ACSH
Partnership for Excellence

Ministry of the Interior and Safety
REPUBLIC OF KOREA

NIA
NATIONAL INFORMATION SOCIETY AGENCY

UNDP

GLOBAL APPROACHES TO AI GOVERNANCE: POLICY, LEGAL, AND REGULATORY PERSPECTIVES

United Nations Development Programme
Astana Civil Service Hub
2025

Global approaches to AI Governance: Policy, Legal, and Regulatory Perspectives

United Nations Development Programme
Astana Civil Service Hub
2025

This publication is the result of the “Capacity Development of Public Servants for Advancing Digital Transformation and Digital Governance” Project, funded by the Government of the Republic of Korea, and jointly implemented by the United Nations Development Programme in Qazaqstan and the Astana Civil Service Hub, in close partnership with the Ministry of the Interior and Safety and the National Information Society Agency of the Republic of Korea. The study focuses on policy, legal, regulatory and ethical aspects of artificial intelligence governance, how these are manifested in international and national legal and regulatory frameworks. It also provides examples of countries around the world examining their approaches to incorporating artificial intelligence in their realms.

Please cite this publication as:
ACSH (2025). Global approaches to AI governance: Policy, Legal, and Regulatory Perspectives. Astana: United Nations Development Programme. – 103 pages.

ISBN 978-601-12-4628-6

Executive summary

The transformative nature of Artificial Intelligence (AI) is having a profound impact on governments and societies nowadays. The ability of machines (software) to perform such tasks as learning, reasoning, problem-solving, and decision-making, as well as image analysis, speech recognition and natural language understanding – tasks that usually require human intelligence – has allowed for AI to penetrate into many and multifaceted aspects of human activities. Hence, AI is rapidly becoming deeply embedded in contemporary societies permeating all walks of life, in effect reshaping industries, economies, and societies and it is influencing the ways humans interact with technology. In daily life, AI shapes how we shop, learn and communicate.

AI is driving automation by reducing routine work; and most probably boosting efficiency and productivity. AI is also enabling the use of tools such as voice assistants and recommendation systems. It also powers advanced systems in finance, transportation, healthcare, education, public administration, etc. For instance, in healthcare, it supports early diagnosis of ailments and provide optimal assistance in developing personalised treatment to fight them. In business, it optimises supply chains, detects fraud, and enhances customer experience. It is also used for deriving predictive analysis and optimising processes. In science, it conducts research by performing big data analyses providing insights from such massive amounts of data that humans could never process alone, thus improving the quality of decision-making processes and outcomes. In public administration, AI is being increasingly integrated in public service delivery performing tasks which range from using AI to improve traffic management, and disaster response to deploying chatbots for citizen services.

Evidently, artificial intelligence is affecting the ways governments operate and interact with citizens. For instance, AI powered chatbots and virtual assistants provide instant responses to citizens' queries, thus considerably reducing bureaucratic delays and making public administration smarter. Furthermore, governments, by analysing large datasets in healthcare, education, traffic management, weather forecasting, crime prevention, etc they can derive better policy-related predictive insights and thus they can achieve better levels of planning and deploying scarce resources. In addition, AI can also assist in detecting irregularities in public spending, tax collection, and welfare programmes, thus reducing corruption and increasing accountability. Moreover, AI has proven very useful in-service delivery, from processing license application to automating social welfare schemes and policies and speeding up routine administrative work. AI has proved useful in disaster management too. AI systems can predict natural disasters, monitor relief distribution, and ensure timely responses in emergencies. In sum, AI is making public administration more efficient, transparent and more citizen-centred.

Obviously, AI has the potential to further enhance, and accelerate human development and improve daily life, and it offers a multitude of opportunities for improving economic and social life, as well as assisting in addressing many global challenges, such as climate change, and the lack of access to quality education and healthcare. On the other hand, however, such remarkable advances also raise challenges that are concerned with job displacement,[1] data

[1] Although, on one hand, new jobs are emerging associated with AI development, data sciences, robotics, and ethics oversight; on the other, routine and low-skill jobs are becoming vulnerable, e.g., data entry, customer support, factory work, etc. This situation is calling for placing considerable emphasis on education and acquisition of new skills and knowledge to offset the loss of lower-skill jobs.

privacy and security, surveillance, ethical use, algorithmic biases,[2] etc that also need to be addressed in this wholesome and rapid transformation process, which call for strict operational guidelines to safeguard privacy and ensure that AI systems do not compromise public trust.

In this context, policy makers should be well-versed in balancing innovation and regulation to ensure that fairness, ethical standards and accountability prevail in AI utilisation. What should always be kept in mind is that artificial intelligence is not a replacement to human decision-making but a powerful tool that should enhance efficiency, accountability, transparency, and trust. AI, if implemented responsibly, it can significantly strengthen governance and improve quality of life for all.

This evolving environment has prompted governments to react by establishing national policies to define boundaries and regulate AI for ensure that AI systems are fair, accountable, transparent and trustworthy. Many countries have embarked on developing and implementing national AI strategies, and governance, policy and regulatory frameworks. Some have been early starters. Canada and Finland developed their national strategies in 2017. They were closely followed by Japan, France, Germany and the United Kingdom in 2018, and the Republic of Korea in 2019. Other countries, such as Brazil, Egypt, Hungary, Poland, and Spain launched a national AI strategy more recently. Several other countries are currently in AI policy consultation and development processes (Galindo et al., 2021).[3]

Countries such as Canada, China, India, and the United States have launched comprehensive strategies that outline ethical principles and frameworks for safe adoption, and targeted investment and funding. The European Union has gone a step further with the introduction of the EU AI Act, the first major attempt at a legal framework for AI, which categorises AI systems by risk level and imposes strict restrictions on high-risk applications such as in healthcare, law enforcement and employment. These initiatives aim at striking a balance between fostering innovation and protecting citizens from potential harms.

This publication has been developed with the aim to provide an overview of the efforts governments around the world are making to establish clear legislative rules, regulations and operational guidelines to ensure that AI is integrated responsibly into government, business and society aimed ultimately at enhancing overall quality of life for humanity. It focuses on policy, legal, and regulatory perspectives, also taking into consideration ethical aspects of the artificial intelligence overall governance domain. It also provides examples of countries around the world examining their approaches to incorporating artificial intelligence in their governance realms.

In this context, the study delves into selective country cases describing and discussing their AI governance models, legal and regulatory frameworks in place or in the making, the development of the necessary infrastructure, and the partnerships between the public and private sector in furthering the development and application of AI, as well as the capacity building and workforce development measures needed to ensure success. Countries covered are Canada, the Republic of Korea, United Kingdom, Qazaqstan and the Philippines, as well as some Project participating countries: Armenia, Azerbaijan, Bangladesh, Cambodia, Georgia, Kyrgyzstan, Lao PDR, Mongolia, Tajikistan and Uzbekistan.

[2] Especially, in sensitive areas, e.g., criminal justice, human resource selection and recruitment, where algorithmic bias can lead to null decisions.

[3] The development of national policies and strategies focusing specifically on AI is a new phenomenon. The OECD AI Policy Observatory (OECD AI) comprises 620 national AI policies from over sixty countries and the European Union (EU), and it is continually updated with changes occurring constantly. https://oecd.ai/en/dashboards/overview

Such countries' experience shows that the effort of governments worldwide to establish AI regulations and guidelines reflect a growing consensus: the future of AI must be shaped thoughtfully and responsibly. Rather than allowing technology to evolve unchecked, it is assumed that policy makers are working to ensure that AI enhances human well-being, protects fundamental human rights, and supports inclusive development. For this reason, governments are also putting safeguards in place to manage the risks associated with AI. These include oversight bodies to monitor compliance with established rules and regulations, mechanisms to address algorithmic bias and discrimination, in order to prevent misuse while encouraging innovation. If they are implemented prudently, these measures will not only safeguard society from potential risks but also harness AI as a powerful tool to improve governance, business, and everyday life, ultimately enhancing the quality of life for humanity.

Such fervent activity cannot be accomplished without global cooperation, as it is essential given AI's borderless nature, where AI development and data flows transcend national boundaries. Without international cooperation, it is plausible that fragmented and contradicting regulations may lead to conflicts, inefficiencies, and loopholes. It is also plausible that lack of cooperation may impact privacy, fairness, freedom of expression, labour rights, etc among other. Particularly in the area of ethics a unified approach is needed to ensure that core human values are respected globally.

Global cooperation can help alleviate the creation of a digital divide between developed and developing nations by nourishing technology transfers and sharing innovations allowing for a fair distribution of benefits across the world. Additionally, global cooperation on AI can help in combatting disinformation, deepfakes and cyberattacks. Setting unified global standards can detect, prevent and penalise misuse. Unified frameworks need to also be standardised using common technical, safety and ethical standards so they can be interoperable and trustworthy. The global effort to regulate AI reflects a recognition that while AI offers enormous potential to improve governance, business, and society, it must be guided by rules that protect human dignity and equity. With proper governance, AI can become a tool not just for technological progress but for advancing social well-being and global development.

This study presents and discusses various schemes of international, regional and intra-national cooperation on AI that exist nowadays. Initiatives such as the Global Partnership on AI (GPAI), or the UN High Level Advisory Body on Artificial Intelligence (HLAB-AI), as well as the Global Digital Compact (GDC) and the Pact for the Future bring together multiple stakeholders – from industry, academia, government organisations and international bodies – to promote responsible AI development and encourage knowledge-sharing across borders. This spirit of cooperation acknowledges that AI challenges and opportunities are global in nature, requiring collective solutions rather than isolated national efforts.

The publication is commendable for its broad examination of global AI governance across the four pillars of policy, legislation, regulation, and ethics. Given that it was produced through collaboration among key international and domestic institutions (UNDP, ACSH, MOIS, NIA), the report carries substantial credibility and policy relevance. A major strength of the report is its balanced treatment of the efficiency and transparency gains from AI use in public administration and public services, alongside the ethical challenges such as data privacy, algorithmic bias, and technological dependence.

It is a publication aimed at policy-makers and practitioners who are interested in widening their knowledge and experience on how AI has been integrating into the policy and legal systems and how they are implemented of numerous countries around the world and what governments are doing to continually engage in a global dialogue that allows them to learn from each and establish universal standards for harnessing artificial intelligence potential benefits and good governance mechanisms to offset potential challenges.

Byungjoon Park
Director
Ministry of the Interior and Safety
Republic of Korea

Alikhan Baimenov
Chairman
Steering Committee
Astana Civil Service Hub

Acknowledgements

This publication is the result of the **"Capacity Development of Public Servants for Advancing Digital Transformation and Digital Governance"** Project, funded by the Government of the Republic of Korea. The Project is jointly implemented by the United Nations Development Programme in Qazaqstan and the Astana Civil Service Hub, in close partnership with the Ministry of the Interior and Safety and the National Information Society Agency of the Republic of Korea. It showcases the collective efforts and shared commitment of all partner institutions to advancing strategic partnerships and promoting evidence-based knowledge on pertinent topics in the public sector.

We extend our deep appreciation to the research team at the Hub, including Project Expert **Tolkyn Omarova**, and Project Associates **Diana Sharipova** and **Amanzhol Yestekov**, with the support of **Panos Liverakos**, Research and Knowledge Management consultant, and **Bakyt Zhexembay**, Project Manager. Their dedicated efforts in research planning, data collection, analysis, and reporting have significantly contributed to the success of this project. Special thanks are also extended to the ACSH management team, who devoted considerable time and effort to coordinate the production of this publication.

A special acknowledgment is extended to the representatives from the Ministry of the Interior and Safety, namely Deputy Director **Taera Choi** and Assistant Director **Junghyeon OH** and from the National Information Society Agency, including Executive Principals **Myungha Hong** and **Changhee Yun**, for contributing with their expertise and technical support throughout the preparation of the study. We are grateful to the Government of the Republic of Korea for its generous financial and substantive contribution for accomplishing this piece of work.

Last, but not least, heartfelt thanks are extended to all collaborators and participants from project beneficiary countries for dedicating their time and interest to enhancing the understanding of AI governance development in their respective countries. Each and every one of them has played a crucial role in turning this initiative into a reality.

About the Ministry of the Interior and Safety (MOIS)

The Ministry of the Interior and Safety (MOIS) is responsible for general affairs of the State Council, promulgation of Acts and subordinate statutes and treaties, government organisation and prescribed number of public officials, awards and decorations, government innovation, administrative efficiency, digital government, personal information protection, management of government buildings and support for elections and referendums. Furthermore, MOIS actively promotes local autonomy and decentralisation by supporting business, finance and taxation of local governments and mediating disputes among local governments. In addition, MOIS takes charge of establishing, supervising, and adjusting policies related to safety and disaster management such as emergency countermeasures, civil defence, and disaster prevention.

With the mission of leading government innovation for a sustainable future by providing more integrated and customised services for the citizens, promoting ethical and efficient use of digital technology, and strengthening private-public partnerships, MOIS has been leading the digital transformation of the public sector and collaborating with its partner countries to build a better digital society for all around the globe.

More information at https://www.mois.go.kr/eng/sub/a02/aboutMinistry/screen.do

About the National Information Society Agency (NIA)

The National Information Society Agency is a public institution founded by the Framework Act on Intelligent Informatisation (Article 12), under the Ministry of Science and ICT and the Ministry of the Interior and Safety of Korea. Since its establishment in 1987, it has played a leading role in promoting national informatisation and the digital government of Korea. It has administered important ICT policies and infrastructure projects, including the Master Plan for the National Basic Information System (1987), the establishment of the Super High Speed Information Network (1995). 11 Initiatives and 31 Tasks for e-Government (2001 and 2003 respectively), the enactment of the Act on Promotion of the Provision and Use of Public Data (2013), and the establishment and promotion of the 5G+ Strategy and the National AI Strategy (2019).

Today, NIA leads Korea's efforts to integrate artificial intelligence into public administration, advancing the implementation of the National AI Strategy and promoting safe, ethical, and human-centered AI use. The Agency develops data-driven public services, supports AI governance and standardization, and fosters trust in emerging technologies. Through both domestic initiatives and international cooperation, NIA continues to position Korea as a responsible and inclusive leader in the global AI ecosystem.

More information at https://eng.nia.or.kr/site/nia_eng/04/10401000000002016093002.jsp

About the United Nations Development Programme (UNDP)

The United Nations Development Programme (UNDP) is the leading United Nations organisation fighting to end the injustice of poverty, inequality, and climate change. Working with its broad network of experts and partners in 170 countries, it helps nations to build integrated, lasting solutions for people and the planet.

More information at www.undp.org

About the Astana Civil Service Hub (ACSH)

The Astana Civil Service Hub is a flagship initiative of the Government of Qazaqstan and the United Nations Development Programme. It was created in 2013 by 5 international organisations and 25 countries: now comprising 45 participating countries. The geographical range of its participants stretches from the Americas and Europe through the CIS, the Caucasus, and Central Asia to ASEAN countries, demonstrating that partnerships for civil service excellence is a constant and universal need for all nations.

Its mandate is to assist in the promotion of public service effectiveness by supporting the efforts of governments of the participating countries in building institutional and human capacity; and thus, contributing to the improvement of civil service systems in the countries of the region and beyond.

The Astana Civil Service Hub is a multilateral institutional platform for the continuous exchange of knowledge and experience in the field of public service development, aiming at supporting government in the region and beyond through fostering partnerships, capacity building and peer-to-peer learning activities, and evidence-based research.

More information at www.astanacivilservicehub.org.

Table of Contents

List of Boxes

List of Figures

List of Tables

Abbreviations & Acronyms

ACSH	Astana Civil Service Hub
ADB	Asian Development Bank
ADGSOM	ASEAN Digital Senior Officials' Meeting
ADM	Automated Decision Making
ADPPF	ASEAN Data Protection and Privacy Forum
AIESIA	Agencia Española de Supervisión de la Inteligencia Artificial
AI	Artificial Intelligence
AIA	Artificial Intelligence Act / Algorithmic Impact Assessment
AICA	AI Competence Alliance
AIDA	Artificial Intelligence and Data Act / Artificial Intelligence Data Annotation
AIDG	Artificial Intelligence and Data Governance
AIRR	AI Research Resource
AISI	AI Safety Institute
AI-4RP	AI-Powered Weather Forecasting for a Resilient Philippines
APEE	Portuguese Business Ethics Association
ASEAN	Association of Southeast Asia Nations
ATRS	Algorithmic Transparency Recording Standard
AU	African Union
AWS	Amazon Web Services
CAD	Canadian Dollar
CADT	Cambodian Academy of Digital Technology
CAISI	Canadian AI Safety Institute
CDEI	Centre for Data Ethics and Innovation
CEIMIA	International Centre of Expertise in Montreal on Artificial Intelligence
CIFAR	Canadian Institute for Advanced Research
CIS	Commonwealth of Independent States
CMA	Competition and Markets Authority
COMELEC	Commission of Elections
DCA	Department for the Coordination of Algorithmic Oversight
DFC	Digital Family Card
DOST	Department of Science and Technology
DPA	Data Protection Authority
DRCF	Digital Regulation Cooperation Forum
DSIT	Department for Science, Innovation and Technology
DTF	Digital Transformation Framework
DTI	Department of Trade and Industry
E-CAIR	Education Centre for AI Research

EDIH	European Digital Innovation Hub
EGDI	E-Government Development Index
EHRC	Equality and Human Rights Commission
ETRI	Electronics and Telecommunications Research Institute
EU	European Union
FCA	Financial Conduct Authority
FIMI	Foreign Information Manipulation and Interference
FOC	Freedom Online Coalition
FRAIA	Fundamental Rights & Algorithms Impact Assessment
GDC	Global Digital Compact
GDP	Gross Domestic Product
GDPR	General Data Protection Regulation
CID	Cambodian Institute for Democracy
GBP	Great Britain Pound
GITA	Georgian Innovation and Technology Agency
GPAI	Global Partnership on Artificial Intelligence
GPT	Generative Pre-training Transformer
GPU	Graphics Processing Unit
HLAB-AI	High-level Advisory Board on Artificial Intelligence
HPC	High Performance Computing
HR	Human Resources
HRIA	Human Rights Impact Assessment
ICO	Information Commissioner's Office
ICT	Information and Communication Technology
IDB	Inter-American Development Bank
IDFI	Institute for Development of Freedom of Information
IEC	International Electrotechnical Commission
IEEE	Institute of Electrical and Electronics Engineers
IRCC	Immigration, Refugees and Citizenship Canada
IRIM	Institute for Regional and International Mobility
ISBN	International Standard Book Number
ISED	Innovation, Science and Economic Development
ISO	International Standards Organisation
ITU	International Telecommunications Union
KISTI	Korean Institute of Science and Technology Information
KRW	Korean Won
KazLLM	Kazakh Language Learning AI Model
KZT	Qazaqstani Tenge
LCO	Law Commission of Ontario
LLM	Large Language Model
MDDIC	Ministry of Digital Development and Communications

MLMD	Machine Learning Enabled Medical Devices
MLOps	Machine Learning Operations
MOIS	Ministry of the Interior and Safety
MoU	Memorandum of Understanding
MPTC	Ministry of Post and Telecommunications
MSME	Micro-, Small-, and Medium-Enterprises
MSP	Managed Service Provider
MTC	Ministry of Technology and Communications
NAIP	National Artificial Intelligence Platform
NAISR	National AI Strategy Roadmap
NATO	North Atlantic Treaty Organisation
NCSC	National Cybersecurity Centre
NEDA	National Economic Development Authority
NHS	National Health Service
NIA	National Information Society Agency
NIT	National Information Technologies
NLP	Natural Language Processing
NMT	Neural Machine Translation
NPC	National Privacy Commission, Philippines
OECD	Organisation for Economic Cooperation and Development
OSFI	Office of the Superintendent of Financial Institutions
PCAIS	Pan-Canadian Artificial Intelligence Strategy
PPP	Public-Private Partnership
R&D	Research and Development
RAG	Retrieval-Augmented Generation
RAM	Readiness Assessment Methodology
RMF	Risk Management Framework
RIO	Regulatory Innovation Office
RTAU	Responsible Technology Adoption Unit
SCC	Standards Council of Canada
SDG	Sustainable Development Goal
SEAMEO	Southeast Asian Ministers of Education Organisation
SIF	Strategic Innovation Fund
SLM	Small Language Model
SME	Small- and Medium-Enterprises
SPARK	Science and Technology (S&T) Policy Analytics Reasoning Model
STEM	Science, Technology, Engineering, and Mathematics
TEF	Testing and Experimentation Facility
TESDA	Technical Education and Skills Development Authority
TIST	Transatlantic Initiative on Sustainable Trade
TTC	Trade and Technology Council

TVETPH	Technical-vocational, Education and Training Philippines
UAE	United Arab Emirates
UN	United Nations
UN DESA	United Nations Department of Social and Economic Affairs
UNDP	United Nations Development Programme
UNESCO	United Nations Educational Scientific and Cultural Organisation
UNIDO	United Nations Industrial Development Organisation
US	United States
USAISI	United States Artificial Intelligence Safety Institute
USD	United States Dollar
WBL	Work-Based Learning
WEF	World Economic Forum
WG-AI	Working Group for Artificial Intelligence
WPSR	World Public Sector Report

Glossary of terms

In this publication several terms are used to convey somewhat similar conceptual connotations, however, their selective usage is guided by certain subtle distinctions that differentiates them.

AI governance frameworks denote structured systems of principles, policies, and practices that guide responsible and ethical development and deployment of artificial intelligence. They guide organisations in developing, deploying, and monitoring AI systems to ensure they are compliant with regulations in place. Overall, they provide guidelines for risk management, ethical decision-making, and data privacy throughout the AI lifecycle.

AI policy frameworks emphasise structured systems that provide guidelines for the responsible development of artificial intelligence based on principles like fairness, accountability, transparency, and security. In other words, they provide the structure for organisations and government to ensure that AI systems are fair, transparent, secure, and compliant with regulations by establishing clear guidelines, assigning accountability, and managing risks. Examples include the OECD AI Principles, the EU AI Act, and the NIST AI Risk Management Framework.

AI regulatory systems refer mostly to the legal frameworks and guidelines established to oversee the development and deployment of artificial intelligence technologies.

AI oversight bodies are structured organisations responsible for monitoring, evaluating, and regulating the development and deployment of AI systems, ensuring responsible and ethical use while safeguarding against risks such as bias, privacy issues, and lack of accountability.

AI safety institutions have a mandate to advance the science of AI safety to enable responsible AI innovation by developing methods to assess and mitigate risks of advanced AI systems. Their work includes creating benchmarks, evaluation tools, and safety guidelines for AI models and applications. Examples are the U.S. AI Safety Institute (USAISI) or the UK AI Safety Institute.

Capacity building is the process of developing and strengthening the skills, knowledge, and resources of individuals and organisations, in order to improve their ability to perform and adapt to evolving conditions. It involves improving individual skills and knowledge, strengthening institutional structures and processes, and developing the systemic policy framework for individuals and organisations to operate within. Common initiatives include training programmes.

Capacity development is the process of individuals and organisations strengthening their ability to achieve their development goals over time. It involves enhancing skills, knowledge, and systems to improve performance and includes elements like technical skills, functional capacities, and transforming mindsets. Capacity development can be a core part of development cooperation and is supported by various international organisations, such as the UN.

1. Global AI governance and cooperation

It is explicit that AI's transformative power makes it not just a technological issue, but a global governance challenge**;** similar to climate change or nuclear safety. AI reshapes how states deliver services and make decisions, making it imperative for new forms of collaboration across borders, sectors, and institutions. Collaboration schemes that promote adaptable governance frameworks to navigate the rapid advancements and complexities of artificial intelligence (A), as its development, use, and impact are global issues extending beyond national borders.

Ongoing cooperation efforts strive to address AI development and the diverse opportunities and challenges it presents. For this purpose, a range of initiatives has sprung up over the past few years at the global, regional and multilateral levels. For instance, at the international level, organisations such as the United Nations, the World Bank, and UNESCO are undertaking efforts to create frameworks and guidelines for safe and trustworthy AI. In fact, the UN convened a multistakeholder High-level Advisory Body on AI (HLAB-AI) to analyse and advance recommendations for the international governance of AI.[4]

At the regional level, the European Union (EU), the Inter-American Development Bank (IDB), and the African Union (AU) are also playing a crucial role. They are creating harmonised AI policies and standards that reflect the unique socio-economic context of their member states. For example, the European regulation on AI and the African Union's proposed continental strategy on AI demonstrate significant progress towards cohesive regional AI governance frameworks.

In addition, multilateral initiatives fill the gaps between global and regional efforts on AI governance. For example, the GPAI, the OECD AI Principles, the Council of Europe Convention on AI, NATO's AI Strategy, the EU-US Trade and Technology Council (TTC), Globalpolicy.AI, G7 and G20 initiatives, AI safety and governance summits, and the AI safety institutes network. These fora facilitate dialogue and collaboration among countries and regions, promoting an integrated approach to AI governance. In 2023, the G7 launched the "*Hiroshima AI Process*" in Japan, marking a milestone in AI governance. The same year, the first AI Safety Summit in the UK resulted in the Bletchley Declaration, signed by 28 countries, focusing on the safety of advanced AI systems. Globalpolicy.AI is an online platform developed to strengthen co-operation between intergovernmental organisations with complementary mandates on AI (OECD, 2025).

1.1 The Role of the UN and its Specialised Agencies

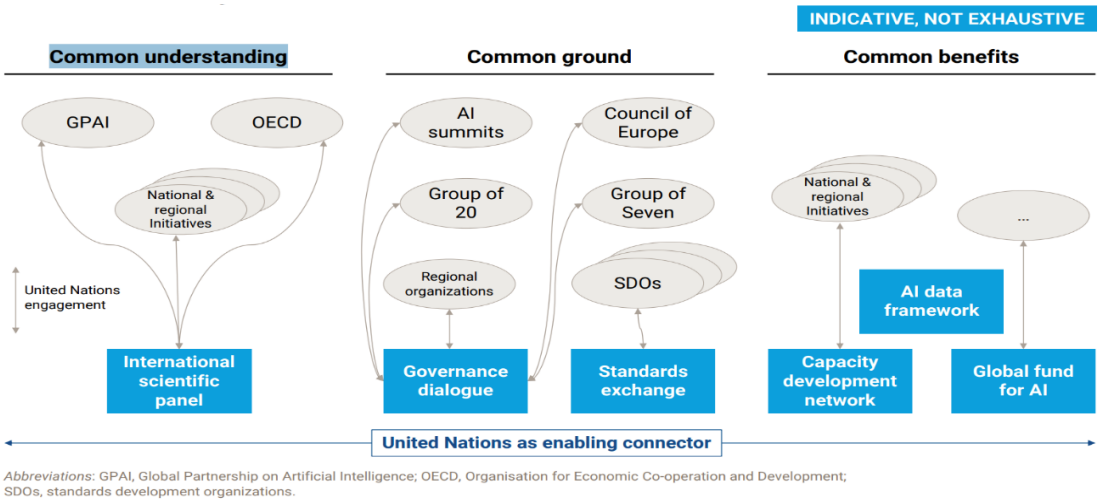
In this evolving AI landscape, the UN and its specialised agencies have become pivotal actors – not only in articulating global ethical standards, but also in supporting countries in addressing the legal, institutional, and ethical complexities of AI governance. In other words, the UN plays a dual role: (i) as a norm-setter, it defines global ethical principles and human rights-based frameworks for AI governance; and (ii) as an implementation enabler, it provides countries, especially developing and emerging economies, with technical assistance, institutional diagnostics, and capacity-building support.

[4] See also sub-section 1.2 for more details on the HLAB-AI.

The UN system acts as a bridge between global principles and national implementation. Its entities provide technical assistance, institutional diagnostics, and hands-on capacity development that align national strategies with globally accepted standards. From universal declarations and human rights frameworks to operational guidance, the UN is working to build inclusive, accountable, and development-focused AI governance ecosystems.

Recognising the unprecedented speed, global diffusion, and fragmented regulatory nature of AI development, the UN Secretary-General has launched a series of strategic initiatives to consolidate international efforts and promote coherence in global governance. These initiatives position the UN system at the forefront of shaping AI governance that is ethical, inclusive, and globally coordinated. Translating global AI governance frameworks into actionable policies requires cross-cutting support from various UN entities. While the above initiatives establish global direction, UN entities and special agencies help countries operationalise principles through targeted capacity-building, diagnostics, and institutional reform.

Figure 1. Role of the United Nations in the International AI Governance System



Source: United Nations (2024), Governing AI for Humanity, p. 19.

Among them, the United Nations Development Programme (UNDP), as the UN's principal development entity, plays a central role in embedding AI within broader digital and development agendas. UNDP works directly with governments to strengthen institutional capacity, rights-based policy design, and digital governance. Its Digital Strategy (2022–2025) promotes a human-centred approach to digital transformation, guiding both national reforms and cross-border cooperation aimed at advancing equitable, responsive, and resilient digital governance.[5]

To translate this Digital Strategy into practice, UNDP developed the Digital Transformation Framework (DTF), a diagnostic tool that assesses national legal, technical, and institutional readiness for emerging technologies, including AI. It identifies key capacity gaps such as regulatory design, data infrastructure, and inter-agency coordination and supports the formulation of ethical, inclusive digital strategies.[6] Several countries have leveraged the DTF to advance their digital and AI governance landscapes in meaningful ways.

[5] United Nations Development Programme (UNDP) Digital Strategy 2022-2025. https://digitalstrategy.undp.org/documents/Digital-Strategy-2022-2025-Full-Document_ENG_Interactive.pdf

[6] UNDP, Digital Transformation Framework. https://www.undp.org/sites/g/files/zskgke326/files/2023-11/%5Bconcept%20note%5D%20digital%20transformation%20framework.pdf

For instance, Montenegro utilised the Framework to evaluate its digital governance landscape in preparation for EU integration. This assessment uncovered critical gaps in interoperability and digital service delivery within the public sector. Guided by these insights, Montenegro revised its digital strategy to invest in upgrading backend systems and enhancing cybersecurity, leading to improved citizen access to e-services and stronger alignment with EU digital standards.[7] Moreover, Moldova applied the diagnostic tool to strengthen inter-ministerial coordination and enhance cybersecurity frameworks. The process highlighted fragmentation in digital governance institutions, prompting legislative reforms and the adoption of rights-based principles in national digital policy.[8] These steps have improved trust in government digital platforms and streamlined service delivery, supporting Moldova's broader public sector reform agenda.

Complementing these efforts, UNDP also supports initiatives such as *AI for Sustainable Development*,[9] the *Digital Capacity Lab*,[10] and the *Digital X Catalogue*,[11] which promote responsible AI deployment, targeted policymaker training, and access to scalable digital solutions aligned with good governance and sustainable development objectives.

As a specialised agency, the United Nations Educational, Scientific and Cultural Organisation (UNESCO) leads the UN's normative work on the ethics of AI. Its 2021 Recommendation on the Ethics of Artificial Intelligence (UNESCO, 2021), endorsed by 193 Member States, defines a globally agreed ethical framework, grounded in human rights, accountability, transparency, and sustainability. To operationalise these commitments, UNESCO developed a Readiness Assessment Methodology that helps governments evaluate their maturity in AI governance and identify gaps in institutions, legislation, and societal protections. These assessments often conducted with UNDP inform targeted reforms and capacity-building measures that promote algorithmic fairness, data protection, and inclusive innovation.

Portugal has engaged actively in developing AI ethics and governance frameworks aligned with UNESCO's global human rights standards. Drawing on UNESCO's 2021 Recommendation on the Ethics of Artificial Intelligence, Portuguese authorities, with support from the United Nations Association of Portugal (UNA Portugal) and the Portuguese Business Ethics Association (APEE), have advanced national dialogue and capacity-building initiatives. These efforts focus on responsible AI deployment in sectors such as health and public administration, fostering transparency, accountability, and inclusive innovation.[12]

Similarly, Finland has integrated UNESCO's ethical AI principles into its national AI strategy, emphasising human-centric AI development, data protection, and broad societal participation. The Finnish government has organised multistakeholder consultations and established ethical guidelines for public sector AI applications, thereby strengthening institutional readiness and citizen trust.[13] In Austria, the Government has collaborated with academic institutions and civil society to adapt UNESCO's AI ethics framework into policy recommendations for fair, transparent, and human-rights-respecting AI governance. Pilot projects in Vienna focus on algorithmic

[7] Montenegro presents its first AI Readiness Assessment Report for the Public Administration. https://www.undp.org/montenegro/press-releases/montenegro-presents-its-first-ai-readiness-assessment-report-public-administration?utm_source=chatgpt.com

[8] UNDP Moldova's Development Results Report for 2023-2024. https://www.undp.org/moldova/publications/undp-moldovas-development-results-report-2023-2024?utm_source=chatgpt.com

[9] UNDP AI for Sustainable Development. https://www.undp.org/digital/ai

[10] UNDP Digital Capacity Lab. https://www.undp.org/digital/capacitylab

[11] UNDP *Digital X Catalogue*. https://digitalx.undp.org/

[12] Ai Ethics Portugal (2023). Ai. Ethics Virtual Session - Ai Ethics Portugal

[13] Finland Balances AI Innovation with Strong Safeguards for Human Rights and Democratic Values (2024) https://www.ai-360.online/finland-balances-ai-innovation-with-strong-safeguards-for-human-rights-and-democratic-values/?utm_source=chatgpt.com

accountability and mitigating bias in AI-driven public services, supported by capacity-building programmes co-designed with UNDP.[14]

These country examples demonstrate how UNESCO's normative framework serves as a foundation for diverse national approaches to operationalising ethical AI governance, tailored to local contexts while adhering to universal human rights principles.

Meanwhile, the United Nations Department of Economic and Social Affairs (UN DESA) supports the integration of AI within broader public administration reform. Through its World Public Sector Report (WPSR), UNDESA analyses the institutional impacts of AI and digital innovation (UNDESA, 2023). Its E-Government Development Index (EGDI), though not AI-specific, tracks national readiness through indicators such as infrastructure, digital services, and human capital – providing insight into the preconditions for ethical AI adoption (UNDESA, 2024).

Among UN specialised agencies, the International Telecommunication Union (ITU) plays a pivotal role in advancing global discourse on the role of AI in sustainable development. Through its flagship AI for Good platform, ITU fosters collaboration between governments, academia, private sector actors, and civil society to ensure that AI technologies support the SDGs. Furthermore, ITU contributes to technical standardisation efforts in AI-enabled communication systems, working to enhance interoperability, safety, and digital inclusion across borders (ITU, 2023). Together, these UN entities and agencies form a coherent support structure that helps operationalise AI governance by embedding global norms into institutional practice.

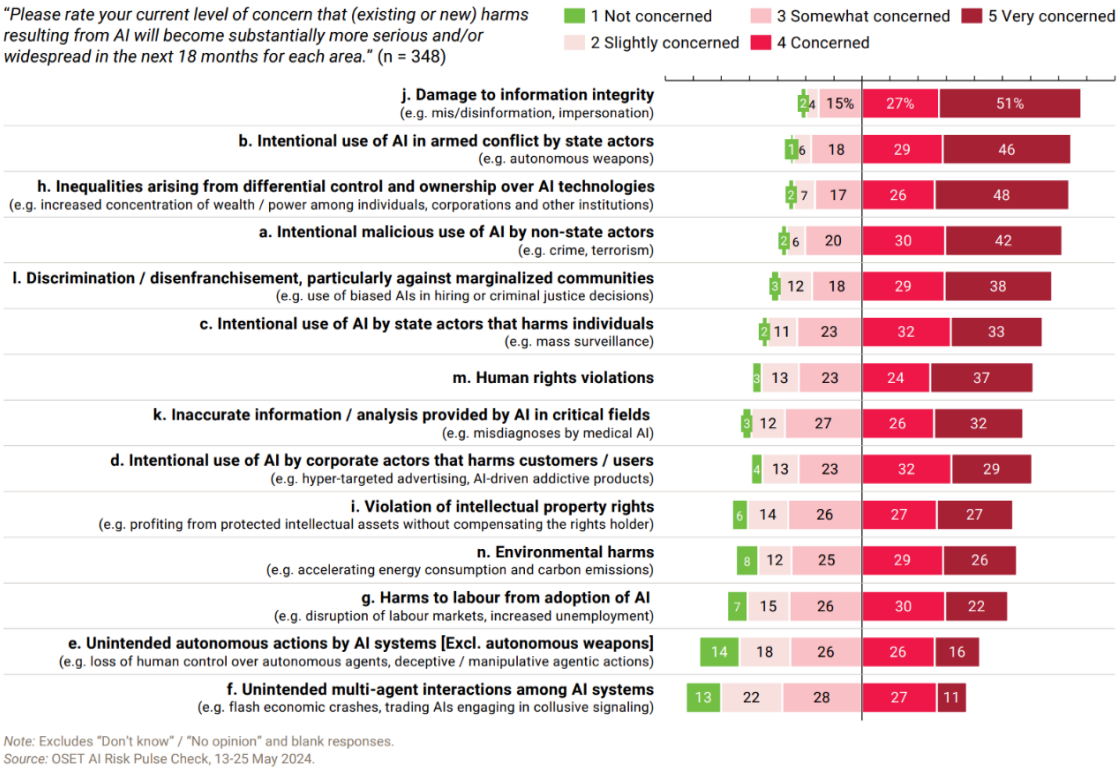
1.2 UN High-Level Advisory Body on Artificial Intelligence (HLAB-AI)

The UN High-Level Advisory Body on Artificial Intelligence (HLAB-AI), launched by the Secretary-General in 2023, was established as a strategic initiative to steer international dialogue and formulate actionable recommendations for the governance of artificial intelligence (HLAB-AI, 2023). Its core mandate was to provide strategic guidance for translating globally shared ethical principles into effective policy and institutional mechanisms, with a focus on protecting human rights, promoting inclusivity, and enabling innovation. Comprising 39 independent experts from 33 countries, spanning public policy, law, technology, human rights, ethics, and the social sciences, the Advisory Body ensured a truly global, multidisciplinary, and impartial perspective. Through a consultative process involving over 2,000 stakeholders from governments, civil society, academia, and industry, the Body gathered global outlook and insights reflecting concerns, sectoral expertise, and collective ambitions (HLAB-AI 2024).

The resulting flagship report, "*Governing AI for Humanity*" identifies a pressing governance gap as AI technologies advance faster than regulatory and ethical frameworks can adapt. The Report calls for coordinated action, warning that without it, the benefits of AI will remain in the hands of a few, while risks such as algorithmic bias, surveillance abuse, and misinformation may destabilise democratic institutions and harm vulnerable populations.

[14] Why we need global governance frameworks for artificial intelligence (2023). https://rudolphina.univie.ac.at/en/governance-frameworks-for-artificial-intelligence?utm_source=chatgpt.com

Figure 2. Experts' levels of concern about AI risks across multiple domains



Source: United Nations (2024), Governing AI for Humanity, p. 29.

This concern is reinforced by the expert assessments presented in the report (Figure 2), where a significant majority of surveyed experts indicated that they were 'concerned' and 'very concerned' about potential harms across multiple domains. Notably, 78% of experts flagged damage to information integrity (*indicator j*), reflecting widespread anxiety over the erosion of trust in information ecosystems. This was followed by 75% expressing concern over the use of AI in armed conflict by state actors (*indicator b*), highlighting the growing threat of militarised AI deployment. Concerns related to societal inequalities were also prominent, with 74% citing the concentration of wealth and power (*indicator h*) and 67% expressing concern about discrimination and disenfranchisement, particularly among marginalised communities (*indicator l*). Additionally, 72% pointed to risks from malicious use by non-state actors (*indicator a*), and 65% were concerned about harmful uses of AI by state actors against individuals (*indicator c*). These findings underscore the global demand for cohesive, human-centred and coordinated governance mechanisms.

To address these multifaceted challenges, the Report outlines a strategic roadmap structured around four interlinked pillars. These include: (i) fostering a common understanding of AI risks and opportunities; (ii) establishing common ground through inclusive policy dialogue and interoperable standards; (iii) delivering common benefits by promoting equitable access to AI tools, data, and capacity development such as through a global AI fund and inclusive governance networks; and (iv) ensuring a coherent effort by creating a dedicated, agile AI office within the UN Secretariat. Together, these pillars aim to cultivate a shared global purpose, strengthen representation, and ensure meaningful participation from all regions and stakeholder groups (Table 1).

Table 1. Institutional Policy Mechanisms for AI Operationalisation

Pillar 1: Common Understanding *To reduce knowledge gaps and foster global trust, it is essential to establish*	**Global Scientific Panel on AI** An independent, multidisciplinary body offering impartial insights and guidance through annual reports, SDG-focused research digests, and rapid assessments of emerging risks and trends to inform global policy-making **Global AI Observatory** A platform to monitor AI technological developments, track governance responses, and identify gaps and best practices globally.
Pillar 2: Common Ground *To align global efforts through shared principles and interoperable rules:*	**Standing Global Dialogue on AI Governance** A permanent multistakeholder forum to build trust, share knowledge, and foster inclusive consensus, cooperation on governance priorities, harmonise standards across jurisdictions. **AI Standards and Data Governance Framework** A harmonised effort promoting ethical, legal, and technical interoperability, grounded in human rights and fairness.
Pillar 3: Common Benefits *To ensure equitable access to AI and development dividends:*	**Global Fund for AI for Sustainable Development** A dedicated financing mechanism supporting digital infrastructure and inclusive AI capacity in low- and middle-income countries. **Global Capacity Development Network** A decentralised delivery model for context-specific training, institutional support, and peer learning.
Pillar 4: Coherent Effort *To ensure strategic coordination and avoid fragmented approaches:*	**AI Office within the UN Secretariat** A light, agile unit under the UN Secretary-General to provide leadership, coherence, and follow-up on implementation. **Partnerships for Inclusive Innovation** Mechanisms to connect governments, academia, civil society, and private sector actors in co-developing AI solutions that serve the global public good.

To operationalise this vision, the report offers a set of concrete institutional and policy mechanisms, built upon four interlinked pillars. It calls for AI governance to be dynamic, collaborative, and anchored in international law. Crucially, it advocates for reducing the global AI divide by strengthening support to developing and emerging economies countries so that no region is left behind.

Figure 3. Overview of Recommendations Addressing Global AI Governance Gaps

Purpose	Enhance representation	Enable coordination	Strengthen implementation
Common understanding International scientific panel on AI	✓	✓	
Common ground Policy dialogue on AI governance AI standards exchange	✓	✓	(✓)
Common benefits Capacity development network Global fund for AI Global AI data framework	✓	✓	✓
Coherent effort AI office within the Secretariat	Advising the Secretary-General on matters related to AI, working to promote a coherent voice within the United Nations system, engaging States and stakeholders, partnering and interfacing with other processes and institutions, and supporting other proposals as required.		

Source: United Nations (2024), Governing AI for Humanity report, p. 47.

Through its actionable recommendations and visionary framework, the HLAB-AI reinforces the United Nations' unique convening power in shaping an ethical and forward-looking digital future. Most notably, the report *Governing AI for Humanity* played a foundational role in shaping two major frameworks adopted during the 2024 Summit of the Future: the *Global Digital Compact* and the *Pact for the Future*.[15] In this way, HLAB-AI's work reflects not only the urgency of AI governance, but also the UN's strategic, system-wide response to managing the risks and harnessing the opportunities of rapidly advancing technologies.

1.3 The Global Digital Compact (GDC) and the Pact for the Future

While '*Governing AI for Humanity*' serves as an expert advisory report, the GDC and the Pact for the Future are intergovernmental political frameworks officially endorsed by Member States. They set high-level principles and implementation pathways that anchor AI within a broader agenda of digital cooperation, inclusion, and human rights. The Pact for the Future and the Global Digital Compact (GDC), adopted by all 193 UN Member States at the 2024 Summit of the Future, represent two foundational frameworks, one articulating a renewed vision for multilateral cooperation, the other outlining a roadmap for inclusive and rights-based digital governance, including AI.

The 'Global Digital Compact', originally proposed in the Secretary-General's 2021 report *Our Common Agenda*, called for a shared vision to shape an open, secure, and inclusive digital future. It emphasised the urgent need to bridge digital divides, safeguard data rights, and promote universal connectivity, while fostering innovation and the responsible governance of emerging technologies. HLAB-AI was established, in part, to provide the expert technical guidance and strategic recommendations that directly informed the Compact's development, particularly its AI-related provisions. Drawing on HLAB-AI's findings, GDC integrates policy tools and institutional mechanisms, including global standards, multistakeholder dialogue, and capacity-building to ensure more equitable and human-centred governance.

[15] Summit of the Future: Outcome Documents. https://www.un.org/en/summit-of-the-future and Global Digital Compact. https://www.un.org/global-digital-compact/en

The '*Pact for the Future*' takes a broader view, committing Member States to revitalised multilateralism across key domains including peace, sustainability, intergenerational equity, and digital transformation. It is a comprehensive intergovernmental agreement designed to strengthen international cooperation. Drawing from HLAB-AI's recommendations, the Pact introduces two key institutional innovations in global AI governance: the establishment of an Independent International Scientific Panel on AI and the launch of a Global Dialogue on AI Governance. In addition, the Pact includes the '*Declaration on Future Generations*', reinforcing the ethical imperative to protect the rights and dignity of future populations. The Declaration situates digital and AI governance within a wider framework of intergenerational justice, long-term foresight, and digital ethics, values championed by HLAB-AI.[16]

Together, the Pact for the Future, the Global Digital Compact, and the Declaration on Future Generations translate HLAB-AI's strategic vision into formal, consensus-based commitments. They mark a pivotal moment in global governance, embedding responsible AI development within multilateral norms, cross-sectoral collaboration, and a shared commitment to universal digital rights. In doing so, the United Nations is establishing itself as a central architect and guiding force for the development of global AI governance.

1.4 Global Partnership on Artificial Intelligence (GPAI)

The Global Partnership on Artificial Intelligence was established in June 2020 by 15 founding members, as the first dedicated multilateral forum focused on advancing the responsible development and use of AI.[17] Conceived as both a political initiative and a practical policy laboratory, the GPAI brings together governments, experts, and international organisations to translate high-level principles into actionable measures. Since its creation, the Partnership has grown steadily: by 2025, membership included 29 countries, alongside the European Union, encompassing a total of 44 states. Its institutional integration with the OECD, in July 2024, gave GPAI a permanent structure and consolidated its normative foundations in the OECD Recommendation on Artificial Intelligence (2019), widely regarded as the intergovernmental standard on trustworthy AI.[18]

1.4.1 Operationalisation of Governance Frameworks and Institutional Arrangements

GPAI operates through a multilayered governance structure designed to bridge normative frameworks with practical implementation. The following:

- A Council of Ministers for high-level strategic direction;
- A Steering Committee with mixed government and non-government representation that oversees operations;
- A Multistakeholder Experts Group of around 100-150 specialists from science, industry, civil society, trade unions, and international organisations, which functions as the primary advisory body; and

[16] United Nations (2024). *Pact for the Future: Outcome Document,* Section II, Declaration on Future Generations. https://www.un.org/en/summit-of-the-future

[17] The Global Partnership on Artificial Intelligence. https://www.dcc.ac.uk/consultancy/GPAI

[18] GPAI's foundational commitment to the OECD Recommendation on Artificial Intelligence provides a shared normative framework emphasising human rights, fundamental freedoms, and democratic values.

– Expert Support Centres in Paris (OECD), Montreal (CEIMIA) and Tokyo,[19] which provide technical expertise, support applied research, and facilitate project delivery.

This structure allows GPAI to act as a policy laboratory, operationalising broad principles through applied projects/ Unlike binding regulatory instruments, GPAI relies on voluntary cooperation, peer learning, and knowledge transfer to advance AI governance.

GPAI's approach to translating international AI policies into practical measures centres on four core working groups addressing (i) Responsible AI; (ii) Data Governance; (iii) Future of Work; and (iv) Innovation and Commercialisation. Together, these groups generate research, recommendations, and pilot projects that help governments and organisations to operationalise ethical principles such as transparency, fairness, and accountability. The Data Governance Working Group has produced guidance and frameworks on data sharing, safeguards against discrimination, and governance models adaptable across various jurisdictions. The Responsible AI Working Group develops methodologies for trustworthy and explainable systems. The Future of Work Working Group examines labour market transitions and reskilling strategies, and the Innovation and Commercialisation Working Group supports SMEs and startups in balancing compliance with competitiveness. Collectively, these groups function as intermediaries between international norms and national practice., ensuring that abstract commitments human rights and democratic values are translated into concrete policy guidance and tools.

1.4.2 Adaptation and Challenges: Governments and the Private Sector

Government adaptation to GPAI frameworks varies significantly across member countries. India's leadership role, serving as Lead Chair in 2024, demonstrated how emerging economies can leverage international partnerships to advance domestic AI strategies, particularly in areas such as data governance, which also contributing to the development of global standards. This was visible in initiatives such as the Global India.AI Summit in 2024, which highlighted responsible and inclusive AI, but also extended to India's broader engagements in GPAI's governance structure and working groups.

Serbia's 2025 chairmanship represents continued engagement from middle-income countries seeking to integrate international best practices in AI governance with national development priorities.[20] Among the founding members, Canada has advanced its Pan-Canadian AI Strategy in parallel with GPAI's participation, reinforcing the emphasis on human-centric and inclusive governance. France, similarly, has used both its national AI strategy and active role in GPAI to promote stronger ethical and regulatory standards, contributing to the diffusion of responsible AI principles in the international policy fora.

For the private sector, GPAI offers a platform to engage directly in governance debates and anticipate future regulatory expectations. Industry representatives, alongside research institutes and civil society actors, participate in the Multistakeholder Experts Group and thematic Working Groups, contributing technical expertise while testing compliance practices. This involvement enables companies to balance three competing priorities: meeting ethical and regulatory standards, fostering innovation, and maintaining competitiveness in global markets. For example, work in the Innovation and Commercialisation Working Group has highlighted the challenges

[19] https://www.oecd.org/en/about/programmes/global-partnership-on-artificial-intelligence.html

[20] The role of international cooperation in the responsible use of AI usage in Serbia and globally (UNDP, 2024). https://www.undp.org/serbia/news/role-international-cooperation-responsible-use-ai-usage-serbia-and-globally

startups and SMEs face in scaling AI solutions under evolving governance frameworks, while the Future of Work Working Group has engaged employers and industry actors in developing approaches for reskilling and adapting to AI-driven labour market transitions.

Despite these efforts, the operationalisation process forces significant limitations. GPAI's institutional arrangements rely primarily on voluntary compliance and peer pressure rather than binding legal instruments. This approach, while preserving national sovereignty, creates challenges for ensuring consistent implementation across diverse jurisdictions with varying regulatory capacities (Engler et al., 2021).

1.4.3 Capacity Building

Capacity building has become an increasingly central dimension of GPAI's mission, particularly in supporting emerging and developing economies. Activities have included producing policy guidance on AI ethics and governance, supporting governments in developing national AI strategies, and elaborating frameworks for data governance. Projects coordinated through the Montreal Centre of Expertise (CEIMIA) have addressed issues such as AI and climate change, pandemic response, and labour market transitions, providing applied models that countries with limited regulatory capacity can adapt. GPAI also promotes peer-to-peer learning by linking experts from advanced economies with institutions in developing contexts, helping to strengthen oversight capacity and improve access to data infrastructure.

1.5 OECD AI Policy Principles

In 2019, the OECD AI Principles became the first inter-governmental standard promoting responsible development and deployment of trustworthy AI.[21] They were revised in 2024, prompted by the rapid advances in generative AI, wider adoption of AI in public services, emerging cybersecurity and data privacy risks, and the urgent need for stronger mechanisms of accountability and regulatory clarity. The updated Principles reinforce their role as a foundation for international cooperation and regulatory interoperability, offering guidance designed to remain relevant in the fast-paced AI landscape. Grounded in universal human rights and democratic values, they are structured around five pillars: (i) inclusive growth, sustainable development, and well-being; (ii) human-centred values and fairness; (iii) transparency and explainability; (iv) robustness, security, and safety; and (v) accountability.

[21] https://www.oecd.org/en/topics/sub-issues/ai-principles.html

Box 1: United Kingdom's Women in Data Science and AI Initiative

Developed by the Alan Turing Institute, the initiative follows a three-tiered approach:

1. Mapping participation of women in data science and AI (UK and globally)

"Where are the Women? Mapping the Gender Job Gap in AI" report (2020) produced the first evidence base that raised awareness of systemic underrepresentation of women in the UK's AI and data workforce (~22%), enabling more inclusive labour policies and organisational strategies to improve women's participation in AI.

2. Investigating diversity and inclusion in workplace cultures (online and physical)

"Where are the Women?" Report (2021) informed employers and policymakers about structural workplace biases, prompting reforms in recruitment, retention, and progression that support fairer access to AI careers for women.

3. Exploring how the gender gap shapes scientific knowledge and technological innovation

"Rebalancing Innovation: Women, AI and Venture Capital" report's (2022) findings shaped UK debates on venture capital diversity and contributed to policies like the £500 million package with £50 million ring-fenced for female-led funds, embedding gender diversity in AI innovation.

Source: https://www.turing.ac.uk/research/research-programmes/public-policy/public-policy-themes/women-data-science-and-ai

Over the past few years, countries have actively translated these Principles into practical policies, reflected in more than 1,000 initiatives across over 70 jurisdictions. Governments are pursuing varied strategies to promote **i**nclusive growth, sustainable development, and well-being (Principle 1), with most national AI strategies and ethics frameworks explicitly referencing this principle. Policy initiatives also emphasise stakeholder participation (e.g., Canada's Quebec AI Forum),[22] including vulnerable groups (e.g., Chile's Participation Process on AI).[23] Or in AI policy design, and support of AI projects that enhance human capabilities and foster creativity (e.g., Türkiye's Breast and Lung Cancer Detection with AI Project),[24] or tackle environmental challenges (e.g., Germany AUTONOM – Performing Arts and AI programme),[25] and reduce economic, social, and gender inequalities (e.g., UK's Women in Data Science and AI initiative). In advancing human-centred values and fairness (Principle 2), governments have mostly relied on non-binding guidelines and initiatives to mitigate AI bias and protect human rights. An example is UNESCO's global recommendation on AI ethics, which emphasises human dignity, fairness, non-discrimination, inclusion, and respect for human rights. Far fewer governments have introduced binding measures, such as certification schemes (e.g., Malta's AI Certification Programme)[26] or Human Rights Impact Assessments (HRIAs). The Netherlands is currently the

[22] https://forumia.quebec/en

[23] https://www.gob.cl/en/news/science-ministry-launches-participatory-process-contribute-developing-national-artificial-intelligence-policy/

[24] https://www.oracle.com/tr/artificial-intelligence/early-detection-cancer-with-oci-vision/

[25] https://www.fonds-daku.de/events-und-diskurs/archiv/autonom-symposium/

[26] https://oecd.ai/en/dashboards/policy-initiatives/ai-certification-programme-5194

only country requiring public institutions to carry out HRIAs for AI systems, building on the Fundamental Rights & Algorithms Impact Assessment (FRAIA) tool.[27] As part of Principle 2 implementation, countries are also adopting policies to safeguard privacy in AI, such as Korea's AI Privacy Risk Management Model for the safe use of AI and data.

Box 2: Republic of Korea's AI Privacy Risk Management Model for Safe Utilisation of AI and Data

In 2024, the Personal Information Protection Commission introduced a risk-based model guiding organisations to *identify*, *assess*, and *mitigate privacy risks across AI lifecycle stages*, embedding privacy considerations into design, deployment, and governance.

The model strengthens accountability and trust by enabling systematic privacy risk management, improving disclosure and user guidance practices, and aligning corporate behaviour with stronger regulatory expectations.

Initial implementation shows that companies providing AI services, including large language model providers, have begun applying the model's risk assessment and mitigation measures, and government agencies are integrating the framework into AI governance practices.

Source: https://www.pipc.go.kr/eng/user/ltn/new/noticeDetail.do?bbsId=BBSMSTR_000000000001&nttId=2745

At the same time, governments are adopting a range of approaches to strengthen AI transparency, from issuing implementation guidelines to establishing oversight bodies. Regulatory authorities increasingly recognise the importance of transparency and explainability (Principles 3), which are already embedded in legislation on data protection, privacy, and consumer rights (e.g., Canada's Directive on Automated Decision-Making,[28] EU's AI Act). Proposed AI-specific regulations also introduce tailored provisions, for example in the workplace. Spain's 2021 Rider Law illustrates this trend by requiring transparency around algorithmic management in the workplace, making it one of the first national laws to regulate AI-driven decision-making in employment.[29] In the public sector, initiatives such as AI registers are being introduced to disclose how AI is used in government services.

[27] https://www.government.nl/documents/reports/2021/07/31/impact-assessment-fundamental-rights-and-algorithms

[28] https://www.tbs-sct.canada.ca/pol/doc-eng.aspx?id=32592

[29] https://oecd.ai/en/wonk/workplace-regulation-2022

Box 3: Finland and The Netherland's Open AI Registers

In 2020, *Finland* and *the Netherlands* launched world's first open AI Registers that list AI/algorithm-based systems used by public sector bodies. These registers disclose what systems are used, by which authority, for what purpose, and often include details for public scrutiny.

In *Helsinki*, the AI Register lists systems such as chatbots for book recommendations, pregnancy and health information, and parking inquiries, as well as an intelligent material management system for libraries. *Amsterdam*'s register includes tools for monitoring parking compliance, categorising citizen reports, and prioritising investigations into possible illegal holiday rentals.

These registers have improved transparency and accountability by making government AI use visible to citizens, enabling risk assessment and public oversight, and pressuring public bodies to disclose systems. They reflect a commitment to human-centred AI grounded in responsibility, transparency, and security. However, audits show that many AI systems remain unregistered, and disclosure detail varies.

Source: https://aiforgood.itu.int/helsinki-and-amsterdam-launch-ai-registers-to-detail-city-systems/

It is important to note, however, that while many AI systems are in use across the Dutch public sector, only about 5% are listed in the national AI/Algorithm Register, highlighting clear gaps in transparency. Across 70 organisations, 433 systems have been deployed, of which 167 remain experimental, and the majority of organisations operate no more than three systems. The Police and the Employee Insurance Agency are the most active users, with 23 and 10 systems, respectively. The Court of Audits review shows that these systems are primarily used for inspection, enforcement, and knowledge-processing tasks. Thus, adoption remains largely experimental, and stronger reporting and governance practices are needed to ensure responsible and transparent use.[30]

Meanwhile, to ensure robustness, security, and safety (Principle 4) in AI, countries are applying a mix of tools, including ethical frameworks, impact assessments, guidelines, and both new and updated laws. Much of this work centres on adapting traffic rules and safety regulations to cover AI systems, e.g., Austria, Germany, Japan, and others. Other measures include developing risk management approaches (e.g., Mexico's Principles and Impact Analysis Guide for the Development and Use of Systems Based on Artificial Intelligence in the Federal Public Administration)[31] and record-keeping practices for data characteristics (e.g., Türkiye's National Data Dictionary).[32]

[30] https://gratanet.com/web_files/users/1529094/7733394852128141466.pdf

[31] https://oecd.ai/en/dashboards/policy-initiatives/principles-and-impact-analysis-guide-for-the-development-and-use-of-systems-based-on-artificial-intelligence-in-the-federal-public-administration-8462

[32] Türkiye, Digital Transformation Office of the Presidency of the Republic of Türkiye, National Data Dictionary Project, https://cbddo.gov.tr/en/projects/nationaldatadictionary/

Box 4: National Legislation on Autonomous Driving

Austria: Amendments to the Motor Vehicles Act and the *Automated Driving Regulation* allow certain driving tasks to be performed by automated systems. The legislation also sets conditions for testing driver-assistance and automated technologies and enables early applications such as autonomous minibuses in urban areas.

Germany: Revisions to the *Road Traffic Act* and the adoption of the *Autonomous Driving Act* legalise automated vehicles on public roads. These measures define safety and technical requirements for highly and fully automated vehicles and clarify liability and insurance obligations.

Other countries, including Denmark, Japan, Lithuania, and the United Kingdom, have also passed laws or regulations to govern the testing and safe operation of self-driving cars.

Source: https://www.ippi.org.il/germany-autonomous-driving-act/#:~:text=Spotlight,can%20intervene%20during%20critical%20situations

To strengthen accountability (Principle 5) in AI, many countries have introduced sector-specific ethical codes, such as Portugal's *Guide to Artificial Intelligence in the Public Administration,*[33] and Singapore's *AI in Healthcare Guidelines*.[34] Proposed AI regulations also require documentation that demonstrates proper system performance across the entire lifecycle. Moreover, some governments have created independent oversight bodies to monitor and audit the use of algorithms, such as the Netherlands' Department for the Coordination of Algorithmic Oversight (DCA).

At the same time, governments are advancing the OECD's five recommendations through a mix of investment, regulation, and cooperation. While approaches vary by context, common trends are emerging: institutionalising AI R&D through dedicated centres, expanding compute and data infrastructures, combining binding regulation with flexible sandboxes, aligning workforce training with labour market transitions, and deepening international cooperation.[35]

[33] https://bussola.gov.pt/Guias%20Prticos/Guia%20para%20a%20Inteligência%20Artificial%20na%20Administração%20Pública.pdf

[34] https://isomer-user-content.by.gov.sg/3/9c0db09d-104c-48af-87c9-17e01695c67c/1-0-artificial-in-healthcare-guidelines-(aihgle)_publishedoct21.pdf

[35] Table 2 highlights vivid national examples that illustrate how these principles are being put into practice

Box 5: The Netherland's Algorithm Supervision Body

The *Department for the Coordination of Algorithmic Oversight (DCA)*, established in 2023 within the Dutch Data Protection Authority, was created in response to 2021–2022 audits revealing widespread non-compliance and controversies such as the childcare benefits scandal. It strengthens independent oversight by auditing algorithmic systems, identifying risks like bias or lack of control, and guiding agencies to improve transparency, governance, and public trust.

Key functions:

- *Coordinated Framework*: Acts as the central authority for algorithms and AI, fostering collaboration among surveillance bodies and supporting legal frameworks.
- *Systemic Risk Analysis*: Identifies high-level risks and publishes the biannual Report AI & Algorithms Netherlands, highlighting persistent challenges and promoting transparency.
- *Tools and Guidance*: Provides practical guidance on responsible algorithm use, including registration and ensuring meaningful human intervention in automated decisions.

The DCA is a key driver of a more transparent, professional, and proactive approach to algorithmic governance.

Source: https://www.autoriteitpersoonsgegevens.nl/en/themes/algorithms-ai/coordination-of-algorithmic-and-ai-supervision/department-for-the-coordination-of-algorithmic-oversight-dca

Overall, six years after adoption the OECD AI Principles have moved from aspirational statements to an active policy-making agenda, as governments worldwide are translating high-level values into funding programmes, standards, impact assessments, regulatory experiments, and institutional oversight. Progress is real and multi-dimensional, but uneven – jurisdictions differ in ambition, legal force, and implementation capacity. To consolidate gains, it is important to prioritise measurable monitoring and enforcement, technical and regulatory interoperability, and international cooperation, including standards and capacity-building, so that the transition from principles to practice yields both innovation and durable public trust.

Table 2. Implementation of OECD Recommendations (Principles 2.1-2.5)

Principle	Implementation Trend	Example(s)	Key Insight
1.1 – Investing in AI R&D	Dedicated funding, national AI institutes, collaborative research networks.	**Canada** – AI Institutes (Amii, Vector, Mila); **France** – National AI Research Programme; **Chile** – National AI Research Centre.	Institutionalising AI R&D through permanent centres creates sustainable ecosystems beyond short-term projects.
2.2 – Fostering a Digital Ecosystem	Expanding compute capacity, sectoral data hubs, and shared infrastructures.	**EU** – High-Performance Computing Joint Undertaking; **France** – Health Data Hub; **US** – National AI Research Resource Task Force.	Ecosystem success depends on compute + trusted governance of data access, not infrastructure alone.
2.3 – Enabling Policy Environment	AI-specific regulations, ethical guidelines, regulatory sandboxes.	**EU** – AI Act; **Norway** – Data Protection Authority Sandbox; **Singapore** – FinTech Regulatory Sandbox.	Binding rules and flexible experimentation work best together to balance innovation and trust.
2.4 – AI Skills, Jobs, Labour Markets	Workforce upskilling, vocational training, talent attraction, job redesign.	**Australia** – Next-Generation AI Graduates Programme; **Singapore** – AI for Industry + Job Redesign Guide; **Kenya** – Digital Literacy Programme.	Strongest strategies combine advanced talent pipelines with sector-specific job redesign for smooth transitions.
2.5 – International Cooperation	Multilateral initiatives, capacity building, global norm-setting.	**GPAI**; **Germany** – FAIR Forward ("AI for All"); **G7** Hiroshima AI Process (2023).	AI governance is shifting from fragmented dialogues toward shared global principles on frontier AI.

Source: OECD (2023).

1.6 European Union 2024 AI Act

With the introduction of the Artificial Intelligence Act (AIA) in February 2024, a unified AI governance system has emerged across the European Union showcasing collaborative governance in practice. The regulation aims to balance the protection of European citizens' rights, health, and safety on the one hand, and the advancement of the internal AI industry on the other.[36] This is mainly done through setting requirements for certain AI systems and prohibiting specific AI uses. To date, the AIA is considered the first-ever and most comprehensive legal framework globally (Outeda, 2024; Quoy et al., 2024) and therefore, is of particular interest for detailed analysis, especially in terms of its operationalisation.

The EU AI Act adopts a cross-sectoral approach, covering all sectors in which AI is deployed, rather than focusing on specific domains such as healthcare, transport, or finance. At the same time, it applies equally to public and private actors, including developers, deployers, and users

[36] The European Union Artificial Intelligence Act (2024). https://artificialintelligenceact.eu/article/1/

of AI systems. This means that AI-based credit scoring systems in banking are subject to the same requirements as AI systems used in predictive policing, since both are classified as "high-risk" under the Act. Furthermore, the Act introduces horizontal regulation, replacing fragmented, sector-specific AI laws with a unified legal framework. This is achieved by emphasising fundamental rights, safety, and transparency; principles that apply to all AI systems regardless of industry. That is, a language model like GPT, used both for chatbots in customer service and for drafting legal documents, is regulated under the same horizontal principles.[37]

Another important characteristic of the AI Act is the risk-based approach which suggests that the higher the risk the more regulatory obligations scale. There are four risk categories: unacceptable, high, limited, and minimal (Table 3). *Unacceptable-risk* AI, such as systems for subliminal manipulation, predictive policing, or mass facial image scraping, is banned outright (Article 5). For example, the US company Clearview AI's facial recognition practices have been banned in the European Union, with multiple national data protection authorities imposing large fines and ordering the company to cease illegal data collection and delete existing data on EU residents.[38] *High-risk* AI, including applications in product safety, education, and law enforcement, is allowed but subject to stringent requirements before entering the EU market and remain subject to monitoring once they are in use. Such an approach ensures that innovation is not stifled by unnecessary bureaucracy.

Table 3. Risk-based regulation in EU

Risk Level	Real Example	AI Act Regulation
Unacceptable	Emotionrecognition in hiring	Banned
	Social scoring systems	Banned
	Real-time facial recognition in public	Banned (with narrow exceptions)
High Risk	CV-screening AI in HR	Requires risk assessments, oversight
	Police biometric identification	Conformity assessments
Limited Risk	Customerservice chatbots	Disclosure required
Minimal Risk	Video games, spam filters, AI art tools	No regulation

Meanwhile, the European Commission is considering delaying some provisions of the EU Artificial Intelligence (AI) Act, particularly those affecting high-risk AI systems, which are scheduled for full implementation in August 2026. While the Act's initial rules banning unacceptable AI practices are already in force, stakeholders have raised concerns about the readiness of technical standards, the complexity of compliance, and potential impacts on innovation, especially for small and medium-sized enterprises. Proposed adjustments include postponing enforcement until technical standards are established, expanding exemptions for SMEs, and creating waivers for simpler AI systems. The Commission is also exploring ways to ensure regulatory consistency across the EU. Although no formal decision has been made, organisations are advised to prepare for compliance while monitoring potential changes in the regulatory framework.

[37] The operational structure of the AI Act is a layered model combining: (i) *EU-level bodies* (AI Office, AI Board, Expert Panel) for guidance and enforcement coordination; (ii) *National authorities*, empowered through law to conduct market surveillance, conformity assessments, audits, and sanctions; and (iii) *Innovation enablers*, via sandboxes and standardisation instruments, fostering compliance and growth.

[38] https://iclg.com/news/21446-dutch-regulator-fines-ai-firm-eur-30-million#:~:text=This%20penalty%20is%20the%20latest,citizens'%20data%20being%20illicitly%20used

1.6.1 National Authorities

One of the key aspects of the AIA implementation at the national level is to designate respective authorities, including national competent authorities (a market surveillance authority and a notifying authority) and national public authorities (Figure 4). While the design of government structures could take various forms across the EU, the establishment should have been completed by 2 August 2025. However, to date, only three EU countries – Lithuania, Luxembourg, and Malta – have clearly designated their national competent authorities to supervise and enforce the AIA. Although over one third of the EU Member States have partial clarity and are in the process of meeting this requirement, with legislative proposals currently pending[39], more than half have yet to establish any of the required competent authorities.[40]

Figure 4. Supervision of the AI Act implementation at the national level

<table>
<tr><th colspan="5">SUPERVISION OF THE AI ACT AT THE NATIONAL LEVEL</th></tr>
<tr><td rowspan="5">National competent authorities
(Articles 3(48) and 70 of the AI Act)</td><td>Notifying authorities
(Articles 3(19) and 28 of the AI Act)</td><td colspan="2">To be designated or established by Member States</td><td>Responsible for designating and monitoring notified bodies</td></tr>
<tr><td rowspan="4">Market surveillance authorities
(Articles 3(26) and 74 of the AI Act)</td><td colspan="2">To be designated at the discretion of the Member States, but for certain high-risk AI systems, the AI Act suggests which authority should be designated:</td><td rowspan="4">Responsible for the post-market surveillance of AI systems covered by the AI Act</td></tr>
<tr><td>National authority responsible for the financial supervision of those institutions
e.g. financial markets authority, national bank</td><td>High-risk systems used by financial institutions regulated by Union financial services law
e.g. credit scoring systems</td></tr>
<tr><td>Data protection authority or any other authority designated pursuant to the same conditions laid down in Directive 2016/680</td><td>High-risk systems used for law enforcement purposes, border management and justice and democracy (+ biometrics in these areas)
e.g. polygraphs</td></tr>
<tr><td>Existing market surveillance authorities under other NLF legislation</td><td>High-risk AI systems covered by the legislation listed in Annex I.A
e.g. toys, lifts</td></tr>
<tr><td></td><td>Authorities protecting fundamental rights
(Articles 77 of the AI Act)</td><td colspan="2">To be designated at the discretion of the Member States</td><td>Has the power to examine whether AI systems violate fundamental rights</td></tr>
</table>

Source: https://doi.org/10.48550/arXiv.2505.02791

Member States are adopting different institutional models to meet the AI Act's obligation to designate competent authorities. The dominant trend is towards a centralised approach, where oversight is concentrated in a single authority. In some cases, this involves the creation of a new institution, as in Spain with the Agency for the Supervision of Artificial Intelligence, or in Poland,[41] where a proposed law envisages the establishment of the Commission on the Development

[39] Denmark, Finland, Germany, Hungary, Ireland, Italy, Poland, Romania, Spain. https://artificialintelligenceact.eu/national-implementation-plans/

[40] Austria, Belgium, Bulgaria, Croatia, Cyprus, Czech Republic, Estonia, France, Greece, Latvia, Netherlands, Portugal, Slovakia, Slovenia, Sweden. https://artificialintelligenceact.eu/national-implementation-plans/

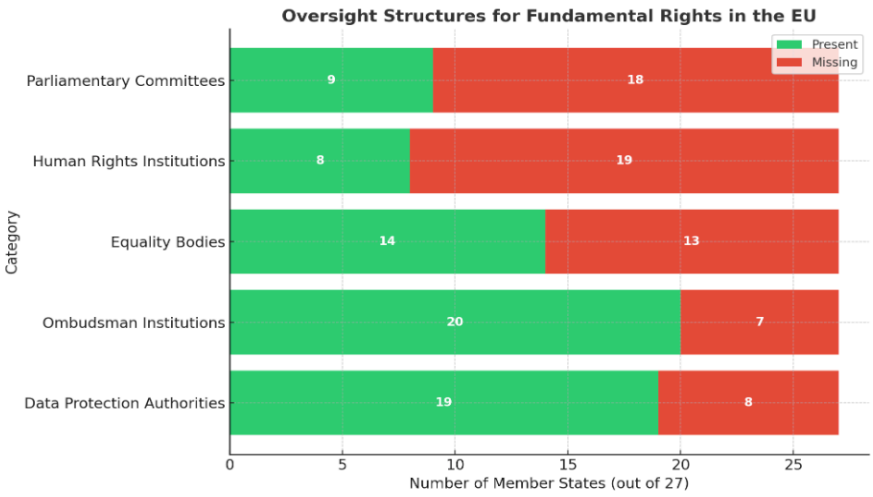
[41] https://www.dudkowiak.com/ai-law/

and Safety of Artificial Intelligence. In other cases, existing institutions have been assigned the task, as in Luxembourg, where the National Commission for Data Protection has been designated as the national market surveillance authority.[42] Apparently, centralisation provides clarity and consistency by giving businesses a single authority to engage with and allowing expertise to build in one place. Yet, it also risks overloading that authority and may leave gaps in sector-specific knowledge needed to assess compliance in areas like health, finance, or critical infrastructure.

Another institutional logic is emerging in countries such as Ireland and Latvia, which have opted for a distributed model. Here, responsibilities are spread across multiple sectoral regulators, with Ireland designating eight separate entities and Latvia entrusting 14 public institutions with elements of AI oversight.[5] This approach draws on existing expertise, allowing regulators with deep knowledge of their sectors to supervise AI applications within their own domain. It also avoids the cost and time involved in creating entirely new structures. However, distribution increases coordination challenges, creating uncertainty over jurisdiction, divergent guidance across sectors, and a risk of regulatory fragmentation if not managed effectively. Ultimately, the effectiveness of either model hinges on strong collaboration within or across Member States, as only close cooperation between supervisory authorities can prevent regulatory fragmentation.[43]

The situation regarding national public authorities tasked with safeguarding fundamental rights is notably better, as nearly all countries – except for Hungary and Italy – have formally designated specific bodies.[44] The analysis of the European Commission's consolidated list of fundamental rights protection authorities shows that there are five common types of such structures; however, their presence across the Union is uneven. While Data Protection Authorities and Ombudsman Offices exist in most countries. Equality bodies are less consistently represented, and Human Rights institutions and Parliamentary Committees appear in relatively few Member States. The countries missing certain types of public authorities in the member states are listed below (Figure 5).

Figure 5. Missing Oversight Structures for Fundamental Rights in the EU



Source: Compiled based on the consolidated list of all identified authorities. https://digital-strategy.ec.europa.eu/en/policies/ai-act-governance-and-enforcement

[42] https://www.clydeco.com/en/insights/2025/05/preparing-for-enforcement-a-guide-to-the-eu-ai-act

[43] https://www.autoriteitpersoonsgegevens.nl/en/current/final-recommendation-on-supervision-of-ai-sector-and-centrally-coordinated

[44] As required by November 2024. Commission consolidated list. https://digital-strategy.ec.europa.eu/en/policies/ai-act-governance-and-enforcement

Even in the absence of finalised national designations, some regulators – mainly data protection authorities – are already asserting their capacity to act against prohibited AI practices, especially where personal data is involved. This underscores that compliance with Article 5 of the AI Act, alongside existing data protection obligations, cannot be postponed. Italy is a case in point. Its Data Protection Authority – the *Garante*[45] – though not formally designated as an AI fundamental rights oversight body, effectively assumed that role when it blocked the AI application DeepSeek. The Garante concluded that the company's explanations regarding data sources, purposes, legal basis, and storage locations were "totally insufficient".[46] By intervening in this way, the *Garante* has demonstrated how existing regulators can already shape the enforcement landscape and fill gaps until formal AI Act structures are in place.

1.6.2 Innovation enablers

While ensuring compliance with the EU AI Act, for the purpose of fostering innovation Member States are required to establish at least one AI regulatory sandbox by 2 August 2026 (Article 57), offering a controlled environment for developing, testing, and validating AI systems before market launch. These sandboxes allow guided experimentation under the supervision of national authorities, who provide regulatory advice, monitor risks, and issue exit reports. Participating providers – especially SMEs and start-ups – can use sandbox documentation as proof of conformity and are shielded from administrative fines if they follow prescribed protocols in good faith, although liability for third-party damages still applies.

The sandbox model aims to balance innovation with compliance. Drawing on fintech-sector evidence from the UK, where sandboxing boosted investment and reduced time-to-market, EU sandboxes anticipate similar benefits: greater legal certainty, streamlined regulatory processes, and strengthened innovation ecosystems. While Member States enjoy flexibility in design – ranging from centralised national programmes (e.g., Denmark, Spain) to regional models (e.g., Belgium, Slovakia) – the Commission and EU-wide initiatives like EUSAiR, Testing and Experimentation Facilities (TEFs), and European Digital Innovation Hubs (EDIHs) aim to foster coordination, share best practices, and prevent fragmentation. As rollout progresses, monitoring by the AI Board and the AI Office will be pivotal to ensuring these sandboxes not only fuel innovation but also strengthen convergence in compliance and governance across the EU.[47]

1.6.3 Member States' adoption of the AI Act

Spain stands out as the most advanced example directly aligned with the EU's regulatory vision. Its pilot, launched in 2022 under a Royal Decree, was explicitly designed to test the requirements of the AI Act, including risk management, technical documentation, and supervisory oversight. Managed by the Agency for the Supervision of Artificial Intelligence together with the national data protection authority,[48] and supported by dedicated funding from the Recovery Plan, Spain's sandbox is effectively a prototype of the future European model.

[45] Garante per la protezione dei dati personali.

[46] https://usercentrics.com/knowledge-hub/eu-regulators-scrutinize-deepseek-for-data-privacy-violations/#:~:text=The%20information%20DeepSeek%20provided%20was,and%20store%20European%20users'%20data

[47] Although the development of AI regulatory sandboxes is progressing unevenly across the EU, several Member States already offer useful lessons for implementation, adopting varied approaches in terms of legal basis, funding models, institutional leadership, and overall alignment with the forthcoming AI Act.

[48] Agencia Española de Supervisión de la Inteligencia Artificial (AESIA).

Germany has taken a systemic path, embedding "experimentation clauses" in sectoral legislation and backing them with substantial funding. These clauses allow temporary derogations from existing rules to test innovative solutions in areas such as mobility and energy. While not yet AI-specific, the German framework demonstrates how sandboxes can be institutionalised through law rather than limited to regulator-led pilot projects. Given its AI strategy explicitly references sandboxes, Germany is structurally well placed to transition toward AI Act-aligned experimentation.

Other Member States have so far concentrated on data protection sandboxes led by their national DPAs. Denmark, France, and Luxembourg have each established sandbox programmes that provide guidance on GDPR compliance. The Danish and Luxembourgish sandboxes, run by their data protection authorities, offer structured, multi-phase processes for testing but no funding or legal exemptions. France's CNIL runs thematic calls and publishes public recommendations from its cohorts. Although these sandboxes do not provide formal regulatory relief, they have built practical expertise in balancing innovation with privacy safeguards. France is somewhat exceptional in that it has also experimented with true legal sandboxing: the 2023 Olympic Games Law temporarily allowed the testing of algorithmic video analytics in public spaces, offering a rare example of experimentation embedded in primary legislation.

For the implementation of the AI Act, these national experiences highlight that the most effective sandboxes are those with a solid legal foundation, adequate funding, and a clear focus on AI-specific regulatory challenges. Spain is already operationalising the Act, Germany is building durable legal structures, and Norway offers tested best practices, while the other Member States are in earlier phases of adaptation.

Eventually, having adequate regulation and an enabling environment in place does not guarantee the institutional or human capacity required for effective implementation. The Brookings analysis of selected EU AI strategies,[49] which benchmarks countries along both people and technology dimensions, reveals disparities that could shape the rollout of the AI Act. France and Germany combine strong innovation ecosystems with high regulatory and skills capacity, positioning them to establish competent authorities, manage sandboxes, and oversee conformity assessments effectively.

Sweden, Denmark, and the Netherlands excel technologically but risk enforcement gaps if supervisory expertise and trained personnel are not expanded. Lithuania, Portugal, and Poland have governance frameworks in place but face limitations in technical expertise and institutional depth, while Italy and Greece, weaker on both dimensions, may struggle to build organisations and staff them adequately. Smaller digital frontrunners such as Estonia and Latvia demonstrate agility in digital services but face sustainability challenges in retaining specialised expertise. The lesson is clear: uneven development of both people and technology risks fragmenting enforcement. To operationalise the AI Act consistently across the Union, Member States must invest not only in governance structures but also in human capital to translate legal frameworks into effective practice.

Norway, though outside the EU but within the EEA, provides the most mature operational sandbox in practice. Since 2020, the Norwegian Data Protection Authority has run a multidisciplinary programme that supports AI developers in health, finance, and other sensitive sectors. The sandbox has stable public funding, has already delivered more than a dozen projects, and now extends to generative AI. Its model – cross-sectoral, well-financed, and thoroughly evaluated – closely resembles the approach envisaged in the AI Act and serves as a reference point for EU Member States.

[49] https://www.brookings.edu/articles/winners-and-losers-in-the-fulfilment-of-national-artificial-intelligence-aspirations/

1.7 EU-US AI Cooperation under the Trade and Technology Council

Launched at the June 2021 U.S. – EU Summit, the Trade and Technology Council (TTC) is the key institutional mechanism for transatlantic cooperation on artificial intelligence (AI). Its objective is to bridge regulatory differences, promote interoperability, and operationalise shared ethical principles in AI development and deployment, while more broadly serving as a forum to strengthen cooperation on global technology, trade and economic issues, and to advance shared competitiveness and market-oriented values. In practice, the TTC initiative helps to connect the EU's binding regulatory approach, embodied in the 2024 AI Act, with the United States' framework-based model, represented by NIST's AI Risk Management Framework (AI RMF 1.0 introduced in 2023). Through this dual-track cooperation, the TTC seeks to reduce regulatory fragmentation and foster innovation while ensuring safety, transparency, and accountability in AI systems.[50]

1.7.1 Operationalisation of Governance Frameworks and Institutional Arrangements

The Trade and Technology Council (TTC) is jointly led at the political level by senior U.S. officials – the U.S. Trade Representative together with the Secretaries of Commerce and State, and by European Commissioners for Trade and Competition on the EU side. Since its launch in 2021, the TTC has convened six ministerial-level meetings, alternating between U.S. and EU venues.[51]

The Council's work is carried forward through ten thematic working groups, which span issues such as technology standards, supply chains, ICT security, climate and clean technologies, and global trade challenges (see Table 4). The cooperation process is carried out through thematic coordination, information exchange, and stakeholder dialogue, engaging representatives from business, labour, and civil society.

Within this framework, AI cooperation is anchored in the groups on technology standards and data governance and platform regulation. At the sixth ministerial in Leuven, in April 2024, this collaboration established the *EU AI Office – U.S. AI Safety Institute Dialogue*, which links the EU's new supervisory authority under the AI Act with NIST's AI Safety Institute. The Dialogue is designed to facilitate cooperation on safety evaluations, testing protocols, and information exchange, thereby strengthening oversight capacity on both sides.

Among the TTC's most tangible achievements are the publication of the *Joint Roadmap on Evaluation and Measurement Tools for Trustworthy AI and Risk Management (2022)* and the *EU – U.S. Terminology and Taxonomy for AI* **(**TTAI, 2024). The Roadmap translates ethical principles such as fairness, robustness, and transparency, into actionable instruments, including benchmarks, testing methodologies, and risk classification criteria. It also creates crosswalks between the EU AI Act's binding risk-based obligations and the voluntary NIST AI RMF, enabling organisations to align compliance efforts across jurisdictions (NIST, 2022).

[50] Joint Statement EU-US Trade and Technology Council of 4-5 April 2024 in Leuven, Belgium. European Commission (2024). https://ec.europa.eu/commission/presscorner/detail/sn/statement_24_1828

[51] U.S.-EU Trade and Technology Council: Background and Issues (2024). https://www.congress.gov/crs_external_products/IF/PDF/IF12575/IF12575.5.pdf

Table 4. TTC Working Groups

#	Working Group Title
1	Technology standards
2	Climate and clean technology
3	Secure supply chains
4	Information and communications technology and services (ICTS) security and
5	competitiveness
6	Data governance and technology platform regulation
7	Misuse of technology threatening security and human rights
8	Cooperation on export controls of dual use items
9	Investment screening cooperation
10	Promoting small- and medium-sized enterprise (SME) access to and use of digital technologies
	Global trade challenges

Source: CRS, based on U.S. Department of State.

Complementing this, the publication of the EU – U.S. Terminology and Taxonomy for AI harmonised over sixty-five technical and regulatory terms, reducing interpretative ambiguities in procurement, audits, and conformity assessments. This effort demonstrates the TTC's role in making governance "operational" by offering a shared language to regulators, businesses, and international organisations.

Through these arrangements, the TTC embeds ethical values into regulatory practice. For instance, the Joint Roadmap advances evaluation methodologies for auditing bias and robustness, while the AI Office – AISI Dialogue creates channels for real-time incident reporting and evaluation of frontier models, enhances information exchange on risks and benchmarks. Furthermore, joint TTC statements have addressed the risks of disinformation and synthetic media, encouraging platforms to apply safeguards against deepfakes and malign AI use. It also commit to combating foreign information manipulation and interference (FIMI) and call on platforms to better protect users, including in election contexts. By incorporating transparency and accountability into regulatory design and procurement rules, TTC initiatives operationalise ethics in both public administration and market supervision.

1.7.2 Adaptation and Challenges: Governments and the Private Sector

TTC cooperation has encouraged both governments and firms to adjust their practices to a more coordinated transatlantic framework. For companies, the *Joint Roadmap* provides a bridge between the EU AI Act and the NIST AI Risk Management Framework, allowing risk controls and documentation developed in one system to be mapped onto the other. This reduces compliance costs for multinational firms and eases cross-border audits and procurement.

Furthermore, the launch of the EU – U.S. *AI for the Public Good Research Alliance* in 2024 further advances this agenda by promoting collaborative research in areas directly linked to the Sustainable Development Goals (SDGs), including climate resilience, sustainable agriculture, public health, energy systems, and disaster preparedness.[52] In parallel, the Transatlantic Initiative on Sustainable Trade (TIST), launched under the TTC, promotes cooperation toward

[52] AI Alliance (2024) AI for Public Good: EU-U.S. Research Alliance in AI for the Public Good. https://digital-strategy.ec.europa.eu/en/library/ai-public-good-eu-us-research-alliance-ai-public-good

"climate-neutral" economies through workstreams on clean energy, critical minerals, sustainable supply chains, and green procurement, as well as potential alignment on conformity assessments.[53]

Despite progress, several challenges persist. First, the *regulatory asymmetry* between the binding nature of the EU AI Act and the voluntary approach of the NIST RMF creates uneven compliance expectations. Second, TTC outcomes depend heavily on *political momentum* and ministerial cycles, raising concerns about continuity and enforceability**,** durability, especially in light of electoral changes. Furthermore, while TTC outcomes are significant on both sides of the Atlantic Ocean, their *global diffusion is limited*. Although the TTC leverages work of ISO, the OECD, and IEEE, developing economies see little direct benefit, raising questions about inclusivity and broader applicability.

Divergences on sensitive issues continue to test the TTC, yet they also underscore its value as a forum for alignment. The U.S. and EU share concerns about China's global influence but differ in their preferred strategies, making the TTC an essential platform to forge a common approach to "de-risking". Likewise, while domestic differences complicate regulation of digital markets and artificial intelligence, the TTC provides space to build guardrails and promote interoperability. Both sides have also stressed the need for deeper coordination, underscoring the transatlantic alliance as a key driver of a 21st Century democratic digital agenda. (Echikson et al., 2024; EC, 2024; Lumenova.ai, 2025).

1.7.3. Capacity-Building Implications

The Trade and Technology Council (TTC) has gradually become not only a forum for regulatory alignment but also a vehicle for transatlantic capacity-building. While its primary mission is to harmonise approaches to AI governance, the Council's initiatives have generated tangible improvements in institutional, human, and technical capacities on both sides of the Atlantic and in partner countries.

One of the earliest structured efforts was the launch of the Talent for Growth Task Force in 2022. By bringing together government officials, businesses, labour unions, and training organisations, the Task Force served as a laboratory for identifying skills gaps in the technology sector and exploring innovative approaches to workforce development. Its activities underscored the interdependence between regulation and human capacity: without a trained workforce, interoperability between the EU's AI Act and the U.S. NIST AI RMF would remain aspirational. The Task Force's recommendations highlighted the need to prepare workers for the green transition, strengthen digital competencies, and embed ethical AI practices into corporate training schemes. Although its mandate was limited to one year, it set a precedent for embedding capacity-building directly into the TTC's policy architecture.

A second major capacity-building channel emerged through the Transatlantic Initiative on Sustainable Trade (TIST), launched at the May 2023 ministerial Meeting. The TIST combined regulatory cooperation with practical learning tools. The publication of a Joint Catalogue of Best Practices on Green Public Procurement equipped both European and American public buyers with concrete strategies to embed sustainability criteria into procurement decisions. This catalogue is not merely a guidance document; it acts as a capacity-building instrument

[53] U.S.-EU Trade and Technology Council: Background and Issues (2024). https://www.congress.gov/crs_external_products/IF/PDF/IF12575/IF12575.5.pdf

by standardising language, providing model clauses, and lowering the knowledge barriers for civil servants and SMEs entering green supply chains. In turn, it strengthens the ability of both regions to steer markets toward climate-neutral goods and services.

The TTC also expanded its focus to stakeholder-driven green transition initiatives, most visibly through the "Crafting the Transatlantic Green Marketplace" event in Washington, D.C., in January 2024. By convening policymakers, industry representatives, and civil society, the event reinforced the TTC's function as a capacity-building hub that goes beyond governments to mobilise entire ecosystems. Proposals generated at this event are now feeding into the TTC's pipeline of collaborative projects, ranging from clean energy cooperation to shared conformity assessment schemes.[54]

In addition, the EU AI Office – U.S. AI Safety Institute (USAISI) Dialogue further reinforces institutional capacity through dedicated funding and resource commitments. In the United States, Congress appropriated USD 10 million in FY2024 to NIST for establishing the AI Safety Institute and developing testing protocols, risk benchmarks, and frontier model evaluation tools.[55] Meanwhile, the EU AI Office,[56] established in 2024 under the AI Act, brings together more than 140 staff from diverse backgrounds, including technical specialists, policy experts, and legal professionals, to oversee implementation of the Act, supervise general-purpose AI models, and coordinate with national authorities. Linking these institutions under the TTC framework transforms parallel efforts into a joint capacity-building mechanism, allowing regulators to pool expertise, reduce duplication, and accelerate the development of robust supervisory practices.

Beyond Europe and the United States, TTC discussions have led to capacity-support for partner countries, particularly in the domain of secure digital infrastructure. Joint EU–U.S. efforts in Tunisia, for instance, provided technical training to public officials, businesses, and IT professionals on cybersecurity standards and 5G deployment. Similar support is being extended to Costa Rica, Jamaica, Kenya, and the Philippines through partnership and investment mechanisms coupled with technical exchanges. These initiatives illustrate how TTC cooperation can be externalised to reinforce resilience in other countries, linking capacity-building to broader geopolitical goals of promoting secure and democratic digital ecosystems.

1.8 ASEAN AI Governance Framework

1.8.1 Building Responsible AI Systems in a Dynamic Global Landscape

The Association of Southeast Asian Nations (ASEAN) stands at the forefront of artificial intelligence adoption, with AI projected to contribute 10-18% to the region's GDP by 2030 (Labrecque, 2024). As ASEAN nations navigate the complex landscape of AI governance while maintaining competitive advantages in innovation, the development of comprehensive frameworks becomes critical for ensuring responsible AI deployment across diverse national contexts and regulatory environments. However, as AI governance involves systems, policies,

[54] White House (2024) U.S-EU Joint Statement of the Trade and Technology Council. https://bidenwhitehouse.archives.gov/briefing-room/statements-releases/2024/04/05/u-s-eu-joint-statement-of-the-trade-and-technology-council-3/

[55] U.S. Senate (2024). First-Of-Its-Kind Funding To Establish A U.S. Artificial Intelligence Safety Institute. https://www.democrats.senate.gov/newsroom/press-releases/majority-leader-schumer-announces-first-of-its-kind-funding-to-establish-a-us-artificial-intelligence-safety-institute-funding-is-a-down-payment-on-balancing-safety-with-ai-innovation-and-will-aid-development-standards-tools-and-tests-to-ensure-ai-systems-operate-safely

[56] The Commission established the AI Office to strengthen EU leadership in safe and trustworthy Artificial Intelligence. https://ec.europa.eu/commission/presscorner/api/files/document/print/en/ip_24_2982/IP_24_2982_EN.pdf
Denmark, Finland, Germany, Hungary, Ireland, Italy, Poland, Romania, Spain.

and practices that foster responsible, ethical, and safe AI development and use; the challenge lies in operationalising these frameworks to turn policies into actionable practices that protect individuals, support innovation, and build trust in AI technologies.[57]

1.8.2 Translating Policy into Practice

The translation of AI policies into practical implementation measures involves diverse regulatory approaches that reflect different national priorities and capabilities. The European Union has adopted the comprehensive, binding EU AI Act as a risk-based legal framework, while ASEAN has taken a more flexible approach with its non-binding Guide to AI Governance and Ethics. This non-binding approach aims to balance innovation with risk mitigation, adapting to the diverse internal regulations and development levels across member states (Toushik Wasi et al., 2025).

International organisations play crucial roles in shaping these approaches. The OECD provides foundational principles that inform national policies, while UNESCO offers AI Ethics Readiness Assessments to help countries identify governance gaps and create national roadmaps (World Bank, 2025). Technical standards such as the NIST AI Risk Management Framework and ISO/IEC 42001 provide practical means for operationalising responsible AI principles like transparency, explainability, and accountability.[58]

ASEAN member states are developing their own strategies that reflect this flexible approach. Singapore has established its Model AI Governance Framework alongside the 'AI Verify' toolkit for testing and auditing AI systems. Malaysia has developed National Guidelines. Thailand has created Ethical Guidelines, and Uruguay has focused on public transparency and accountability for AI in government applications. These varied approaches demonstrate how the region's principle-based framework allows for contextual adaptation while maintaining common ethical foundations.[59]

1.8.3 Institutional Arrangements and Enforcement

Effective AI governance requires integrated ecosystems that combine institutional structures, legal instruments, and enforcement mechanisms. At the global and regional levels, institutional arrangements are evolving rapidly. The EU AI Act has established the European AI Office, the European Artificial Intelligence Board, and the European Centre for Algorithmic Transparency, creating a comprehensive governance architecture.[60] ASEAN continues to develop its guidance framework, working on additional guidelines for generative AI applications.

National approaches to institutional arrangements vary significantly. Some countries are creating dedicated AI agencies, such as Malaysia's proposed National AI Authority, while others leverage existing regulators like data protection authorities, etc. Countries like China and Nigeria have proposed AI Ethics Committees to oversee development and conduct audits, while Estonia demonstrates effective central coordination through its Government CIO Office and Ministry of Economic Affairs and Communications (World Bank, 2025).

[57] *AI governance: What it is & how to implement it* (K. Farnham, 2025). https://www.diligent.com/resources/blog/ai-governance

[58] Ditto.

[59] *Beyond the matrix: AI governance gaps in Southeast Asia* | CSIS (H. T. Htoo (2025). https://www.csis.org/blogs/new-perspectives-asia/beyond-matrix-ai-governance-gaps-southeast-asia

[60] *Governing AI: Evolving institutional responses. Istanbul Innovation Days* (UNDP, 19 August 2025). https://istanbulinnovationdays.org/governing-ai-evolving-institutional-responses/

The legal instruments supporting these institutional arrangements span from binding legislation to voluntary guidelines. Hard law approaches include the EU AI Act's comprehensive risk-based framework and China's detailed regulations for specific AI applications. Soft law approaches, exemplified by ASEAN's Guide and international principles from the OECD and UNESCO, offer flexibility while providing important normative guidance. Technical standards bridge these approaches by providing specific implementation guidance that can demonstrate compliance with mandatory requirements.

Enforcement mechanisms are becoming increasingly sophisticated and adaptive. Regulatory sandboxes provide controlled environments for testing new AI approaches while fostering innovation, with the EU AI Act mandating such sandboxes by August 2026.[61] Singapore's AI Verify toolkit exemplifies voluntary self-testing against ethical and technical compliance standards. Regular audits and oversight, including mandatory Algorithmic Impact Assessments in countries like Canada and Uruguay, ensure ongoing compliance and accountability (World Bank, 2025).

The most effective approaches emphasise continuous monitoring and agile governance. AI systems require ongoing performance and ethical compliance monitoring, including tracking for model drift and maintaining incident response protocols.[62] Clear accountability structures, potentially including Chief AI Officers or AI Governance Committees, establish fundamental responsibility frameworks. Organisations are increasingly adopting "governance sprints" – short, focused cycles where cross-functional teams rapidly address emerging risks – and establishing "AI Risk Labs" that bring together diverse experts to co-design governance solutions.[63]

1.8.4 Embedding Ethical Principles

The Expanded ASEAN Guide on AI Governance and Ethics for Generative AI (2025) establishes seven core principles that member states must adapt to their diverse contexts: transparency, fairness, security, human-centricity, privacy, accountability, and robustness. Unlike universal AI governance approaches, ASEAN's framework addresses unique regional challenges including linguistic diversity across ten member states and varying AI readiness levels.

Transparency and explainability in ASEAN requires culturally sensitive approaches, demonstrated by region-specific models like *ThaiLLM*, VinAI's *PhoGPT*, and Singapore's *SEA-LION* that address local languages under-represented in Western-developed AI systems. Accountability mechanisms must navigate complex value chains across ASEAN's diverse regulatory environments, with Accenture's six-country implementation showing how multinational governance structures can span the region.

Fairness and bias mitigation becomes critical given ASEAN's cultural diversity, where standard Western bias metrics may miss regional nuances. Singapore's Project Moonshot exemplifies this by incorporating ASEAN-specific benchmarks for cultural relevance. Privacy protection leverages existing frameworks like the ASEAN Framework on Personal Data Protection and ASEAN Data Protection and Privacy Forum (ADPPF) to coordinate across member states while respecting sovereignty. This integrated approach supports ASEAN's goal of capturing the estimated

[61] *AI Regulatory Sandbox Approaches: EU Member State Overview | EU Artificial Intelligence Act.* https://artificialintelligenceact.eu/ai-regulatory-sandbox-approaches-eu-member-state-overview/

[62] *Fostering AI innovation while Ensuring Responsible Governance*. OCEG (L. Dittmar, 17 October 2024). https://www.oceg.org/fostering-ai-innovation-while-ensuring-responsible-governance/

[63] *AI governance organisations must evolve for a new era.* OCEG (L. Dittmar, 20 November 2024). https://www.oceg.org/ai-governance-organizations-must-evolve-for-a-new-era/

USD 6 trillion Gen AI economic opportunity through coordinated, culturally-aware governance (ASEAN Secretariat, 2025).

1.8.5 International Organisations Shaping AI Governance

International and multilateral organisations play vital roles in shaping and applying AI governance, particularly in development contexts where they work to ensure AI serves sustainable development goals and human rights (World Bank, 2025). The UN system has made significant commitments to leveraging AI for achieving Sustainable Development Goals while protecting human rights and ensuring global benefits.

Specialised UN agencies contribute distinct expertise to this effort. UNDP supports AI projects and data foundations for equitable digital responses while managing ethical challenges in development contexts. UNESCO champions ethical AI in education and culture through its AI Ethics Readiness Assessment Methodology, which has helped countries like Lao PDR identify governance gaps and develop national roadmaps. The International Telecommunication Union focuses on bridging the global AI skills gap through initiatives like the AI Skills Coalition, which provides training particularly targeted at developing regions (Goel et al., 2025).

These organisations play crucial roles in norm-setting and policy harmonisation. UNESCO's Recommendation on the Ethics of AI represents the first global ethical standard adopted by 193 member states. The OECD AI Principles provide foundational guidelines for trustworthy AI that have influenced frameworks like the EU AI Act. Many experts favour hybrid governance models that combine soft law approaches with binding regulations, and voluntary AI standards are increasingly referenced in binding laws.

The integration of AI ethics into global development programmes ensures equitable and human-centric outcomes. This integration is anchored in human rights and Sustainable Development Goals, with ethical guidelines promoting transparency, accountability, and privacy. Proposed UN AI institutions could establish funding mechanisms for AI projects addressing SDGs in critical areas like healthcare and education, while existing multilateral donors increasingly prioritise digital governance projects that incorporate AI ethics.[64]

Capacity building and digital divide mitigation represent crucial areas where international organisations can make significant impacts. Programmes like China's "AI Capacity-Building Action Plan" target underdeveloped countries to build infrastructure and share expertise.[65] The ITU's AI Skills Coalition provides training to bridge the global AI skills gap,[66] while UNCTAD's "AI for Development" programme assesses policy readiness and supports capacity building for least developed countries.[67]

[64] A UN Institution for AI Governance: why and how it could be a good idea – The Digital Constitutionalist. *The Digital Constitutionalist – The Future of Constitutionalism*. (F. M. Abrusci, 22 May 2024). https://digi-con.org/a-un-institution-for-ai-governance-why-and-how-it-could-be-a-good-idea/

[65] *Artificial Intelligence Capacity-Building Action Plan for good and for all to support global AI development and governance.* International Research Center for AI Ethics and Governance (Yizengcasia, 27 September 2024). https://ai-ethics-and-governance.institute/2024/09/26/artificial-intelligence-capacity-building-action-plan-for-good-and-for-all-to-support-global-ai-development-and-governance/

[66] *Press release* (ITU, 20 January 2025). https://www.itu.int/en/mediacentre/Pages/PR-2025-01-20-AI-education-to-close-the-AI-skills-gap.aspx

[67] Ditto.

1.8.6 Government and Private Sector Adaptation

Governments worldwide are rapidly adapting their institutional structures and policy frameworks to address AI governance challenges. This adaptation involves diverse organisational models reflecting different national contexts and priorities. Some countries like Spain and Malaysia have established dedicated AI agencies to centralise policy coordination, while others like the United Arab Emirates have appointed dedicated ministers for AI to integrate governance at the cabinet level (Toushik et al., 2025).

The European Union has created a complex multi-layered governance ecosystem under the EU AI Act, while the United States has adopted a decentralised approach requiring federal agencies to designate Chief AI Officers and establish AI Governance Boards. Countries are also adapting existing regulatory frameworks, with some jurisdictions delegating AI regulation enforcement to existing Data Protection Authorities.

Governments are moving toward dynamic policy frameworks that support agile and adaptive governance. This includes using regulatory sandboxes for controlled testing and adopting outcome-based approaches that offer flexibility while maintaining safety standards. Public procurement policies are increasingly leveraged to promote responsible AI by setting ethical requirements in government contracts, recognising that governments are often AI consumers rather than developers.[68]

Private sector adaptation reflects the dual challenge of harnessing AI's potential while navigating complex regulatory landscapes. Organisations are embracing dynamic governance structures that match AI's evolving nature through governance-by-design approaches that embed oversight throughout the AI lifecycle. This includes automated fairness audits during development, real-time monitoring during deployment, and continuous compliance assessment during operations.[69]

Executive leadership structures are evolving with many companies establishing Chief AI Officer roles or AI Governance Committees to oversee strategy, ethics, and risk management. Cross-functional teams including legal, policy, compliance, engineering, and operations personnel work together to map AI system lifecycles, identify obligations, and implement safeguards including human-in-the-loop oversight.

The most successful private sector approaches view governance as a strategic advantage rather than merely a compliance burden. Strong governance builds trust with customers, regulators, and partners, opening doors to new markets and partnerships, particularly in highly regulated industries. Organisations adopt risk-tiered governance where high-risk systems receive rigorous oversight while lower-risk applications have lighter controls to foster innovation (etc).

1.8.7 Challenges and Global Governance Gaps

The global AI governance landscape faces significant regulatory, institutional, and capacity-related challenges that hinder effective coordination and implementation. The fundamental pacing problem sees technological advancement consistently outpacing regulatory development, with AI evolving at unprecedented rates that make static governance frameworks quickly

[68] *Governing AI: Evolving institutional responses.* Istanbul Innovation Days (UNDP, 19 August 2025). https://istanbulinnovationdays.org/governing-ai-evolving-institutional-responses/

[69] *AI governance organisations must evolve for a new era.* OCEG (L. Dittmar, 20 November 2024). https://www.oceg.org/ai-governance-organizations-must-evolve-for-a-new-era/

obsolete. This creates substantial governance gaps where regulations lag behind innovation, particularly challenging with generative AI's exponential development pace.

The "black box" problem compounds these challenges, as many advanced AI models operate with decision-making processes that are difficult to interpret even for experts. This opacity erodes trust, complicates accountability, and makes effective governance challenging, especially in critical applications like healthcare or finance where understanding reasoning is essential. Current legal frameworks struggle to define accountability for autonomous systems that make independent decisions.

Regulatory fragmentation represents another major challenge, with little global consensus on how AI should be regulated. Different countries adopt varied frameworks reflecting distinct political, economic, and cultural contexts, leading to a confusing global landscape. The EU's strict comprehensive approach contrasts sharply with the US's decentralised, sector-specific model, while countries often emphasise different underlying values in their regulatory approaches.[70]

Institutional challenges include the fragmented international system that exists as a "regime complex" – a loosely connected network of partially overlapping institutions without central authority. This leads to institutional ambiguity and difficulties establishing clear governance pathways. Many existing institutions were not designed for AI governance, with AI mandates added without sufficient reconfiguration of institutional capacities, resulting in duplicative efforts and coordination fatigue (Toushik et al., 2025).

The concentration of power among large technology companies creates risks of regulatory capture, where regulations align with corporate interests or favour voluntary frameworks over binding laws. This can marginalise voices from civil society and smaller businesses while limiting effective interdisciplinary collaboration needed for comprehensive AI impact assessments (etc).

Capacity-related challenges particularly affect developing nations, which face insufficient infrastructure, limited technical expertise, and financial constraints that risk marginalisation in the AI race. The global AI skills gap and digital literacy challenges slow AI diffusion, while inconsistent documentation and data silos in organisations obscure transparency and accountability.

1.8.8 Geopolitical and Economic Asymmetries

Geopolitical asymmetries significantly shape current governance shortcomings, with AI becoming a pivotal geopolitical instrument central to national sovereignty and global influence. The intensifying competition between major powers creates "innovation blocs" and strategic alliances based on technological capabilities and regulatory philosophies, often leading nations to prioritise competitiveness over safety concerns (Toushik et al., 2025).

The rise of "*minilateralism*" sees small, strategically aligned coalitions attempting to govern specific AI aspects, but these efforts risk deepening global fragmentation and creating competing regulatory regimes with limited interoperability. Diplomatic mistrust complicates efforts to foster collaboration, while divergent philosophical underpinnings of "human-centric" AI can turn ethical frameworks into tools for asserting normative dominance.

Economic asymmetries concentrate wealth and power among a few countries and companies due to the high threshold for building foundational AI models. This concentration exacerbates

[70] *Transatlantic AI governance – Strategic implications for U.S. – EU compliance* (King and Spalding, 2025). https://www.kslaw.com/news-and-insights/transatlantic-ai-governance-strategic-implications-for-us-eu-compliance

the divide between technology "haves" and "have-nots," creating risks of "AI colonialism" where powerful nations leverage AI development to reinforce dominance over developing countries (Wang et al., 2024).

Technological asymmetries create significant disparities through unequal access to digital infrastructure, computing resources, and skilled workforces. The geographic concentration of critical AI components creates supply chain vulnerabilities, while export controls on advanced semiconductors become geopolitical levers that disrupt AI supply chains and prompt the formation of parallel AI ecosystems.

1.8.9 Capacity Building for Developing Countries

Capacity-building initiatives are crucial for developing economies to effectively govern AI and harness its potential for sustainable development. These initiatives address interconnected gaps through comprehensive approaches that combine technical assistance with institutional development models.

International initiatives demonstrate varied effectiveness in supporting AI governance in developing countries. The UN system provides multiple pathways, with proposed institutions that could establish funding mechanisms for SDG-aligned AI projects and facilitate knowledge-sharing platforms. The ITU AI Skills Coalition provides training on generative AI and machine learning while developing specialised government training in AI governance and ethics (UNSCEB, 2024).

UNDP's Digital Strategy includes nurturing AI projects, investing in data foundations, and managing ethical issues in international development. UNESCO's AI Ethics Readiness Assessment Methodology helps countries identify legal and governance gaps, with Chile successfully incorporating these insights into its updated National AI Policy.[71]

Regional initiatives show promise in building collective capacity. The African Union's AI Policy Framework, supported by UNDP and UNESCO, emphasises inclusion, equity, and sovereignty while encouraging interoperable laws and shared institutional infrastructure (Vujičić, 2023). ASEAN seeks to strengthen cooperation through regional capacity-building initiatives, though the non-binding nature of its guidance may limit consistency across varying national readiness levels (Prilliadi, 2022).

National initiatives demonstrate diverse approaches to capacity building. India's 'AI for All' program provides self-learning opportunities to raise public awareness, while Singapore's AI Apprenticeship Programme focuses on building technical talent alongside its governance frameworks. These programmes address priority capacity gaps including digital infrastructure, human capital development, legal expertise, data construction capabilities, public literacy, and institutional coordination (World Bank, 2025).

The most promising models combine multistakeholder collaboration with practical experimentation. Regulatory sandboxes offer controlled environments for testing new approaches while building regulatory capacity. Regional cooperation allows smaller countries to pool expertise and present unified positions in international forums. Community-driven governance initiatives ensure AI oversight reflects local languages, cultural values, and regional priorities.

[71] A UN Institution for AI Governance: why and how it could be a good idea – The Digital Constitutionalist (F. M. Abrusci, 22 May 2024). *The Digital Constitutionalist – The Future of Constitutionalism*. https://digi-con.org/a-un-institution-for-ai-governance-why-and-how-it-could-be-a-good-idea/

2 COUNTRY CASES

2.1 National AI Governance and Implementation Approaches

Countries pursue different governance models to coordinate the implementation of their national AI policies. In some cases, they assign oversight of the development and implementation strategies to an existing ministry, department or other government entity. Among existing ministries or agencies tasked with developing on implementing AI strategies, most often are the following: (i) information technology and communication ministries; (ii) economics or finance ministries; or (iii) education, science and technology and innovation ministries. In other cases, counties choose to create a new governmental or independent AI coordination entity. And in some other cases, they establish AI expert advisory entities, which are generally multistakeholder entities comprising AI experts tasked with identifying and reporting on current and future opportunities, risks and challenges arising from the use of AI in society. These entities provide recommendations to the government including on issues of data ethics and personal data protection. A selection of national AI governance varied approaches are presented below.

Figure 6. Selection of National AI Governance Approaches

Assigning oversight to an existing ministry or department	The White House Office of Science and Technology Policy oversees the **United States**' national Artificial Intelligence strategy. **Estonia**'s Ministry of Economic Affairs and Communications created the National Artificial Intelligence Strategy. **France** coordinates AI policy implementation from within the Prime Minister's Office.
Creating a new governmental or independent body for AI	AI policy in the **United Kingdom** is coordinated by the UK Government's Office for Artificial Intelligence. The **United States** White House established the National AI Initiative Office. **Singapore** created a National Artificial Intelligence Office to coordinate the implementation of its national AI strategy.
AI expert advisory entities	**Austria**'s Council on Robotics and Artificial Intelligence. **Canada**'s Advisory Council on Artificial Intelligence. **Spain**'s Artificial Intelligence Advisory Council. The **United States**' Select Committee on Artificial Intelligence under the National Science and Technology Council.
Oversight and advisory bodies for AI and data ethics	**Germany**'s Data Ethics Commission. The Data Ethics Advisory Group in **New Zealand**. The **United Kingdom**'s Centre for Data Ethics and Innovation (CDEI). **Singapore**'s Advisory Council on the Ethical Use of AI and Data.

Source: OECD (2021).

This part of the publication constitutes an overview of the AI ecosystem, i.e., AI Governance practices, regulatory frameworks, oversight mechanisms, and operationalisation practices in selective countries around the world, e.g., Canada, Republic of Korea, Philippines, U.K., and

Qazaqstan.[72] This overview also encompasses such areas as the state of infrastructure in place, capacity development initiatives for creating an AI-savvy workforce, the degree of penetration of AI technologies in the public sector, and the extent to which public-private partnerships are involved in implementation activities. It concludes with a short overview of the status of AI development and implementation in such countries as Armenia, Azerbaijan, Bangladesh, Cambodia, Georgia, Kyrgyzstan, Lao PDR, Mongolia, Tajikistan, and Uzbekistan.[73]

2.2 CANADA

2.2.1 AI Governance

Canada holds a distinctive role in the global AI landscape: it was the first country to launch a national AI strategy in 2017: the *Pan-Canadian Artificial Intelligence Strategy (PCAIS)*. The Strategy combines world-class research with innovative approaches to governance and policy experimentation and is coordinated by the Canadian Institute for Advanced Research (CIFAR) operating under the policy umbrella of the Innovation, Science and Economic Development (ISED), which acts as the main body for digitalisation and innovation (Attard-Frost et al., 2024). This governance model reflects a distinctive federated approach, balancing strong federal leadership with strong provincial and institutional autonomy, enabling regions to cultivate their own AI strengths while contributing to national priorities.

ISED provides strategic oversight, sets national priorities, and ensures that Canada's AI agenda is aligned with broader innovation and economic goals. CIFAR functions as the central coordinating body, while implementation relies heavily on Canada's three globally recognised AI hubs: *Mila* (Montréal) – a world leader in deep learning research; *Vector Institute* (Toronto) – specialising in machine learning and applied AI; *Amii* (Edmonton) – focusing on reinforcement learning and health-related AI applications (CIFAR, 2025).

CIFAR's functions include: advancing world-class AI research by funding cutting-edge projects and creating Canada CIFAR AI Chairs, prestigious research leadership positions designed to anchor global talent in Canada. Over 100 CIFAR AI Chairs (research professorships) have been created, providing long-term support to leading scientists while fostering both frontier AI research and the training of future generations of specialists (CIFAR, 2024). Beyond this, CIFAR strengthens the AI talent pipeline through graduate scholarships, postdoctoral fellowships, and international recruitment efforts, effectively reinforcing Canada's position as a premier destination for AI expertise; promoting cross-sector AI innovation, particularly in domains of public value such as healthcare, climate action, agriculture, and government services; and convening multi-stakeholder dialogue between government, academia, and industry to experiment with AI policy approaches and develop governance frameworks that reflect both innovation needs and ethical considerations.[74]

[72] Purposeful sampling was used to ensure representation of countries from the global AI governance landscape. Guided by the Global AI Vibrancy Rankings (2023) - https://hai.stanford.edu/ai-index/global-vibrancy-tool - two countries from the top ten were selected based on their relatively high scores in the Policy and Governance pillar; particularly relevant for the scope of this study. Accordingly, the United Kingdom (9.67/10), Canada (7.4/10), and the Republic of Korea (7.05/10) were selected due to their robust AI policy and governance frameworks. These countries also scored relatively well on both "Scale" and "Intensity" measures, https://www.tortoisemedia.com/data/global-ai#rankings. Two beneficiary countries of the joint RoK/UNDP/ACSH inter-regional project were also selected, the Philippines for its robust performance in the AI domain along with Qazaqstan. The latter was also chosen as it is not a member of any non-UN global initiatives – unlike other countries mentioned earlier. As such, it represents a broader group of countries that are not engaged in any major efforts on enhancing AI governance and implementation (UN, 2024).

[73] These are the countries that are part of capacity development project implemented jointly by UNDP/ACSH and MOIS/NIA. https://www.astanacivilservicehub.org/articles/view/second-joint-regional-project-on-digital-transformation-and-digital-governance

[74] By embedding governance functions into a research-led framework, Canada has avoided a purely top-down model and instead promotes co-creation of policies with research and innovation actors.

The Strategy is underpinned by three core pillars: (i) Commercialisation – supporting national AI institutes and industry-led innovation clusters to translate research into applications; (ii) Standards and Responsible AI – advancing safety, accountability, and interoperability frameworks through the Standards Council of Canada; and (iii) Talent, Research, and Compute – consolidating leadership through Canada CIFAR AI Chairs, expanding advanced research programmes, and investing in computing infrastructure (ISED, 2025). Together, these phases and pillars illustrate Canada's shift from scientific leadership toward societal and economic transformation, ensuring that cutting-edge research produces tangible benefits for citizens, industries, and public services.

Institutionally, Canada has established an Advisory Council on Artificial Intelligence (2019),[75] composed of academics, industry leaders, civil society representatives, and government officials, to provide independent and strategic advice on national AI priorities. Although the Council's mandate has evolved over time, its creation clearly signalled Canada's commitment to multi-stakeholder input, ensuring that governance frameworks reflect not only innovation priorities but also ethical and societal considerations. More recently, in May 2025, Canada appointed its first Minister of Artificial Intelligence and Digital Innovation, marking a significant step toward formalising AI governance at the federal level (IAPP, 2025). Yet, Canada continues to rely on a federated governance model, where ISED, research institutions such as CIFAR, and provincial governments collaborate in tandem to design and implement policies. This layered approach balances centralised leadership with distributed expertise, reinforcing Canada's principles-based and adaptive regulatory style.

Internationally, Canada amplifies its governance role by acting as a founding member and host of the Secretariat of the Global Partnership on AI (GPAI) in Montréal, while also contributing actively to the OECD AI Policy Observatory, G7, G20, and United Nations dialogues. These platforms allow Canada to shape global and regional AI norms, ensuring its domestic governance aligns with international cooperation, best practices.

2.2.2 AI Regulation

Canada's regulatory approach to AI is evolving into a multi-layered governance model that integrates federal strategies, operational directives, sectoral standards, and provincial measures. Although the absence of a single binding national AI law has resulted in some fragmentation, Canada has developed one of the world's most principles-based, adaptive, and ethically grounded public-sector AI frameworks. This approach allows flexibility across jurisdictions while safeguarding transparency and innovation. Rather than imposing rigid, top-down controls, Canada relies on risk-based regulation, transparency, and infrastructure governance as key instruments to balance innovation with public trust (LCO, 2021).

[75] *The Government of Canada creates Advisory Council on Artificial Intelligence. (2019).* https://www.canada.ca/en/innovation-science-economic-development/news/2019/05/government-of-canada-creates-advisory-council-on-artificial-intelligence.html

Table 5. Layers of Canada's AI Governance

Layer	Key Regulations / Instruments	Purpose
Federal Strategies & Legislation	Pan-Canadian AI Strategy (2017, renewed 2022); Artificial Intelligence and Data Act – AIDA (Bill C-27, 2022 – lapsed 2024); Canadian AI Safety Institute – CAISI (2024).	National innovation priorities, AI research, and infrastructure; drive funding, attempted federal legal framework for "high-impact" systems.
Operational Directives (Public Sector)	Directive on Automated Decision-Making (2019); Algorithmic Impact Assessments – AIAs (2019, updated 2025).	Ensure accountability, transparency, and proportional risk assessment in federal government use of AI.
Sectoral & Provincial Measures	Québec Law 25 (2022, in force 2023-25); Ontario Bill 194 (2024, in force 2025); OSFI (AI finance regulator) Model Risk Management Guideline (2024/25); Health Canada AI guidance (2025).	Sector-specific and provincial regulation of AI in health, finance, public administration, and data protection.
Ethical & Soft-Governance Frameworks	Voluntary Code of Conduct on Generative AI (2023); Standards Council of Canada – AI/ISO alignment (ongoing); OECD AI Principles (2019, adopted by Canada).	Non-binding guidance, ethics principles, and technical standards to align industry and civil society practices.

2.2.2.1 Federal Oversight: Strategies and Legislation

At the strategic level, Canada's AI governance has been anchored in the **Pan-Canadian Artificial Intelligence Strategy (PCAIS),** launched in 2017 as the first national AI strategy worldwide. Its primary goals were to strengthen Canada's position as a global leader in AI research, attract and retain top academic talent, and stimulate innovation ecosystems around key hubs such as Toronto, Montréal, and Edmonton.

The **Pan-Canadian Artificial Intelligence Strategy (PCAIS)** is structured into two long-term phases:

- **Phase I (2017–2022): Building Research Excellence and Critical Mass**
  The first phase prioritised consolidating Canada's role as a global leader in AI research by investing **CAD 125 million**. It focused on attracting leading scholars, retaining domestic talent, and establishing the three AI institutes as hubs of excellence (CIFAR, 2022). By 2022, Canada had succeeded in hosting approximately **10% of the world's leading AI researchers** and ranked consistently among the **top three G7 countries in AI publications per capita.**[76]

- **Phase II (2022–2031): Expanding Impact through Innovation and Infrastructure**
  The second phase broadened Canada's AI agenda beyond frontier science to ensure that research excellence translates into economic competitiveness and societal benefit. Supported by over CAD 443 million, it emphasises commercialisation of AI technologies through partnerships with industry and innovation clusters, the development of digital

[76] The Impact of the Pan-Canadian AI Strategy (2023). https://cifar.ca/wp-content/uploads/2023/11/aican-impact-2023-eng.pdf

infrastructure, including advanced computing capacity under the Digital Research Alliance of Canada, workforce development to equip Canadians across sectors with AI-relevant skills, and the integration of responsible AI principles to align innovation with public trust. Notably, about CAD 40 million was earmarked specifically to expand computing capacity, underscoring the recognition that infrastructure had become a critical bottleneck compared with other G7 peers.[77] (Gov-Canada, 2022; Alliance-Can, 2024).

Complementing strategic investment, the Government sought to regulate high-risk AI applications through the Artificial Intelligence and Data Act (AIDA), tabled in 2022 as part of Bill C-27. AIDA was designed as Canada's first horizontal, economy-wide AI law, targeting "high-impact systems" in domains such as employment, credit scoring, law enforcement, and healthcare. Its objective was to create safeguards while supporting responsible innovation. To achieve this, the draft law introduced four core mechanisms (Vani et al., 2023):

- Risk management programmes for developers and deployers of high-impact AI, requiring testing, monitoring, and incident reporting.
- Transparency obligations, including plain-language disclosures to affected users and detailed documentation to ensure auditability.
- Governance and oversight through a federal AI and Data Commissioner empowered to investigate, order corrective actions, and monitor compliance.
- Penalties for non-compliance, ranging from administrative fines to potential criminal charges in cases of severe harm or misconduct.

Despite advancing through parliamentary study, AIDA ultimately stalled and lapsed in early 2025. Its failure leaves Canada without a comprehensive federal AI statute, relying instead on a patchwork of operational directives, sectoral regulations, and voluntary standards. Policymakers are now weighing whether to reintroduce AIDA in a narrower form, refine the scope of "high impact" systems, and align future enforcement with global peers such as the EU AI Act (Scorer et al., 2025).

2.2.2.2 Operational Directives

At the operational level, Canada has developed some of the most transparent and enforceable AI governance instruments in the public sector. The flagship initiative is the Directive on Automated Decision-Making (ADM),[78] introduced by the Treasury Board of Canada in 2019 and subsequently revised in 2021, 2023, and again in June 2025 (Gov-Canada, 2025). The stated objectives of the Directive are threefold: (i) to *ensure accountability* by making automated decisions explainable and reviewable; (ii) to *strengthen public trust* by granting citizens the right to understand how AI systems affect them; and (iii) to *standardise practices* by requiring all departments to follow a coherent set of procedures.

[77] Government of Canada launches second phase of the Pan-Canadian Artificial Intelligence Strategy. https://www.canada.ca/en/innovation-science-economic-development/news/2022/06/government-of-canada-launches-second-phase-of-the-pan-canadian-artificial-intelligence-strategy.html

[78] Directive on Automated Decision-Making. https://www.tbs-sct.canada.ca/pol/doc-eng.aspx?id=32592

The Directive applies across all federal departments and makes the use of automated decision-making systems conditional on compliance with a series of safeguards. The Directive requires that AI systems used in government services undergo an Algorithmic Impact Assessment (AIA)[79] – a mandatory, open-source risk-classification tool with 65 risk and 41 mitigation questions that evaluates potential impacts on citizens' rights, health, safety, and economic interests.[80] Depending on the risk level assigned (from low to very high), federal departments must meet increasingly strict requirements in terms of transparency, human oversight, documentation, and peer review before deployment.

In practice, this framework has positioned Canada as a global pioneer. As of mid-2025, 29 federal AI systems, spanning domains such as immigration, social development, health, and public services, have been publicly assessed under the ADM Directive across more than 10 federal departments (Open-Gov, 2025). These tools are publicly disclosed on the open-data portal, giving citizens visibility into where and how the government uses algorithmic systems. Some of the assessments led to either redesign or enhanced mitigation measures, such as bias audits or human-review thresholds, before deployment. Collectively, these measures have established the ADM Directive as a globally recognised benchmark for algorithmic transparency and accountability in the public sector.

Despite being globally recognised, the framework has faced implementation challenges. A 2022 evaluation noted that some departments lacked the technical expertise to correctly apply the AIA, underscoring the need for continuous training and capacity building. At the same time, the framework has evolved to address these gaps: the 2025 revision introduced stricter requirements for senior-level accountability, expanded transparency obligations, and strengthened monitoring procedures. Internationally, the Directive and its tool are now frequently cited in OECD, EU, and broader global policy discussions as a benchmark for public-sector algorithmic transparency.

The 2022 evaluation also highlighted that some departments lacked the technical expertise to correctly apply the AIA, underscoring the need for continuous training and capacity building. The review further underlined the Directive's limited scope (excluding internal services affecting federal employees), ambiguity around what constitutes a "meaningful explanation", and weak transparency in peer review and disclosure of AIAs (SSRN, 2022). At the same time, the framework has evolved: the 2024–25 review of the Directive on Automated Decision-Making has proposed revisions introducing stricter senior-level accountability (requiring Deputy Head sign-off), expanded transparency obligations (clearer timing for the release of AIAs and more detailed compliance reporting), and strengthened monitoring, assessment of impacts and oversight, strengthen client protections (Treasury Board, 2024). Internationally, the Directive and its Algorithmic Impact Assessment tool are now frequently cited in OECD, EU, and global policy discussions as a leading benchmark for algorithmic transparency in the public sector.

Canada has also signalled its intent to ensure interoperability with the EU AI Act. Efforts focus on aligning risk-based classification, conformity assessment, and technical standards, without adopting identical legal structures.

[79] Algorithmic Impact Assessment tool. https://www.canada.ca/en/government/system/digital-government/digital-government-innovations/responsible-use-ai/algorithmic-impact-assessment.html

[80] Open-Gov Portal (2025). Completed AIAs. https://search.open.canada.ca/opendata/?collection=aia&page=1&sort=date_modified+desc

2.2.2.3 Sectoral and Provincial Instruments

Beyond the federal ADM Directive, Canada's regulatory landscape has been strengthened by sectoral and provincial instruments that establish enforceable guardrails in specific domains. Public service delivery has been an early focus. In 2024, Ontario enacted the Strengthening Digital Security and Artificial Intelligence Act 2024 (Bill 194),[81] the first provincial law to create binding AI governance obligations for public bodies. The Law requires transparency in the use of algorithmic systems, mandates impact assessments, and sets cybersecurity and governance standards. It also integrates responsible AI principles into procurement, ensuring that accountability is considered at the earliest stages of system design and contracting. This framework has begun to standardise AI governance across Ontario's extensive civil service, creating clearer expectations for documentation, oversight, and citizen-facing transparency.

Québec's Law 25 complements these efforts by modernising privacy rules with direct implications for AI. It requires privacy impact assessments for all new systems that handle personal information, strengthens consent, privacy default, and transparency provisions, and introduces significant penalties for violations – up to CAD 25 million or 4% of global turnover for severe violations.[82] As a result, public and private entities operating in Québec have been forced to upgrade data governance practices, directly shaping how AI is trained, evaluated, and deployed.

Healthcare regulation has also advanced significantly. Health Canada has issued pre-market guidance for machine-learning–enabled medical devices (MLMDs),[83] establishing safety, robustness, and lifecycle performance as explicit regulatory requirements. Hospitals and manufacturers are now required to provide pre-market evidence, manage updates to adaptive algorithms, and conduct continuous post-market monitoring. By mandating that AI-enabled diagnostics be governed under existing Medical Devices Regulations, supplemented by AI-specific provisions such as quality management systems and safeguards against algorithm drift, Health Canada has institutionalised stronger accountability in the healthcare sector. This has led to a culture shift in hospitals and among device makers, where AI tools are scrutinised with the same rigor as traditional medical technologies.

In the financial sector, the Office of the Superintendent of Financial Institutions (OSFI) has introduced *Guideline E-23 on Model Risk Management*, which applies to banks and insurers using complex models, including AI and machine learning. The Guideline mandates board-level accountability, comprehensive model inventories, independent validation, and testing for bias and stability, supported by strong documentation and challenge functions. Implementation is being phased in between 2024 and 2026, giving institutions time to adapt (Molino and Langevin, 2024). Already, federally regulated banks and insurers have begun embedding these requirements into their risk frameworks, strengthening auditability, resilience, and confidence in financial AI applications.

Taken together, these measures create a layered governance environment and set strong guardrails on data-driven AI. Although Canada still lacks a horizontal federal AI statute, the existing frameworks already provide enforceable, domain-specific standards that enhance safety, accountability, and public trust.

[81] Bill 194, Strengthening Cyber Security and Building Trust in the Public Sector Act (2024). https://www.ola.org/en/legislative-business/bills/parliament-43/session-1/bill-194

[82] OGS (2025). Quebec's Privacy Law 25: What You Need to Know. https://www.outsidegc.com/blog/quebecs-privacy-law-25-what-you-need-to-know

[83] MLMDs (2025) Pre-market guidance for Machine Learning-enabled Medical Devices. https://www.canada.ca/en/health-canada/services/drugs-health-products/medical-devices/application-information/guidance-documents/pre-market-guidance-machine-learning-enabled-medical-devices.html

2.2.2.4 Ethical and Soft-Governance Frameworks

Alongside binding regulations, Canada has relied on ethical and soft-governance frameworks to build trust, shape market behaviour, and ensure alignment with international norms. These frameworks, while voluntary, play a critical role in extending responsible AI practices into the private sector and public procurement well ahead of comprehensive federal legislation.

A key milestone was the launch of the Voluntary Code of Conduct for Advanced Generative AI, which sets commitments around safety testing, secured deployment, data governance, transparency, and human oversight.[84] Major Canadian AI developers and enterprise users have signed on, and the Code is increasingly referenced in procurement and partnership criteria. This voluntary adoption has raised baseline standards across industries, making risk controls and transparency practices more routine even outside the public sector.

Canada has built a robust AI standards ecosystem through the Standards Council of Canada (SCC).[85] Internationally, SCC represents Canada in ISO/IEC JTC 1/SC 42, the main global body setting standards for AI safety, risk, and interoperability. Domestically, SCC hosts the Artificial Intelligence and Data Governance (AIDG) Standardisation Hub,[86] an online platform that helps adapt global standards to Canadian needs and provides practical resources for businesses, particularly SMEs and Indigenous communities. In 2023, Canada became the first country to certify organisations under the new ISO/IEC 42001 standard for AI management systems, with ATB Financial serving as a pilot.[87] These initiatives give Canadian companies clearer pathways to demonstrate compliance with international benchmarks, making it easier to trade abroad, while also equipping governments with trusted and measurable tools for procurement and oversight.

Canada's ethical orientation is also visible in the ADM Directive, which requires transparency and impact assessments for federal AI use. Complemented by disclosure requirements and emerging procurement guidance, these mechanisms embed accountability, contestability, and risk management across the public service even before comprehensive legislation is in place. Internationally, Canada reinforces this through leadership in GPAI, OECD, and UNESCO processes, ensuring that its governance models remain interoperable while exporting principles of openness and accountability.

2.2.3 AI Infrastructure

Canada has recognised that sustaining its leadership in AI research and innovation requires parallel investments in sovereign digital infrastructure to ensure long-term competitiveness. While the PCAIS initially focused on advancing research excellence and nurturing talent, its second phase introduced initiatives to strengthen access to advanced computing resources alongside continued support for research and skills development. Yet Canadian startups and universities and startups have often depended on foreign commercial cloud providers, raising concerns about equitable access to resources and the resilience of national digital sovereignty.

[84] Antaki, Naim Alexandre and Alycia Riley (2023). Canada publishes Voluntary Code of Conduct on Generative AI Systems. https://gowlingwlg.com/en/insights-resources/articles/2023/canada-voluntary-code-of-conduct-on-ai-systems

[85] Standards Council of Canada. https://en.wikipedia.org/wiki/Standards_Council_of_Canada

[86] Standards Council of Canada and Statistics Canada launch AI and Data Governance Standardisation Hub. https://scc-ccn.ca/resources/news/standards-council-canada-and-statistics-canada-launch-ai-and-data-governance

[87] SCC (2025). AI accreditation pilot. https://scc-ccn.ca/areas-work/digital-technology/ai-accreditation-pilot

As a result, Canadian universities and startups have often relied on commercial cloud providers, a dependence that raises questions about equitable access to resources and the resilience of national digital sovereignty.

To address this, the federal government announced in April 2024 a CAD 2.4 billion investment package to strengthen national infrastructure, secure sovereign AI computing capacity and reduce dependence on foreign technology providers. The package includes up to CAD 300 million for the AI Compute Access Fund to help Canadian SMEs purchase compute resources and meet near-term needs; up to CAD 700 million to support projects from industry, academia, and the private sector to build Canadian AI-ready data centres; and up to CAD 705 million for a new Canadian-owned AI supercomputing system, complemented by a smaller secure system for government and national security R&D. In addition, up to CAD 200 million will augment existing public computing infrastructure to address immediate gaps. Together, these measures aim to broaden equitable access for innovators, reduce dependence on foreign providers, and embed governance, security, and sustainability standards into the very foundations of Canada's AI infrastructure.[88] The initiative is not only economic but geopolitical: it aims to reduce reliance on foreign cloud and chip providers, safeguard national security, and ensure that Canadian SMEs and researchers can compete on equal footing.

These investments mark a turning point in Canada's AI governance, where infrastructure is treated not only as an enabler of innovation but also as a tool of regulation and sovereignty. By controlling allocation rules and linking compute access to ethical standards, Canada has begun embedding oversight mechanisms into the very foundations of its AI infrastructure. This approach positions Canada as one of the few G7 countries seeking to align digital sovereignty with AI competitiveness, though its success will depend on timely implementation and coordination with provincial stakeholders.

In 2023, Canada's computing capacity was under 0.7% of the global total , leaving researchers reliant on U.S. cloud providers (Shah et al., 2024). New investments now tie access to environmental safeguards, using Canada's mostly renewable grid. By mid-2025, funds had already supported genomics researchers in Quebec (USD 7 million) and SMEs AI for agriculture in Saskatchewan (USD 3.4 million). The next challenge is to expand access beyond major urban hubs in a sustainable and inclusive way, advancing SDG 9 (Industry, Innovation, and Infrastructure) and SDG 13 (Climate Action).

In 2024, the Government launched the Canadian AI Safety Institute,[89] working with CIFAR and national AI institutes to advance safety research and provide guidance to both government and industry. While modelled in part on the UK's institute, Canada's approach emphasises a stronger research orientation.

2.2.4 Workforce Development and Capacity Building

Human capital has been the cornerstone of Canada's AI strategy since 2017. The first phase of PCAIS anchored global AI expertise by creating Canada CIFAR AI Chairs, which established over 150 long-term research positions (AI Chairs) at Mila, Vector Institute, and Amii. These Chairs

[88] Government of Canada opens applications for the AI Compute Access Fund (2025). https://www.canada.ca/en/innovation-science-economic-development/news/2025/06/government-of-canada-opens-applications-for-the-ai-compute-access-fund.html

[89] Canadian Artificial Intelligence Safety Institute (CAISI). https://www.canada.ca/en/innovation-science-economic-development/news/2024/11/canada-launches-canadian-artificial-intelligence-safety-institute.html?utm_source=chatgpt.com

not only attracted world-class scholars but also created robust training ecosystems for graduate students, postdoctoral fellows, and industry partners – ensuring continuity of expertise across generations (CIFAR, 2019).

Canada's AI talent pipeline has expanded significantly under the PCAIS: by 2023, over 140,000 Canadians worked in AI roles, representing a 29% year-over-year increase. Canada also led the G7 in the growth of women in AI careers, with a 67% annual increase, signalling progress toward more inclusive participation. In Phase II, the emphasis broadened to workforce development across sectors. With CAD 443 million allocated, programmes now target interdisciplinary training, applied AI skills, and commercialisation pathways. This includes scholarships, professional upskilling programmes, and specialised courses for civil servants. The three national AI institutes now run industry-focused training initiatives, where the Canada School of Public Service has become a key actor, training federal employees in algorithmic literacy, bias testing, and ethical AI design – building capacity inside government to both deploy and regulate AI systems.

Government capacity-building has become a priority in its own right. The AI Strategy for the Federal Public Service (2025–2027) was launched to address persistent digital skills gaps, notably the 30% vacancy rate in IT and data roles. Measures include establishing an AI Centre of Expertise, expanding AI-focused curricula and trainings at the Canada School of Public Service, and piloting secondment programmes with industry and academia.[90] Early evidence suggests these efforts are bearing fruit: Agriculture and Agri-Food Canada, for instance, has embedded in-house AI teams trained under these initiatives, enabling faster, less outsourced development of digital tools such as Ag-Pal for farmers.

2.2.5 AI Technologies in the Public Sector

Canada has been an early mover in testing and deploying artificial intelligence (AI) across federal government services. These technologies are embedded within a governance framework, notably the Directive on Automated Decision-Making and its mandatory Algorithmic Impact Assessment (AIA) to ensure transparency, accountability, and fairness. Several practical deployments illustrate how AI is already reshaping service delivery.

- Immigration and Visa Processing – AI-assisted triage, used by Immigration, Refugees and Citizenship Canada (IRCC),[91] to screen millions of annual visa applications. The system flags routine, low-risk applications for faster processing, while high-risk or complex cases remain subject to full officer review.
- Ag-Pal (Agriculture and Agri-Food Canada) – a machine-learning tool developed by Agriculture and Agri-Food Canada. It matches farmers with more than 70 funding federal, provincial, and territorial agricultural programmes. By simplifying navigation and reducing application errors, AgPal increases programme uptake and saves time for both farmers and administrators.[92]
- Translation Bureau has deployed neural machine translation (NMT) tool to manage growing demand for translation between English and French. While routine content

[90] AI Strategy for the Federal Public Service 2025-2027: Priority areas. https://www.canada.ca/en/government/system/digital-government/digital-government-innovations/responsible-use-ai/gc-ai-strategy-priority-areas.html

[91] Use of AI in Decision-Making at IRCC. https://www.canada.ca/en/immigration-refugees-citizenship/corporate/transparency/committees/cimm-nov-29-2022/question-period-note-use-ai-decision-making-ircc.html

[92] AgPal portal. https://agpal.ca/en/home

is handled by AI systems, human translators remain involved in sensitive, technical, or high-profile texts. This hybrid model helps manage surges in demand, reduce costs, and maintain quality.[93]

Canada's approach remains deliberately cautious: most deployments are assistive or triage-based, not fully autonomous, reflecting a balance between efficiency gains and the protection of rights. This incremental path has strengthened public trust while aligning with SDG 16 (Peace, Justice, and Strong Institutions) by promoting transparent, accountable governance.

2.2.6 Public-Private Partnerships on AI

Public–private partnerships (PPPs) are central to Canada's AI ecosystem, linking research excellence with industrial application. The three national AI hubs – Mila (Montréal), Vector Institute (Toronto), and Amii (Edmonton) – were founded on PPP models combining federal, provincial, and corporate funding. They connect academia, industry, and government, hosting over 130 Canada CIFAR AI Chairs and fostering talent, commercialisation, and policy innovation. Its hybrid structure allows Canada to translate academic discoveries into industry solutions while maintaining independence from any single corporate stakeholder.

Partnerships have yielded tangible outcomes. Mila's collaboration with pharmaceutical firms has advanced drug discovery using deep learning; Vector's work with Canadian banks has strengthened fraud detection and risk management systems; and Amii has pioneered applications in health informatics and climate modelling. These cases highlight Canada's ability to scale frontier research into domains of public and economic value. Beyond the AI research hubs, Canada strategically deploys Scale AI, a federally supported innovation cluster co-funded at a 2-to-1 private-to-public ratio, which has recently invested CAD 98.6 million in 23 applied AI projects spanning sectors such as aerospace maintenance, supply chains, manufacturing, logistics, and recycling. These initiatives integrate AI ethics and security considerations from project inception and have already yielded concrete outcomes.[94]

Complementing this, the Strategic Innovation Fund (SIF) provides substantial support for innovation, including in scaling digital industries and AI-driven technologies.[95] As of 2025, SIF had contributed more than CAD 10.4 billion across 143 projects, advancing public–private collaborations and helping SMEs scale digital solutions (SIF, 2025). At the international level, Canada anchors the Global Partnership on AI (GPAI), which brings together governments, researchers, and industry partners to co-develop responsible AI frameworks and applications in areas such as ethics, data governance, the future of work, and sustainable development.

Collectively, these PPPs showcase Canada's ability to align academic independence, public investment, and private sector dynamism in advancing AI innovation and adoption. Yet despite these successes, challenges remain: partnerships are unevenly distributed, with hubs concentrated in three provinces, and concerns persist about dependency on large multinational firms. Still, the PPP model is widely regarded as a uniquely Canadian strength – blending academic independence, public resources, and private initiative into a governance approach that fosters both innovation and trust.

[93] From Wikipedia (2025).

[94] Canada's AI Adoption Accelerates, Driving Growth for Industries: Nearly $100M Invested in 23 New Projects Through SCALE AI's Latest Funding Round. https://www.scaleai.ca/100m-invested-in-23-new-projects-through-scale-ais-latest-funding-round

[95] SIF (2025). Investments: Strategic Innovation Fund. https://ised-isde.canada.ca/site/strategic-innovation-fund/en/investments

2.3 REPUBLIC OF KOREA

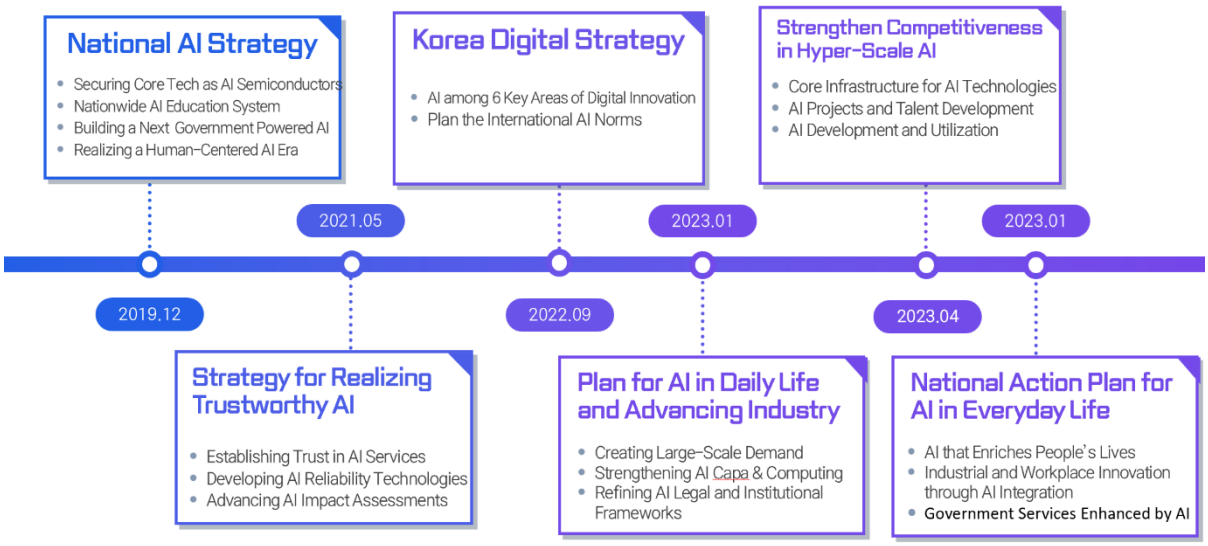
2.3.1 Overview of the AI Policy Framework (2019-2025)

Since the establishment of the National AI Strategy in 2019, Korea has pursued a comprehensive policy agenda up to 2025. In December 2019, the Moon Administration introduced the National Artificial Intelligence Strategy as a national vision to elevate the country from an IT powerhouse to an AI leader. The strategy is structured around three key areas – talent, data, and infrastructure – comprising nine strategic priorities and 100 specific action items. Its objectives include fostering an AI ecosystem through the full-scale opening of public data, nationwide AI education, and investments in AI infrastructure, with the overarching goal of maximising AI utilisation and securing global competitiveness by 2030. With the rapid diffusion of AI technologies, the Government announced the AI Ethics Guidelines in December 2020. Building upon this, in 2021, the Presidential Commission on the Fourth Industrial Revolution adopted the Strategy for Realising Trustworthy AI. This strategy introduced practical measures such as establishing a reliability support system covering the entire AI lifecycle (development, validation, and deployment), defining high-risk AI and mandating user notification obligations, and introducing AI impact assessments.

In January 2023, the Government announced the Plan for Mainstreaming AI in Daily Life and Advancing Industry, outlining policies to embed AI into citizens' everyday lives and accelerate its integration across all sectors. The plan featured initiatives such as large-scale demand creation through nationwide AI mainstreaming projects, the expansion of AI solutions in key domains including manufacturing, healthcare, and education, and the development of next-generation breakthrough technologies supported by AI semiconductor-based infrastructure. Through these measures, Korea declared its ambition to become a "world-leading AI powerhouse".

In April of the same year, in response to the rapid emergence of generative AI models such as ChatGPT, the Government introduced the Measures to Strengthen Competitiveness in Hyper-Scale AI. Backed by an investment of approximately KRW 390 billion, the initiative focused on building high-quality Korean-language datasets, advancing technologies for logical reasoning and model optimisation, establishing a domestic AI semiconductor–based K-Cloud, launching flagship AI projects, creating a new private-sector consultative body, and fostering global talent. Together, these actions signalled the full-scale mobilisation of national efforts to enhance Korea's competitiveness in hyper-scale AI.

Figure 7. Timeline of the National AI Strategy



Source: NIA AI Policy Research Team.

In April 2024, the Government unveiled the AI-Semiconductor Initiative as a national strategy to address the growing demand for AI computing power. Building on Korea's established leadership in memory semiconductors, the initiative sought to secure a competitive edge in next-generation AI semiconductors through joint government–industry efforts in R&D investment, talent cultivation, and ecosystem development – representing an expansion of the so-called "K-Semiconductor Miracle".[96]

Subsequently, in September 2024, the inaugural meeting of the National AI Committee was convened, during which the National AI Strategy Policy Directions were newly articulated. This policy framework reviewed and reinforced the progress of existing initiatives while incorporating private-sector consultations to establish an innovation roadmap. The vision was clearly set: to position Korea as one of the world's top three AI powers (AI G3) within the next five years.[97]

In February 2025, at its third meeting, the National AI Committee announced the National AI Capacity Enhancement Plan. At its core, the plan emphasises the expansion of AI computing infrastructure through the procurement of 10,000 advanced GPUs, support for private-sector GPU leasing, and the introduction of domestic AI semiconductors. Complementary measures include the launch of the World Best LLM Project to accelerate the development of domestic AI models, pilot projects for AI data centres, support for on-device AI, talent acquisition initiatives, and the organisation of global AI challenges.[98]

In parallel, a public–private AI Innovation Fund of KRW 200 billion was established to nurture AI start-ups. Furthermore, the institutional framework was strengthened with the promulgation in March 2025 of the AI Basic Act – formally designating AI as a national strategic technology – alongside amendments to the Special Taxation Act.

[96] Ministry of Science and ICT (2024). AI-Semiconductor Initiative. https://www.msit.go.kr/bbs/view.do?sCode=user&mId=307&mPid=208&bbsSeqNo=94&nttSeqNo=3184413

[97] National AI Committee (2024). Policy Directions for the National AI Strategy. https://eiec.kdi.re.kr/policy/materialView.do?num=257789

[98] National AI Committee (2025) National AI Capacity Enhancement Plan through Expansion of Computing Infrastructure. https://svakorea.org/datas/157

As a result, Korea's AI policy has evolved through distinct phases from strategic formulation to the establishment of ethical principles, to industrial advancement, and finally to the refinement of infrastructure and legal frameworks. As of 2025, the nation has entered a transformative phase aimed at realising its vision of becoming one of the world's top three AI powers (AI G3) while building a sustainable AI ecosystem.

2.3.2 Key Principles and Current Status of Core AI Legal Frameworks

Korea's AI legal framework is primarily structured around the Framework Act on the Promotion of Artificial Intelligence and the Establishment of a Trust-Based Environment (AI Basic Act) and the E-Government Act. Enacted in December 2024, the AI Basic Act is the world's second comprehensive legislation on artificial intelligence, following the EU AI Act, and will enter into force in January 2026. The Act aims to foster AI advancement while establishing a trust-based environment, clearly defining the scope and boundaries of AI in a rapidly evolving technological landscape, and assigning legal rights and responsibilities to relevant stakeholders. Notably, it introduces the concept of "high-risk AI", requiring separate oversight of systems that could significantly affect human safety and fundamental rights. Its foundational principles – human dignity, the public good, and the assurance of technological safety and reliability – translate the philosophy of Korea's existing AI Ethics Guidelines into binding law.

Under the Act, the government bears a dual responsibility: promoting the AI industry while safeguarding public trust. To this end, a Presidential-level National AI Committee has been established to deliberate and coordinate national strategies and mid-to-long-term policies. Article 27 mandates that the government prepare, publish, and implement AI ethics standards, while Article 31 obliges AI providers to ensure transparency, safety, and user notification. Providers of generative AI or high-risk AI face additional responsibilities, such as preventing manipulation or plagiarism of outputs. To protect users, the Act introduces rights such as access to AI-related information, remedies for harm, and dispute resolution mechanisms, with administrative penalties including fines for non-compliance. In doing so, the AI Basic Act seeks to balance industrial promotion with risk prevention, creating an institutional environment where citizens can use AI with confidence.

The E-Government Act, by contrast, has long served as the cornerstone for Korea's digital transformation in public administration. Since its enactment in 2001, it has provided the legal foundation for electronic administrative processes, including the digitalisation of government services, recognition of the legal validity of electronic documents, electronic approval systems, and the establishment of administrative information-sharing mechanisms. Amendments in 2014 introduced provisions for demand-driven, customised services, and in 2017, 24 July was designated as "E-Government Day" to share achievements with the public and encourage civic participation.

More recently, in response to the AI era, the 2021 amendment added provisions for "intelligent e-government services", enabling public institutions to legally deploy AI and big data in predictive administration, chatbot-based public services, and automated case reviews. The Ministry of the Interior and Safety has also prepared accompanying guidelines and ethical standards for AI-based service adoption.

The Digital Platform Government strategy is likewise being advanced on the basis of the E-Government Act.[99] The 2022 amendment strengthened the foundations for data-driven governance by introducing the designation of national reference data and the integrated management of information resources on cloud infrastructure. These measures have accelerated the development of AI-based platforms for administrative innovation.[100]

In addition, related legislation – including the Framework Act on Data, the Personal Information Protection Act, and the Act on Promotion of Information and Communications Network Utilisation and Information Protection – complements Korea's AI governance framework. In particular, the Framework Act on Data, enacted in 2021, established principles for the management, sharing, and opening of data, thereby laying the groundwork for the active utilisation of data as a core resource for AI development. Meanwhile, the Personal Information Protection Act, through provisions on the use of pseudonymised information, seeks to balance data use with privacy protection in the AI era.

Taken together, Korea's AI-related legal framework constitutes a multi-layered structure encompassing industrial promotion, ethics, safety, and digital transformation. This framework provides the foundation for Korea to position itself as an AI democratic nation in the evolving process of shaping global AI norms.

2.3.3 AI Governance Framework and Key Institution[101] [102]

The Ministry of the Interior and Safety (MOIS) serves as the lead agency for integrating AI technologies and advancing digital transformation in public administration. It is building a government-wide common AI platform to establish standardised environments that enable AI adoption across agencies, while spearheading the introduction of generative AI–based administrative services such as chatbots, document automation, and summarisation and content generation tools. By expanding open access to public data and strengthening data governance, MOIS has reinforced the foundations for AI development, while also applying AI to address societal challenges, including disaster and safety management and support for vulnerable populations. In cooperation with the Digital Platform Government Committee, it is broadening citizen-centred digital innovation in governance. Notably, its 2025 work plan identifies "laying the foundation for full-scale AI adoption" as a core task, focusing on automating administrative processes and improving civil service efficiency. MOIS thus acts as the central ministry driving AI transformation within the government.

The National AI Committee, launched in September 2024, is the highest-level presidential body for AI governance, bringing together government and private-sector leadership. Tasked with guiding Korea's rise as one of the world's top three AI powers, the committee is chaired by the President, with a civilian expert serving as Vice Chair. Its membership includes ministers from relevant ministries as well as the heads of the Korea Communications

[99] Ministry of the Interior and Safety and Digital Platform Government Committee (2024). Guidelines 2.0 for the Introduction and Utilisation of Public-Sector Foundation Models. Joint distribution. https://nia.or.kr/site/nia_kor/ex/bbs/View.do?cbIdx=99852&bcIdx=26677&parentSeq=26677

[100] Ministry of the Interior and Safety (2022). Designation of National Reference Data and Cloud-Based Integrated Information Resource Management. Policy announcement. https://www.mois.go.kr/cmm/fms/FileDown.do?atchFileId=FILE_00115909fNabSPV&fileSn=0

[101] Ministry of the Interior and Safety (2024). Pilot Operation Plan for Government-Dedicated Generative AI Service. Press release. https://www.mois.go.kr/frt/bbs/type010/commonSelectBoardArticle.do;jsessionid=AikmYQO1+N616SKhqFVzxqtQ.node20?bbsId=BBSMSTR_000000000008&nttId=109883

[102] Ministry of the Interior and Safety (2023). MOU on the Development of Government-Dedicated AI Administrative Support Services. Press release. https://www.mois.go.kr/cmm/fms/FileDown.do?atchFileId=FILE_00118499soTL1oT&fileSn=1

Commission and the Personal Information Protection Commission. The committee deliberates and decides on core policies such as the National AI Master Plan, R&D investment strategies, and data infrastructure expansion. It also addresses a broad agenda ranging from regulatory reforms and the dissemination of ethical standards to frameworks for managing high-risk AI. Specialised subcommittees and advisory groups enhance its expertise and responsiveness. By consolidating the knowledge and creativity of both public and private stakeholders, the committee functions as a control tower for cross-government coordination. Its establishment marks a significant step forward in the maturity of Korea's AI governance framework.

These institutions, while dividing responsibilities across planning, execution, and coordination, work in close collaboration to strengthen the driving force of Korea's AI policy. Through this cooperative structure, they ensure a balanced advancement along three pillars: technological innovation, public service innovation, and the establishment of an ethical and trust-based foundation.

2.3.4 Government–Public Institutions–Private Sector Cooperation Mechanism in AI

Korea's AI governance operates on a structure of close collaboration between government, public institutions, and the private sector. The Government incorporates private-sector innovation and expertise into policy, while industry advances technological development and the responsible use of AI under institutional and financial support from the State. This complementary mechanism has been translated into concrete cooperative practices across multiple domains.

First, the AI semiconductor sector exemplifies strategic collaboration between government and industry. Since 2022, the Ministry of Science and ICT has convened an AI Semiconductor Strategic Dialogue with major companies such as Samsung Electronics and leading research institutes to discuss technology roadmaps and align government investment priorities. In April 2024, the AI-Semiconductor Initiative formally designated AI semiconductors as a national strategic industry. Through this initiative, joint public–private R&D and infrastructure investment plans were established. The Government provides the institutional framework and funding, while industry drives technological innovation and commercialisation, accelerating the development of an AI semiconductor ecosystem.

Second, public–private cooperation has also been central to building the hyper-scale AI ecosystem. In 2023, the Ministry of Science and ICT launched a public–private council involving major domestic firms such as Naver, KT, and LG AI Research, as well as start-ups, to share progress on model development, provide infrastructure support, and explore application domains.

In 2024, this effort expanded into the Hyper-scale AI Promotion Committee, a large-scale platform of over 100 participating organisations established to provide policy advice and enable regular consultations. Through this platform, the Government has institutionalised support for R&D, ethics standards, and talent development, while the private sector has strengthened technological competitiveness and driven industrial innovation.

Third, joint efforts to establish AI ethics represent another important dimension of cooperation. In 2020, the Presidential Commission on the Fourth Industrial Revolution and the Ministry of Science and ICT convened over 100 experts from academia, industry, and research to formulate

human-centred AI Ethics Guidelines. These guidelines outlined three core principles – human dignity, the public good, and technological purposefulness – along with ten requirements including transparency, safety, and accountability, intended for voluntary compliance by companies and developers.[103] Building on this, a public–private AI Ethics Council was created to share industry practices and monitor implementation. In 2022, the council developed ethical guidelines for generative AI; and in 2023, the Korea Communications Commission, in cooperation with major companies, issued User Protection Guidelines for Generative AI that underscored principles of human-centredness, fairness, and safety. These joint standards have provided the basis for strengthening corporate responsibility and securing user trust.

In conclusion, Korea's AI governance is anchored in a collaborative structure in which the government and private sector cooperate across strategy, technology, and ethics. This framework strengthens both technological progress and social acceptability. The government plays the role of coordinator and enabler, while the private sector acts as the driver of innovation. Together, their collaboration serves as a central pillar for enhancing the sustainability and global competitiveness of the AI ecosystem.

2.3.5 The AI Safety Research Institute & the National AI Policy Centre

To address the rapid proliferation of AI technologies and the associated risks, the Korean Government has been establishing specialised research and policy institutions to build a safe and systematic AI governance framework. Two representative examples are the AI Safety Research Institute, launched in 2024, and the forthcoming National AI Policy Centre, scheduled to be inaugurated in January 2026.

The AI Safety Research Institute was established in November 2024 in Pangyo, Gyeonggi Province, as an affiliated institute of the Electronics and Telecommunications Research Institute (ETRI). Following the President's announcement at the 2023 AI Global Summit, it was created by benchmarking leading safety institutes in major countries, with the goal of becoming a representative AI safety hub in the Asia-Pacific region. The Institute operates through three laboratories – policy, evaluation, and technology – focusing on the scientific identification of AI risks and the development of relevant policies and regulatory frameworks. Its specific functions include defining and assessing AI-related risks, developing safety verification tools and datasets, researching technologies to prevent misuse of AI, and advancing methods for deepfake detection and model control.[104]

The Institute also manages a domestic AI Safety Partnership in collaboration with academia, industry, and research communities, while engaging internationally with safety bodies in the United States, the United Kingdom, Japan, and other countries to contribute to global norm-setting and research exchange. The Government intends to use the institute's analyses as a foundation for evidence-based AI safety policies and to integrate its findings into legislative and regulatory reforms. In essence, the AI Safety Research Institute functions as a core institution advancing specialised research and international cooperation to ensure the trustworthiness and safety of AI technologies.

[103] Ministry of the Interior and Safety (2025). 2025 Work Plan: Establishing the Foundation for Full-Scale AI Adoption. Policy briefing material. https://www.korea.kr/news/policyNewsView.do?newsId=148938617

[104] ETRI (2024). Plan for the Establishment and Operation of the AI Safety Institute. Research institute press release and policy material. https://www.korea.kr/docViewer/iframe_skin/doc.html?fn=d5de06b9eedbe86f265fbafb04b96a6a&rs=/docViewer/result/2024.11/27/d5de06b9eedbe86f265fbafb04b96a6a

2.3.6. Strategy for Building Public-Sector Foundation Models

As global competition in artificial intelligence intensifies, the Korean Government has launched an initiative to develop its own large-scale foundation models as part of a broader strategy to secure technological sovereignty and reduce dependence on foreign technologies. Securing a national-level foundation model is essential not only for delivering AI services tailored to the Korean language and domestic data but also for creating a safe and reliable environment for government and public-sector use of AI. This approach goes beyond mere technological development, framing AI as a public good to be shared and as a strategic driver of digital transformation across society.

To this end, under the National AI Capacity Enhancement Plan announced in 2025, the government initiated the World Best LLM (WBL) Project in collaboration with the private sector. The project aims to achieve performance at least 95% of leading global commercial models and is being implemented over a three-year period through comprehensive support for GPUs, datasets, and talent. It also establishes a sustainable operational pipeline that enables training and retraining centred on domestic data, while adopting an open development structure that combines private-sector technological capabilities to ensure both sovereignty and practical utility.

This policy direction is being realised through a government-wide initiative to build a foundation model–based AI platform for public use. The platform will allow civil servants in central ministries to utilise generative AI in their work without security concerns, supported by shared resources such as standardised training datasets and national foundation models.

The Ministry of the Interior and Safety is leading the platform development through a public–private partnership (PPP) model. Within the National Information Resources Service Data Centre in Daegu, a joint cloud zone is being established where private cloud companies provide platform-as-a-service solutions. Major participants include Samsung SDS, KT Cloud, and NHN Cloud, which together form a Managed Service Provider (MSP) consortium responsible for building and operating the infrastructure and platform. Each consortium is also integrating private-sector large-scale AI models – including Naver's *Hyper-CLOVA X*, KT's *Midm*, and LG CNS's *EXAONE* – thereby ensuring technical diversity and competitiveness.

In sum, the government is integrating GPU infrastructure, training datasets, foundation models, and MSP operational systems into a unified government-wide public AI platform. This platform will provide civil servants with a secure and efficient AI-enabled work environment, promote the use of private infrastructure, and foster the development and autonomy of Korea's AI ecosystem. In the short term, the strategy aims to advance AI adoption in the public sector; in the longer term, it will serve as a critical foundation for securing Korea's AI sovereignty and enhancing its global competitiveness.

2.3.7 Public-Sector Use Cases and Public–Private Implementation Mechanisms

In Korea's public sector, government agencies are actively exploring the development of domain-specific large language models (LLMs) and AI services to enhance administrative efficiency and leverage specialised knowledge. These applications are expected to go beyond simple task automation, enabling more accurate decision support in advanced professional domains and transforming information search processes. Several public institutions have already begun adopting such initiatives proactively.

For example, the Korea Institute of Science and Technology Information (KISTI) has developed KONI, a 13-billion-parameter LLM specialised in science and technology big data. Trained on scholarly papers, patents, and research datasets, KONI applies retrieval-augmented generation (RAG) techniques to integrate its database knowledge into responses, thereby reducing hallucinations and improving reliability. The model is now being distributed to institutions such as the National Assembly Library and the Army Logistics Command, providing a secure and cost-effective AI solution for organisations unable to adopt commercial AI services due to budgetary or security constraints.

The Korea Institute of Science and Technology Planning and Evaluation (KISTEP) operates SPARK, a small language model (SLM) tailored to science and technology policy, along with its on-premises AI service, KISTEP-GEN. This system supports Q&A, summarisation, and drafting of policy documents and is designed for use in closed networks, enabling the secure handling of sensitive policy information. SPARK has also been released as open source, laying the foundation for broader adoption and collaboration across public agencies.

The Korea Institute of Geoscience and Mineral Resources has developed the Geo-AI platform, a web-based system for the integrated analysis and visualisation of spatiotemporal data across six domains, including minerals, marine geology, and earthquake hazards. By separating data collection from analysis and consolidating domain-specific AI models, the platform enhances flexibility, applicability, and scalability in real-world contexts.

In the nuclear energy sector, the Korea Atomic Energy Research Institute has developed AtomicGPT, the world's first nuclear-specialised LLM, which has been released as open source. Trained on nuclear research papers, regulations, and reports, the model provides high-level domain expertise and is already being applied in closed network environments for report drafting and technical review. Future plans include linking reactor operation data to enable AI-based virtual operator functions.

The Korea Institute of Energy Research has introduced Gyubom-i, a RAG-based chatbot specialised in regulatory search, which has significantly reduced processing times for regulatory inquiries and generated substantial cost savings in administrative processes.

These public-sector applications are being realised within a governance framework rooted in public–private collaboration. As noted earlier, the Ministry of the Interior and Safety is building the public foundation model platform through a public–private partnership, with private cloud providers serving as Managed Service Providers (MSPs) responsible for infrastructure operation. Meanwhile, the Ministry of Science and ICT has established joint development structures by convening consortia of industry, academia, and research institutions, providing integrated support for infrastructure, data, and talent to advance public-sector foundation models. This structure links private-sector innovation capacity with government policy goals in a mutually reinforcing manner.

In conclusion, Korea's public foundation model initiatives are advancing on a governance model in which government-led institutional frameworks are combined with private-sector expertise. This approach simultaneously secures public value and innovation, establishing a practical model of governance for sustainable AI development in the public sector.

2.3.8 Agile Regulation and the AI Regulatory Sandbox

To keep pace with the rapid advancement of AI technologies, the Korean Government has moved away from rigid, traditional regulatory approaches toward more flexible and adaptive frameworks – an approach referred to as Agile Regulation. Whereas revising existing legislation often requires lengthy processes, AI-driven innovations emerge quickly alongside technological progress, creating an urgent need to bridge this regulatory gap. In response, the Ministry of Science and ICT has sought to establish an experimental governance framework capable of responding swiftly to technological change, most notably through the operation of the AI Regulatory Sandbox.

The AI Regulatory Sandbox allows innovative AI products and services to be tested in real-world environments, even if they conflict with existing laws, by temporarily exempting them from certain regulatory requirements. Through deliberation by the ICT New Technology and Service Review Committee, companies may be granted either temporary permits or special exemptions for demonstration. Relevant ministries then review the safety and legal legitimacy of the projects before approval. This system enables businesses to test emerging technologies without excessive regulatory burdens, while allowing the government to use demonstration results as a basis for legislative and regulatory improvements – thereby fostering a virtuous cycle between innovation and regulation.

Since its introduction in 2019 in the ICT sector, the sandbox has designated a total of 231 regulatory exemptions as of June 2024, including 70 temporary permits and 161 demonstration exemptions. More recently, the government has launched planned sandboxes targeting strategic fields such as generative AI, bio-health, and green technologies. Under this mechanism, ministries proactively identify areas for regulatory easing through inter-ministerial coordination. For instance, restrictions on the use of original video data under privacy laws were partially lifted through the sandbox, enabling the development of AI-based CCTV systems, while financial institutions were permitted to adopt generative AI cloud applications within internal networks.

One representative case is the Soft V2X service, a smartphone-based traffic safety solution. Initially constrained by the Location Information Act and the Personal Information Protection Act, the service was unable to achieve commercialisation. However, regulatory exemptions under the sandbox allowed it to proceed, becoming a widely recognised example of how flexible legal interpretation can reconcile regulatory requirements with the societal utility of technology.

The sandbox has thus emerged as a primary instrument of agile regulation, enabling evidence-based validation of new technologies while facilitating the removal of unnecessary regulatory barriers. In 2024, additional measures were under discussion, including streamlined documentation requirements and exemptions from redundancy testing, particularly for generative AI, with the aim of strengthening accountability while promoting innovation. Going forward, the results of these demonstrations are expected to serve as key reference points for refining Korea's AI legal and regulatory framework.

Table 6. Core Features of Korea's AI Regulatory Sandbox

System	Background (Rationale)	Key Content (Operation)	Implementation (Achievements)	Representative Cases (Impact)
Agile Regulation	With rapid advances in AI technology, existing regulations struggle to keep pace. Flexible and adaptive regulation is needed to balance technological progress and public safety.	Focuses on pre-emptive regulatory improvement and experimental application through guidelines. Legal interpretations, and stakeholder participation, reflecting feedback quickly.	Principe or approach, with no separate outcomes.	Difficult to quantify as a general regulatory innovation case.
AI Regulatory Sandbox	AI representative tool of agile regulation. Temporary regulatory exemptions are needed to demonstrate innovative services.	The Review Committee for New Technologies and Services approves regulatory exemptions. Temporary permits – allowed for a certain period – and demonstration exemptions – testing in limited environments – are granted. Legal restrictions are waived under safety conditions.	From 2019 to end of June 2024: 231 exemptions were issued; 70 temporary permits; and 161 demonstration exemptions. February 2025: Thematic Regulatory Sandbox introduced. Five key projects selected including Generative AI.	Soft V2X traffic safety solution: Commercialised through exemptions on personal and location data. AI CCTV original video training: Use of raw data permitted under conditions of secure personal data processing.

Source: Compiled by the research team.

In conclusion, Agile Regulation and the regulatory sandbox function as institutional mechanisms that seek to balance technological innovation with social acceptability. By doing so, the government is fostering an AI ecosystem that is ethical, inclusive, and innovative, grounded in the principles of safety, accountability, and trust.

2.3.9 Institutionalising AI Social Impact Assessments

As AI technologies rapidly permeate all areas of society, the Korean Government is advancing the institutionalisation of social impact assessments to prevent potential harms – particularly those arising from high-risk and generative AI – and to ensure their trustworthy use. Large language models, autonomous driving systems, and medical AI are considered domains that can

directly affect human life, safety, and fundamental rights, reinforcing the necessity of systematic pre-assessment and oversight.

Beginning in 2026, the AI Basic Act will provide the legal foundation for defining both high-impact AI and generative AI, establishing distinct obligations and management standards for these categories. High-impact AI refers to systems deployed in sensitive sectors such as healthcare, transportation, law enforcement, education, and finance – areas with significant implications for citizens' rights and safety. Generative AI, which produces text or images from large-scale data, has also been designated as subject to oversight due to its broad societal impact. Under this framework, providers of such technologies are required to ensure transparency, comply with notification obligations, and conduct safety verifications. Public institutions, in particular, are encouraged to prioritise certified and validated technologies when adopting high-impact AI systems.

A central element of this policy direction is the introduction of an AI impact assessment mechanism. Article 35 of the AI Basic Act recommends that providers offering products or services involving high-impact AI conduct prior assessments of potential effects on fundamental rights. This procedural safeguard is designed to identify risks such as discrimination, rights infringements, and other adverse social impacts at the early stages of AI development and deployment, and to require mitigation measures before full implementation. While not a binding legal obligation, providers are encouraged to prepare and retain self-assessment reports. Furthermore, public procurement guidelines give preference to products that have undergone such impact assessments, thereby ensuring the practical effectiveness of the evaluation system.

Through this approach, Korea aims to embed accountability and rights protection into the governance of AI, promoting not only innovation but also responsible and socially sustainable adoption.

2.3.10 Red-Teaming and Safety Evaluation at the AI Safety Institute (2024)

As AI technologies continue to proliferate globally, ensuring safety and trustworthiness has emerged as a critical international priority. In response, the Korean Government established the AI Safety Institute in November 2024 in Pangyo, Gyeonggi Province, as a dedicated research body to address these challenges. This Institute is the sixth national AI safety body in the world and aspires to serve as the leading AI safety hub in the Asia-Pacific region. Following the government's official commitment at the 2023 Seoul AI Summit, the institute was launched under the Electronics and Telecommunications Research Institute (ETRI).

The Institute's mandate is to conduct scientific risk analyses and to secure safety technologies, with a particular focus on building AI model safety evaluation systems and conducting red-teaming tests. Red-teaming is a simulation-based technique designed to probe the vulnerabilities of AI systems by posing adversarial or manipulative queries to elicit abnormal outputs and identify potential risks. Through these methods, the Institute is accumulating real-world risk scenarios and test datasets and is developing evaluation standards for high-risk AI systems based on both domestic and international case studies. For instance, it identified the risk of weapons-related information leakage from the Chinese foundation model DeepSeek and recommended blocking its domestic service as a precautionary measure.

In parallel, the Institute conducts research on AI safety policy, monitoring global regulatory and normative developments and integrating them into Korea's domestic framework. By analysing recommendations from bodies such as the OECD and UNESCO, and by collecting risk data from domestic and international sources, the institute provides key policy input for establishing certification criteria for high-risk AI. It also assesses the broader social, economic, and labour market impacts of AI, generating policy alternatives that aim to minimise negative consequences while maximising positive outcomes.

On the technical side, the Institute pursues research to enhance AI system resilience, including adversarial defence methods for deep learning models, content filtering techniques, and real-time risk monitoring systems. It also provides consulting on safety-by-design principles to companies. Furthermore, it is building testbed environments and developing open benchmark datasets for safety evaluation, enabling joint use by academia, industry, and research institutions.

It is operated on an open, collaborative basis. Domestically, it leads the Korea AI Safety Consortium, a partnership of 24 institutions across industry, academia, and research, to conduct joint studies. Internationally, it participates in networks with AI safety institutes in the United States, the United Kingdom, Japan, and Singapore, and contributes to the international standardisation of safety evaluation scenarios.

Currently, the Institute is conducting pilot red-teaming experiments and developing evaluation datasets for large language model safety. In 2025, it plans to release a draft AI Risk Map and publish prototype evaluation tools. Functioning as a support body rather than a regulator, it seeks to assist private-sector companies by providing science-based assessments that enhance both the safety and the global competitiveness of AI technologies.

2.4 UNITED KINGDOM

2.4.1 AI Governance and Regulation

The United Kingdom has taken a principles-based, pro-innovation approach to AI regulation, relying on existing regulators rather than creating a single authority.[105] Its framework is non-binding and built around five cross-sectoral principles: (i) safety; (ii) transparency; (iii) fairness; (iv) accountability; and (v) contestability (DSIT, 2024). Although there is no comprehensive AI law in place, several legislative initiatives are moving forward, such as the Artificial Intelligence Regulation Bill and the Public Authority Algorithmic Decision-Making Bill. Existing laws like GDPR, the Data Protection Act, the Equality Act, consumer rights, and competition law also apply where relevant.[106] Alongside legislation, voluntary measures and standards have been promoted, including the commitments from the AI Safety Summit, NCSC guidelines for secure system development, and the Algorithmic Transparency Recording Standard (ATRS) (DSIT, 2024).

Despite these efforts, the coherence and strategic direction of UK AI governance remain under debate. The reliance on multiple regulators creates risks of overlap, gaps, and inconsistency,

[105] A New UK Labour Government: A Fresh Approach to AI. Dechert.com (2024). https://www.dechert.com/knowledge/onpoint/2024/7/a-new-uk-labour-government--a-fresh-approach-to-ai-regulation.html

[106] AI law, regulation and policy - highlights from 2024 and what to look forward to in 2025.(24 April 2025). Burges Salmon. https://www.burges-salmon.com/articles/102jr1b/ai-law-regulation-and-policy-highlights-from-2024-and-what-to-look-forward-to/

contrasting with the EU's more comprehensive AI Act. While the Regulatory Innovation Office has been established to support regulators, it does not replace the need for a centralised body. At the strategic level, the National AI Strategy (2021) and AI Opportunities Action Plan (2025) set ambitious goals for the UK to become an "AI maker" rather than an "AI taker".[107] However, issues such as a serious shortage of computing infrastructure, siloed departmental mandates, and the absence of binding commitments for major projects raise questions about the UK's ability to deliver on these goals. Critics also argue that the principles-based approach, while initially designed for flexibility, has become too vague and is falling behind international regulatory developments. Even the government acknowledges that binding legislation will eventually be required as understanding of AI risks matures (Seger, 2024).

To oversee compliance with AI laws and standards, several formal processes and bodies have been established or repurposed:

- Digital Regulation Cooperation Forum (DRCF): This voluntary forum, comprising the CMA, FCA, ICO, and Ofcom, aims to foster cooperation on digital regulation, prevent overlaps, and enhance coordination. Its AI and Digital Hub pilot provided informal cross-regulatory advice to innovators.
- Department for Science, Innovation & Technology (DSIT): This department has overall responsibility for AI policy and regulation. It houses a multidisciplinary team for cross-sectoral risk monitoring.
- Regulatory Innovation Office (RIO): Proposed by the Labour government, this office will consolidate government functions and streamline approval processes for innovative products and services related to AI. It is designed to set targets for technology regulators and guide them according to industrial strategy and may introduce binding regulations for powerful AI models. It is explicitly stated that it will not be a new AI regulator.
- AI Safety Institute (AISI): Established to advance the world's knowledge of AI safety, it develops and conducts evaluations on advanced AI systems, drives foundational AI safety research, and facilitates information exchange to inform the UK's regulatory framework. The Labour Government intends for it to become a statutory body.
- A Steering Committee with government and regulator representatives is being established to support coordination across the AI governance landscape.
- The AI Energy Council brings together industry and government to ensure the UK's energy infrastructure can sustainably support AI, address computational demands and promote renewable energy solutions.
- The Responsible Technology Adoption Unit (RTAU), formerly the Centre for Data Ethics and Innovation (CDEI), develops tools and techniques for responsible AI adoption in both the private and public sectors. It also conducts public attitudes research to align regulatory approaches with public values.
- An Inter-Ministerial Group and lead AI Ministers across all government departments have been established to drive effective coordination on AI issues and oversee the implementation of frameworks and guidelines for public sector AI usage.

[107] Isambard-AI, the UK's Most Powerful AI Supercomputer, Goes Live. NVIDIA Blog. B. Caulfield (17 July 2025). https://blogs.nvidia.com/blog/isambard-ai/

2.4.2 AI Infrastructure and Investment

The United Kingdom is making substantial investments in foundational Artificial Intelligence (AI) infrastructure while also addressing the sustainability challenges associated with its rapid expansion. The Government sees AI as a strategic opportunity to establish the country as a global leader in safe and responsible AI and has committed billions of pounds to build capacity in supercomputing, data centres, research, and skills (DSIT, 2024).

Since 2014, the UK has invested over GBP 2.3 billion in AI, with a further GBP 100 million to support innovation and regulation. By 2030, it aims to expand sovereign computing capacity twentyfold, supported by more than GBP 1.5 billion for public sector supercomputers (HM Government, 2021). A landmark GBP 1.34 billion investment is funding two advanced systems — *Isambard-AI* in Bristol and *Dawn* in Cambridge – under the AI Research Resource (AIRR). Isambard-AI, operational since 2024, is the UK's most powerful supercomputer, designed for both high performance and energy efficiency. Dawn, scheduled to come online at the end of 2024, focuses on energy-efficient computing and scientific AI training. These efforts are complemented by broader investments of GBP 900 million into computing infrastructure.[108]

Data infrastructure is also a priority. The UK already ranks third globally in the number of data centres, and the Labour government plans to speed up new projects by classifying them as Nationally Significant Infrastructure. New AI Growth Zones are being created, starting with a large site in Culham, and similar projects are expected in Scotland and Wales. Alongside this, long-term funding commitments for research and development include a National Data Library, nine new AI research hubs, and strengthened partnerships with the US.[109]

Investment in AI skills has reached GBP 290 million since 2018, with over 1,500 PhD students trained through AI Centres for Doctoral Training. The private sector has also committed heavily, with more than GBP 25 billion in new UK data centre projects announced since mid-2023, including a GBP 2.5 billion Microsoft investment (DSIT, 2024).

The rapid expansion of AI infrastructure presents major energy and environmental challenges. Data centres currently consume 1–2% of Britain's electricity, a figure projected to rise to 10% by 2050. To manage this, the UK has created an AI Energy Council co-chaired by the Technology and Energy Secretaries. This forum brings together government, regulators, and industry leaders such as Microsoft, Google, and National Grid to ensure that AI growth is aligned with clean energy goals, consumer interests, and sustainable practices (Energy UK, 2025).

Supercomputing projects are designed with sustainability in mind. Isambard-AI ranks among the world's most energy-efficient systems, powered entirely by zero-carbon electricity and equipped with liquid cooling that recycles waste heat. Dawn also prioritises energy efficiency. More broadly, reforms to the grid connection process could free up significant capacity, while Ofgem is investing nearly GBP 9 billion to expand Britain's high-voltage network. The Government has committed over GBP 15 billion to nuclear energy and GBP 9.5 billion to carbon capture, while also expanding renewable energy support through mechanisms like Contracts for Difference.[110]

[108] Isambard-AI launches as UK's most powerful supercomputer. UKAI (21 July 2025). The trade association for UK AI Businesses. https://ukai.co/isambard-ai-launches-as-uks-most-powerful-supercomputer/

[109] AI Opportunities Action Plan Government Response. GOV.UK (2025). https://assets.publishing.service.gov.uk/media/678639913a9388161c5d2376/ai_opportunities_action_plan_government_repsonse.pdf

[110] UK AI supercomputing investment gets £1.34 billion boost. AI CERTs News (21 July 2025). https://www.aicerts.ai/news/uk-ai-supercomputing-investment-gets-1-34-billion-boost/

Long-term policy measures are being developed to ensure sustainability and security. These include mandatory carbon reporting for data centres, guidance on planning and grid connections, and integrating AI energy demand into national energy strategies. Co-location of data centres with energy assets is being explored to reduce grid strain. The Responsible Technology Adoption Unit is also engaging the public to ensure that regulatory approaches align with social values and build trust (Energy UK, 2025). Through this combination of large-scale investment, sustainability initiatives, and strategic planning, the UK is seeking to balance its ambition of becoming a global AI superpower with the environmental and energy realities of supporting such an infrastructure.

2.4.3 AI Ethical Standards and Principles

The UK has developed a comprehensive framework for ethical AI, with a strong focus on the public sector. Its pro-innovation approach seeks to balance growth with safety, fairness, and ethics, guided by the same five principles as for regulation – safety, transparency, fairness, accountability, and contestability – that may soon become statutory duties for regulators. The Government positions the public sector as a model for ethical AI, led by the Responsible Technology Adoption Unit, which develops governance tools and standards.

Transparency is advanced through the Algorithmic Transparency Recording Standard (ATRS), requiring departments to publish details on AI use. Sectoral initiatives include procurement guidelines and the NHS AI Lab's national AI strategy for health and social care. Explainability is central: regulators push for clear communication, user notifications, and disclosure of decision processes, though challenges remain with complex models and administrative burdens. ATRS and safety guidance aim to improve model reporting, content identification, and deepfake risks.[111] Accountability is reinforced through efforts to clarify responsibility across AI supply chains, adapt legal frameworks for general-purpose AI, and strengthen oversight via a new Regulatory Innovation Office, a central risk-monitoring function, and the AI Safety Institute. Fairness is addressed through initiatives like the Fairness Innovation Challenge and HR guidelines, supported by data protection laws and the Equality Act. Regulators are urged to monitor bias, with discussions on allowing sensitive data collection for bias detection.

Human rights are embedded across the framework, aligned with democratic values and international standards. The Equality Act and participation in global agreements, such as the Council of Europe's AI Convention, provide safeguards. Public engagement is emphasised to maintain trust. Overall, the UK's model blends ethical principles, sectoral strategies, and institutional reforms, aiming to adapt to rapid AI advances while keeping transparency, accountability, fairness, and rights at its core.

2.4.4 Public–Private Collaboration Models in AI

The UK government's "pro-innovation" approach to AI regulation champions collaboration with industry, academia, and civil society to drive economic growth and improve public services. Key examples of such collaboration include:

[111] Implementing the UK's AI Regulatory Principles. GOV.UK (2024). https://assets.publishing.service.gov.uk/media/65c0b6bd63a23d0013c821a0/implementing_the_uk_ai_regulatory_principles_guidance_for_regulators.pdf

- The AI Safety Institute (AISI): This government-backed organisation, the first of its kind globally, advances AI safety in the public interest by evaluating and testing advanced AI systems. Leading AI companies, including Google DeepMind, Microsoft, OpenAI, and Anthropic, have supported the UK's approach and pledged to provide the Institute with priority access to their systems for testing. The AISI also engages with academia and civil society to provide independent expert perspectives.
- The AI and Digital Hub (Digital Regulation Cooperation Forum – DRCF): This pilot advisory service, launched by the DRCF (comprising the Information Commissioner's Office (ICO), Ofcom, the Competition and Markets Authority (CMA), and the Financial Conduct Authority (FCA), offers free and informal advice to innovators on cross-regulatory queries involving AI and digital technologies. It aims to support innovation and streamline navigation of the regulatory landscape. The DRCF also hosts joint workshops focusing on AI transparency and accountability and shares approaches to auditing AI systems.
- The AI Energy Council: Co-chaired by the Technology and Energy Secretaries, this council brings together industry heavyweights from the energy and technology sectors (including Google, Microsoft, Amazon Web Services, ARM, Equinix, National Energy System Operator, Ofgem, and National Grid) to address the energy demands of AI infrastructure. Its objectives include ensuring sustainable energy supply and promoting the safe and secure adoption of AI across the energy system.
- Partnership with OpenAI: The UK government has a formal Memorandum of Understanding (MoU) with OpenAI to accelerate the responsible use of AI across public services, infrastructure, and national growth zones. This includes exploring joint investments in regional AI growth zones, sharing technical insights with the AI Safety Institute, and expanding OpenAI's UK operations. OpenAI's models are already being used in Whitehall tools for administrative tasks like processing public consultation responses and assisting small businesses with GOV.UK services. Similar cooperation agreements exist with Google DeepMind and Anthropic.
- The NHS AI Lab: Launched in 2019, the NHS AI Lab fosters collaborations to accelerate the safe and effective adoption of AI in healthcare. It partners with companies like Faculty to build machine learning tools for improving service delivery and patient care, such as predicting A&E demand. The Lab is also creating a National Strategy for AI in Health and Social Care with input from policymakers and frontline healthcare providers.
- AI Infrastructure Development: The government is investing heavily in supercomputing capabilities, such as Isambard-AI (Bristol) and Dawn (Cambridge), as part of the AI Research Resource (AIRR). These initiatives involve collaboration with universities (Cambridge, Bristol), and private companies (NVIDIA, Intel, Dell, HPE) to provide public AI infrastructure accessible to researchers, startups, and SMEs. This infrastructure aims to support sovereign, domain-specific AI models aligned with national regulatory frameworks.
- The Responsible Technology Adoption Unit (RTAU), formerly CDEI: This unit within the Department for Science, Innovation and Technology (DSIT) works with public and private sectors to enable trustworthy innovation using data and AI. It collaborates with industry, academia, and legal experts to develop tools, governance frameworks, and

standards, and demonstrate positive uses of data and AI. This includes joint work with tech-UK on the Portfolio of AI Assurance Techniques and the "Fairness Innovation Challenge" with Innovate UK and regulators to address bias (DSIT, 2024).

Safeguards are built into these partnerships. On data ownership and privacy, the UK relies on the Data Protection Act 2018, UK GDPR, and forthcoming legislation such as the Data Protection and Digital Information Bill, alongside initiatives like the National Data Library. Sensitive data use is being considered to monitor bias, and projects such as the Nightingale model use anonymised NHS data under strict controls. Algorithmic transparency is advanced through the Algorithmic Transparency Recording Standard, now mandatory across government, and through guidance on AI safety, content labelling, and deepfake risks. Accountability and governance are reinforced by clarifying responsibility within AI supply chains, appointing departmental AI ministers, and establishing a central oversight function within DSIT. The AI Safety Institute provides evidence for regulation, while contestability and redress mechanisms aim to ensure individuals can challenge harmful outcomes. Fairness and bias mitigation are further supported by reviews, innovation challenges, and HR guidance, with emphasis on diverse datasets and teams (DSIT, 2024). Overall, the UK's strategy combines innovation with safeguards, ensuring that public-private collaborations advance AI in ways that remain transparent, fair, accountable, and aligned with democratic values.

2.4.5 Workforce Development and Capacity Building

The UK is pursuing a multi-faceted strategy to build AI expertise and talent pipelines, combining long-term national initiatives with practical training for public servants. The National AI Strategy aims to secure leadership in science and AI by addressing global competition for skills, data, compute, and finance. To close the AI skills gap, the government supports advanced training programmes such as Turing AI Fellowships, Centres for Doctoral Training, and AI Conversion Courses, alongside flexible career pathways for software engineers, data scientists, and product managers. Employers and employees are encouraged to upskill through initiatives like Skills Bootcamps, while schools benefit from AI-focused curricula via the National Centre for Computing Education. Clear career guidance and pathways are being developed to inspire wider participation. To attract international talent, the UK offers streamlined visa routes, including the "Global Talent", "High Potential Individual", "Scale-up", "Innovator", and "Graduate" routes, as well as the "Global Entrepreneur Programme". These are designed to bring leading researchers, practitioners, and founders into the UK's AI ecosystem.

Significant R&D and infrastructure investments underpin this effort. The UK Research and Innovation programme connects academia and industry, while new AI research hubs and supercomputing resources – such as Isambard-AI (Bristol) and Dawn (Cambridge) – are expanding national compute capacity. The government is also creating a National Data Library and ensuring AI energy demands are met sustainably through the AI Energy Council. The UK advances international collaboration through agreements with the US, Japan, South Korea, and Singapore, with the AI Safety Institute playing a key role in cross-border evaluation and governance.

Within the public sector, AI adoption follows a "Scan > Pilot > Scale" approach, supported by new procurement guidelines and the AI Management Essentials scheme. The Cabinet Office is tripling AI engineers to establish a government AI Incubator, while the Department

for Science, Innovation and Technology coordinates regulator expertise through a central risk analysis function and a GBP 10 million funding package. The Digital Regulation Cooperation Forum and the Responsible Technology Adoption Unit support cohesive regulatory oversight and governance testing.

Transparency and ethics remain central. The Algorithmic Transparency Recording Standard is now mandatory across government, requiring disclosure of algorithmic tools. Guidance on AI safety, ethics, and HR use – developed with The Alan Turing Institute – equips public servants with practical frameworks for responsible AI deployment. Overall, the UK is combining skills development, talent attraction, infrastructure investment, regulatory innovation, and ethical safeguards to ensure a sustainable AI workforce and responsible public sector adoption.

2.4.6 AI Technologies in the Public Sector

The UK Government is actively integrating Artificial Intelligence (AI) across the public sector, guided by a "pro-innovation" regulatory approach that seeks to balance technological advancement with risk management and public trust. With the domestic AI market projected to exceed USD 1 trillion by 2035, the strategy aims to position the UK as a global leader in safe and responsible AI. AI is already reshaping government operations. Procurement processes now include ethical safeguards and transparency requirements, supported by new policy notes and the AI Dynamic Purchasing System.

Service delivery platforms are being transformed in healthcare, education, civil service administration, transport, and defence: the NHS is deploying diagnostic AI tools and training a health foundation model, schools are exploring AI in personalised learning, civil servants are piloting AI chatbots and consultation analysis tools, and departments such as Defence and Transport are embedding AI into energy management and automated vehicle policy. These efforts follow a "Scan > Pilot > Scale" model, designed to accelerate adoption while ensuring interoperability and reusability. Importantly, decision-support tools remain focused on augmenting human judgment rather than replacing it.

The UK's regulatory approach emphasises agility through a cross-sector, principles-based framework built on safety, transparency, fairness, accountability, and contestability. This is supported by a network of oversight bodies: (i) the Department for Science, Innovation and Technology sets overall policy; (ii) the AI Safety Institute evaluates frontier models; and (iii) the Digital Regulation Cooperation Forum that promotes cross-regulator collaboration. The new Labour Government has announced a Regulatory Innovation Office to strengthen oversight and introduce binding rules on high-risk systems, while measures such as the Algorithmic Transparency Recording Standard have become mandatory across government. Risks relating to bias, privacy, security, and accountability are addressed through guidance from regulators like the ICO and EHRC, the National Cyber Security Centre's standards for secure development, and proposed statutory duties on AI developers.

Beyond government, AI promises significant economic and social impact. It could add GBP 10 billion annually to public sector productivity by the end of the current parliament, with broader contributions to GDP growth and global competitiveness. At the same time, it raises questions around job disruption, skills gaps, data access, and market concentration. The government is investing in supercomputing capacity, a National Data Library, AI Growth Zones, and training initiatives such as Skills Bootcamps to close these divides.

Ensuring that AI benefits are widely shared remains central, reinforced by international commitments such as the Council of Europe's AI treaty on human rights, democracy, and the rule of law. Overall, the UK's strategy combines adoption, regulation, and investment to drive innovation while embedding safeguards, transparency, and ethics. By aligning public trust with technological progress, it seeks to ensure AI becomes a sustainable driver of economic growth and public good.

2.4.7 Barriers and Enablers for AI Adoption in the Public Sector

The United Kingdom aims to position itself as a global leader in Artificial Intelligence (AI) and become an "AI maker, not an AI taker," with the domestic market expected to surpass GBP 1 trillion by 2035. To achieve this, the government is integrating AI into public services to boost productivity, enhance service delivery, and drive economic growth. Yet political, economic, legal, and technical barriers remain significant (DSIT, 2024).

Politically, challenges include fragmented departmental mandates, uncertainty around the regulatory framework, and ongoing debate between advocates of light-touch, "pro-innovation" regulation and those pressing for stronger safeguards (Tony Blair Institute for Global Change, 2025). Concerns have also been raised about over-reliance on US-based AI companies and limited transparency in agreements affecting public data use. Economically, high infrastructure costs, elevated electricity prices, and limited fiscal headroom constrain large-scale AI development, while skills shortages and competition for talent remain acute. Smaller firms face difficulties accessing funding and managing compliance burdens, and liability costs deter individuals from pursuing redress for AI-related harms.

Legally, the absence of AI-specific legislation leaves the UK dependent on a principles-based approach enforced through existing regulators, raising concerns about enforceability, consistency, and accountability across the AI lifecycle. Questions about liability in complex supply chains remain unresolved, while existing frameworks for data protection and intellectual property often fail to address AI-specific risks. Technical barriers compound these issues: the UK lags in computing capacity, faces energy demands from data centres, and must grapple with the opacity of advanced models, challenges of data access, and the speed of technological change outpacing regulation.

To address these barriers, the Government has advanced a series of enablers. The National AI Strategy (2021) and the AI Opportunities Action Plan (2025) outline a vision for an AI-enabled economy, anchored by cross-sectoral principles of safety, transparency, fairness, accountability, and contestability. Central coordination by the Department for Science, Innovation and Technology (DSIT) aims to ensure regulatory coherence, while GBP 1.34 billion is being invested in supercomputers such as Isambard-AI and Dawn to expand national compute capacity. AI Growth Zones, starting with a pilot at Culham, are designed to accelerate infrastructure with enhanced power access and expedited planning, supported by an AI Energy Council to manage sustainability challenges.[112]

On data and transparency, initiatives include the creation of a National Data Library, the mandatory Algorithmic Transparency Recording Standard, and ongoing exploration of meaningful transparency measures such as AI labelling and explainability requirements. Regulatory capacity

[112] AI Energy Council to ensure UK's energy infrastructure ready for AI revolution. GOV.UK (7 April 2025). https://www.gov.uk/government/news/ai-energy-council-to-ensure-uks-energy-infrastructure-ready-for-ai-revolution

is being strengthened through the Digital Regulation Cooperation Forum, new cross-regulator advice services, and a GBP 10 million package to support technical expertise. At the same time, an AI assurance ecosystem is being developed to provide risk-assessment tools, audits, and certification (DSIT, 2024).

Skills development remains central to the Strategy, with investments in apprenticeships, bootcamps, conversion courses, and global visa routes to attract talent. The Government is also positioning the public sector as an exemplar for safe and ethical AI adoption, applying a "Scan > Pilot > Scale" approach and working in partnership with leading AI companies such as OpenAI. Ethical considerations are embedded in efforts to ensure fairness, inclusivity, and bias mitigation, while international leadership is advanced through the AI Safety Summit, the Bletchley Declaration, and collaboration with G7, UN, and the Council of Europe to shape global governance frameworks.

Overall, the UK's approach combines investment, governance reform, and international engagement to overcome structural barriers and promote cost-effective, customisable, and inclusive AI models, seeking to balance rapid innovation with public trust and democratic values. The UK is pursuing a comprehensive strategy combining skills development, talent attraction, infrastructure investment, regulatory innovation, and ethical safeguards. This multi-faceted approach aims to ensure sustainable AI workforce development and responsible public sector adoption while positioning the UK as a global AI leader.

2.5 QAZAQSTAN

2.5.1 AI Governance

In Qazaqstan, just recently the former Ministry of Digital Development, Innovation, and Aerospace Industry has been reorganised by Presidential Decree into the Ministry of Artificial Intelligence and Digital Development. Earlier, in May 2024, a dedicated authority for artificial intelligence was established – the Artificial Intelligence and Innovation Development Committee (Committee) – was created under the Ministry to foster a robust AI ecosystem, promote the safe deployment of emerging technologies, and enhance human capital.[113] Operating under the Ministry of Digital Development, Innovation, and Aerospace Industry (Digital and Space Ministry), the Committee oversees the development and regulation of artificial intelligence, innovation, the electronics sector, and digital assets. Its core functions include shaping and implementing state policies in these fields, encouraging innovation and scientific research, and supporting technological advancement. The Committee is also tasked with regulatory and legal oversight, cross-sectoral coordination, and providing advisory support within its areas of responsibility. It also coordinates the activities of the Astana Hub International Technology Park, a key platform for nurturing IT startups and driving innovative projects in Qazaqstan.

The Commission on Digitalisation also serves as Qazaqstan's top-level strategic authority for steering the integration of artificial intelligence nationwide, within the broader digital transformation of the country. Led by the Prime Minister and composed of key ministers across government, the Commission holds a cross-ministerial mandate to align policies, set overarching

[113] Government of the Republic of Qazaqstan. Decree No. 342 (27 April 2024): "On certain issues of the Ministry of Digital Development, Innovations and Aerospace Industry of the Republic of Qazaqstan". https://adilet.zan.kz/rus/docs/P2400000342

national goals, and ensure cohesive execution across sectors. Its role is primarily strategic and supervisory, it guides the direction of AI and digital initiatives, identifies and removes systemic barriers, and drives comprehensive, multi-sectoral transformation. By providing high-level oversight and facilitating coordination among government bodies, the Commission ensures that Qazaqstan's AI integration aligns with broader digital modernisation efforts and contributes effectively to the nation's long-term development objectives.

Inspired by international best practices, including Singapore's experience, Qazaqstan is in the process of developing Digital Transformation Roadmaps within government institutions as part of its broader governance modernisation agenda.[114] These roadmaps are expected to serve as policy instruments aimed at streamlining and accelerating core administrative processes through the application of artificial intelligence and advanced digital technologies. They prioritise reengineering outdated procedures, digitising datasets, and responding to citizen requests received via platforms such as e-*Otinish*, the 109 hotline, and social media channels. As integral components of Qazaqstan's evolving governance framework, these tools are designed to enhance institutional efficiency, foster innovation, and create a more citizen-centric public administration. However, as implementation is still at the planning stage, their effectiveness in driving systemic digital transformation remains to be seen.

As part of Qazaqstan's institutional framework for artificial intelligence, the state leader announced plans to establish an International Advisory Council on Artificial Intelligence, which would operate under his direct supervision.[115] The envisioned entity is designed as a strategic platform for public-private dialogue and knowledge exchange with leading global experts in AI. By bringing together engineers, researchers, and entrepreneurs, it aims to develop recommendations for national AI policy and to strengthen international cooperation in this field. However, these efforts are relatively new and not yet solidified into sustained, strategic international cooperation platforms, although they clearly highlight Qazaqstan's intention to position itself as an active participant in global AI governance.

Qazaqstan's international cooperation on AI is advancing through capacity-building, global market integration, and cross-sectoral innovation support.[116] While not a member of GPAI, Qazaqstan's innovation-led initiatives and multi-stakeholder collaborations align with GPAI principles, indicating readiness for deeper engagement. However, the country's role in shaping international AI governance remains undefined, as it has yet to articulate a formal foreign policy stance or institutional mandate focused on global AI norms, coordination, or standards-setting.

Meanwhile, Qazaqstan is actively engaging in bilateral cooperation on AI governance, particularly through its growing partnership with the United Arab Emirates. In early 2025, the two Governments agreed to collaborate on the development of regulatory sandboxes, with artificial intelligence identified as an initial area of focus. This initiative aims to establish adaptive regulatory frameworks that strike a balance between effective oversight and the promotion of innovation in emerging technologies. A tangible outcome of this cooperation is a training

[114] Qazaqstan to create AI infrastructure based on supercomputer. EL.KZ. (18 July 2025). https://el.kz/en/Qazaqstan-to-create-ai-infrastructure-based-on-supercomputer_400030258/

[115] В Казахстане появится совет по искусственному интеллекту под руководством Токаева [Qazaqstan to establish AI council under Tokayev's leadership]. Zakon.kz. (15 January 2025). https://www.zakon.kz/politika/6463464-v-Qazaqstane-poyavitsya-sovet-po-iskusstvennomu-intellektu-pod-rukovodstvom-tokaeva.html

[116] The Astana Hub collaborates with the World Bank and Draper University on the Hero Training programme, which has already yielded early investments, such as CITIX securing USD 1 million. Partnerships with Alchemist-X and the Silicon Valley Residency help local startups enter U.S. markets, while the Silkroad Innovation Hub Office in Astana and overseas IT hubs in Singapore, Riyadh, and the U.S. strengthen global connectivity. The Silkway Accelerator, run with Google for Startups, has enabled over 60 startups to raise nearly USD 28 million and export to 85 countries.

programme for Qazaqstani civil servants, delivered in the UAE, with thematic emphasis on AI, information security, and digital transformation.

A central pillar of the partnership is the Executive Leadership Programme, jointly implemented by the Mohammed bin Rashid School of Government and Qazaqstan's Academy of Public Administration. Under this programme, Qazaqstan plans to send vice-ministers and senior officials engaged in digital policy matters to the UAE for capacity-building in advanced public sector management, with a specific focus on AI applications in governance. Although still in its formative phase, this bilateral engagement underscores Qazaqstan's commitment to strengthening institutional capacity and aligning national AI governance with international best practices.

2.5.2 AI Regulation

Following the establishment of a designated body, the Qazaq government adopted the Concept for Artificial Intelligence Development for 2024–2029 (AI Concept) to enhance the efficiency of public administration, drive economic growth, and improve the quality of life through the application of AI technologies.[117] The document provides the legal and strategic foundation for the ethical deployment of AI, the development of standards, and regulatory oversight. It sets out key objectives, including the creation of infrastructure for AI, ensuring access to data, developing human capital in the field of AI, and supporting research and development in this domain. Thus, Qazaqstan's AI Concept focuses on national competitiveness, while the EU AI Act is centred on a human-centric approach. Priority areas for AI application identified in the document include public administration, industry, energy, transport, logistics, water supply, and agriculture.

The Concept recognises the critical importance of ethical standards in AI governance, outlining broad commitments to prevent discrimination, uphold human rights, and ensure that AI systems reflect societal values. It also anticipates the establishment of ethical norms and the formation of an advisory council to guide their implementation. Yet, although the AI Concept emphasises technical standards such as quality, accuracy, and safety, it does not explicitly incorporate ethical norms into the proposed future conformity assessment system. The Concept currently lacks a detailed roadmap for converting international best practices into concrete governance tools. This gap weakens the coherence between the international ethical frameworks it references, and the actionable national measures proposed for AI ethics and accountability.

Nonetheless, the Concept shows meaningful progress across essential AI readiness pillars – governance, infrastructure, and workforce development, and aligns well with globally recognised frameworks for ethics (UNESCO), governance (OECD), and inclusive development (SDGs). To fully align with global benchmarks, however, the Concept would benefit from incorporating clearer interim KPIs, structured timelines, and public monitoring dashboards, ensuring the strategic goals translate into measurable outcomes and sustainable policy learning. These additions would help ensure that strategic goals evolve into measurable outcomes and facilitate ongoing, evidence-based policy refinement.

[117] Government of the Republic of Qazaqstan. Концепция развития искусственного интеллекта на 2024–2029 годы [Concept for Artificial Intelligence Development for 2024-2029]. 26 June 2024. https://govtec.kz/assets/media/kontseptsiya-26062024-1500.pdf

Currently, the AI Concept's implementation is in its early stages, with foundational projects like the national AI platform and building supercomputer capacity actively under development. Therefore, tangible progress remains to be demonstrated, and the successful realisation of these initiatives will depend heavily on effective institutional coordination, sustained long-term funding commitments, and the integration of robust performance-based evaluation mechanisms.

Many of the current gaps in regulatory clarity and institutional roles are expected to be partially addressed through Qazaqstan's forthcoming dedicated legislation. On 14 May 2025, the Mazhilis (the lower chamber of Parliament) approved in the first reading a draft law "On Artificial Intelligence", which represents a major step toward establishing a legal framework for AI governance and regulation. Drawing significantly from the EU AI Act it introduces risk-based classification of AI systems, prohibitions on harmful use cases, and defines responsibilities of users and system holders. It creates institutional roles, mandates a National AI Platform for centralised data management, and outlines the state's coordination function. If adopted after the second and third readings, Qazaqstan will become one of the first countries in the world, alongside the European Union and the Republic of Korea, to implement a dedicated AI law rather than relying on sectoral regulations or voluntary guidelines.

That said, the law remains largely declarative and lacks clear mechanisms for enforcement, oversight, and technical standards (GRATA International, 2025). Critical actors such as AI developers are undefined, and key elements, such as risk management procedures and criteria for banned systems, require further regulatory elaboration. Several elements of the draft remain ambiguous, potentially hindering consistent interpretation and enforcement. As a result, the practical implementation of the law is likely to depend on the timely development of supplementary regulations and the establishment of appropriate technical standards. Without comprehensive secondary legislation and practical guidance, effective implementation will remain uncertain, potentially delaying Qazaqstan's alignment with global AI governance norms.

In the absence of technical standards, there is currently no formal conformity assessment mechanism to monitor compliance with AI requirements or to ensure their consistent implementation, the Passport of the Case serves as the main procedural tool for approving AI-based projects, including AI agents in public governance. This form is not grounded in formal legislation but derives from internal procedures of the Digital and Space Ministry. Approval follows an administrative process in which the Artificial Intelligence and Innovation Development Committee reviews submissions, with final decisions issued by the Vice Minister. While the template requests information on the system's function, data sources, KPIs, and projected economic impact, it entirely overlooks ethical, legal, and societal considerations. Issues such as data privacy, algorithmic bias, human oversight, and alignment with international AI governance norms are not addressed. As a result, the framework prioritises technical and performance metrics over the responsible and trustworthy deployment of AI.

2.5.3 AI Infrastructure

Qazaqstan is actively translating its AI Concept into operational capacity through strategic infrastructure development. These efforts reflect a deliberate national strategy to build sovereign digital capabilities that support domestic innovation, reduce foreign dependence, and enable inclusive access to artificial intelligence (Omirgazy, 2025).[118] A cornerstone of this strategy is

[118] *Qazaqstan's new supercomputer marks breakthrough in national tech independence*. The Astana Times. (Omirgazy D., 10 July 2025). https://astanatimes.com/2025/07/Qazaqstans-new-supercomputer-marks-breakthrough-in-national-tech-independence/

the USD 50 million *Alemcloud* National Supercomputing Centre, launched in July 2025 and currently the most powerful facility of its kind in Central Asia (Brkic, 2025).[119]

Even so, the supercomputer system underpins *QazCompute*, a forthcoming national high-performance computing (HPC) access programme that aims to democratise AI infrastructure by offering free access to startups, universities, R&D institutes, and government agencies, thereby removing a major cost barrier to developing advanced AI models. A tiered pricing mechanism governs private sector use: local IT firms and international companies contributing to domestic job creation receive significant discounts, while foreign entities with limited national integration are charged market rates. This pricing architecture ensures broad access while supporting financial sustainability and aligning economic incentives with Qazaqstan's digital development goals.

Critically, this national computing infrastructure also supports the National Artificial Intelligence Platform (NAIP), a centralised system operated by National Information Technologies (NIT) to integrate generative AI into public administration.[120] NAIP aims to create a unified ecosystem for AI development across government, facilitating use cases such as automated administrative processing, predictive analytics, and citizen-facing virtual agents. By enabling model training, fine-tuning, and data automation at scale, NAIP operationalises QazCompute within mission-critical state functions. This alignment allows Qazaqstan to retain control over data flows, latency, and cybersecurity, while reducing strategic vulnerabilities linked to reliance on commercial cloud services like AWS or Google Cloud.

Supporting this infrastructure is a rapidly expanding data ecosystem. The Digital and Space Ministry has commissioned a Tier III-certified data centre in Astana that currently houses the national supercomputer. A larger facility – Central Asia's first Tier IV data centre – is scheduled to become operational in the first quarter of 2026. With capacity for 4,000 server racks, this facility will double Qazaqstan's total computer infrastructure and meet global standards for uptime, with only 26 minutes of allowable annual downtime. These reliability benchmarks are essential for sectors such as finance, governance, and critical infrastructure, where service continuity is non-negotiable. To accelerate infrastructure growth, the government offers custom duties exemptions and VAT offsets on imported computing equipment (GRATA International, 2025).

To complement these central infrastructure efforts, the National Information Technologies (NITEC) has launched its first digital modernisation project funded through external borrowing, securing KZT 34.3 billion via a credit line from the Development Bank of Qazaqstan.[121] Unlike the centralised supercomputing facilities, this project targets the infrastructure layer within government agencies themselves, enabling them to effectively interface with national platforms such as QazCompute and NAIP. The funding is being used to procure and install state-of-the-art server equipment, thereby upgrading the internal ICT systems that support public service delivery, data exchange, and secure information storage. Currently in the implementation phase,

[119] Built on an NVIDIA H200 GPU cluster, the system delivers up to 2 exaflops of peak performance using FP8 precision (Omirgazy, 2025), making it purpose-built for AI and machine learning tasks. While this capability is notable regionally, it does not equate to exaflop-scale performance under FP64 benchmarks, which define global rankings of high-performance computing systems. The advertised "2 exaflops" refers to AI-specific performance using FP8 (8-bit floating point), which is fast but less precise. For scientific computing, such as climate modelling, physics simulations, or cryptography, FP64 (double precision) is the standard used in global supercomputer rankings like the TOP500. On that tougher FP64 benchmark, the actual performance of the Alemcloud supercomputer is around 12 petaflops – more than 160 times lower than the FP8 figure.

[120] https://www.nitec.kz/ru/proekty/aitec

[121] *NITEC to attract KZT 34 billion of borrowed funds for development of server infrastructure of government agencies*. NITEC. National Information Technologies (NIT) (1 February 2025). https://www.nitec.kz/en/news/nitec-attract-kzt-34-billion-borrowed-funds-development-server-infrastructure-government

the project is designed to enhance resilience, data protection, and service reliability across government bodies. In doing so, it not only strengthens the broader digital infrastructure needed to support AI deployment but also reduces reliance on outdated or fragmented legacy systems.

Together, QazCompute, NAIP, and the emerging data centre ecosystem form a cohesive foundation for Qazaqstan's AI ambitions. This integrated model reflects a shift from fragmented digital projects to a coordinated, sovereign AI ecosystem, where computational capacity, policy direction, and application design are strategically synchronised. It enables the country not only to localise AI solutions in Qazaq and Russian but also to foster an innovation-friendly environment for both domestic and foreign actors. Moreover, by ensuring public ownership and regulatory control over critical infrastructure, Qazaqstan enhances its digital sovereignty, aligns infrastructure development with national policy priorities, and strengthens its position in the global AI landscape.

2.5.4 Workforce Development and Capacity Building

Qazaqstan has announced the establishment of Alem.ai, an international artificial intelligence centre designed to serve as the flagship ecosystem for AI innovation, talent development, and economic diversification. Positioned at the intersection of education, research, and technology, Alem.ai is tasked with cultivating a critical mass of AI professionals and supporting the creation of globally competitive applications, agents, and products. The initiative recognises that human capital is the key resource in the digital economy and aims to mobilise a broad range of participants, including school and university students, researchers, civil servants, corporate experts, Big Tech professionals, and investors, within a collaborative, innovation-driven environment. This collective approach aims to cultivate world-class talent and generate AI solutions with the potential to make a global impact.

Alem.ai is also envisioned as a cornerstone for Qazaqstan's transition to an AI-oriented economic model and its ambition to become a competitive player in the global AI economy. It is expected to contribute significantly to the national goal of reaching USD 5 billion in IT service exports by 2029, up from USD 691 million in 2025, according to the National Bank of Qazaqstan.[122] It should be noted, however, that Alem.ai is still in the planning phase, and its true capacity to deliver on these far-reaching objectives will depend on effective execution, sustained investment, and robust governance.

Located in Astana, Alem.ai will comprise eight functional zones. These include the TUMO Centre, which provides extracurricular education in generative AI, animation, and robotics for youth aged 12–18, and the Tomorrow School, offering peer-based learning in AI and software development for older students. The Centre will also host co-working and accelerator space for up to 100 startups per year, research labs, a GovTech sandbox for piloting AI applications in government, and investor support services. Institutional partners include the Ministry of Digital Development and the Astana Hub (Bakpaeva, 2025).

This initiative aligns with the broader Generative Nation concept, which aims to build a digitally intelligent society capable of using AI to enhance productivity and economic growth. Over the next five years, the Government plans to train 1 million people in AI-related competencies, including 500,000 school students, 300,000 university students, 90,000 civil servants,

[122] From the interview of the Director-General of the Astana Hub International Technopark Magzhan Madiyev. https://www.youtube.com/watch?v=XyrEpsRHORg

80,000 private sector professionals, and 30,000 individuals from other target groups.[123] Several initiatives have been launched to support this agenda. “AI Corporate” focuses on upskilling professionals in the private sector, while “AI People” offers foundational AI literacy to the broader population, supported by Telegram-based microlearning and dedicated bootcamps.

To date, more than 16,000 civil servants have completed training in ethical and practical uses of AI. As part of the *TechOrda* Programme, Qazaq students can study at leading private IT schools using a voucher system.[124] However, significant implementation challenges remain. Institutional capacity to manage data quality, model oversight, and coordination across government is still evolving. The commercialisation of government-supported AI solutions remains under-developed, and ensuring deeper private sector involvement – from both domestic SMEs and global technology partners – will be essential to sustain momentum and innovation. Ultimately, the effectiveness of Alem.ai and its affiliated programmes will ultimately be judged by their ability not only to deliver education and tools, but to enable self-sustaining innovation and economic value across public and private sectors.

2.5.5 AI Technologies in the Public Sector

Qazaqstan’s public sector demonstrates a multifaceted integration of artificial intelligence (AI) technologies, which can be categorised into three primary functional roles: (i) AI for decision-making; (ii) AI for public service delivery; and (iii) AI for operational effectiveness and decision support. This framework highlights the nuanced ways in which AI contributes to governance, social protection, citizen engagement, healthcare, and law enforcement while preserving necessary human oversight.

In the category of AI for Decision-Making, Qazaqstan’s Digital Family Card (DFC) stands out as a pioneering example. The DFC system employs AI-driven analytics, including decision-tree algorithms and predictive models, to evaluate family vulnerabilities by aggregating data from over 100 socio-economic indicators across government databases. Crucially, the AI autonomously determines eligibility for social benefits and automatically issues notifications via SMS to vulnerable households, thereby eliminating the need for application submissions.

This fully automated decision-making capability enables the DFC to enhance social protection with minimal human intervention, reducing service delivery times from seven days to one and covering more than 20 million citizens since its 2022 pilot launch. Furthermore, the system supports policymakers by providing predictive insights on emerging social risks, exemplifying AI’s role not only in direct decision execution but also in strategic governance. By transforming public administration and enhancing citizens’ quality of life, the DFC earned international recognition in the “Inclusive Digital Transformation” category of the 2024 GovTech Prize in Dubai.[125]

In contrast**,** AI for Public Service Delivery encompasses technologies that directly interact with citizens to automate communication and transactional services, improving accessibility

[123] *Qazaqstan to build Central Asia’s first Tier IV data centre*. (DKnews.kz. 17 June 2025). https://dknews.kz/en/articles-in-english/362659-Qazaqstan-to-build-central-asia-s-first-tier-iv-data

[124] Omirgazy, D. *Qazaqstan Opens Silkroad Innovation Hub Office to Promote Regional Startups*. (The Astana Times, 13 September 2023). https://astanatimes.com/2023/09/Qazaqstan-opens-silkroad-innovation-hub-office-to-promote-regional-startups/

[125] Qazaqstan’s Digital Family Card awarded the 2024 GovTech Prize. (UNDP in Qazaqstan, 13 February 2024). https://www.undp.org/Qazaqstan/press-releases/Qazaqstans-digital-family-card-awarded-2024-govtech-prize

and efficiency without making autonomous policy decisions. Notable applications include conversational AI assistants such as e-agent and e-Gov AI, integrated into the national e-Gov platform to address citizen inquiries, manage over 60 public services, and process core transactions like record filings and certificate issuance. Other tools like Tax Helper and E-translator facilitate tax filing and provide multilingual communication assistance to citizens.

By 2029, the Government plans for AI-assisted services to cover 20% of all public service delivery, reflecting wide-scale integration ambitions (Zhazetova, 2024). To date, Qazaqstan has launched over 50 AI projects involving 16 government bodies and reaching 8 million citizens, demonstrating tangible progress toward AI-powered public services (Akhmetkali, 2025). Most are still at the pilot or initial deployment phase. The value proposition is broadly framed around improved user experience, automation, and faster service delivery, but concrete performance indicators (such as user adoption rates, response time reductions, or cost savings) have not yet been released publicly.

The third category, AI for Operational Effectiveness and Decision Support, involves AI solutions designed to augment human work by delivering analytical insights, enhancing workflows, and supporting but not replacing complex decision-making. The AI-therapist pilot in healthcare exemplifies this role by assisting clinicians in preliminary diagnoses, with up to 80% accuracy, and reducing documentation time by approximately 40%, thereby improving diagnostic efficiency while preserving clinical judgment. Within the law enforcement domain, the Digital Investigator Assistant automates case classification, suggests investigative strategies, analyses interview transcripts for inconsistencies, and drafts investigation documents incorporating judicial precedents.

Complementing this, a crime prediction model uses extensive historical data to forecast likely crime locations and times, enabling proactive resource allocation. The AI-powered asset tracing platform further aids operational efficiency by supporting the recovery of illicitly transferred assets. While these tools have contributed to tangible outcomes, such as detecting over 2,500 wanted individuals and debtors, they function primarily as decision-support systems, with final investigatory and prosecutorial decisions retained by human authorities. Notwithstanding their potential, these deployments currently lack comprehensive public reporting on performance metrics and necessitate stronger ethical and accountability frameworks.

This categorisation accentuates the breadth and balance of AI roles within Qazaqstan's public sector, from autonomous decision-making in social welfare through the DFC to facilitative public service delivery and supportive operational tools in healthcare and justice. Such differentiation illustrates Qazaqstan's strategic approach in leveraging AI's benefits, enhancing efficiency, responsiveness, and inclusiveness, while safeguarding human oversight, particularly in sensitive domains. Recognising where AI acts independently versus collaboratively within government processes also provides a valuable lens for guiding future evaluation, governance, and ethical considerations; as Qazaqstan continues to develop its AI-powered public administration. However, there is still a lack of transparency regarding algorithmic accountability, data privacy protections, and oversight mechanisms within the programme (UNESCO, 2023).

2.5.6 Public-Private Partnerships on AI

Qazaqstan's policy and institutional landscape regards public–private partnerships (PPPs) as a central pillar for advancing national AI capacity, innovation, and commercial application. The collaborative development of the Qazaq-language AI model (KazLLM), led by Nazarbayev University's Institute of Smart Systems and Artificial Intelligence, exemplifies this approach. This initiative harnesses the expertise of research institutes and private partners to process local-language data, train contextually appropriate models, and ensure the alignment of AI systems with Qazaqstan's linguistic and societal needs. Furthermore, Qazaqstan's emerging national AI platform and supercomputing infrastructure underscore the commitment to PPPs, granting both businesses and researchers access to key technological resources, thus reinforcing shared innovation and knowledge transfer.

The Almaty Hub Industrial AI Accelerator,[126] coordinated by the Ministry of Digital Development, Innovations and Aerospace Industry, exemplifies the public-private partnership (PPP) model by collaborating with relevant ministries and private enterprises to pilot and deploy AI solutions across key sectors like mining, energy, healthcare, fintech, and the public sector. The Accelerator offers both technical and regulatory support, making it easier for private companies to develop and apply AI technologies in practical settings. For example, Qala AI, a leading Urban-Tech company, has created the Resilient City Platform and Resilient Business Platform, which use geographic data and advanced analytics to help city officials and businesses make smarter decisions related to urban planning and resiliency, such as managing risks from natural disasters or optimising business locations. Another example is Onco-Zero AI which leverages neural networks (a kind of AI inspired by the human brain) combined with an easy-to-use interface to quickly detect cancer and support ongoing patient treatment monitoring. These examples highlight how the PPP framework nurtures innovative and scalable AI solutions that address real-world challenges across multiple industries.

Qazaqstan also uses PPPs to address the need for skilled AI talent, notably through programmes such as Tech-Orda, which funds private training centres with the goal of training 20,000 IT professionals by 2029, the corporate-led "AI Corporate" initiative, and youth-focused TUMO Centres. These programmes align workforce skills with industry demands and offer direct channels for private sector participation in curriculum development and training delivery. Strategic direction and legal certainty are provided by Qazaqstan's 2024-2029 AI Development Concept and the impending AI Law, which collectively foster an enabling environment for private sector engagement.

However, Qazaqstan's AI governance framework is still emerging. According to a 2025 KPMG study, just 20% of companies have implemented AI-specific policies, while 40% still use only general data governance frameworks (KPMG, 2025). This suggests that progress is made toward structured AI oversight but also reveals a gap in organisational awareness of evolving regulatory requirements. Notably, even among early adopters of AI, more than 60% lack dedicated AI regulations, indicating that issues of compliance and governance are only gradually being integrated into corporate strategies. Despite these shortcomings, just 13% of companies voice concern over government AI restrictions, a figure that contrasts with anxiety levels in highly regulated environments like the European Union. This reflects a regulatory posture in Qazaqstan that is perceived as enabling rather than constraining AI innovation, positioning PPPs as both a practical and strategic driver for the country's inclusive, globally competitive AI ecosystem.

[126] https://industrial-ai-acceleration.kz/eng#programme

2.6 THE PHILIPPINES

2.6.1 AI Governance

The governance of artificial intelligence (AI) in the Philippines is led by a set of interrelated agencies, each bringing a distinct focus and mandate to national AI policy. The Department of Trade and Industry (DTI), the Department of Science and Technology (DOST), the Department of Information and Communications Technology (DICT), and the National Economic Development Authority (NEDA) collectively shape strategic priorities, regulatory frameworks, and research agendas. Despite the active involvement of these multiple agencies, the Philippines lacks a single, dedicated lead agency responsible for coordinating and consolidating the national approach to AI policy implementation. According to UNESCO's Readiness Assessment Methodology (RAM), this siloed approach has hindered the practical implementation of AI strategies: while stakeholders are aware of existing roadmaps, strategies, and initiatives, the absence of a central coordinating body makes it difficult to ensure consistency, coherence, and effective policy execution.

DTI first introduced the National AI Strategy Roadmap (NAISR 1.0) in May 2021, establishing the country's initial AI priorities, with a focus on industry adoption, innovation, and workforce development. Building on this foundation, DTI launched NAISR 2.0 in July 2024, incorporating recent advancements in AI, such as generative AI, while integrating ethical considerations into the governance frameworks. NAISR 2.0 identifies seven strategic imperatives, including improving internet connectivity, enabling data access, developing AI talent, and fostering research and development. Implementation is being operationalised through the Centre for AI Research (CAIR), which engages with regional and international partners and addresses structural challenges such as the digital divide, low R&D investment (0.3% of GDP), shortage of AI professionals, and early-stage regulatory alignment.

DOST, in contrast, leads the scientific, governance, and ethical aspects of AI development, spearheading the recently introduced *National AI Strategy for the Philippines* (NAIS-PH), aimed at making innovation inclusive, ethical, and responsive. The strategy focuses on five core areas – infrastructure, workforce, innovation, ethics and policy and deployment – with implementation planned from 2024 to 2028. While DTI drives AI adoption for business growth and digital transformation, DOST ensures that AI integration aligns with national priorities, research capacity, and responsible innovation principles. DOST actively collaborates with universities and research institutions to advance AI solutions, positioning the Philippines as a potential AI innovation hub by 2040.

Box 6: National AI Strategy for the Philippines (NAIS-PH)

Objectives

- Foster inclusive AI innovation and digital transformation.
- Strengthen AI governance, ethics, and data protection.
- Promote globally competitive industries through responsible AI integration.

Key Focus Areas:

- **Infrastructure**: Expansion of high-performance computing (HPC) power by 26 times by 2028, with regional HPC sites to support AI research.
- **Workforce**: Upskilling and reskilling through AI micro-credentials and platforms such as Coursera and SPARTA.
- **Innovation**: Establishment of the **AI Factory** and **AI Refinery** to connect research and industry.
- **Data Governance and Policy**: Development of clear standards for ethical and transparent AI use.
- **AI Deployment**: Decentralisation through regional offices and creation of an **AI Hub (i-Hub)** to provide tailored AI solutions for MSMEs and local industries.

Implementation:
Guided by the Philippine AI Programme Framework (2024-2028), the strategy targets key sectors – agriculture, education, smart cities, creative industries, and national security. It encourages cross-sector collaboration to build a cohesive AI ecosystem grounded in innovation, ethics, and public trust.
Expected Outcomes:
By 2028, NAIS-PH envisions an AI-powered economy that promotes inclusive growth, strengthened governance, and sustainable digital transformation, positioning the Philippines as a regional leader in ethical and human-centred AI development.

Source: https://archive.opengovasia.com/2025/05/22/the-philippines-president-approves-comprehensive-ai-roadmap/

DICT functions as the primary ICT policymaking and administrative body, integrating AI governance into broader digital government initiatives. Its role includes ethical guidance, standard-setting, capacity building, and international engagement. DICT has developed frameworks such as the Joint Memorandum Circular on Ethical AI Use, aligned with ASEAN and OECD guidelines, and fosters skills development through the Philippines Skills Framework for Analytics and AI. It also manages digital platforms like e-GOVPH to enhance citizen access and government efficiency, while participating in international fora on AI governance and safety.

This multi-actor system provides agility, enabling sectoral bodies to advance AI initiatives without waiting for centralised approval. However, it also fragments accountability. Without a centralised regulator, coherence between economic, ethical, and legal dimensions of AI remains uncertain. Moreover, according to UNESCO's Readiness Assessment Methodology for the Philippines (2025), the country's AI policy and strategy implementation suffers from a siloed approach. This fragmentation has hindered the effective operationalisation of AI initiatives. While various stakeholders are aware of existing roadmaps and strategies, the absence of a dedicated lead agency has made it challenging to ensure consistency and coordination in policy

execution. This lack of centralised oversight impedes the establishment of a cohesive national AI governance framework.

Despite these challenges, the *2024 Government AI Readiness Index*, where the Philippines ranked 56th out of 188 countries – above the global average – indicates that the country's AI governance is advancing, demonstrating significant potential to drive economic growth, accelerate technological adoption, and promote citizen-centred public services. At the same time, the Philippines actively participates in regional and multilateral AI fora, including ASEAN, ITU, and UNDP platforms. Bilateral cooperation with partners such as South Korea, Japan, and the European Union provides technical assistance, capacity building, and opportunities for policy benchmarking. These engagements partially compensate for domestic gaps in expertise and embedding the Philippines within a broader international AI ecosystem.

2.6.2. AI Regulation

The Philippines currently does not have a comprehensive, legally binding framework for regulating AI. Instead, the Government has adopted a cautious, "light-touch" approach intended to balance innovation with ethical and social considerations. While pragmatic in light of rapidly evolving technologies, this approach also creates gaps in coordination, enforcement, and policy coherence (Chua et al., 2025). Several bills are pending in Congress that aim to establish rules for different dimensions of AI governance.[127] Other proposals seek to establish a Philippine Council on AI to provide national oversight for human-centred AI adoption.[128] While these legislative efforts reflect rising interest, none has yet been enacted, leaving the regulatory landscape fragmented and incomplete.

At present, AI regulation depends on a patchwork of broader laws and frameworks related to data management, digital transformation, cybersecurity, and innovation. While these measures offer partial safeguards and incentives, they do not directly address AI-specific risks such as algorithmic accountability, transparency, or bias. To address these gaps, the country is actively advancing its AI governance framework through initiatives led by the National Privacy Commission (NPC) and updates to the Data Privacy Act. Recent NPC advisories provide detailed guidance on applying data protection principles to AI systems, emphasising transparency, accountability, fairness, and respect for data subject rights. These measures require organisations handling personal data to adopt responsible and ethical practices throughout the AI system lifecycle, from design to deployment and operation. Collectively, these efforts aim to ensure that AI adoption aligns with legal standards, strengthens regulatory compliance, safeguards privacy, and fosters public trust.

Meanwhile, the absence of comprehensive legislation has also led to sector-driven regulatory activity. The Commission on Elections (COMELEC), for example, has issued campaign guidelines banning deepfakes as tools of disinformation. Meanwhile, the Supreme Court is drafting an *AI Governance Framework for the Judiciary*, under its Strategic Plan for Judicial Innovations 2022–2027. This framework sets standards for AI applications in administration, finance, legal research, case management, and courtroom procedures, with pilot projects already underway in transcription and AI-enabled research tools.

[127] These include general AI regulation (HB 7396, HB 10385), generative AI (HB 10751), ethical and responsible use (HB 1177), AI integration in government services (HB 10845), labour protection in the context of automation (HB 9448, HB 10460), and deepfake accountability (HB 9425, HB 10567).

[128] HB 7913, and HB 10944.

These initiatives demonstrate the country's growing awareness of AI's risks and opportunities, yet progress remains uneven. Without a cohesive national framework, sectoral approaches risk duplication, gaps, and inconsistent enforcement. Moving forward, the Philippines will need to consolidate these dispersed efforts into a comprehensive regulatory structure that ensures legal clarity, protects citizens, and fosters an enabling environment for responsible, innovation-driven AI adoption.

Efforts to shape AI policy in the Philippines increasingly recognise that ethical considerations cannot be treated as secondary to economic growth. While national development goals often prioritise speed of innovation and market expansion, concerns such as diversity, inclusion, equity, environmental impact, and public health remain underexplored. This imbalance risks side-lining ethics in favour of short-term competitiveness. Encouragingly, discussions led by industry and civil society are beginning to broaden the agenda toward a more holistic approach that places social and cultural aspects at the core of AI governance.

However, these conversations unfold against the backdrop of a persistent digital divide. Millions of Filipinos still lack reliable and affordable internet access, with costs significantly higher than in neighbouring ASEAN countries and broadband penetration rates lagging behind. Limited competition in the telecommunications sector, underinvestment in infrastructure, and stark disparities between urban and rural areas exacerbate exclusion. The consequences are not only economic but also social, reinforcing inequities between income groups, genders, and marginalised communities. Women, for instance, face higher risks of online abuse and financial exclusion, while persons with disabilities remain largely overlooked in digital policy and system design. Non-profit initiatives and academic institutions are beginning to address these challenges, yet progress is constrained by structural barriers and limited resources.

2.6.3 Infrastructure

The Philippines has made significant progress in developing AI infrastructure, evidenced by steady investments in data centres and expanding internet connectivity. This reflects both foreign and domestic interest in positioning the country as a regional digital hub capable of supporting advanced AI technologies and cloud computing services. Leading international players, such as Equinix,[129] and Alibaba Cloud,[130] have expanded operations in the Philippines, while local providers – including PLDT,[131] Converge,[132] and SpaceDC[133] – are scaling their data centre capacities. Together, these investments strengthen the nation's ability to support high-performance computing, large-scale AI model training, and the delivery of AI-driven services, forming a critical foundation for broader digital transformation.

Despite these developments, the Philippines relies heavily on cloud services and external vendors for large-scale AI operations, highlighting its limited sovereign computing capacity. Unlike countries that have invested in national supercomputers and dedicated AI platforms, the Philippines lacks domestic hardware and platforms capable of training frontier AI models. While partnerships with global cloud providers facilitate rapid deployment, they also introduce risks such as vendor lock-in, data sovereignty concerns, and restricted capacity to develop and

129 https://www.reuters.com/markets/asia/data-center-firm-equinix-enters-philippines-eyes-southeast-asia-2024-07-23/

130 https://www.reuters.com/world/asia-pacific/alibaba-cloud-announces-new-data-centres-malaysia-philippines-2025-07-02/

131 https://www.philstar.com/business/2024/07/11/2369212/pldt-completes-biggest-data-center

132 https://www.bworldonline.com/corporate/2024/09/16/621569/converge-plans-annual-data-center-expansion

133 https://spacedc.com/spacedc-to-build-the-largest-hyperscale-data-center-in-the-philippines/

scale indigenous AI solutions. Enhancing domestic computing infrastructure will be essential to strengthen the country's resilience, foster innovation, and ensure sustainable growth in AI capabilities.

To address the issue, the DOST is leading the efforts to build a strong national AI foundation, with planned investments exceeding PHP2.6 billion (about US$44 million) by 2028.[134] Guided by its AI R&D Framework 2019–2029, DOST focuses on enhancing national infrastructure, research capacity, and data systems. Major initiatives include the Computing and Archiving Research Environment (COARE), a high-performance computing facility that supports local research, and the upcoming AI Virtual Hub – a national centre designed to connect start-ups, researchers, and industry for collaborative innovation.

Alongside computing capacity, the accessibility and quality of digital networks are critical for AI adoption. Internet penetration is gradually improving, with fixed broadband and mobile networks reaching more users each year. However, challenges persist in upload speeds, reliability, and affordability, particularly in rural areas. These constraints affect households and micro, small, and medium enterprises that depend on digital infrastructure to integrate AI systems and optimise operations. The urban-rural connectivity gap represents a structural challenge that may limit inclusive participation in AI-driven innovation and economic development beyond metropolitan centres.[135]

Energy supply represents another key factor in AI infrastructure. Data centres and AI workloads demand significant and reliable electricity, yet the Philippines faces high energy costs and occasional service disruptions.[136] The Government is pursuing energy reforms, including renewable energy projects and exploring small modular nuclear reactors, to balance rising demand with sustainability objectives. Stable and affordable power, combined with robust digital networks, is critical to integrating AI into public services and industrial workflows. Without improvements in these areas, AI adoption could be constrained, and disparities in access to digital innovation may persist.

Therefore, while the Philippines has advanced its AI infrastructure through growing data centre capacity and increasing internet coverage, fully realising AI's benefits requires continued investment. Priority areas include enhancing network speeds and reliability, improving affordability, expanding energy capacity, and developing sovereign computing resources to support large-scale AI operations and develop national digital resilience.

2.6.4 Workforce Development and Capacity Building

The Philippines has made substantial strides in developing its AI workforce, combining upskilling, reskilling, and longer-term educational reforms to prepare for an AI-driven economy. Programmes such as the Department of Science and Technology's (DOST) AI PINAS and the SPARTA initiative have collectively trained over 49,000 individuals in AI and data science competencies. These programmes emphasise not only technical skills but also awareness of ethical AI use and governance principles, addressing UNESCO's call for a workforce capable of supporting responsible AI adoption. Alongside workforce development, the Philippines has

134 https://opengovasia.com/the-philippines-significant-investment-to-advance-ai-ecosystem/?c=kz

135 https://documents1.worldbank.org/curated/en/099011824231036851/pdf/P502027179f71d08418678193f2fabcdbec.pdf

136 https://www.trade.gov/market-intelligence/philippines-data-centers

supported more than 100 AI research and development projects between 2018 and 2024, reinforcing its innovation capacity and technical expertise.[137]

Government efforts increasingly foster collaboration across academia, industry, and public agencies to create robust learning ecosystems that promote research, innovation, and practical AI applications. The establishment of the Centre for AI Research (CAIR), in 2024, exemplifies this approach, linking research excellence with skills development and applied projects in sectors such as healthcare, agriculture, education, and manufacturing. These initiatives serve as practical platforms for cultivating AI capabilities while encouraging interdisciplinary collaboration focused on leveraging AI to drive socio-economic growth, address industry and societal challenges, and promote inclusive development.

Recognising the structural inequalities in access to education and digital skills, national strategies also focus on inclusive workforce development. UNESCO's RAM framework emphasises bridging regional and socio-economic disparities, and the Philippines is actively working to expand AI literacy to under-served populations, including rural communities and smaller enterprises. Programmes aim to ensure that AI's benefits are broadly distributed, mitigating the risk of exclusion and fostering equitable participation in innovation ecosystems.

Despite these advances, challenges remain in the education domain. Structural weaknesses in primary and secondary education, such as under-resourced teaching, low salaries, and limited professional development, constrain student learning outcomes and hinder the development of foundational ICT skill development. Surveys indicate persistent gaps in digital literacy among Filipino youth and adults, exacerbating inequality in access to AI opportunities. Research and development capacity is also limited, with relatively low numbers of full-time researchers and modest patent outputs, constraining the country's AI innovation potential.

To address these gaps, government agencies are updating curricula at all levels to embed STEM competencies, entrepreneurship, and AI literacy, while also recognising the value of Arts, Humanities, and Social Sciences in shaping responsible AI. Initiatives under TESDA, such as TVETPH 4.0 and the Artificial Intelligence Data Annotation (AIDA) Training Programme, target workforce readiness and inclusion, particularly for women and marginalised groups. Universities are also taking proactive steps: the University of the Philippines has developed Principles for Responsible and Trustworthy AI, established the multidisciplinary AI Advancement Committee, and created the UP AI Advisory Board to guide AI teaching, research, and policy. Silliman University has similarly introduced institutional policies and guidelines to integrate AI into its educational framework responsibly.[138]

Collectively, these efforts demonstrate a multi-layered approach to workforce development in the Philippines, combining immediate upskilling opportunities with foundational reforms that prepare future generations. Success will depend on sustained investment, expanded outreach to underserved populations, and continued integration of ethical, technical, and governance dimensions into all AI-related education and training programmes.

[137] https://www.dost.gov.ph/knowledge-resources/news/86-2025-news/3927-dost-s-solidum-highlights-ph-pivot-to-advance-ai-at-unesco.html

[138] https://unesdoc.unesco.org/ark:/48223/pf0000393860?posInSet=1&queryId=cb72b22d-9dd3-44cd-9090-c4c89328a09c

2.6.5 AI Technologies in the Public Sector

The Philippines has begun integrating AI solutions across multiple public sector domains, combining operational efficiency, service delivery, and decision-support capabilities. In the digital government domain, the national e-government platform, *eGovPH*, together with the embedded *eGovAI* system, offers seven active AI-enabled services accessible nationwide: identity verification via *PhilSys*, document extraction and validation, automated workflow routing for government transactions, complaint and feedback categorisation, fraud detection and anomaly alerts, resource allocation recommendations for agencies, and service delivery prioritisation based on usage patterns. These systems enhance operational efficiency and provide government decision-makers with actionable insights, enabling more responsive, evidence-based public service delivery.

In education, the Education Centre for AI Research (E-CAIR), established in 2025, represents the country's first hub dedicated to AI applications in the sector. E-CAIR focuses on practical AI solutions that address systemic challenges in teaching, learning, and school administration. Early outputs include AI tools for optimising education voucher distribution, computer vision systems to identify malnutrition and disabilities among learners, AI-driven hazard mapping of schools, and automated assessment of school leadership qualifications. These initiatives illustrate how AI is being deployed not as abstract experimentation, but as targeted interventions designed to enhance inclusion, equity, and administrative efficiency.

The judiciary has also piloted AI solutions to support operational decision-making. The Supreme Court's *Scriptix* transcription system, tested between 2023 and 2024 in several courts including the Sandiganbayan, automatically transcribes hearings in bilingual (Taglish) proceedings. By reducing the time required to produce accurate transcripts, Scriptix allows judges and clerks to review case content more quickly and focus on case analysis, enhancing the efficiency and reliability of judicial decision-making.

Beyond routine operations, several AI-based initiatives in the Philippines public sector are actively supporting evidence-based decision-making. In disaster risk management, the DOST-led AI-Powered Weather Forecasting for a Resilient Philippines (AI-4RP) integrates artificial intelligence into weather prediction and dam management processes, enhancing resilience and reducing flood risks in vulnerable communities.[139] Initiatives like *SkAI-Pinas* use remote sensing and big data for environmental monitoring, while the *Accelerated Earthquake Multi-Hazards Mapping and Risk Assessment* (ACER) programme supports nationwide disaster preparedness.[137]

In public health, Project AEDES, developed by *CirroLytix*, leverages AI to predict dengue outbreaks by analysing climate data, digital indicators, and satellite imagery, providing early warnings to local health authorities for targeted interventions. Recognised as a *Digital Public Good* by the Digital Public Goods Alliance, the platform is publicly accessible and primarily used by local government units and health agencies, with support from UNICEF and DOST, underscoring its significance and practical impact.[140] These operational deployments illustrate how AI is being applied in practice to inform proactive policymaking, improve resource allocation, and support timely interventions in critical public services.

[139] https://opengovasia.com/the-philippines-ai-powered-weather-forecasting-bolsters-preparedness/

[140] https://www.unicefventurefund.org/story/aedes-platform-dengue-prediction-using-climate-and-health-data-epidemic-management

Collectively, these initiatives illustrate that AI in the Philippine public sector is being deployed not as abstract experimentation but as targeted, sector-specific tools that enhance administrative efficiency, service delivery, and enable decision-making. By combining operational support with practical decision insights, these applications help build trust in digital government services and set a foundation for broader, responsible AI adoption.

2.6.6 Public-Private Partnerships on AI

Public-private partnerships (PPPs) play a vital role in advancing AI adoption across local enterprises, small and medium-sized enterprises (SMEs), and educational institutions in the Philippines. At the grassroots level, the Department of Science and Technology (DOST) actively supports AI integration for MSMEs, enabling innovations such as AI-based grading systems for durian crops used by agro-entrepreneurs in Davao to optimise harvest timings and pricing decisions. Similarly, in Cebu, MSMEs leverage generative AI for crafting marketing campaigns and conventional AI approaches to streamline supply chains, demonstrating AI's role as a complementary tool enhancing productivity and strategic decision-making without replacing human interaction.

The E-CAIR exemplifies a collaborative framework that links government agencies, schools, and private partners. Through initiatives such as AI-based data mapping under the Adopt-A-School Programme, E-CAIR is aligning public and private sector contributions with actual infrastructure needs. Regionally, its partnership with the Southeast Asian Ministers of Education Organisation (SEAMEO) reflects an ambition to position the Philippines as a leader in AI-driven education solutions in Southeast Asia.[141] While the long-term impact still requires systematic measurement, the Centre has already established itself as a credible mechanism for advancing AI adoption in education by focusing on concrete outputs, inclusiveness, and the potential for regional scaling.

Beyond specific sectors, PPPs initiatives led by the Department of Science and Technology, in collaboration with private technology providers and multilateral institutions such as the Asian Development Bank (ADB), support broader national strategies centred on research, capacity building, and infrastructure development. DOST has committed significant investments in AI projects spanning healthcare, education, mobility, and disaster risk reduction, while the ADB has provided funding to enhance infrastructure critical for AI deployment. These collaborations accelerate AI adoption, help close digital gaps, and develop a workforce equipped with the skills needed for emerging technologies. By integrating government resources, private sector expertise, and academic research, the Philippine PPP ecosystem is fostering a more coordinated, sustainable, and inclusive environment for AI innovation.

[141] https://aibc.world/news/philippines-opens-ai-center-for-education/

2.7 OTHER PROJECT COUNTRIES' AI GOVERNANCE FRAMEWORKS: OVERVIEW

2.7.1 Armenia

Armenia is in the early stages of developing a dedicated AI governance framework, as current regulation relies mainly on general data protection rules and lacks AI-specific legislation.[142] In December 2024, the Government approved joining the Council of Europe's Framework Convention on Artificial Intelligence, Human Rights, Democracy, and the Rule of Law, signalling its commitment to align domestic legislation with international standards.[143]

Leading policy efforts, the Ministry of High-Tech Industry launched the Virtual AI Institute in 2025,[144] in collaboration with AWS and Mistral AI, to foster AI research, innovation, workforce development, and product commercialisation, as well as to provide local researchers, startups, and public institutions with access to advanced AI tools and training. Armenia has also announced a landmark USD 500 million public–private partnership to build the first AI supercomputing data centre in the Caucasus,[145] scheduled to open in 2026 and aiming to position the country as a regional AI hub.

Meanwhile, Armenia is part of the EU4Digital Programme, aimed at strengthening cross-border cooperation, and is rolling out AI-driven initiatives in several sectors under its 2021–2025 digital transformation strategy. For instance, the Ministry of Labour and Social Affairs applies machine learning tools to help social workers more effectively identify and support vulnerable groups.[146] The Corruption Prevention Commission uses AI to enhance the review of public officials' asset declarations, with an algorithmic tool designed to flag anomalies and adapt its detection methods over time.[147] In partnership with the Foundation for Armenian Science and Technology, the Armenian Government has piloted a Generative AI programme in 16 public high schools across the country, introducing students to the fundamentals of artificial intelligence.[148]

2.7.2 Azerbaijan

Azerbaijan has moved rapidly over the past year to establish the foundations of AI governance. Until late 2024, its efforts in this area were marked by uncertainty, despite strong digital infrastructure and human capital. A decisive shift came in December 2024 with the adoption of two international AI standards – ISO/IEC TR 24028:2024 on trustworthiness and ISO/IEC TR 24372:2024 on computational approaches – anchoring national policy in principles of reliability, transparency, and interoperability. In March 2025, the Government approved the Artificial Intelligence Strategy for 2025–2028, outlining governance frameworks, data infrastructure, education and skills development, public-private partnership, and research and innovation. This was followed in April 2025 by an Implementation Plan that assigned institutional responsibilities,

142 Hetq (2024). Armenia to sign first international artificial intelligence treaty. https://hetq.am/en/article/171802

143 Armenia joins Council of Europe Convention; commits to ethical AI (2025). EU4Digital. https://eufordigital.eu/armenia-joins-council-of-europe-convention-commits-to-ethical-ai/

144 Virtual AI Institute, https://ai.gov.am/en/about-the-project.html

145 *Landmark $500 million AI project to launch in Armenia*. EU4Digital (2025). https://eufordigital.eu/landmark-500-million-ai-project-to-launch-in-armenia/

146 *Armenia's AI advancements, gaps & opportunities.* Civil Net (2023). https://www.civilnet.am/en/news/749620/armenias-ai-advancements-gaps-opportunities/

147 *Perspectives: Artificial intelligence boosts anti-corruption efforts in Armenia*. Eurasia net (2023). https://eurasianet.org/perspectives-artificial-intelligence-boosts-anti-corruption-efforts-in-armenia

148 https://fast.foundation/en/program/4379/2024/general-information

timelines, and accountability mechanisms, creating a sequenced policy architecture and greater coherence for AI development.

Meanwhile, Azerbaijan is beginning to operationalise AI across several priority sectors, with concrete use cases already piloted and ready for deployment. In the energy sector, AI is applied for predictive maintenance of generation assets, demand forecasting, and smart-grid management to improve efficiency and integrate renewable sources. In public services, AI-powered chatbots and virtual assistants are being tested in pilot mode to support e-government platforms, while healthcare pilots explore diagnostic support and triage tools. Education initiatives are testing adaptive learning systems and analytics for early dropout detection, and agriculture pilots use computer vision on satellite imagery for crop monitoring and land-use management. These applications reflect the government's focus on embedding AI into decision-making, service provision, and sectoral modernisation, positioning the country to move rapidly from experimentation to scaled implementation.

Despite these advances, gaps remain. Safeguards for fundamental rights such as privacy, non-discrimination, and freedom of expression are acknowledged only in principle, with no binding rules or operational mechanisms like mandatory impact assessments, algorithmic transparency requirements, or citizen redress. While the adoption of international standards provides a basis for voluntary compliance checks and pre-deployment assessments, these remain non-binding and have not yet evolved into enforceable oversight practices. Institutional arrangements also lack clarity, with no designated regulatory authority for AI supervision. Progress will further depend on cloud solutions and effective data sharing, supported by strong data governance, regulation of cloud usage, and improved data availability and quality. Finally, while the country has notable human capital in digital skills, sustained investment is needed to expand AI expertise and ensure sufficient compute capacity for research and deployment.

Thus, as of now, Azerbaijan has transitioned from uncertainty to a structured governance framework built on international standards, a national strategy, and an implementation roadmap. The next step is to consolidate these commitments into enforceable laws, clear institutions, and demonstrable practices that safeguard rights, strengthen compliance mechanisms, and build public trust in the responsible use of AI.

2.7.3 Bangladesh

Currently, Bangladesh is in the process of developing a national AI strategy or policy. The existing National Strategy for Artificial Intelligence, prepared in 2019-2020 with UNDP support, was never formally endorsed, reflecting "political distractions, bureaucratic inertia, and lack of executive ownership".[149] A subsequent draft National AI Policy 2024 was circulated and publicly discussed but has stalled in government approval process. Civil society actors criticised the process for excluding rights-based stakeholders and concentrating authority within government, raising concerns about legitimacy and accountability (Hoque, 2025).

With no dedicated budgetary commitments or a specialised authority, AI governance had been embedded in a broader ICT strategy – *Smart Bangladesh 2041*.[150] This vision, introduced by the previous government and now no longer active, emphasised ICT-led growth, but does

[149] *From paper to policy: My personal journey with Bangladesh's national AI strategy.* M. S. Alam (2025). https://shofiul.com/bangladeshs-national-ai-strategy/

[150] https://file-dhaka.portal.gov.bd/uploads/3cd0d9a5-2bbb-4f5e-9796-6a68880d403d/63f/706/211/63f7062119efb056903235.pdf

not establish enforceable safeguards like algorithmic accountability, citizen redress, or impact assessments.[151]

In practice, AI activity in Bangladesh is emerging through experimental projects and service delivery rather than formal regulation. The ICT Division,[152] under the Ministry of Posts, Telecommunications, and Information Technology, oversees initiatives such as a2i's iLab,[153] and the ICT Innovation Fund,[154] which support small-scale pilots in areas like natural language processing, education technology, and agriculture, including AI-driven crop advisory and disease detection tools.

The most visible deployment is the 333 National Help Desk, which has processed over 84 million citizen calls since its launch and. In 2023, it began piloting *SmartSathi*, an AI-powered bot for handling voice and text queries.[155] While such initiatives improve efficiency and accessibility in public services, they remain incremental and fragmented. Without an adopted national strategy and a dedicated body to guide implementation, Bangladesh's AI trajectory risks remaining project-driven, with limited scalability and insufficient attention to long-term ethical, legal, and accountability frameworks.

Launched in early 2024, *Ekush* LLM is the first large language model specifically trained on Bangla (Bengali), marking a significant milestone in Bangladesh's AI landscape. Developed by Intelsense AI, the model is designed to enhance natural language understanding and generation in Bangla, enabling more effective human-computer interactions in the language. Yet, it is still in the early stages of development and deployment.

2.7.4 Cambodia

Cambodia's AI landscape is rapidly evolving, and significant progress is being made. The *Cambodia Digital Economy and Society Policy Framework 2021–2035,*[156] and the *Cambodia Digital Government Policy 2022-2035* outline various policy measures and priority actions to expand the use of AI for data sharing,[157] enhance user-friendly digital services, develop policies and standards, and promote key digital skills. In addition, the *Cambodia Digital Skills Development Roadmap 2024-2035* aims to train 100,000 digital professionals within 10 years across 7 skill tracks,[158] including AI and Data Science. In July 2025, UNESCO, in collaboration with the Cambodia Academy of Digital Technology (CADT) and supported by the Ministry of Post and Telecommunications (MPTC), launched the *Cambodia AI Readiness Assessment Report* using UNESCO's Readiness Assessment Methodology (RAM).[159]

MPTC, with the mandate from the Royal Government of Cambodia to lead AI governance and AI backbone infrastructure, drafted the *National Artificial Intelligence Strategy 2025-2030.*[160]

[151] *The future can't wait: Why the budget falls short for ICT.* The Business Standard. M. R. Sohel (2025). https://www.tbsnews.net/thoughts/future-cant-wait-why-budget-falls-short-ict-1167536

[152] https://ictd.portal.gov.bd/

[153] https://a2i.portal.gov.bd/site/page/89b709df-c33a-489c-aff8-e69d5cd6ef1a/-

[154] http://ims.ictd.gov.bd/

[155] *333's SmartSathi to further simplify receiving govt services.* The Business Post (2023). https://businesspostbd.com/national/333s-smartsathi-to-further-simplify-receiving-govt-services

[156] https://go.gov.kh/dgc/cdespf-21-35-eng

[157] https://go.gov.kh/dgc/cdgp-en

[158] https://go.gov.kh/dgc/dsdr-24-35

[159] https://go.gov.kh/mptc/ramkh

[160] https://go.gov.kh/mptc/g8f

The draft strategy has been completed and put to multiple rounds of consultations, including a National Workshop in early July 2025. The final strategy is expected to be submitted to the Digital Government Committee and the National Digital Economy and Society Council later in 2025. The Draft NAIS also highlights the significance of AI governance by promoting ethical and responsible AI as one of its strategic measures. This is especially important given that civil society organisations have advocated for stronger democratic governance by promoting digital democracy, citizen engagement, and public accountability through dialogue on AI and governance.[161]

Cambodia's AI governance is at an emerging stage. MPTC is developing a complementary *National AI Governance Framework*, aligning with the *ASEAN Guide on AI Governance and Ethics,*[162] *and ASEAN Expanded Guide on AI Governance and Ethics – Generative AI.*[163] In addition, MPTC is working on several important documents: (i) *Draft Data Governance Policy 2025-2035;* (ii) *Draft Cloud-First Policy*; (iii) *Draft Law on Personal Data Protection*; and (4) *Draft Law on Cybersecurity*.

At both the international and regional levels concerning AI, MPTC represents the government in the ASEAN Working Group on AI Governance (WG-AI) under the ASEAN Digital Senior Officials' Meeting (ADGSOM). Lastly, MPTC is also a member of the Hiroshima AI Process Friends Group.

At the implementation level, AI is being deployed and piloted in finance, education, agriculture, logistics and transport, and health; though these efforts remain fragmented and small-scale. Without a comprehensive governance system and a clear institutional anchor, Cambodia risks a project-driven trajectory where oversight and ethical safeguards lag behind ambition.

2.7.5 Georgia

Georgia currently lacks a comprehensive national AI policy, although some sector-specific regulations have begun to emerge. In 2020, the National Bank of Georgia introduced a risk management framework for artificial intelligence and machine learning models in the financial sector, aiming to strengthen transparency and accountability in their application.[164]

However, in the absence of a unified legislative framework, significant challenges remain in ensuring the ethical deployment of AI across different domains. This gap is particularly visible in law enforcement. The Ministry of Internal Affairs operates more than 5,000 surveillance cameras, while legislation on facial recognition remains vague and insufficient. Inspections have already documented systemic violations in data processing, and the absence of activity logs increases the risk of misuse, including unauthorised access to personal data (Eristavi and Davituri, 2021).

The Government has supported the startup ecosystem primarily through the Georgian Innovation and Technology Agency (GITA),[165] a state body promoting innovation and entrepreneurship more broadly rather than focusing exclusively on AI. Through its grant programmes, GITA provided funding to companies such as Pulsar AI prior to its landmark acquisition by a U.S.

[161] https://www.cid.asia/digital-democracy-and-ai-for-citizens-engagement-and-good-governance-in-cambodia-november-21-2024/

[162] https://go.gov.kh/dgc/asean-guide-23

[163] https://go.gov.kh/mptc/expanded-guide-25

[164] *The National Bank of Georgia approves the regulation on "Data-Driven Statistical, Artificial Intelligence, and Machine Learning Model Risk Management".* https://nbg.gov.ge/en/media/news/the-national-bank-of-georgia-approves-a-new-regulation-to-establish-a-high-standard-of-ris

[165] https://gita.gov.ge/en

firm, illustrating how targeted public investment can help local ventures reach international markets. *Theneo*, also supported by GITA, has established themselves globally through AI-powered API management solutions, winning the Web Summit's 2022 Startup Competition. Other ventures, including Enagram, are advancing AI applications for the Georgian language, addressing barriers faced by low-resource languages in the digital sphere. Nevertheless, the initiatives operate within an environment marked by governance gaps that create uncertainties for businesses' AI solutions, coupled with shortages of skilled talent, a small domestic market, and limited institutional support, underscoring the need for a more robust national framework to sustain momentum.

Complementing these domestic dynamics, Georgia has also demonstrated commitment at the international level. In September 2024, it signed the Council of Europe's Framework Convention on Artificial Intelligence and human rights, democracy and the rule of law, a first-ever international legally binding treaty that anchors AI development in human rights, democracy, and the rule of law. Additionally, as a member of the UN family of nations, Georgia is engaged in all the latest UN instruments on the ethical application and regulation of AI. Furthermore, Georgia is a member of and remains actively engaged in AI-focused initiatives through the Freedom Online Coalition (FOC), a group of countries committed to the human rights and fundamental freedoms proclaimed in the Universal Declaration of Human Rights.

While this underscores Georgia's willingness to align with international norms, domestic governance mechanisms remain underdeveloped. The absence of a centralised oversight body and standardised regulations continues to hinder the creation of a cohesive AI governance framework, highlighting the urgent need for a comprehensive national AI strategy. Such a strategy should set out clear ethical guidelines, regulatory mechanisms, and institutional structures, while leveraging Georgia's strengths in digital reform, openness to innovation, and international partnerships to ensure the responsible and transparent use of AI technologies.

2.7.6 Kyrgyzstan

Kyrgyzstan has entered a decisive stage in shaping its artificial intelligence (AI) ecosystem. In 2025, the country adopted the landmark legal document, the *Digital Code,*[166] which consolidates regulation of data, telecommunications, and digital services and, for the first time, establishes the definition of artificial intelligence system. It will come into force in early 2026, providing the first overarching legal framework for regulating the digital environment, encompassing the use of artificial intelligence. The Digital Code introduces a risk-based AI framework in which non-hazardous systems face minimal regulation, while high-risk applications must undergo impact assessments and comply with requirements for safety, transparency, data quality, human oversight, and disclosure of significant outcomes. Compliance can be demonstrated flexibly through self-declaration, adherence to industry ethics codes, or international certification. The Code introduces a special regulatory mechanism, similar to regulatory sandboxes, that allows temporal exceptions from certain regulations to test innovative digital solutions under regulatory oversight, including for AI. This approach creates flexible conditions for experimentation while ensuring the protection of citizens' rights and adherence to safety principles.[167]

[166] The full text of the Digital Code of the Kyrgyz Republic is not yet directly available to the public. It was adopted by Parliament on 18 June 2025, and it was signed into legal force by the President on 4 August 2025. https://24.kg/english/338421_President_of_Kyrgyzstan_Sadyr_Japarov_signs_Digital_Code/.

[167] https://internetpolicy.kg/wp-content/uploads/2023/04/%D0%9A%D0%BE%D0%BD%D1%86%D0%B5%D0%BF%D1%86%D0%B8%D1%8F-%D0%A6%D0%9A_eng-1.pdf

In parallel, the Government has been preparing a *National AI Strategy*, which is currently undergoing exploration and analysis.[168] To coordinate these developments, the Cabinet of Ministers, established in 2025,[169] a National Council for the Development of AI, tasked with guiding policy, fostering cross-sectoral cooperation, and ensuring safe adoption of AI technologies. These measures build on the Digital Transformation Concept for 2024-2028),[170] which positions AI as a tool to modernise public services, expand new economic sectors, and improve quality of life. The Concept prioritises the creation of a national AI platform and computing centre, expansion of high-performance computing, and development of Kyrgyz-language technologies, including voice assistants and cultural heritage applications. In early 2025, the Chairman of the Cabinet of Ministers and Head of the Presidential Administration of the Kyrgyz Republic confirmed plans to procure a supercomputer to strengthen computational resources, particularly for natural language processing.

Survey data shows that AI is already being piloted in public service centres to optimise administrative processes. Intelligent chatbots are being designed to assist citizens with routine government service queries, providing faster and more accessible interactions with state institutions. These applications illustrate the Government's pragmatic focus on service delivery while laying the groundwork for broader adoption across the public sector. However, several barriers remain. Open data access is limited, datasets are fragmented and uneven in quality, and advanced hardware costs remain prohibitive. Challenges related to the internet infrastructure and stability also remain (Eferin et al., 2025).

Alongside state initiatives, Kyrgyzstan is strengthening its international engagement. The country participates in regional dialogue, including the Turkic States AI Summit, and cooperates with organisations such as UNDP to advance responsible AI governance. These partnerships help raise awareness of international standards, notably the UNESCO Recommendation on the Ethics of AI and the OECD AI Principles, while reinforcing domestic capacity.

The private sector and professional associations are also shaping the emerging AI landscape. The AI Developers Association of Kyrgyzstan has drafted a Code of Ethics to encourage responsible AI use, though it is not yet officially adopted by the government.[171] Businesses in Kyrgyzstan are actively experimenting with practical AI applications. For instance, MBank has become one of the first in the country to deploy AI-powered customer verification and automation systems, with its app achieving over 10 million downloads and reaching roughly 80% of the adult population.[172] Meanwhile, local developers, most notably the Cramer Project, have created *AkylAI*, the first AI assistant to understand and speak Kyrgyz, with functional web chat and prototype smart speakers in development.[173] Furthermore, the startup *NineNineSix* developed *KaniTTS*, a Kyrgyz-language speech synthesis model that rivals the quality of solutions from OpenAI. Google, and Microsoft.[174]

[168] *Kyrgyzstan to adopt National Strategy on Artificial Intelligence soon.* (Trend News Agency, 31 January 2025). https://en.trend.az/casia/kyrgyzstan/4000012.html

[169] Cabinet of Ministers of the Kyrgyz Republic. (2025). *Первое заседание Национального совета по вопросам развития искусственного интеллекта прошло* [The first meeting of the National Council on the development of artificial intelligence took place]. https://www.gov.kg/ru/post/s/24926-zasalma-intellektti-onukturuu-maseleleri-boyunca-uluttuk-kenestin-birinci-oturumu-bolup-ottu

[170] Presidential Decree of the Kyrgyz Republic No. 90/2024. *Concept of digital transformation of the Kyrgyz Republic for 2024-2028.* https://digital.gov.kg/wp-content/uploads/2024/06/concept-digital-transformation-of-the-kyrgyz-republic-for-2024-2028.pdf

[171] https://www.aida-ai.org/

[172] *MBANK leads in KYC and digital innovation.* (Press release, Financial Content, 2025). https://markets.financialcontent.com/ms.intelvalue/article/marketersmedia-2025-2-22-mbank-leads-in-kyc-and-digital-innovation

[173] *История AkylAI – как создавали первый кыргызоязычный искусственный интеллект* [The story of AkylAI – how the first Kyrgyz-language artificial intelligence was created]. Economist.kg. Akhmatova, I. (2023). https://economist.kg/novosti/2023/10/30/istoriia-akylai-kak-sozdavali-piervyi-kyrghyzoiazychnyi-iskusstviennyi-intielliekt/

[174] The model has gained international recognition, with more than 15,000 downloads on the Hugging Face platform.

The National Commission on State Language and Language Policy under the President of the Kyrgyz Republic created the National Platform *AiTil*, the country's first state-owned natural language processing (NLP) system for the Kyrgyz language. The platform integrates modules for speech synthesis and recognition, machine translation, document processing, and a generative chat assistant trained on Kyrgyz texts.[175] This comprehensive solution strengthens digital support for the Kyrgyz language, and it is regarded as a strategic component of the nation's linguistic and technological sovereignty.

Despite this progress, the domestic AI market remains nascent. To ensure the sustainability of the AI ecosystem, universities are updating their educational programmes in machine learning, data engineering, MLOps, and data architecture. The practical components of training are being enhanced, while international internship opportunities are expanding, enabling graduates to integrate more effectively and rapidly into AI development projects but research capacity is weak, with few patents, limited publications, and scarce opportunities for advanced study (Eferin et al., 2025). The planned *National AI Platform*, encompassing a repository of models and data, AI accelerators, competence centres, and an AI Computing Centre, could serve as a catalyst for innovation by addressing infrastructure gaps and linking academic, entrepreneurial, and government efforts.

2.7.7 Lao PDR

Laos is laying the groundwork for artificial intelligence (AI) governance through emerging policies, institutional initiatives, and pilot projects, though a comprehensive national strategy and dedicated legal framework are still in development. The regulatory basis remains indirect, anchored in the *Law on Electronic Data Protection (2017)*,[176] *E-Transactions Law (2022)*,[177] and *Cybercrime Law (2015)*,[178] which address aspects of data protection and digital security but leave gaps in accountability, transparency, and redress for AI-specific risks. Broader policy direction comes from the "*20-Year National Digital Economy Vision (2021–2040)"* and related digital development plans; survey respondents also reported a recently adopted *"10-Year Digital Government Strategy 2025–2035"* that is expected to set the enabling framework for AI adoption across government services.

The Ministry of Technology and Communications (MTC) is the Centre of gravity for national efforts. The MTC's Digital Government Centre implements e-government platforms and civil-service capacity building, including training on ethical AI use, data validation and dataset preparation. At the same time, Laos has pursued international cooperation to accelerate capabilities: in 2025 the Laos–China AI Innovation Cooperation Agreement was signed to establish a joint AI Innovation Centre in Vientiane for applied research and technology transfer,[179] and the National Data Centre (under MTC) signed an Memorandum of Understanding with Silicon Tech Park (Lao) to carry out a feasibility study for AI infrastructure powered by green energy, a potential PPP model for an AI Special Economic Zone.[180]

[175] Generative chat assistants are AI chatbots with with generative AI capabilities offering more enhanced functionality with their understanding of common language and complex queries, their ability to adapt to a user's style of conversation and use of empathy when answering users' questions.

[176] https://natlex.ilo.org/dyn/natlex2/r/natlex/fe/details?p3_isn=107518

[177] http://laopremier.com/e-transactions-law/#:~:text=The%20E%2DTransactions%20Law%20defines,Contract%20and%20Tort%20as%20well

[178] The Law on Prevention and Combating Cyber Crime (2015). https://dig.watch/resource/laos-law-on-prevention-and-combating-cyber-crime-2015

[179] https://english.news.cn/20250224/17c802d3fa214f068248e43fa805322d/c.html

[180] https://www.vientianetimes.org.la/freefreenews/freecontent_104_laossings_y25.php

In the public sector, AI is already showing practical value. With UNDP support,[181] the *Gov-X* one-stop mobile platform centralises citizen access to e-services: licenses, welfare checks, business registration, learning resources, creating an operational backbone for future AI enhancements. The Government has also been rolling out digital ID initiatives,[182] which, together with centralised citizen databases, can enable AI-driven identity verification and fraud detection once governance safeguards and data standards are in place. The Ministry of Finance and other line ministries have started feasibility work on data-driven AI approaches for revenue and service delivery.

International cooperation is central to progress. Laos aligns with ASEAN AI Governance and Ethics Guidelines and benefits from partnerships with China, UNESCO, and UNDP in areas of innovation, ethics, and rural inclusion. A 2025 capacity-building programme on AI governance underscored the country's need to balance experimentation with safeguards and to strengthen cross-sectoral coordination.[183]

Survey respondents confirm that Laos remains at a nascent stage of AI development. Priority needs include improving dataset quality, creating a Lao-language large language model (LLM), and developing a holistic AI roadmap to integrate infrastructure, governance, and skills development. Persistent barriers exist – including low digital literacy, low internet penetration in the rural areas, high infrastructure costs, and limited data centres, as well as the fragmentation of information and data that need to be converted from paper to digital format to be able to support AI development – which constrain scale and equity in adoption. These needs mirror the structural constraints documented in national assessments (infrastructure, skills, and coordination).[184]

The private sector and research ecosystem are small but emerging. Events such as Lao Digital Week and other public-private convenings are building dialogue and nascent startups;[185] universities are beginning to add data and AI topics. However, innovation finance, ICT workforce scale-up and research outputs remain limited. Advancing Laos's AI agenda will require continued public investment and PPPs for green, low-carbon AI infrastructure, stronger data governance and standards, targeted skills programmes for civil servants and researchers, and careful, phased regulation that balances experimentation with rights protection.

2.7.8 Mongolia

Overall, Mongolia's AI governance framework is being significantly shaped by key international collaborations. UNDP has played a pivotal role in strategy development and capacity building; the WEF collaboration underscores Mongolia's intention to strengthen AI-informed policymaking;[186] and the AI Futures foresight initiative illustrates emerging governance maturity.[187] Together, these partnerships reflect Mongolia's pragmatic blend of global expertise and domestic ambition.

[181] https://www.undp.org/laopdr/press-releases/gov-x-one-stop-mobile-application-public-e-services

[182] National digital ID rollout announced for 2024–25 (October 2025). https://laotiantimes.com/2025/08/29/laos-to-launch-digital-id-cards-nationwide-in-october-2025/

[183] https://laopdr.un.org/en/300785-remarks-capacity-building-programme-ai-governance

[184] https://thedocs.worldbank.org/en/doc/c01714a0bc2ca257bdfe8f3f75a64adc-0070062022/original/Positioning-The-Lao-PDR-for-a-Digital-Future-11-10-22.pdf

[185] https://www.aseanaccess.com/news/1616-lao-digital-week-2025.html

[186] *Mongolia and the World Economic Forum to cooperate on developing "Strategic Intelligence Platform"*. Montsame News Agency (2025). https://montsame.mn/en/read/360135

[187] *Anticipating Tomorrow(s): Why Mongolia Needs Futures Thinking*. https://www.undp.org/mongolia/blog/anticipating-tomorrows-why-mongolia-needs-futures-thinking; and *AI and Mongolia: A chance to close gaps, not widen them*. https://www.undp.org/mongolia/blog/ai-and-mongolia-chance-close-gaps-not-widen-them. UNDP Mongolia (2025).

Key policy documents, such as Mongolia's long-term development policy "Vision 2050", the 2021-2030 action plan for implementing "Vision-2050", the Government of Mongolia's 2024-2028 Action Programme, and the "Digital Nation" policy, incorporate 49 specific provisions to accelerate sectoral development through AI solutions. These provisions represent a critical step in defining Mongolia's long-term vision for AI before the formalisation of a national strategy.

The Ministry of Digital Development, Innovation and Communications (MDDIC) developed a draft National Strategy for Big Data and Artificial Intelligence in 2025, conducting a 10-month consultation process across sectors to incorporate stakeholder feedback. On 24 September 2025, the Government of Mongolia officially adopted the "*National Strategy for Big Data and Artificial Intelligence*".[188]

The Strategy outlines 66 projects and initiatives to be implemented in two phases from 2026 to 2030, based on four key pillars: Pillar 1: Integrate national big data into the economic cycle to create "The Knowledge Yurt"; Pillar 2: Prioritise ethical and responsible AI development to ensure digital balance; Pillar 3: Enhance productivity in priority socio-economic sectors using AI to align with "The Earth's Pulse"; and Pillar 4: Compete in the global AI infrastructure market to establish a "Digital Data Node Centre".

Implementation of the Strategy focuses on building legal and institutional foundations, including the establishment of a National Council on Artificial Intelligence, the creation of a GPU cluster-based AI Centre, and the launch of a National Big Data Depository.[189] Implementation also focuses on: (i) the provision of training to 300 masters and doctoral studies in AI through scholarship programmes; (ii) the introduction of AI curricula in 500 general education and 70 vocational and technical institution; (iii) the training of 60,000 public servants in AI skills; (iv) supporting the development of 100 AI research and development projects covering priority economic sectors; (v) promoting the export of 10 AI-based products and services in sectors such as mining, fintech, and renewable energy.

In September 2025, the Government of Mongolia also approved the "DIGITAL FIRST" policy recommendation.[190] The "DIGITAL FIRST" policy emphasises prioritising digital technologies at all levels of government decision-making, policy formulation, and service delivery, moving away from traditional, paper-based, and multi-step processes. By adhering to this policy across all government operations, Mongolia aims to transform the Big Data and AI Strategy into tangible outcomes, achieving its set objectives.

Legal foundations for data and privacy are also in place and positioned as the basis for AI accountability. Mongolia's *Personal Data Protection Law* was adopted in 2021, creating modern data-handling rules that underpin algorithmic safeguards (PwC, 2022). Regulators are also experimenting with innovation-friendly tools such as regulatory sandboxes (already used in financial services),[191] and proposals for tax incentives and cybersecurity strengthening aim to balance experimentation with consumer protection. Ethical issues, including risks of AI-driven disinformation, are explicitly on the policy agenda, with multilateral partners, e.g. UNESCO, advising on media and trust-related reforms (Singhal, 2025).

[188] *National Strategy on Big Data and Artificial Intelligence presented*. Montsame News Agency (2025).

[189] The MDDC announced the establishment of an AI infrastructure GPU cluster, while the private startup Fibo Cloud announced plans to build a foreign-financed GPU cluster centre in 2026.

[190] With Prime Minister G. Zandanshatar declaring 2026 as a year to accelerate digital transformation and IT development during the Tech-week 2025 event.

[191] Regulation on Sandbox Environment (2021). https://www.frc.mn/resource/frc/Document/2023/01/10/c0zszjpdskfvfyn3/Regulation%20on%20Sandbox%20Environment.pdf

Mongolia aims to leverage its unique strengths to develop a big data and AI ecosystem, addressing critical social, economic, and environmental challenges. The Strategy sets ambitious goals to position Mongolia as a competitive player in the AI domain while fostering ethical and responsible AI. Mongolia aspires to rank among the top 10 AI-ready countries in Asia and the top 50 globally in the AI Readiness Index by 2030.

Practical use cases and local innovation are beginning to take shape. Launched in 2020 with the digitisation of 181 reference and certificate services from 21 government institutions, the E-Mongolia system has evolved into a platform offering 1,268 services from 87 government entities. Now in its fifth improved version, it has become Mongolia's most widely used platform, with 2.1 million users. The last two versions introduced AI-based solutions, including algorithms that recommend subsequent services based on users' service history and an AI assistant designed to simplify access to government services.[192]

A notable public-private partnership is Chimege Systems, a Mongolian AI company developing speech-to-text and text-to-speech technologies for applications such as virtual assistants, voice-controlled kiosks, and government services. The Chimege Writer system, used in court recordings, is now being integrated into the e-Mongolia platform to enhance accessibility for users with disabilities.

The growth and adoption of E-Mongolia have fostered a culture of digitalisation and ambition across all levels of government. However, the AI trends, hype, and advancements over the past three years have, on one hand, created a bubble of expectations and, on the other, driven demand for innovative products, services, and solutions. This has significantly raised expectations among society, the public, citizens, and organisations. To meet this high demand and fulfil expectations, Mongolia must make its data openly accessible to relevant stakeholders while establishing clear regulations for data collection, classification, and protection of sensitive data. In a nutshell, AI can only fulfil its potential when big data is readily available and properly managed.

While budget constraints reported by survey participants limit AI development, structural factors both enable and constrain Mongolia's AI trajectory. The country's youthful and highly literate population provides a strong base for AI adoption, yet rural–urban disparities in connectivity, gaps between educational outcomes and employer needs, limited venture financing, and constrained high-performance compute capacity remain significant barriers to inclusive advancement. So do the concerns from society and tech companies regarding potential risk, such as inconsistent government policies, unstable support, and insufficient funding. These challenges must be addressed to ensure successful execution.

In 2024, Mongolia advanced five places in the University of Oxford's "AI Readiness Index," ranking 98th out of 188 countries, demonstrating progress in artificial intelligence (AI). This achievement is largely attributed to years of research and development by private sector entities such as Chimege Systems and Global, which have successfully introduced AI-based products and services tailored to Mongolia's unique characteristics. Notably, the introduction of a Mongolian-language Large Language Model (LLM) has laid the foundation for offering products like text-to-speech and speech-to-text in the native language across various sectors.

[192] Government of Mongolia introduces AI-based chatbot. Montsame News Agency (2024). https://www.montsame.mn/en/read/354863

2.7.9 Tajikistan

Tajikistan broke new ground in 2022 by adopting its "National AI Strategy to 2040" – the region's first long-term strategy, making it both the earliest Central Asian state and the first low-income country worldwide to commit to a dedicated AI policy framework.[193] Framed as a conceptual document, the Strategy sets ambitious targets for economic growth, public sector modernisation, and the integration of AI into education and healthcare, but it lacks concrete projects, timelines, and implementation mechanisms, and as a result, remains largely aspirational three years after adoption. Even so, the move highlights how resource-constrained governments are beginning to institutionalise AI governance, using high-level strategies to signal ambition, position themselves as forward-looking partners, and shape narratives that can facilitate donor engagement and investment.

Oversight rests with the Agency for Innovation and Digital Technologies under the President, which centralises coordination of policy, IT-park development, and international cooperation. The strength of this model lies in decisive agenda-setting, though the concentration of authority raises questions about accountability and inclusiveness. Legally, the framework is anchored in the Law on Personal Data, which regulates the collection and use of personal and biometric data. While this provides a baseline, enforcement mechanisms remain weak and AI-specific regulations are not in place, creating uncertainty for innovators and partners.

Implementation of AI in Tajikistan remains at an early stage but is gradually taking tangible form. The Government relies heavily on donor-financed programmes, with the World Bank's *Digital Foundations* project and UN-backed education initiatives strengthening digital public infrastructure and training civil servants.[194] E-government platforms are beginning to experiment with AI-assisted features, particularly in administrative automation and citizen service delivery, though most remain at the pilot stage. In July 2025, Tajikistan launched *SoroLLM*,[195] its first publicly available large language model tailored to the Tajik language and dialects, developed by zehnlab.ai with government support. *SoroLLM* provides speech recognition, text generation, and translation capabilities, marking the country's first operational AI language model and illustrating progress toward implementing the National AI Strategy.

A notable domestic initiative is *TajRupt*,[196] a non-profit AI laboratory established in 2022 with support from UNDP and other partners. *TajRupt* focuses on training and youth empowerment, offering supervised machine learning curricula, Python programming, statistics, and AI entrepreneurship courses to students in Dushanbe and Khujand. Complementing these public-sector and nonprofit initiatives, Zypl.ai,[197] a Tajik fintech startup, collaborates across Central Asia, partnering with banks and payment systems in Uzbekistan and Qazaqstan, and has established a strategic AI-focused partnership in the UAE with the Emirate of Ras Al Khaimah.

Despite these advances, workforce development remains the most critical bottleneck. Training initiatives exist, but Tajikistan faces a shortage of AI professionals and continued brain drain. Infrastructure constraints, limited funding, and the absence of AI-specific regulations further slow the translation of strategy into large-scale deployment. Collectively, these developments show

193 https://dig.watch/resource/strategy-for-the-development-of-artificial-intelligence-in-the-republic-of-tajikistan-for-the-period-up-to-2040

194 https://www.worldbank.org/en/news/press-release/2024/12/17/world-bank-to-strengthen-digital-public-infrastructure-and-digital-skills-in-tajikistan

195 https://asiaplustj.info/en/news/tajikistan/society/20250707/tajikistan-unveils-first-national-ai-language-model-sorollm

196 https://innovation.eurasia.undp.org/tajik-youth-driving-ai-assisted-development

197 https://zypl.ai/blog/zypl-ai-spearheading-ai-driven-financial-innovation-and-regional-integration-in-central-asia

that Tajikistan is moving from aspirational policy toward practical implementation, combining civil society education, domestic innovation, pilot projects, and international collaboration. Scaling these efforts successfully will require addressing human, legal, and technological capacity gaps.

Internationally, Tajikistan has positioned itself as a diplomatic first mover, advancing a UN General Assembly resolution on AI and proposing the establishment of a Regional AI Centre in Dushanbe.[198] This complements domestic reforms and signals regional ambitions but also heightens the need for credible, enforceable domestic rules. Overall, Tajikistan's trajectory reflects a mix of bold vision and fragile implementation capacity. Immediate opportunities lie in leveraging donor support to scale low-risk AI pilots in government services, operationalising secondary regulations under the data protection law, and developing a sustainable workforce pipeline. Progress in these areas will determine whether Tajikistan can convert its early policy move into tangible AI adoption and regional influence.

2.7.10 Uzbekistan

Uzbekistan was among the first countries in Central Asia to begin regulating artificial intelligence, linking early policy measures to its broader "*Digital Uzbekistan – 2030"* agenda. In 2021, the Government adopted a resolution to accelerate AI adoption and mandated the creation of institutional capacity.[199] Subsequently, the former Ministry of Information Technologies and Communications subsequently established the Digital Technologies and Artificial Intelligence Research Institute, tasked with supporting applied research and coordinating national projects. In August of the same year, authorities launched an AI regulatory sandbox at the IT Park,[200] giving domestic developers a controlled environment to test and pilot AI-based solutions. Those early steps created a legal and institutional basis for pilots across multiple sectors, but UNDP's research finds that the sandbox regime has not yet delivered the expected outcomes,[201] and many pilot projects remain incomplete or uncertain, prompting calls to design a more effective, functional testing and deployment mechanism.

Building on these foundations, Uzbekistan formalised its long-term vision by approving the *AI Strategy* until 2030 in October 2024.[202] Unlike many regional peers, the Strategy provides detailed mechanisms, financing plans, and timelines, including the creation of a dedicated AI development centre, rollout of sectoral projects in agriculture, tax and customs administration, banking, energy, and healthcare, as well as investment in institutional and workforce capacity.

In 2022, the Government created the Advisory Council on Artificial Intelligence under the Ministry of Innovative Development, which brings together government, academia, business, and civil society. Its role is to review pilot projects, provide policy advice, and ensure AI aligns with human-centred values. While the Council has increased multi-stakeholder participation, most of its work remains consultative, and systematic impact assessments for AI systems are not yet in place. Ethical oversight is guided by general data protection and cybersecurity laws rather than dedicated AI legislation, leaving room for stronger accountability frameworks.

198 https://timesca.com/un-adopts-ai-resolution-for-central-asia-at-tajikistans-initiative/

199 Resolution of the President of the Republic of Uzbekistan PP-4996/17.02.2021 "On Measures to Create Conditions for the Accelerated Implementation of Artificial Intelligence Technologies".

200 Decree PQ-5234 "On Measures to Introduce a Special Regime for the Use of Artificial Intelligence Technologies".

201 https://www.undp.org/sites/g/files/zskgke326/files/2025-05/uz_digital-economy-study_eng.pdf

202 Resolution of the President of the Republic of Uzbekistan RP-358/14.10.2024 "On the Approval of the Strategy for the Development of Artificial Intelligence Technologies until 2030". https://lex.uz/en/docs/7159258

Infrastructure and data capacity remain the main technical constraints. Uzbekistan has rapidly expanded fibre-optic networks and mobile base stations and built new data centres, including a TIER-3 facility for e-government, yet it still lacks national supercomputing capacity sufficient for large-scale model training. In response, the Government has secured foreign investment commitments (e.g., a Data-Volt project announced in 2024 to build "green" supercomputing data centres, with plans for modular centres and NVIDIA GPUs to operate high-performance computing resources to support AI workloads by 2026. Even so, access to curated, high-quality datasets is limited: the government open data portal (data.egov.uz) has expanded but the number of AI-ready datasets remains small and practical data access for developers is constrained.

Another priority is the development of Uzbek-language datasets to enable machine learning applications, marking the country's first structured effort to support natural language processing, though Uzbekistan has not yet set out ambitions for a large-scale language model. On the human resources side, capacity building is a declared priority: the Strategy places strong emphasis on education, AI-focused laboratories, international training and academia-industry links to cultivate talent. UNDP flags, however, that the public sector still suffers from a shortage of AI expertise, and that closing this skills gap, particularly among policymakers and civil-service implementers, is essential for the strategy to translate into public-sector adoption at scale.

Practical applications are already emerging across several domains. In public safety, AI-powered surveillance systems employing facial recognition and behavioural analysis enhance urban security. In healthcare, diagnostic tools leveraging AI are used to interpret medical images, including X-rays and MRIs, to detect conditions such as tuberculosis and cancer at earlier stages. The agricultural sector benefits from AI-driven monitoring of crop health through satellite imagery and sensor data, optimising yields and resource use. In education, adaptive platforms and AI chatbots are being introduced to personalise learning pathways, while the financial sector applies machine learning for fraud detection by analysing transaction patterns and anomalies. AI is also being piloted in smart traffic management systems to reduce congestion in major cities. More than 60 AI-related projects have been approved for 2025-2026, highlighting the breadth of experimentation underway.

Public–private partnerships play an active role, including cooperation with AIM Lab and the AI Competence Alliance (AICA), which link government institutions with local developers and international actors. Internationally, Uzbekistan engages in bilateral and multilateral cooperation, including within the CIS Interparliamentary Assembly on model laws, and works with UNDP and UNESCO to align with global ethical frameworks.

Authorities emphasise both practical utility and digital sovereignty, framing AI adoption as central to economic competitiveness and administrative efficiency. Taken together, these developments position Uzbekistan as a regional frontrunner, which has moved from experimental measures toward a comprehensive AI governance framework, backed by regulatory instruments, investment plans, and practical applications.[203] The challenge ahead lies in bridging the gap between ambitious targets and consistent implementation, particularly by addressing persistent skills shortages, ensuring access to high-quality data, and embedding Uzbek-language AI tools into domains that deliver tangible value for government, business, and society.

[203] Uzbekistan's progress is also reflected internationally. The country advanced from 87th place in 2023 to 70th in the Oxford Insights Government AI Readiness Index 2024, signalling gains in policy coherence, infrastructure, and institutional readiness.

3. Conclusions

Artificial intelligence is reshaping the global landscape, offering transformative opportunities for sustainable development, innovation, and public sector efficiency, while also posing complex challenges related to ethics, equity, and institutional readiness. Furthermore, as AI evolves faster than government policy and regulatory systems can adapt, an urgent need for globally coherent, inclusive, and future-proof governance is apparent. It is encouraging to observe that common approaches are emerging in regulating and overseeing AI development and applications. The evolution of several major international AI governance initiatives reveals a notable attempt to establish a common ground at the international, and by extension, at the national level.

However, different countries adapt different approaches in exploring, formulating, or implementing policy and regulations that reflect their cultural, legal and political contexts.[204] The first approach is principles-based anchored in the *OECD AI Principles* adopted in 2019. These principles offer flexibility and adaptability, based on broad guidelines that evolve with technological change. Conversely, this approach is voluntary, thus countries and organisations may selectively follow or ignore the principles, prioritising profits over ethics, potentially causing harm. Furthermore, these broad principles lack the specificity needed in addressing complex technical and legal challenges such as privacy breaches, bias in algorithms and accountability in autonomous systems. Thus, AI frameworks should be more precise in order to enforce clear, preventive compliance rules that could assist in managing AI-related risks, safeguard public interest and build trust in AI technologies (UN, 2025).

The second approach is risk-based. It assumes that AI systems often function as black boxes with little indication of what is happening inside. Thus, it strives to identify and mitigate potential harms before AI technologies are deployed. For instance, the European Union AI Act classifies AI applications by levels of risk, namely, unacceptable, high, limited and minimal. High-risk applications, such as biometric identification, involve strict regulations aimed at preventing harm before applications reach the market.[205] This approach attempts to address the complexity and unpredictability of AI systems. With the application of pre-emptive regulation, organisations can only deploy AI systems that meet compliance standards. In this manner, safety and ethics standards are encouraged from the outset, reducing potential collective harm. There is also emerging a third approach based on liability, an approach that promotes fairness and predictability by applying uniform rules and standards. By holding developers and deployers accountable for their AI systems, this approach encourages organisations to prioritise safety, reliability, and ethics from the outset. This can ensure more trustworthy and robust AI.

All at once, the global AI governance landscape exposes several critical gaps hindering effective implementation worldwide. First, geographical representation remains severely limited, with 118 countries, primarily from the Global South, excluded from major international AI governance initiatives. This exclusion creates a fundamental legitimacy problem for frameworks claiming global applicability (UN, 2025). Second, the voluntary nature of international agreements limits enforcement mechanisms. While, for instance, GPAI provides valuable platforms for

[204] Some jurisdictions, such as the European Union, have enacted comprehensive, enforceable regulatory frameworks; others – Canada, South Korea, and the United Kingdom – prioritise adaptive, sector-specific guidelines, or non-binding principles to stimulate experimentation alongside oversight.

[205] However, risk-based approaches are also not without limitations. For instance, categorising AI technologies can be highly subjective and challenging, particularly self-modifying AI systems that evolve over time.

dialogue and knowledge sharing, the absence of binding commitments reduces incentives for comprehensive implementation. Countries can selectively adopt elements of international frameworks while avoiding more challenging requirements. And third, coordination across multiple initiatives creates confusion and duplication. The proliferation of AI governance frameworks from different organisations and regional bodies results in competing standards and conflicting requirements. This fragmentation particularly affects multinational organisations attempting to ensure compliance across diverse regulatory environments.

Indeed, international and multilateral organisations are playing an important role in shaping global AI governance, acknowledging the far-reaching implications for human rights, development, and security. Efforts by the United Nations, UNESCO, OECD, and others increasingly define ethical norms and policy guidance that transcend national borders, encouraging responsible AI use. For instance, countries such as the Philippines, Bangladesh, Cambodia, Laos, Uzbekistan, and Georgia leverage UNESCO's Readiness Assessment Methodology to evaluate their AI preparedness and inform strategic policymaking, exemplifying structured, evidence-based governance aligned with global standards.

Despite these noteworthy efforts AI regulation remains complex, and the involvement of multiple stakeholders and the existence of varied geopolitical interests make a single global framework unlikely. Regulatory action is expected to stay rooted at the national level, with international efforts offering supplementary guidance rather than unified solutions. Still, foundational principles, such as those articulated by the OECD, are converging, emphasising human rights, accountability, and a human-centric approach to AI design and deployment.

On the other hand, successful AI integration also entails a considerable capacity building effort across several equally critical aspects. Current capacity building initiatives reveal both promise and limitations in supporting AI governance, primarily in developing countries. For instance, the UN Resolution A/78/311, led by China and co-sponsored by over 140 countries, demonstrates significant interest in AI capacity building from the Global South.[206] The Plan articulates a comprehensive approach addressing infrastructure, industry development, talent cultivation, and safety governance. However, capacity building efforts remain inadequate relative to identified needs. Research indicates that developing countries face critical gaps in digital infrastructure, legal expertise, technical capabilities, and institutional frameworks for AI oversight.

The Government AI Readiness Index reveals that most developing countries score below global averages in the technology sector capabilities, and the data infrastructure pillars (UNIDO, 2024). Priority capacity gaps include inadequate computational resources, shortage of qualified personnel in AI law and regulation, limited capacity for AI system evaluation and monitoring, and weak governance structures for AI oversight. Addressing these gaps requires substantial investment in education, infrastructure, and institutional development that exceeds, however, current international assistance levels.

In general, countries operationalise AI through diverse governance models, tailored to their circumstances and readiness. Similarly, implementation of AI policies varies significantly across countries reflecting the differences in strategy, capacity and resources. Mature AI ecosystems such as Canada, Korea and the UK, feature strong multi-stakeholder collaboration, agile risk-based regulation, regulatory sandboxes, and substantial public investment in sovereign

[206] International Research Centre for AI Ethics and Governance. *Artificial Intelligence Capacity-Building Action Plan for good and for all to support global AI development and governance*. International Research Center for AI Ethics and Governance. https://ai-ethics-and-governance.institute/2024/09/26/artificial-intelligence-capacity-building-action-plan-for-good-and-for-all-to-support-global-ai-development-and-governance/

AI infrastructure and talent. They also exemplify advanced AI ecosystems, combining research excellence, regulatory foresight, industry engagement, and operational safeguards – from dedicated AI agencies to mandatory impact assessments and investment in computing. Furthermore, their strong ethical, human-centric, and collaborative approaches contrast with the nascent, fragmented, or less formalised frameworks observed in many Project countries.

Conversely, Project beneficiary countries focus on emerging strategies, policy experimentation, and capacity building, yet often face fragmented governance, insufficient infrastructure, and enforcement challenges. Across Central Asia, the Caucasus, Asia-Pacific, and Mongolia, AI adoption levels diverge. Qazaqstan and Uzbekistan are developing robust infrastructure and policy regimes, aspiring to regional leadership. By contrast, Tajikistan, Lao PDR, and Cambodia show limited AI activity, while Bangladesh and the Philippines are using AI to support public service delivery.[207] Armenia and Azerbaijan focus on supercomputing and data solutions, while Mongolia, Kyrgyzstan, Uzbekistan, and the Philippines target public services and economic sectors. Only Qazaqstan and Uzbekistan have launched special regulatory sandboxes to spur innovation.

Tajikistan was early to set an AI strategy, its framework remains general, compared to the clearer, industry-aligned approaches of Qazaqstan, Azerbaijan, and Uzbekistan, who created more detailed frameworks later. Georgia and Mongolia are completing their strategies, with an eye towards global AI integration. AI leaders tend to adopt product-centric models, focusing on building AI products and solutions, while Central Asian, Caucasus, and Mongolian countries are more regulation-focused, with underdeveloped practical deployment.

Persistent structural constraints hinder regional integration into the global AI ecosystem: technological dependence, infrastructure gaps, skill shortages, financial limitations, insufficient datasets, and inconsistent digitalisation. Enhanced interstate cooperation and international partnerships are needed. Regional bodies, such as the Commonwealth of Independent States, Organisation of Turkic States, and ASEAN, could accelerate technology transfer, capacity building, and joint initiatives, exemplified by the Qazaqstan-Uzbekistan AI Technology Centre launched in 2024.

In sum, benchmark countries like Canada, Korea, and the UK showcase the path from policy to implementation: establishing AI agencies, setting operational mandates, building sovereign compute resources, and deploying red-teaming for safety. Canada's CIFAR strategy, Korea's AI Basic Act, and the UK's AI regulatory initiatives demonstrate maturity, especially with their ethical, transparent, and collaborative frameworks. By contrast, project countries remain in initial phases, highlighted by fragmented coordination in the Philippines and the more centralised, but oversight-dependent AI programmes in Qazaqstan and Uzbekistan.

Globally, governments are shifting from blanket bans or voluntary codes toward risk-based classification, operational mitigation through impact assessments and conformity checks, and use of sandboxes and pilot certifications. Ethical considerations, once merely aspirational, are now increasingly enforced, becoming the backbone of governance alongside robust compute and data capacity. While benchmarks have made major investments in supercomputers and national clouds, project countries, including the Philippines, Uzbekistan, and Qazaqstan, continue to address gaps through public–private partnerships.

[207] The Philippines' *eGovPH* platform typifies this, integrating AI tools – from chatbots to identity verification across government services.

Successful national AI strategies will be measured by more than GDP or model output; they depend on whether states can align AI with democratic governance, social equity, and sustainable growth. Nations that embed ethics in infrastructure, integrate innovation with public trust, and pursue global cooperation, while protecting sovereignty, are best positioned to thrive. The future of AI governance will not be a binary choice between innovation and regulation, but a convergence toward systems that are globally interoperable and nationally tailored, driving transformation that is both technologically advanced and socially inclusive.

To close the gap, project countries should prioritise a pragmatic sequence: adopting enforceable safeguards (like mandatory impact assessments for public sector AI), investing in regional compute and data commons, and leveraging public procurement to drive standards. By targeting public-sector applications with clear social benefits, such as in agriculture, education, and administration, governments can build trust, catalyse markets, and create lasting institutional capacity for inclusive AI transformation.

Table 7. AI governance status and implementation in the project countries

Country	Strategy	Laws	Responsible Agency	Super-computing	LLM
Armenia			Ministry of High-Tech Industry	Expected in 2026	
Azerbaijan	AI Strategy for 2025-2028		Ministry of Digital Development and Transport	Under completion, with a computing power of 1.2 PFLOPS	A prototype
Bangladesh	Draft National AI Policy		Ministry of Posts, Telecommunications, and Information Technology		Ekush LLM
Cambodia	Draft National Artificial Intelligence Strategy 2025-2030		Ministry of Post and Telecommunications		In progress
Georgia			Ministry of Economy and Sustainable Development (through its Innovation and Technology Agency)		In progress
Qazaqstan	AI Concept for 2024-2029	AI Bill in its 1st reading	Ministry of Artificial Intelligence and Digital Development	Launched in 2025, with a speed of up to 2 exaflops	Qaz-LLM

Kyrgyzstan	Digital Code 2025 will come into force in 2026; Draft National AI Strategy	Sub-legislation under development	Ministry of Digital Development and Innovation Technologies; National Council for the Development of AI	Expected	AkylAI
Lao PDR	On-going drafting of the National AI Strategy		Ministry of Technology and Communications		Underway Lao AI (Lao LLM) is under development
Mongolia	Draft National Strategy for Big Data and Artificial Intelligence under review		Ministry of Digital Development, Innovation and Communications		Egune LLM
Philippines	NAISR 2.0, 2024		DTI, DOST, DICT, NEDA	Super Jojie, with a computing power of 1.2 PFLOPS	FiLLM, iTanong
Tajikistan	National AI Strategy to 2040		Agency for Innovation and Digital Technologies		SoroLLM
Uzbekistan	AI Strategy until 2030	AI Bill past its 1st reading	Ministry of Digital Technologies	Expected in 2026	In progress

4. Policy Recommendations for Ethical, Inclusive, and Effective AI Governance

To harness AI's promise while safeguarding societal values, national strategies should prioritise:

- Instituting risk-based and adaptive legal frameworks that clearly define governance responsibilities, enforce accountability, and allow agile response to emerging challenges;
- Building sovereign digital infrastructure including shared compute resources and curated data ecosystems to reduce foreign dependency and support local AI innovation tailored to national languages and needs;
- Facilitating multi-sector, multi-stakeholder governance platforms integrating government leadership, private innovators, academia, civil society, and marginalised communities to ensure inclusive policy design and oversight;
- Scaling capacity-building across civil service and society to raise AI literacy, embed ethical awareness, and foster digital skills aligned with evolving AI ecosystems;
- Leveraging public procurement, transparent AI impact assessments, and open data policies as strategic tools to promote responsible AI use and market incentives;
- Advancing regional and international cooperation to harmonise standards, share best practices, develop interoperable legal frameworks, and mitigate geopolitical tensions;
- Embedding ethical principles, including fairness, transparency, privacy, human dignity, and sustainability, at every stage of the AI lifecycle to build trust and social acceptability.

5. Future Perspectives for International Collaboration

AI governance necessitates a renewed global partnership reflecting shared responsibilities and diverse contexts. Future collaboration should be grounded in solidarity, transparency, and practical cooperation, aiming to:

- Develop interoperable governance frameworks anchored in shared principles but flexible for local adaptation;
- Facilitate equitable access to AI infrastructure, skills, and markets via global capacity-building programmes and technology transfer;
- Establish inclusive multilateral platforms for dialogue, norm-setting, and joint research on AI safety and ethics;
- Promote open data flows and responsible cross-border data governance that respects sovereignty and human rights;
- Coordinate rapid response mechanisms to emerging AI risks, safeguards against misuse, and evolving ethical dilemmas.

Such collaboration will help leverage AI as a global public good, advancing sustainable development goals, equity, and democracy, while managing the risks intrinsic to complex, autonomous technologies.

6. References

Alliance-Can (2024). *Pan-Canadian AI Compute Environment (PAICE).* https://alliancecan.ca/en/services/advanced-research-computing/pan-canadian-ai-compute-environment-paice

Akhmetkali, A. (2025). *Qazaqstan and South Korea explore AI use in public sector*. The Astana Times (21.04.2025). https://astanatimes.com/2025/04/Qazaqstan-and-south-korea-explore-ai-use-in-public-sector/

ASEAN Secretariat (2025). Expanded ASEAN Guide on AI Governance and Ethics – Generative AI. https://asean.org/wp-content/uploads/2025/01/Expanded-ASEAN-Guide-on-AI-Governance-and-Ethics-Generative-AI.pdf

Bakpaeva, A. (2025). *ИИ, стартапы, роботы и госуслуги будущего: чем удивит новый центр Alem.AI в Астане* [AI, startups, robots and future public services: what will surprise at the new Alem.AI Centre in Astana]. BAQ.kz. 20 June 2025. https://rus.baq.kz/ii-startapy-roboty-i-gosuslugi-buduschego-chem-udivit-novyy-tsentr-alemai-v-astane_300016221/

Blair Attard-Frost, Ana Brandusescu, and Kelly Lyons (2024). The governance of artificial intelligence in Canada: Findings and opportunities from a review of 84 AI governance initiatives. *Government Information Quarterly Journal*, Volume 41, Issue 2, June 2024, 101929. Elsevier. https://www.sciencedirect.com/science/article/pii/S0740624X24000212

Brkic, B. (2025). *The most powerful supercomputer in Central Asia launches in Qazaqstan in bid for AI boost*. Euronews. 20 July 2025. https://www.euronews.com/next/2025/07/20/the-most-powerful-supercomputer-in-central-asia-launches-in-Qazaqstanin-bid-for-ai-boost

Chua, L., Tham, P., and Okolo, C. T. (2025). AI safety governance, the Southeast Asian way. Brookings Institution. https://www.brookings.edu/wp-content/uploads/2025/08/GS_08252025_AISA_report.pdf

CIFAR (2025). *The Pan-Canadian AI Strategy.* https://cifar.ca/ai/?utm_source=chatgpt.com

CIFAR (2024). *Impact Report 2023 – 2024.* https://cifar.ca/impact-report-2023-2024

CIFAR (2022). *CIFAR announces plans for second phase of the Pan-Canadian Artificial Intelligence Strategy.* https://cifar.ca/cifarnews/2022/06/22/cifar-announces-plans-for-second-phase-of-the-pan-canadian-artificial-intelligence-strategy

CIFAR (2019). *Annual Report of the CIFAR: Pan-Canadian AI Strategy.* https://cifar.ca/wp-content/uploads/2020/04/ai_annualreport2019_web.pdf

DSIT (2024). A pro-innovation approach to AI regulation: Government Response. Department for Science, Innovation and Technology, GOV.UK. https://www.gov.uk/government/consultations/ai-regulation-a-pro-innovation-approach-policy-proposals/outcome/a-pro-innovation-approach-to-ai-regulation-government-response

Echikson, William, Dimitar Lilkov, and Clara Riedenstein (2024). Transatlantic Trade and Technology: Partners or Rivals? https://cepa.org/comprehensive-reports/transatlantic-trade-and-technology-partners-or-rivals/

Eferin, Y., Gromova, K., Anderson, R., and Esengeldiev, U. (2025). *Artificial intelligence in the Kyrgyz Republic: A silent transformation in the making?* World Bank. https://blogs.worldbank.org/en/opendata/artificial-intelligence-in-the-kyrgyz-republic--a-silent-transfo.

Energy UK (2025). Powering the Cloud: How data centres can deliver sustainable growth. https://www.energy-uk.org.uk/wp-content/uploads/2025/06/Energy-UK-Powering-the-Cloud.pdf

Engler, A., Renda, A., Kerry, C. F., Meltzer, J. P., and Fanni, R. (2021). Strengthening international cooperation on AI. *Brookings*. https://www.brookings.edu/articles/strengthening-international-cooperation-on-ai/

Eristavi, D., and Davituri, G. (2021). *The use of artificial intelligence systems in Georgia: Legislation and practice*. Tbilisi: Institute for Development of Freedom of Information (IDFI). https://idfi.ge/public/upload/Article/ENG%20The%20Use%20of%20Artificial%20Intelligence%20Systems%20in%20Georgia_Legislation%20and%20Practice_19_02_2021.pdf

Galindo, L., K. Perset, and F. Sheeka (2021). "An overview of national AI strategies and policies", *OECD Going Digital Toolkit Notes*, No. 14, OECD Publishing, Paris. https://www.oecd.org/en/publications/an-overview-of-national-ai-strategies-and-policies_c05140d9-en.html

Goel, B., Nilkanth, D., and Singh, S. (2025). *AI governance beyond 2025: UN pathways and implications*. Geneva Graduate Institute. https://www.graduateinstitute.ch/sites/internet/files/2025-09/ARP35_Report--2----Shreya-Singh.pdf

GRATA International (2025). *Overview of Qazaqstani draft laws on artificial intelligence and its provisions on protecting personal data*. https://gratanet.com/publications/overview-of-Qazaqstani-draft-laws-on-artificial-intelligence-and-its-provisions-on-protecting-personal-data

HLAB-AI (2023). "*About the UN Secretary-General's High-level Advisory Body on AI*". https://www.un.org/en/ai-advisory-body/about#:~:text=To%20foster%20a%20globally%20inclusive,the%20international%20governance%20of%20AI.

HLAB-AI (2024). *Governing AI for Humanity* – High-Level Advisory Body on Artificial Intelligence. https://www.un.org/en/ai-advisory-body and https://unctad.org/system/files/non-official-document/ppt_ai-g_01_woo.pdf and https://www.un.org/sites/un2.un.org/files/governing_ai_for_humanity_final_report_en.pdf and

https://informationsecuritybuzz.com/un-report-on-governing-ai-for-humanity/ and

https://www.cdomagazine.tech/aiml/un-advisory-bodys-7-key-recommendations-for-global-ai-governance

HM Government (2021). National AI Strategy. Command Paper 525. https://assets.publishing.service.gov.uk/media/614db4d1e90e077a2cbdf3c4/National_AI_Strategy_-_PDF_version.pdf

Hoque, S. (2025). *Reforming AI laws and regulation in Bangladesh: Current harms and possible future(s)*. Tech Global Institute. https://techglobalinstitute.com/research/reforming-ai-laws-and-regulation-in-bangladesh-current-harms-and-possible-futures/

IAPP (2025). *We cannot be left behind:' How Canada is balancing AI regulation, innovation.* https://iapp.org/news/a/we-cannot-be-left-behind-how-canada-is-balancing-ai-regulation-innovation

IRIM (2025). *Mongolia AI Readiness Assessment successfully completed*. Institute for Regional and International Mobility [IRIM]. https://www.irim.mn/a/2142.

ISED (2025). *Pan-Canadian Artificial Intelligence Strategy.* https://ised-isde.canada.ca/site/ai-strategy/en

ITU (2023). *AI for Good Global Summit Report*. Geneva: International Telecommunications Union. https://aiforgood.itu.int/

KPMG (2025). Qazaqstan AI readiness. KPMG Qazaqstan. https://assets.kpmg.com/content/dam/kpmg/kz/pdf/2025/03/Qazaqstan-AI-Readiness-eng.pdf

Labrecque, C. (2024). Insight: Southeast Asia - *ASEAN issues guidelines for artificial intelligence*. Asia Pacific Foundation of Canada. https://www.asiapacific.ca/sites/default/files/publication-pdf/Insight_SEA_Mar06_EN_V2.pdf

LCO (2021). *Regulating AI: Critical Issues and Choices*. Law Commission of Ontario. https://www.lco-cdo.org/wp-content/uploads/2021/04/LCO-Regulating-AI-Critical-Issues-and-Choices-Toronto-April-2021-1.pdf

Lumenova.ai (2025). AI Governance Frameworks Explained: Comparing NIST RMF, EU AI Act, and Internal Approaches.

https://www.lumenova.ai/blog/ai-governance-frameworks-nist-rmf-vs-eu-ai-act-vs-internal/

Molino, Tracy and Liliane Langevin (2024). *Modernizing Financial Risk Management: OSFI's Draft Guideline on AI Model Risk Management*. https://www.blakes.com/insights/modernizing-financial-risk-management-osfi-s-draft-guideline-on-ai-model-risk-management/

NIST (2022). TTC Joint Roadmap on Evaluation and Measurement Tools for Trustworthy AI and Risk Management. https://www.nist.gov/system/files/documents/2022/12/04/Joint_TTC_Roadmap_Dec2022_Final.pdf

OECD (2025). Steering AI's Future: Strategies for Anticipatory Governance. OECD Artificial Intelligence Papers, No. 32. Paris: OECD Publishing. https://www.oecd.org/content/dam/oecd/en/publications/reports/2025/02/steering-ai-s-future_70e4a856/5480ff0a-en.pdf

OECD (2023). *The state of implementation of the OECD AI Principles four years on*. Paris: OECD Publishing. https://www.oecd.org/content/dam/oecd/en/publications/reports/2023/10/the-state-of-implementation-of-the-oecd-ai-principles-four-years-on_b9f13b5c/835641c9-en.pdf.

OECD/UNESCO (2024). G7 Toolkit for Artificial intelligence in the public sector. https://www.oecd.org/en/publications/g7-toolkit-for-artificial-intelligence-in-the-public-sector_421c1244-en.html

Outeda, C. C. (2024). The EU's AI Act: a framework for collaborative governance. Internet of Things, 101291.

Prilliadi, H. (2022). *Harnessing AI for ASEAN's Future: Governance, Adoption, and Sustainability under the DEFA*. https://www.eria.org/uploads/Harnessing-AI-for-ASEAN%E2%80%99s-Future.pdf

PwC (2022). *Mongolia introduces Personal Data Protection Law effective from 1 May 2022* [Tax

and Legal Alert No. 02/2022]. PricewaterhouseCoopers Mongolia. https://www.pwc.com/mn/en/tax_alerts/pdf/tax_alert_en_02_2022.pdf.

Singhal, M. (2025). *AI regulations in Mongolia*. Just-AI. https://www.justai.in/ai-regulations-in-mongolia/.

SSRN (2022). 3rd Review of the Treasury Board Directive on Automated Decision-Making. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4087546

Toushik Wasi, A., Haque Eram, E., Afroz Mitu, S., and Ahsan, M. (2025). *Generative AI as a Geopolitical Factor in Industry 5.0: Sovereignty, access, and control*. https://arxiv.org/html/2508.00973v1

Treasury Board (2024). *4th review of the Directive on Automated Decision-Making Overview and proposed modifications*. https://wiki.gccollab.ca/images/f/f1/Overview_4th_review_of_the_Directive_on_Automated_Decision-Making.pdf

Quoy N., Duisberg A., and Boullet A. (2024). European Union: leading the way with a legislative framework. https://www.ashurst.com/en/insights/european-union-leading-the-way-with-a-legislative-framework

Scorer, Kay, Ryan Bernard, Kathryn Gullason and Bob Tarantino (2025). A*rtificial intelligence trends to watch in 2025: Regulation of AI*. https://www.dentons.com/en/insights/newsletters/2025/january/23/global-regulatory-trends-to-watch/dentons-canadian-regulatory-trends-to-watch-in-2025/artificial-intelligence-trends-to-watch-in-2025

Shah, Ronak, Courtney Wong, and Tushar Sharma (2024). *Canada's AI Compute Quest: Powering an AI Enabled Future*. https://cassels.com/insights/canadas-ai-compute-quest-powering-an-ai-enabled-future

Tony Blair Institute for Global Change (2025). Sovereignty, Security, Scale: A UK Strategy for AI Infrastructure. https://institute.global/. https://institute.global/insights/tech-and-digitalisation/sovereignty-security-scale-a-uk-strategy-for-ai-infrastructure

TTAI (2024). EU–U.S. Terminology and Taxonomy for AI (2nd edition).

https://digital-strategy.ec.europa.eu/en/library/eu-us-terminology-and-taxonomy-artificial-intelligence-second-edition

UN (2024). Governing AI for Humanity: Final Report. AI Advisory Board. https://www.un.org/sites/un2.un.org/files/governing_ai_for_humanity_final_report_en.pdf

UN (2025). Global collaboration for inclusive and equitable AI. *Technology and Innovation Report 2025*. https://unctad.org/system/files/official-document/tir2025ch5_en.pdf

UNDESA (2024). *United Nations E-Government Survey* (2024)*: The Future of Digital Government.* https://publicadministration.un.org/egovkb/en-us/Reports/UN-E-Government-Survey-2024

UNDESA (2023). *World Public Sector Report. Transforming Institutions for a Greener, More Inclusive and More Resilient Future*. https://desapublications.un.org/sites/default/files/publications/2023-10/World%20Public%20Sector%20Report%202023.pdf

UNESCO (2025). Philippines. Global AI Ethics and Governance Observatory. https://www.unesco.org/ethics-ai/en/philippines.

UNESCO. (2023). *UNESCO emphasises importance of AI Ethics in Qazaqstan*. UNESCO. 22 May 2023. https://www.unesco.org/en/articles/unesco-emphasizes-importance-ai-ethics-Qazaqstan

UNESCO (2021). *Recommendation on the Ethics of Artificial Intelligence*. https://unesdoc.unesco.org/ark:/48223/pf0000381137_eng

UN HLAB (2024). *Governing AI for Humanity: Final Report*, UN, https://www.un.org/sites/un2.un.org/files/governing_ai_for_humanity_final_report_en.pdf

UNIDO (2024). Bridging the AI Divide: Empowering Developing Countries through Manufacturing. https://www.unido.org/sites/default/files/unido-publications/2024-10/IID%20Policy%20Brief%2012_0.pdf

UNSCEB (2024). United Nations System White Paper on AI Governance: An analysis of the UN system's institutional models, functions, and existing international normative frameworks applicable to AI governance. https://unsceb.org/sites/default/files/2024-04/United%20Nations%20System%20White%20Paper%20on%20AI%20Governance.pdf

Vani, Marc, and George R. Wray (2023). *The AIDA companion document: A roadmap for Canada's artificial intelligence law.* https://www.blg.com/en/insights/2023/03/the-aida-companion-document-a-roadmap-for-canadas-artificial-intelligence-law

Vujičić, J. (2023). *Strategic legal frameworks for artificial intelligence: Why smaller countries must Act Now*. https://wjaets.com/sites/default/files/WJAETS-2024-0325.pdf

Wang, J., Milne, C., Zichen Hu, J., and Khan, F. (2024). *OXGS Report | Navigating geopolitics in AI governance*. Oxford Global Society. https://oxgs.org/2024/04/08/oxgs-report-navigating-geopolitics-in-ai-governance/

World Bank (2025). Global Trends in AI Governance. In *worldbank.org*. https://documents1.worldbank.org/curated/en/099120224205026271/pdf/P1786161ad76ca0ae1ba3b1558ca4ff88ba.pdf

Zhazetova, Zh. (2024). *Every fifth public service in Qazaqstan is going to be provided with the help of AI by 2029*. Kursiv Media. 18 March 2024. https://kz.kursiv.media/en/2024-03-18/every-fifth-public-service-in-Qazaqstan-is-going-to-be-provided-with-the-help-of-ai-by-2029/

Zhou, V. (2025). *The Mongolian startup defying Big Tech with its own LLM*. Rest of World. https://restofworld.org/2025/mongolia-egune-ai-llm.

United Nations Development Programme
One United Nations Plaza
New York, NY 10017
www.undp.org

Astana Civil Service Hub
UN Building
14, Mambetov Street, Astana, Z11D8T7
www.astanacivilservicehub.org

ISBN 978-601-12-4628-6