for regulation, not merely the regulation of AI. However, such use of AI in rule-making triggers complex jurisprudential questions about delegated cognition. For example, to what extent can predictive analytics influence legislative drafting without eroding democratic legitimacy? The legal tradition will likely evolve mechanisms to ensure human interpretive supremacy within algorithm-assisted governance.

Additionally, the Dubai State of AI Report, published in April 2025, outlines the city's commitment to shaping international AI governance through global forums and initiatives like the Dubai AI Acceleration Taskforce, which invites groups to co-develop frameworks.

Moreover, the Dubai Centre for Artificial Intelligence has introduced the "Dubai AI Seal," a verification system designed to accelerate the growth of the emirate's AI industry. Through this initiative, legally operating AI businesses of any size can submit their details via an online application process. Each application is assessed by the DCAI team using the Dubai AI Business Activity Classification System.

Approved businesses receive a personalized Dubai AI Seal, which includes a tier ranking and a unique serial number at no cost. The seal features six tiers that reflect the level of economic contribution: S, A, B, C, D, and E. Tier S represents the highest impact on Dubai's AI economy. The program aims to strengthen business credibility, protect public and private entities from irrelevant suppliers and AI-washing, and streamline access to trusted AI providers in Dubai.

The UAE Media Council also illustrates practical AI integration through its agreement with Presight to launch the Unified Media AI and Analytics Platform, designed to assess and regulate media content prior to publication.

Future outlook

As the UAE continues to build on its Artificial Intelligence Strategy 2031, the focus is shifting from strategic planning to operational execution, emphasizing the integration of AI into public services, infrastructure expansion, and the institutionalization of governance mechanisms that ensure responsible deployment. The next phase is expected to prioritize workforce development, cross-sector collaboration, and international partnerships to consolidate the country's global leadership in AI innovation.

At GITEX Global 2025, the Ministry of Human Resources and Emiratisation unveiled Eye, an AI-powered system to automate work permit processing. Leveraging intelligent document verification for passports and academic credentials, the system minimizes manual intervention and accelerates approvals — an example of how AI agents are being operationalized within core government functions to enhance efficiency and reduce costs across the labor ecosystem.

Looking ahead, the UAE's AI agenda includes expanding sovereign digital infrastructure, exemplified by the Stargate supercomputing cluster in Abu Dhabi, which will host large-scale national AI models and bolster computational capacity. The country is also heavily investing in upskilling programs, strategic partnerships, and the operationalization of responsible AI frameworks through initiatives such as the UAE Charter for the Development and Use of Artificial Intelligence and emirate-level ethical guidelines.

As AI becomes increasingly embedded in governance, commerce, and social systems, the UAE's next steps will likely involve refining regulatory coherence, institutionalizing ethical accountability, and ensuring alignment with evolving international AI governance norms. Key emerging priorities include modular AI legislation targeting high-risk systems, the Gulf Cooperation Council's Guiding Manual for the Ethics of AI Use, human capital and regulator upskilling, and cross-border data governance alignment with the GDPR and Asia-Pacific Economic Cooperation standards.

The UAE's AI journey reflects a deliberate evolution from policy vision to structured governance. By embedding responsible AI principles across institutional frameworks and advancing measurable AI maturity models, the country is setting global benchmarks for trustworthy, human-centric AI. Its future trajectory points toward a hybrid model, balancing innovation with robust ethical oversight, interoperability with global AI regulations, and transparency through governance audits and digital assurance. The UAE's growing participation in international AI forums, such as the World Government Summit and UNESCO's Policy Dialogue on AI Governance, reinforces its commitment to shaping global AI norms.

Ultimately, the UAE's approach is transforming AI governance into an enabler of trust, economic diversification and responsible digital transformation, creating a resilient foundation for an inclusive, safe, and globally respected AI-driven economy.

Published October 2025

Global AI Governance Law and Policy: UK

By Joe Jones, Alexander Milner-Smith, Lee Ramsay and Sundip Athwal

Though the U.K. does not have any legislation specific to the regulation or governance of artificial intelligence, it does have an AI Security Institute and a variety of relevant principles-based soft law and policy initiatives, as well as binding regulations in other domains like data protection and online safety. The AI Security Institute, which started life as the AI Safety Institute, launched at the world's first global AI Safety Summit held in the U.K. in November 2023. However, in February 2025 AISI changed its name to reflect its change in focus to serious AI risks with security implications, such as using AI for developing weapons, as opposed to safety issues and risks, e.g., bias and discrimination.

The U.K. has taken a decentralized, principles-based approach with cross-sector regulators expected to set binding guidelines and enforce the core principle set by the U.K. government. The development, integration and responsible governance of AI is a strategic priority across U.K. policymaking and regulatory capacity building with a focus on enabling existing regulators to enforce their core principles. An AI bill was announced in the King's Speech in July 2025, but it would only regulate the most powerful AI models. The timing and scope of such a bill has since changed, with no formal bill expected until the next King's Speech, reportedly in May 2026.

History and context

The U.K. has long played an important role in the development of AI. In the 1950s and 60s, the potential of AI-generated enthusiasm and expectation led to the formation of several major AI research centers in the U.K. at the universities of Edinburgh, Sussex, Essex and Cambridge. Even today, the U.K. is regarded as a center of expertise and excellence regarding AI research and innovation.

Fast forward to September 2021, when the U.K. government's National AI Strategy announced a 10-year plan "to make Britain a global AI superpower." That plan set the stage for ongoing consideration as to whether and how to regulate AI, noting, with emphasis, AI is not currently unregulated by virtue of other applicable laws. Since 2018, the prevailing view in U.K. law and policymaking circles has been that "blanket AI-specific regulation, at this stage, would be inappropriate" and "existing sector-specific regulators are best placed to consider the impact on their sector of any subsequent regulation which may be needed."

A consequence of the U.K. leaving the EU is that the EU AI Act does not directly apply in the U.K. as it does to the remaining 27 EU member states. However, the act does have extra-territorial scope that will certainly impact U.K. businesses. Indeed, the EU AI Act has accelerated and amplified independent U.K. policy development on whether, how and why AI should or could be regulated further and in ways more targeted than what exists, via the application of existing laws to AI.

The U.K. continues to forge its own path, instead focusing on flexibility, innovation and sector-specific regulatory expertise when it comes to AI regulation. The aim is to take a proportionate approach to regulation, with the government tracking AI development and only legislating where they deem it necessary.

In Tortoise Media's September 2024 Global AI Index, which benchmarks nations on their level of investment, innovation and implementation of AI, the U.K. maintained its ranking in fourth place, below the U.S., China and Singapore. The U.K. is "strong on commercial AI" and research but other countries are catching up fast, e.g., France, currently in fifth place, "now outperforms [the U.K.] on open-source [large language model] development and in other key areas including public spending and computing." This is, therefore, likely one of the many reasons why the U.K. agreed to the Tech Prosperity Deal with the U.S., obtaining an investment of over USD41 billion from U.S. businesses into U.K. AI infrastructure.

Approach to regulation

As general context, there is no draft or current U.K. legislation that specifically governs AI, except for a Private Member's Bill in the House of Lords, although such bills rarely become law. Instead, the U.K. government has relied on the existing body of legislation, which doesn't specifically regulate AI but undoubtedly applies to its development and deployment. For instance, the U.K. General Data Protection Regulation and Data Protection Act 2018 apply to AI. The government has also focused its efforts on soft law initiatives, e.g., cross-sector regulatory guidelines, to adopt an incremental, pro-innovation approach to AI regulation.

As already mentioned, an AI bill was announced in July 2024. However, due to the protracted legislative passage of the Data (Use and Access) Act 2025 — which was held up by unsuccessful attempts to include provisions relating to the use of copyright material to train AI — new assurances were sought on AI and copyright. The act includes a requirement for the secretary of state to report on the use of copyright works in the development of AI systems. The secretary of state must also report on the economic impact of the policy options proposed in the copyright and AI consultation paper by 19 March 2026. Any AI bill, expected in the second half of 2026 at the earliest, will likely deal with copyright matters as well as the most powerful AI models.

White paper on AI regulation and consultation response

In March 2023, the former U.K. Conservative party government published its white paper "A pro-innovation approach to AI regulation" for consultation, setting out policy proposals regarding future regulation.

Notably, the document does not define AI or an AI system but explains the concepts are characterized by adaptivity and autonomy, aligning with commonly accepted definitions of AI, as used in the Organisation for Economic Co-operation and Development and the EU's AI Act definitions of an AI system. It goes on to describe that the U.K.'s AI regulatory framework should be based on the following five cross-sectoral nonbinding principles: safety, security and robustness; appropriate transparency and explainability; fairness; accountability; and contestability and redress. Finally, the white paper does not propose the creation of a new AI regulator; instead, it advocates for the empowerment of existing regulators.

In February 2024, the government published its response to the white paper's consultation, which largely reaffirmed its prior proposals with one important caveat. The response indicated future legislation is likely to "address potential AI-related harms, ensure public safety, and let us realize the transformative opportunities that the technology offers." However, the government will only legislate when it is "confident that it is the right thing to do."

AI Opportunities Action Plan

In January 2025, the U.K. launched its AI Opportunities Action Plan, a strategic initiative aimed at leveraging the transformative capabilities of AI across multiple sectors, with the objective of establishing the U.K. as an AI superpower.

The plan is structured around three pillars: laying the foundations to enable AI and investing in AI infrastructure; promoting the adoption of AI particularly across the public sector, positioning it as the "largest customer and as a market shaper"; and securing the future of homegrown AI by positioning the U.K. as "national champions at the frontier of economically and strategically important capabilities."

However, the plan says very little about regulation and instead is much more focused on investment and infrastructure to encourage innovation and support and the growth of AI. As mentioned above, the recently announced Tech Prosperity Deal with the U.S. will fund some of the proposed investments into U.K. infrastructure.

More recent developments indicate the U.K. is moving away from the EU and its legislative approach and more towards the U.S. and an innovation approach with limited safeguards. With the economic rewards of AI at stake, this might not be entirely surprising. That said, the U.K., U.S. and EU are all signatories to the Council of Europe Framework Convention on Artificial Intelligence and Human Rights, Democracy and the Rule of Law, meaning equality must be respected and discrimination prohibited throughout the AI lifecycle. The developments over the next twelve months merits close attention.

UK regulator guidelines

U.K. regulators have continued to produce guidelines related to their own sectors. There has also been some cross-functional work on AI issues, such as with the Digital Regulation Cooperation Forum, which consists of the Information Commissioner's Office, Competition and Markets Authority, Ofcom and the Financial Conduct Authority. The DRCF is responsible for ensuring greater regulatory cooperation on online issues.

Data protection

The ICO has been actively regulating how data is used in connection with AI for a long time, updating their AI and data protection guidance in March 2023. In November 2024, the ICO published their audit outcomes report and recommendations for AI providers and developers of AI-powered sourcing, screening and selection tools used in the recruitment process. Following their consultation series on generative AI and data protection, the ICO published their outcomes report in December 2024.

In June 2025, the ICO announced their AI and biometrics strategy to ensure that AI and biometric technology is developed and deployed lawfully, responsibly and in ways that maintain public trust. The ICO recognizes the significant opportunities for innovation such technology presents but emphasizes that it must be used in ways that protect personal data and uphold individual rights. A number of guidelines are expected on key topics as part of this strategy.

Following the enactment of the Data (Use and Access) Act 2025, there are also a number of ongoing consultations and revisions to guidance. ICO's automated decision-making and profiling guidance is of particular interest; the consultation for which is expected to launch in fall 2025, with final guidance expected to be published in spring 2026.

Online safety

In March 2025, Ofcom released its guidance on applying the Online Safety Act to generative AI and chatbots in the form of an open letter.

Competition and markets

The CMA and ICO issued a joint statement regarding foundation model approaches in March 2025. The joint statement expressed the organizations' ongoing commitment to collaborate on various initiatives that enhance user autonomy and control, ensure fair access to data, and distribute accountability appropriately across the foundation model supply chain.

Other UK AI governmental/ parliamentary initiatives

Despite the U.K. government's continued presence at the global AI Action Summit, the focus has moved away from AI safety and towards "strengthening international action towards artificial intelligence." While safety is still on the agenda, it is no longer the primary focus of these summits. The U.K. government opted not to sign the official agreement produced at the February 2025 Paris AI Action Summit, expressing concerns around "global governance" and national security. The policy direction adopted at the next Summit, scheduled to be held in India in late 2025, will be of significant interest.

It is also worth noting that while only certain provisions of the EU AI Act currently apply in Northern Ireland, the European Commission has proposed to add the EU AI Act to the Windsor Framework, making it directly applicable as a whole to Northern Ireland. This process is ongoing, and it will be important to keep track of developments.

Separately, Conservative Peer Lord Holmes of Richmond re-introduced a Private Members' Bill, the Artificial Intelligence (Regulation) Bill, in March 2025. This bill is identical to the previous version introduced in the last parliamentary term. This compact document advocates for the formation of a standalone AI regulator and the new role of an AI officer for organizations that develop, deploy or use AI. Crucially, it is rare for Private Members' Bills to be passed into law. Instead, they are often intended to provide constructive policy recommendations or apply legislative pressure.

In January 2025, the Department of Science, Innovation and Technology published a voluntary Code of Practice for the Cyber Security of AI that sets the "baseline cyber security principles to help secure AI systems and the organizations which develop and deploy them," protecting them from cyber risks arising from "data poisoning, model obfuscation, indirect prompt injection and operational differences associated with data management." This was accompanied by a practical implementation guide. The U.K. government plans to submit the code and guide to the European Telecommunications Standards Institute to "be used as the basis for a new global standard … and accompanying implementation guide."

The Government Digital Service, which sits within DSIT, is also establishing a new Responsible AI Advisory Panel to help shape the U.K.'s approach to "building responsible AI in the public sector." The panel aims to ensure safe, ethical and responsible AI development by bringing together AI expertise from a wide range of organizations with a diverse skillset.

The U.K. government also launched an AI playbook in February 2025 to offer guidance and support to government departments and public sector organizations to safely, effectively, and responsibly harness the power of a wider range of AI technologies.

Wider regulatory environment

While the U.K. does not have legislation specifically governing AI, various broader statutes and case law applies to the area.

Data protection

From a data protection perspective, the U.K. legal system comprises the U.K. General Data Protection Regulation, the Data Protection Act 2018, the Privacy and Electronic Communications (EC Directive) Regulations 2003 (SI 2003/2426) and the Data (Use and Access) Act 2025, which amends the preceding legislation and brings into force the U.K.'s long awaited data reforms. In addition, the EU GDPR has an extra-territorial effect and likely applies to U.K. entities that process personal data relating to EU individuals.

The use of AI systems raises many compliance questions and potential trade-offs under U.K. data protection law. These issues range from establishing the roles of the data processing entities to ensuring the accuracy of personal data used in training, while also adhering to requirements for profiling, automated decision-making, and the data minimization principle.

It is important to note the provisions in the Data (Use and Access) Act 2025 on automated decision-making are perhaps one of the biggest changes this act will bring into force. While the provisions were not in effect as of press time, that is expected to change in 2026. Unless special category data is involved, the U.K. will move from a regime based on prohibition with exceptions for automated decision-making — currently only limited lawful bases can be relied upon for ADM — to one that is based on permission but with safeguards. Broadly speaking, all and any lawful bases can be used for ADM provided that safeguards are in place, such as the right to human intervention, the ability to contest the decision, and transparency about the logic and criteria used.

In addition, the act also clarifies what is meant by solely automated, from a U.K. perspective, at least and therefore possibly brings a few decisions previously thought of as ADM out of scope. This change, if enacted, could be a significant game changer for the use of AI decision making tools in the U.K. — making it easier and possibly therefore more widespread.

Intellectual property

The main types of intellectual property rights in the U.K. are registered and unregistered trademarks, patents, registered and unregistered designs, copyright, and trade secrets. The key U.K. intellectual property statutes are the Patents Act 1977, Copyright, Designs and Patents Act 1988, and Trade Marks Act 1994.

Copyright questions are relevant to AI, given the training data may include copyrighted works, e.g., books, news, academic articles, web pages, photographs or paintings. An AI system itself may create works that could potentially be protected under copyright, albeit there is uncertainty on this.

In January 2023, Getty Images began U.K. court proceedings against Stability AI, claiming infringement of various different intellectual property rights such as trademark, passing off, database rights, and multiple types of copyright. Getty alleged Stability AI scraped millions of images from its websites without consent and unlawfully used them to train and develop its deep-learning AI model, thereby infringing Getty's intellectual property. Getty dropped various claims at trial, with judgment expected sometime in fall 2025.

As discussed, the AI and copyright debate continues. Following assurances given so that the Data (Use and Access) Act could complete its parliamentary passage, it is expected that these issues will be addressed — initially in the secretary of state reports mentioned above, followed by the proposed AI (Regulation) Bill expected later in 2026.

Patent questions are also very relevant, including whether an AI system can be considered an "inventor" for the purposes of the Patents Act 1977. In December 2023, the U.K. Supreme Court dismissed an appeal from Stephen Thaler, affirming the Comptroller-General of Patents, Designs and Trademarks' decision that a machine, which embodies an AI system, could not be an inventor under the law. In September 2025, the High Court also dismissed Thaler's appeal against a U.K. Intellectual Property Office decision, ruling that it was not the judge's place to rule on "whether provision needs to be made requiring an AI-generated invention to be identified as such."

Online safety

In October 2023, the Online Safety Act entered into law. The act is intended to address two fundamental issues: the tackling of illegal/harmful online content and the protection of children online. It does so by imposing obligations, known under the law as duties of care, on a sliding scale for a broad range of online entities, such as social media networks, search engines, video-sharing platforms and marketplaces or listing providers.

Many of the OSA's substantive obligations, such as duties to protect users from illegal content, child protection duties and age assurance measures, are now in force after a phased implementation. The law imposes extensive requirements that will impact AI systems, including the monitoring for and takedown of AI-generated content that could be illegal or harmful, an increasing challenge when the use of AI by the general public is becoming commonplace.

Employment

From an employment law perspective, the Equality Act 2010 prohibits discrimination by employers on the basis of any protected characteristics, such as age, disability, race or sex.

Due to the nature of its training data and other factors, unless mitigation steps are taken, some AI systems have the potential to exhibit and/or perpetuate biases. The use of such systems for recruiting decisions and/or performance management could therefore raise U.K. employment law compliance considerations.

It is also important to note that the Trades Union Congress announced its pro-worker AI innovation strategy in August 2025, building on its draft AI (Regulation and Employment Rights) Bill and seeking to empower workers and promote responsible AI regulation where innovation is embraced alongside workers' rights. It remains to be seen whether the U.K. government's policy will be influenced by the TUC's work in this area.

Consumer protection

In terms of consumer protection, the U.K. has a patchwork of laws including the Consumer Rights Act 2015 and the Consumer Protection from Unfair Trading Regulations 2008 (SI 2008/1277). These interact with numerous AI use cases, like the information or guidance provided by chatbots to consumers or the sales contract terms between an organization and consumer for AI-related products and services. The CMA is active in this area, working in conjunction with other regulators as appropriate.

Product liability

From a product liability perspective, the key source of law is Part 1 of the Consumer Protection Act 1987. This implements the strict liability regime set out in the EU Product Liability Directive (2024/2853). In addition, individuals may have rights under the common law of tort. Complex issues are likely to arise regarding duties of care and liability assessments for defective AI systems. Early examples of how to tackle such issues were seen in connection with autonomous vehicles in the Automated and Electric Vehicles Act 2018.

Agentic AI

While the government has not specifically addressed agentic AI in any of the released policies, guidelines or laws, the broader principles on responsible and trustworthy AI are likely to be used to apply to the development of agentic AI.

Published September 2025

Global AI Governance Law and Policy: US

By C. Kibby, Richard Sentinella and Anthony Hilton

The U.S. lacks an omnibus federal law that specifically targets artificial intelligence governance. A market-driven approach of self-regulation has been traditionally preferred over government intervention when addressing emerging risks of privacy, civil rights and antitrust, reflecting an effort to foster competitive innovation.

As such, federal involvement in AI policy has mainly come from the issuance of agency guidance opinions when interpreting existing statute in the context of the usage of AI technology. Additionally, executive orders issued by the recent several presidential administrations have directed federal government policy and practice on AI governance, catalyzing a series of agency regulations focused on government use of AI.

The U.S. established the Center for AI Standards and Innovation, housed within the National Institute of Standards and Technology and aided by a consortium of over 280 AI stakeholders who support its mission.

Numerous states have proposed and, in some cases, enacted AI laws. Colorado was the first to enact comprehensive, state-level, AI regulation that focuses on algorithmic discrimination. California has enacted a series of legislation to address several of the key concerns that have risen since the advent of AI. Federal agencies, including the Federal Trade Commission, have made it clear their existing legal authorities extend to the use of new technologies, including AI.

History and context

The formal inception of AI as a field of academic research can be traced to Dartmouth College in Hanover, New Hampshire. In 1955, a group of scientists and mathematicians gathered for a summer workshop to test the idea that "every aspect of learning or any other feature of intelligence can be so precisely described that a machine can be made to simulate it."

Several broad strategic drivers guide the U.S.'s approach to federally regulating AI. At a national policy level, Congress and, to some extent, the current administration's agencies have deliberately taken a light-touch and business-friendly approach.

This is founded on three key motivations. The first is a desire to see U.S. companies retain and expand their global AI leadership, particularly in competition with China. The second is the thought that innovation, development and deployment are stifled by governmental involvement. The third motivation is a philosophical belief that market-driven solutions are better suited to identifying and addressing market concerns than government intervention.

The AI Action Plan released in July 2025 seeks to advance these inclinations, implementing policies that accelerate AI innovation in the U.S. by dismantling regulatory obstacles, building American AI infrastructure with leaner permitting and funding incentives to foster construction and skills training, and leading in international AI diplomacy and security by promoting AI exports to allies as a default and prioritizing military and cybersecurity AI innovation for rapid government adoption.

Tortoise Media's September 2024 Global AI Index ranked the U.S. first in the world for its AI talent, infrastructure, research and development, and commercial investment. The U.S. earns the silver medal in two metrics: first, it lags slightly behind Italy in AI operating environment category, which measures AI-related public opinion, labor mobility and treatment in legislative proceedings. Second, only Saudi Arabia has publicly announced more government spending on AI. However, in the time after the report's publication, attitudes in the public and private sectors have changed significantly. Lawmakers are working to develop strategies around emerging AI technologies in ways that keep the U.S. at the forefront of AI development and deployment.

Approach to regulation

The U.S. federal approach to regulating AI has primarily come from actions taken by the executive and legislative branches, supplemented by increasingly active state-level initiatives. The executive branch has focused on two primary strategies: the promulgation of guidelines and standards through federal agencies and industry self-regulation, referred to as regulatory sandboxes to foster flexible and innovative development.

For the most part, Congress has relied on existing legislation to adapt to the new challenges AI poses. This includes integrating AI concepts and applications into existing laws, such as civil rights, consumer protection, and antitrust, and bridging gaps as they go, rather than enacting an entirely new regulatory framework. However, states enacting their own AI legislation create a statutory patchwork of varying cross-jurisdictional rules and regulations for the private sector to navigate.

Executive Actions

On 23 July 2025, the Trump administration released America's AI Action Plan, a broad policy document focused on fostering U.S. AI development and innovation. This plan builds on multiple previously issued AI-related executive orders, the first of which came out in 2019 during the first Trump administration.

The plan lays out the Trump administration's vision for how the U.S. can win the global AI race, such as building energy infrastructure to power new data centers and supply chains necessary to run computationally intense models. The Trump administration sees AI as an economic engine; the website for the AI Action Plan states that "whoever has the largest AI ecosystem will set the global standards and reap broad economic and security benefits." While the plan lays out the government's vision for AI's economic impact and the support it needs, it is relatively limited on regulatory governance. Many of the executive orders signed by the president expand and clarify the administration's vision for AI.

During his first administration, Trump signed Executive Order 13859, Maintaining American Leadership in Artificial Intelligence. This executive order highlights AI's importance to national security, the economy, and public trust and establishes the American AI Initiative to guide policy. The plan focuses on driving research and development, improving access to federal data and computing resources, developing technical standards, training the workforce, promoting international and intersectoral cooperation, and protecting U.S. advantages from foreign threats. It tasks federal agencies with prioritizing AI in budgets, research, education and regulation, all under coordinated oversight.

In December 2020, Trump signed Executive Order 13960, Promoting the Use of Trustworthy Artificial Intelligence in the Federal Government. It encourages federal agencies to use AI "to improve Government operations and services in a manner that fosters public trust, builds confidence in AI, protects our Nation's values, and remains consistent with all applicable laws, including those related to privacy, civil rights, and civil liberties." The executive order promotes responsible use through principles like accuracy and resiliency and task-oriented leadership in agencies like the Office of Management and Budget and the Federal Chief Information Officers Council to develop guidance, criteria and use plans for AI.

Former President Joe Biden also signed several executive orders relating to AI, representing a preference for a push toward regulatory oversight, imposing requirements for safety testing and reporting. Executive Order 14141, Advancing United States Leadership in Artificial Intelligence Infrastructure, directs the Department of Defense and Department of Energy to identify sites on federal land on which to build AI data centers.

At the beginning of his second term, Trump signed broader executive orders containing policies and directives pertinent to AI. These orders repealed some of the prior administration's policies, such as Executive Order 14148, Initial Rescissions of Harmful Executive Orders and Actions, which retracts earlier executive orders with the aim of reducing regulatory burdens across sectors.

Trump later signed Executive Order 14275, Restoring Common Sense to Federal Procurement, which significantly reduces the Federal Acquisition Regulation with the goal of making procurement more efficient. Executive Order 14277, Advancing Artificial Intelligence Education for American Youth, creates a task force on AI education, creates a presidential AI challenge to encourage student adoption of AI, provides for AI training and professional development for teachers, and instructs the secretary of labor to develop AI-related registered apprenticeships.

On 23 January, Trump signed Executive Order 14179, Removing Barriers to American Leadership in Artificial Intelligence, which called for members of several agencies to create the AI Action Plan. On 23 July, the day the AI Action Plan was released, Trump signed additional executive orders that put some of the plan's key points into action.

Executive Order 14318, Accelerating Federal Permitting of Data Center Infrastructure, aims to streamline environmental permitting for AI data centers by simplifying steps and removing rules in the process. Executive Order 14320, Promoting the Export of the American AI Technology Stack, creates the American AI Exports Program to promote export of "full-stack AI technology packages," which include all of the hardware and software necessary to deploy AI from start to finish, like graphics cards, the model itself and training data.

The OMB

The OMB issued two AI memos in response to Executive Order 14179. According to a White House fact sheet, the first memo, Accelerating Federal Use of AI through Innovation, Governance, and Public Trust, "gives agencies the tools necessary to embrace AI innovation, while maintaining strong protections for Americans' privacy, civil rights, and civil liberties." It instructs federal agencies to increase their use of AI to innovate, "cut down on bureaucratic bottlenecks," and make the government run smoothly and efficiently while implementing risk management practices. To guide how the government acquires AI, the second memo, Driving Efficient Acquisition of Artificial Intelligence in Government, emphasizes prioritizing competition and American-made AI systems.

The FTC

At the U.S. AI Summit 2025 in June, FTC Commissioner Melissa Holyoak delivered a keynote speech about the rapid developments in the AI space and how they present novel antitrust enforcement challenges. The first step in determining if a company has a monopoly in a particular market is to define what the market is. She noted that AI is "a technology that is projected to be both a critical input to and potentially a competitor with almost every firm in the economy."

Holyoak further noted AI's widespread presence in nearly all markets makes it difficult to determine what companies or products it competes with. She raised concerns that AI companies are trying to draw in users and developers to their platforms with low cost or free access. Once those users have already locked into the AI's infrastructure, these companies could raise those prices in the future. Holyoak emphasized that the FTC and DOJ do not make proactive regulations, but instead intervene after a violation, enforce the laws, and issue guidance on how other companies can avoid breaking laws in the future.

In early 2025, the FTC released a preliminary staff report on large AI partnerships and investments, which provides insights on the "corporate partnerships and investments" that connect cloud computing companies with AI companies. The report finds that these partnerships could impact competition by, for example, attracting a large portion of AI talent and giving them access to "sensitive technical and business information that may be unavailable to others."

The FTC has filed many complaints and settlements with companies that use AI in allegedly anticompetitive ways. For example, in fall 2024 the FTC announced Operation AI Comply, an enforcement sweep against companies they claim misused "AI hype" to defraud consumers. These companies include several "passive income" e-commerce operations like Ascend Ecom that allegedly sold AI-powered software, inventory, and a spot in online marketplaces. Ascend Ecom promised to help the consumer earn tens of thousands of dollars a month in passive income, which never materialized. The FTC has settled with Ascend and others, including FBA Machine and its owner Bratislav Rosenfeld. The settlements usually permanently forbid the company or person from operating any kind of similar business.

Other enforcement actions similarly focus on companies falsely representing the capabilities of their AI. DoNotPay, the company promising to replace human lawyers with AI, faced a complaint and consent order prohibiting it from stating or implying that its products operate like a human lawyer. Another, Ryter, offered a service where consumers could generate unlimited product reviews that allegedly misled online shoppers; it has since been barred from selling any similar service.

Congress

While the U.S. lacks a comprehensive law designed to regulate AI, Congress has been active on the AI front. It has introduced targeted AI-related bills including the NO FAKES Act of 2024 and passed a raft of legislation including the AI Training Act, the National AI Initiative Act of 2020, the AI in Government Act of 2020, and the TAKE IT DOWN Act of 2025. While each of these has often been a lesser component of larger appropriations bills, their presence remains noteworthy. The scope of these measures mirrored executive branch actions designed to facilitate AI adoption within the federal government and achieve coordination among federal agencies in its application.

The NO FAKES Act, first introduced in 2024 and recently re-introduced on 9 April 2025 as S.1367, seeks to protect the voice and visual likeness of individuals from unauthorized digitally generated recreations, such as through the use of generative AI. The law, which would preempt state legislation in the same area, would require internet gatekeepers to remove unauthorized recreations or replicas of audiovisual works, images or sound recordings.

The AI Training Act requires the director of the OMB to create an AI training program for employees of executive agencies. The National AI Initiative Act of 2020, included within a larger budget law, creates the National AI Initiative Office, which oversees and implements the U.S. national AI strategy. The AI in Government Act of 2020, also within a budget law, creates the AI Center of Excellence, which facilitates AI adoption in the federal government. The TAKE IT DOWN Act of 2025 prohibits the online publication of nonconsensual intimate visual depictions, including computer-generated images, requiring online platforms to remove them within 48 hours of notification.

The federal approach contrasts sharply with state-level initiatives, as demonstrated by congressional consideration of imposing preemptive law in the One Big Beautiful Bill Act. Originally, the act included a moratorium on all state-level AI legislation enforcement for 10 years, further indicating the federal preference for self-regulation, but the Senate removed the provision with a vote of 99-1. The moratorium would have targeted laws that impose AI-specific duties on developers and deployers, including model registration, risk assessments, watermarking and disclosure rules, audits, and private rights of action.

Through the Senate's AI Insight Forum and bipartisan framework on AI legislation and the House of Representative's bipartisan Task Force on AI, members of Congress have continued to explore how the legislature should address the promises and challenges of AI. The proposals have ranged from establishing a licensing regime administered by an independent oversight body to holding AI companies liable for privacy and civil rights harms via enforcement and private rights of action. They additionally call for mandatory disclosures by AI developers to regarding information about the training data, limitations, accuracy and safety of their models.

State-level regulation

States have taken action to propose and implement comprehensive legislation to fill in the gaps where the federal government has elected to employ temperance or abstinence. The consequence has been a mosaic of differing and overlapping rules and regulations with varying degrees of minimum and maximum effect and limitations.

Colorado’s Artificial Intelligence Act, enacted in May 2024, represents the most comprehensive state level AI regulation to date. Initially slated to take effect 1 Feb. 2026, the date has been pushed back to 30 June 2026, pending governor approval, it requires developers and deployers of high-risk AI systems to implement risk management practices and conduct impact assessments to prevent algorithmic discrimination in consequential decisions that would affect housing, employment, education, health care, and other critical areas.

Other states, like California and New York, have taken a sectoral approach rather than a comprehensive one to AI regulation, targeting specific industries rather than having an umbrella regulatory scheme. In 2024, California Governor Gavin Newsom signed several legislative packages around AI, defining "artificial intelligence" (California Assembly Bill 2885) and addressing many of the risks arising from its use. For example, California lawmakers sought to ensure transparency through measures such as watermarking (SB 942) and the obligation for developers to publish documentation on training data for AI systems made available publicly on the internet.

The distribution of certain AI creations were criminalized, such as nonconsensual, intimate, or deepfake images (SB-926 and SB-981) and child sexual abuse materials (AB-

1831 and SB-1381). California also took steps to protect the acting profession and political transparency, obligating the entertainment industry to obtain consent from actors or their estates to replicate their image (AB-2602 and AB-1836). Bills also passed requiring the disclosure of AI-generated content in political advertisements during election periods (AB-2355 and AB-2839). Also enacted where series of consumer protection laws requiring the disclosure of AI-generated voices used for robocalls (AB-2905) and health care communications (AB-3030).

Continuing the sectoral approach, in January 2025, New York state enacted legislation, amending it's already existing General Business Law to impose safety regulation on AI companions, systems simulating ongoing human-like interactions. In January 2024, New York legislators passed legislation requiring state agencies to assess and oversee their own AI usage without human oversight. At press time, New York is currently considering more expansive legislation such as the RAISE Act, which would regulate "frontier AI models," establishing safeguards, reporting and disclosure obligations, and other requirements for large developers of frontier AI models.

Illinois has also targeted worker protection, enacting House Bill 3773 in August 2024, amending the Illinois Human Rights Act to include regulation for AI use in employment decisions. Effective 1 Jan. 2026, the law requires employers to provide notice when using AI for hiring, promotions, or terminations, and prohibits AI systems that discriminate based on protective characteristics.

The scope of state action is becoming extensive. In 2024, 700 AI legislative proposals were made, and 45 states, Puerto Rico, Washington D.C. and the U.S. Virgin Islands introduced AI bills; thirty-one states, Puerto Rico and the U.S. Virgin Islands enacted legislation or adopting resolutions. Such proactive legislation is not limited only to the state level as local municipalities weighed in as well. For instance, in 2023, New York City enacted NYC Local Law 144, which requires bias audits for AI tools used in employment decisions.

Self-regulation

In line with the U.S.'s long history of favoring a self-regulatory approach to industry, informal commitments have been a key policy pool in its regulatory approach to AI. In July 2023, Amazon, Google, Meta, Microsoft and several other AI companies convened at the White House and pledged their voluntary commitment to principles around the safety, security and trust of AI. These principles include ensuring products are safe before introducing them onto the market and prioritizing investments in cybersecurity and security-risk safeguards.

NIST's AI Risk Management Framework

Perhaps the strongest example of the U.S.'s approach to AI regulation within the paradigm of industry self-regulation is NIST's AI Risk Management Framework, released in January 2023. The AI Risk Management Framework aims to serve as "a resource to the organizations designing, developing, deploying or using AI systems to help manage the many risks of AI." To facilitate implementation of the framework, NIST subsequently launched the Trustworthy and

Responsible AI Resource Center, which provides operational resources, including a knowledge base, use cases, events and training.

NTIA's AI Accountability Policy

The National Telecommunications and Information Administration's Artificial Intelligence Accountability Policy also falls into the self-regulation category. The report provides guidance and recommendations for AI developers and deployers to establish, enhance and use accountability inputs to provide assurance to external stakeholders.

Agentic AI

The autonomous nature of agentic AI, used as automated tools for project and operations management, creates unique and regulatory challenges, particularly around accountability and liability. Traditional regulatory frameworks struggle to relevantly address any potentially harmful agentic AI decisions and actions because models are more efficient and better able than humans are to coordinate and manage multiple tasks across varying functions at once.

This raises questions about human oversight requirements and responsibility chains. At both the federal and state levels, the U.S. does not currently have specific legislation targeting agentic AI as a technology. Sector-specific legislation will likely apply to AI agents, especially as they might be used in highly regulated industries, such as finance, insurance, medicine, or employment. This has also been true of state laws in practice that apply to AI in these areas. U.S. agencies working on standards and regulation for AI will likely include considerations for agentic AI, such as when NIST will revise the AI Risk Management Framework.

Wider regulatory environment

This section covers regulatory actions and discussions from before the implementation of the AI Action Plan, which promises to pivot towards a more limited, market-driven approach to AI oversight. The material here remains relevant as context and record, but it reflects a different regulatory climate than the one shaping policy today.

Intellectual property

In the realm of intellectual property, efforts undertaken by the U.S. Patent and Trademark Office have centered on incentivizing innovation and inclusivity within AI and emerging technologies. The AI and Emerging Technologies Partnership program brings the USPTO together with the AI and emerging technologies communities from academia, industry, government and civil society. The partnership hosts listening sessions and provides public symposia and guidance at the intersection of AI and intellectual property.

The thorny copyright law and policy issues raised by AI have been on the radar of the U.S. Copyright Office for several years. Since its AI initiative launched in 2023, the office has held numerous public listening sessions and webinars. The Copyright Office also issued a notice of inquiry on copyright and AI to inform its future guidance.

Several important cases have been decided on or are in the courts. Regarding AI and the fair-use doctrine, the Dow Jones & Company along with NYP Holdings — publishers of the Wall Street Journal and New York Post — sued Perplexity AI for copyright infringement and trademark violations. The lawsuit claims Perplexity scraped copyrighted material without authorization, which the news outlets indicate harms their advertising and subscription profits. The news outlets also claim the results from Perplexity's system often include the exact text from the original news articles or attribute false information to the publications. The outcome of this case could set legal precedents regarding fair use and the application of copyright law to generative AI systems.

In a lawsuit brought against Anthropic by a group of authors, the courts ruled that although purchased books can be used to train AI models, pirated books may not as they do not fall under the fair use doctrine. The opinion likens machine learning models digesting books to a young author learning to write by reading books to better their technique. With the lack of clarity from Congress about the use of copyrighted materials to train AI, it is likely these lawsuits will have a significant impact on the applicability of the fair use doctrine. If lawyers decide the use of copyrighted materials without licensing infringes on the copyright, companies will either have to pay large fees for the infringement and/or retrain their models on datasets of licensed materials.

Employment

The Equal Opportunity Employment Commission's AI and Algorithmic Fairness Initiative was launched to ensure AI, machine learning and other emerging technologies comply with federal civil rights law. Through the initiative, the EEOC provided the public with information and guidance on the use of AI in making job decisions for people with disabilities, mitigating discrimination and bias in automated systems, and assessing the adverse impacts of the technology in employment decisions.

One of the first landmark cases to demonstrate the complexity of applying traditional anti-discrimination regulatory frameworks to algorithmic decision-making was EEOC v. iTutorGroup. The EEOC successfully sued iTutorGroup, a Chinese company hiring U.S.-based tutors to provide English language tutoring to adult students in China. The company used an automated hiring system that would reject female applicants aged 55 or older and male applicants aged 60 or older. The case resulted in a USD365,000 settlement, establishing liability obligations for employers using AI tools for employment purposes.

The second significant case, Mobley v. Workday, currently in litigation as of press time, is a class-action lawsuit originally filed by a worker that claimed he was discriminated against based on race by Workday's AI-powered job applicant screening system. This case is particularly significant because it addresses not just the employer, but the liability of the vendor providing the AI tools. While the judge did not find any intentional discrimination by the software provider, she did not rule out that the software did not discriminate against applicants and allowed the lawsuit to go forward. If the applicants succeed in their case against Workday, it could substantially increase the duty of care for human resources software providers who use AI in the hiring process.

These cases highlight a fundamental tension in AI regulation for employment protections. AI can perpetuate or amplify historical and systemic biases previously practiced by an organization, whether known or unrealized, as a consequence of the training data it receives.

AI in legal practice

AI technology presents a unique challenge for the legal profession, raising questions about ethical obligations around attorney-client privilege and confidentiality. The efficiency that AI offers, in areas such as mergers and acquisitions and litigation, is becoming more prevalent and too difficult to ignore. However, the issues go beyond just the complexities of general AI governance.

Attorney-client privilege traditionally protects confidential communication between lawyers and clients. Some matters which are extremely complex for a person — often due to heavy paper load requiring extensive amounts of time and effort to review and correlate — can be more efficiently processed and coordinated by AI systems and tools. The results are attractive cost savings for the client. The allure of such AI tools is therefore high.

AI systems, however, are not currently air-gapped and often require transmitting client information to third-party servers or cloud services. This is especially true for generative AI systems like ChatGPT or Claude. This can potentially breach the privilege right under the current rules, requiring legal professionals to navigate between the efficiency gains AI provides and the risk of inadvertent disclosure of privileged information.

Client information entered into AI prompts may be processed and used to train future models. In other words, the information is not stored in isolation; it becomes integrated into the AI's knowledge base and reasoning pattern. This raises the risk of unconscious application of insights gained from one client by the AI system when advising another client on matters involving competing interests.

For instance, AI excels at pattern recognition across the large data sets. By processing information from multiple firms, proprietarily developed legal strategies for deal structures and risk assessment could be unconsciously applied for the benefit of one client to the detriment of another simply because the firms involved are using the same AI system. An AI system could use this information to recognize industry trends, negotiation strategies and/or legal vulnerabilities that should have remained compartmentalized. The AI system then unwittingly applies it to the benefit of one party in competition with another, which represents an unfair intelligence transfer. The AI system could potentially combine privileged litigation strategy that was shared by a firm with public court filings and inadvertently reveal adverse parties' tactical approaches to trial and/or settlement positions.

Continuing the patchwork approach in the U.S., several state bar associations have issued guidance on AI use in legal practice. For example:

The Florida Bar issued Ethics Opinion 24-1 that states lawyers are allowed to use generative AI if they obtain a client's consent for use with their confidential information, investigate the AI systems security measures and retention policies to ensure privilege is maintained, and maintain direct oversight of all AI-generated work product for reliability and accuracy.

The State Bar of California issued written guidance for lawyers, requiring anonymization of client information when using AI systems, diligent security review with consultation from IT professionals of AI systems used, terms of use review to ensure client information is not used for training, and oversight of AI generated work product for reliability and accuracy.

The New York City Bar Association issued Formal Opinion 2024-5, providing guidance similar to that offered by California.

Until agentic AI can be properly and economically air-gapped, the concerns of privilege and conflict will remain for attorneys.

International strategy

In February 2025, the G7 countries created a voluntary AI reporting framework "to encourage transparency and accountability among organizations developing advanced AI systems." The framework came from the Hiroshima AI Process, a collaboration between the G7 to provide low-friction tools that can scale without binding regulation. The reporting framework invites developers of advanced systems to publish standardized reports tied to the HAIP code of conduct.

In his remarks at the Paris AI Action Summit on 11 February 2025, Vice President JD Vance urged countries to avoid "excessive regulation" and emphasized U.S. ambitions for AI growth; the U.S. and U.K. subsequently declined to sign the summit declaration focused on "inclusive and sustainable artificial intelligence."

In parallel, NIST’s Center for AI Standards and Innovation is coordinating technical work through a 280-plus member consortium on testing and standards and has cooperation agreements with leading model developers to support safety research. In January 2025, NIST and its Center for AI Standards and Innovation hosted a workshop for AI experts to "provide a comprehensive taxonomy" of agentic AI tools. NIST published "lessons learned" from the workshop in August, identifying two potential taxonomies of AI tools: one based on "what they enable the model to do," and the other focusing on what constraints limit the tool’s capabilities.

In May, the Department of Commerce rescinded the Biden-era AI Diffusion Rule, which limited exports of AI model weights and advanced chips based on a tiered country classification system. It required licenses for exporting to most countries, with potential exceptions for allied countries and presumptive license denial for countries like China and Russia. Instead, DOC stated that it would issue a replacement rule in the future with fewer sweeping regulations.

Latest developments

In the U.S., the few law and policy developments related to AI are in the acceleration phase. Here's a limited preview of what to expect in the near future.

- New AI Risk Management Framework: The AI Action Plan instructs NIST to revise the AI Risk Management Framework and develop a 2025 National AI Research and Development Strategic Plan. The period for comment on this new plan has closed.
- Congress watchlist: The 119th Congress has proposed several bills that impact AI, including the following:
  - The CREATE AI Act would increase access to AI research and development tools.
  - The No Adversarial AI Act would bar federal use of AI from adversary countries.
  - The TEST AI Act would set up NIST AI testbeds.
  - The NO FAKES Act would create a federal right against unauthorized AI replicas of one's voice or likeness.
- OMB timelines: The OMB memos require CFO agencies to publish an AI strategy and file public compliance plans within 180 days of 3 April 2025; agencies must then continue to update these plans every two years until 2036. The agencies must also update internal data privacy policies and issue AI use policies within 270 days. They must maintain public AI use cases annually.

Future outlook

The U.S. federal government's market-driven approach is intended to encourage rapid innovation and competitiveness in the world AI market. While other jurisdictions forge forward with comprehensive rules and requirements, like the EU AI Act, the U.S. has elected to leave the issues of systemic risk management to voluntary self-regulation. The practical impact of these different approaches will become clearer as industry practices evolve and as policymakers assess whether existing frameworks adequately address emerging challenges.

Contacts

Joe Jones
Director of Research and Insights, IAPP
jjones@iapp.org

Will Simpson
Westin Research Fellow, IAPP
wsimpson@iapp.org

Heather Domin
Vice President, Head of Office of Responsible
AI and Governance, AIGP
heather.domin@hcltech.com

For further inquiries, please reach out to research@iapp.org.

Follow the IAPP on social media

Published December 2025.

IAPP disclaims all warranties, expressed or implied, with respect to the contents of this document, including any warranties of accuracy, merchantability, or fitness for a particular purpose. Nothing herein should be construed as legal advice.