TAKSHASHILA INSTITUTION

State of AI Governance, 2025

A report analysing AI governance measures across countries, companies, and multi-stakeholder gatherings

Arindam Goswami, Bharath Reddy, Sridhar Krishna, Anwesha Sen, Shobhankita Reddy

Takshashila Report 2026-03. Version 1.0, February 2026.

About the State of AI Governance Report

AI is rapidly changing economies, job markets, and global power structures. As countries compete to lead in AI, the rules and systems for managing it have become more complicated and divided. Our annual report serves as a compass for navigating the rapidly changing AI governance landscape, giving policymakers, analysts, and interested citizens evidence-based insights into what's working, what's shifting, and what's coming next.

This is the second State of AI Governance Report from the Takshashila Institution. Building on last year's edition (State of AI Governance, 2024), this report looks at how AI governance has changed in 2025 across countries, companies, and international groups. We review major policy changes, check how accurate our past predictions were, and highlight new trends in rules, innovation, and global competition.

To find out more about our work, visit the Takshashila website.

Some of our work on AI governance is linked below:

- A Pathway to AI Governance
- Takshashila Policy Advisory - Working Paper on Generative AI and Copyright (Part 1)
- Building India's Data Centres
- AI Adoption - Think Tasks, not Jobs
- A Primer on AI Chips
- A Survey of Military AI Technologies

Executive Summary

The year 2025 was defined by high volatility in AI policy. As industry leaders balanced "bubble" anxieties and circular investments against the promise of massive productivity gains, governments have pivoted toward a high-stakes geopolitical innovation race. The race to lead in AI is reshaping global power, and the rules governing it are fracturing along geopolitical fault lines. In 2025, every major power chose innovation over accountability, betting that winning the AI race matters more than governing it responsibly. This report analyses how the AI governance landscape evolved in 2025 amid the rapidly evolving technological landscape.

Countries

In the past year, major countries have handled AI in different ways. In the US, the Trump administration rolled back most Biden-era safety rules and focused more on infrastructure and global competition. Meanwhile, individual states took over as main regulators, resulting in a patchwork of laws across the country.

The EU has continued building the framework for implementing its wide-ranging AI Act, while also committing resources and outlining strategies to promote European innovation and technological sovereignty. Pushback from industry and some member states has led to calls for simplification of rules through the Digital Omnibus proposal.

China kept refining its governance model, aiming to balance government control and economic growth. 2025 has seen a massive push for diffusion of AI across sectors and also building domestic capabilities in critical parts of the AI supply chain.

India has prioritised innovation through self regulation and voluntary disclosures. There are also initiatives to build long term resilience across the semiconductor value chain, incentives for setting up data centres, and compute subsidies for priority use cases. There are ongoing efforts in establishing the AI Governance Group (AIGG) and the AI Safety Institute (AISI). The focus is on multilingual models and applications that can bridge state capacity limitations in delivering public services.

Companies

AI companies are choosing to use governance frameworks, such as publishing principles, doing risk assessments, and setting up oversight structures. Still, reporting standards are not consistent, and outside scrutiny differs a lot. The tension between speed of innovation and robustness of safety mechanisms more often than not prioritises speed in a bid to be the first to market. The This report looks at governance practices among key players in the AI industry.

International Forums

AI Summits and partnerships like GPAI continue to support discussions, but they do not lead to binding agreements. Even big announcements often lack full backing. For instance, the US and EU did not sign key agreements at the February 2025 AI Action Summit. These forums mostly serve as spaces to share concerns instead of making real changes.

The report also offers predictions about how AI governance could change in 2026.

Table of Contents

Note: The icons will be used as navigational tools throughout the document.

How Did Our Predictions For Last Year Fare?

#	Prediction	Confidence	Region	How did the prediction fare?
1.	Compute thresholds for enforcing regulations will no longer be relevant. The effectiveness of these thresholds might be challenged as a measure of capability as inference computing begins to scale and smaller models become more efficient. The US and EU have 10^26 and 10^25 flops as training compute thresholds for enforcing certain regulations.	High	Global	This prediction has proven partially accurate. In the US, the revocation of Executive Order 14110 eliminated the disclosure requirements for large models (those trained using >10^26 FLOPs). The EU AI Act retains the 10^25 flops threshold to identify models with systemic risks. However, the Act also incorporates supplemental criteria for classifying models with systemic risks, such as capability evaluations, user base size, scalability, and access to tools.
2.	Investments in sovereign cloud infrastructure will increase, driven by geopolitical considerations.	High	Global	This prediction has proven accurate, as seen by a global trend of aggressive, state-supported investment in AI infrastructure. In the EU, the state is investing in sovereign AI factories. China is expanding its data centre capacity through access to subsidised power and land. The US has focussed on fast tracking environmental clearances and building out power infrastructure. Indian efforts include 20+ year tax holidays for data centres and subsidising compute access for priority sectors.

Confidence reflects how likely the authors believed their prediction will be accurate in the coming year.

#	Prediction	Confidence	Region	How did the prediction fare?
3.	Open-source and open-weight models will continue to be pushed by China and EU as a pathway to strategic autonomy and technology leadership. DeepSeek and Mistral will remain open-weight or open-source.	High	China, EU	This prediction has proven accurate. Open weight model are seen as a battlefield for global AI influence. The United States and China have identified dominating the open weight model ecosystem as a priority. The EU and India also champion open models as a pathway to strategic autonomy. Most of the leading open weight models are released by Chinese labs.
4.	The compute capacity created under the IndiaAI mission aimed towards incentivising startups and building indigenous models will be underutilised. This is due to lack of demand that meets the criteria to qualify for the subsidies as well as due to friction in the bureaucratic process involved.	High	India	This prediction has not proven accurate. The GPUs offered via empanelled vendors has scaled up to 38000 and is growing. Over ₹1200 crores has been allocated under the compute mission, to over 300 users, with most of it going towards training indigenous models. While the process to apply for compute via this portal is likely to involve friction, there seems to be sufficient demand from model developers. This is on par for the budget allocated for this pillar over five years.
5.	AI governance regulations at the state level will continue to prioritise innovation over encouraging transparency, accountability, and societal well-being. In other words geopolitical considerations will trump protection of individual rights as a governance priority.	Moderate	Global	This trend seems to hold true. Sovereignty across the AI value chain is an important governance priority and most regions have opted for a light touch regulatory approach. The EU AI Act which had a focus on safety and accountability has also seen pushback from industry and some member states, resulting in simplification efforts under the Digital Omnibus for AI.

Confidence reflects how likely the authors believed their prediction will be accurate in the coming year.

#	Prediction	Confidence	Region	How did the prediction fare?
6.	US chip restrictions on China will not escalate further. This is because Deep Seek makes owning newer chips less relevant.	High	US	This prediction turned out to be partially accurate. The situation is nuanced. Chip restrictions are at significantly lower levels that during the Biden administration. However, this is driven more by diplomatic/economic concerns than DeepSeek's technical achievements.
7.	Federal laws focused on monitoring AI safety and federal agency assessment of AIs for discrimination and bias will be made defunct or watered down significantly. By the end of the year, AI safety guardrails will be driven by private firms.	Moderate	US	This prediction turned out to be accurate. The Trump Administration has revoked Biden's EO 14110 which mandated AI safety monitoring, eliminated federal agency assessments for discrimination and bias, NIST AI RMF being revised to remove bias mitigation guidance, and the OMB guidance on "Preventing Woke AI". AI safety is now primarily driven by private company initiatives and state-level regulations (though under federal challenge).
8.	EU's comprehensive regulatory framework, including penalties for non-compliance, would result in a few companies not releasing their AI models/features in the EU. This might lead to a milder enforcement of the regulations. As per the declared timeline, rules on notified bodies, general purpose AI models, governance, confidentiality, and penalties start to apply from August, 2025.	Moderate	EU	This prediction has proved to be accurate. Apple and Meta have restricted the launch of key AI features/models in the European Union, citing regulatory uncertainties. Over 40 of Europe's most influential companies wrote an open letter asking for a pause or delay of the EU AI Act's most stringent requirements. The Digital Omnibus on AI is proposed to simplify and ease compliance under the AI Act.

Confidence reflects how likely the authors believed their prediction will be accurate in the coming year.

#	Prediction	Confidence	Region	How did the prediction fare?
9.	The US Diffusion Framework will not stop state-of-the-art AI models coming out of China, at least not in the next year. This is because Deep Seek makes owning newer chips less relevant, and China has built up an overcapacity of data centres over the past few years.	Moderate	China	This prediction is partially accurate. The US Diffusion Framework itself was repealed, in favour of other restrictions on the most advanced AI chips. Chinese AI labs have adapted through a mix of algorithmic or efficiency innovations, bypassing restrictions to access chips but have pivoted to open weight models. All frontier models are developed by US labs.
10.	Regulatory focus will be on protecting the information ecosystem from content deemed harmful for the government or Indian society.	Moderate	India	This prediction has held true. The IT Rules Amendment on synthetically generated information, announced in 2025, is a move aimed at minimising the impact of misleading information generated using computer resources.
11.	Governance of AI within companies will become a bigger requirement as governments firm up on their positions regarding AI. 'Chief Responsible AI Officer' will be a new role at companies seeking to deploy AI solutions at scale, whose duty it will be to ensure AI is deployed in a manner that will, at the very least, protect them from litigation.	Moderate	Corporate	This has proven partially accurate. Due to global AI innovation race, regulation is tentative and lags behind innovation. Most countries have preferred self regulation and voluntary disclosures, reserving reserving strict mandates only for a narrow sliver of high-risk use cases. The friction between speed-to-market and safety often ends up with safety being a lower priority. The Chief Responsible AI Officer role has not become very popular. The expectations of the Chief Data Officer has expanded to ensure regulatory compliance and adherence to organisations governance principles.

Confidence reflects how likely the authors believed their prediction will be accurate in the coming year.

Predictions For This Year

Sr. No.	Prediction	Confidence	Region
1.	Indications of an AI bubble will manifest more at the model layer rather than at the application layer. The model layer will face a profitability crisis, while the application layer will see increased action leading to further investments in inference optimised data centres.	High	Global
2.	Frontier models will face a profitability crisis leading to some of them relaxing the guardrails in a bid to capture market share. This could take the form of attracting users seeking non-neutral political discourse or NSFW content, leading to regulatory pushback.	High	Global
3.	AI diffusion will lead to restructuring of the workforce in high exposure sectors such as customer service, administration and entry-level IT services. There will be increased calls for AI impacted hiring and firing to remain transparent, fair and accountable.	High	Global
4.	Major federal-state legal battles will define AI governance, and it will converge towards light touch federal regulations. For instance, the December 2025 Executive Order 14365 will trigger extensive litigation, with multiple states (likely including California, Colorado, New York) filing suit challenging federal overreach, Constitutional questions on Tenth Amendment, Commerce Clause, and federal preemption, preliminary injunctions likely sought to block federal enforcement actions.	High	US
5.	Genesis Mission will deliver early demonstrable results, because they are building on existing DOE infrastructure (not starting from scratch); it has clear leadership and accountability structure; industry partnerships already secured (24 initial MOUs); specific 270-day and 1-year milestones in Executive Order.	High	US

Confidence reflects how likely the authors believe their prediction will be accurate in the coming year.

Sr. No.	Prediction	Confidence	Region
6.	There will be continued pressure to weaken or dilute some of the provisions of the AI Act from companies and some member states as discussion on the Digital Omnibus for AI progress.	High	EU
7.	Chinese AI labs will continue to dominate leaderboards for open weight AI models. Due to concerns about data sovereignty with Chinese models, open weight will be a pathway to gain wider usage.	High	China
8.	Following the tax holidays on data centres, foreign companies will set up Indian data centres that serve global clients.	High	India
9.	Applications and Indic language models continue to be financed and built but downstream adoption will be low.	High	India
10.	Integration of advertisements in generative AI outputs enables hyper-personalisation at scale. This is likely to lead to governance measures focused on algorithmic transparency and conspicuous disclosures of ad content.	Moderate	Global

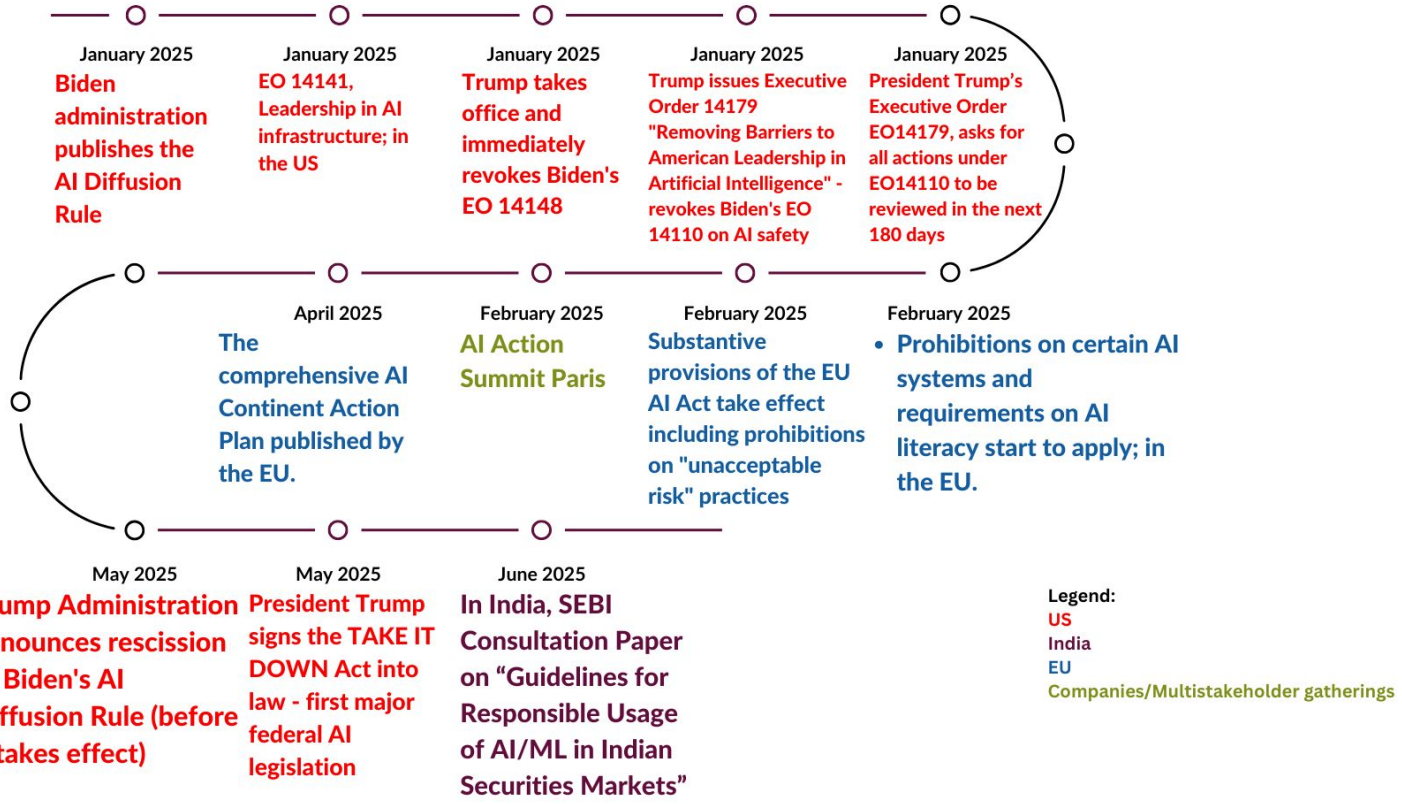
Confidence reflects how likely the authors believe their prediction will be accurate in the coming year.

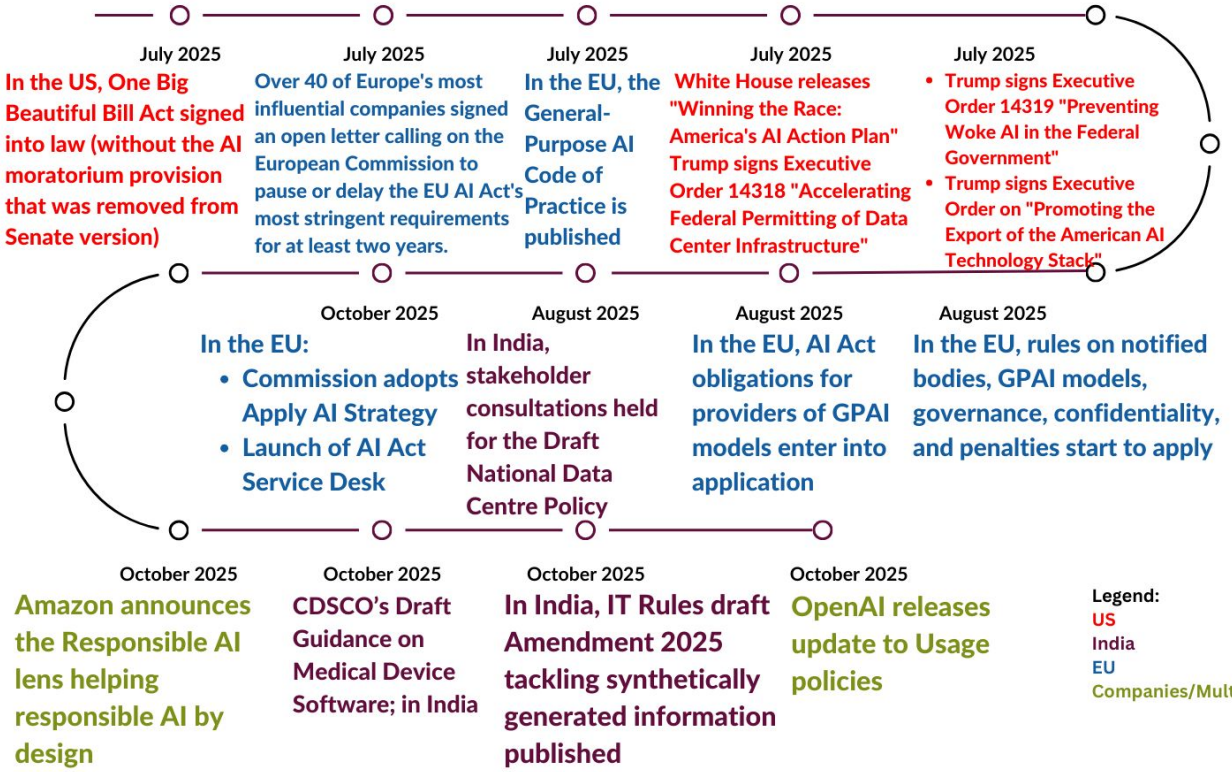
Sr. No.	Prediction	Confidence	Region
11.	Policy will continue to balance the need for strategic autonomy with the short term goal of being globally competitive. For instance, this is visible in banning advanced US AI chips but creating strategic exemptions for some of its largest companies such as Alibaba, Tencent, and ByteDance.	Moderate	China
12.	An expansion of the regulatory oversight on consumer AI, specifically controls on human-like chatbots will intensify	Moderate	China
13.	A state-led AI safety framework, focused on advanced application risks will be formalised spanning the entire AI deployment lifecycle.	Moderate	China
14.	Improvements in initiatives for data digitisation, collection, curation, and annotation since the availability of Indian data for AI training and fine-tuning is currently a bottleneck for deploying AI.	Moderate	India
15.	FTC enforcement of TAKE IT DOWN Act will be inconsistent, because FTC capacity is weakened by Commissioner terminations, industry pushback on compliance costs, potential litigation over enforcement discretion, etc.	Low	US

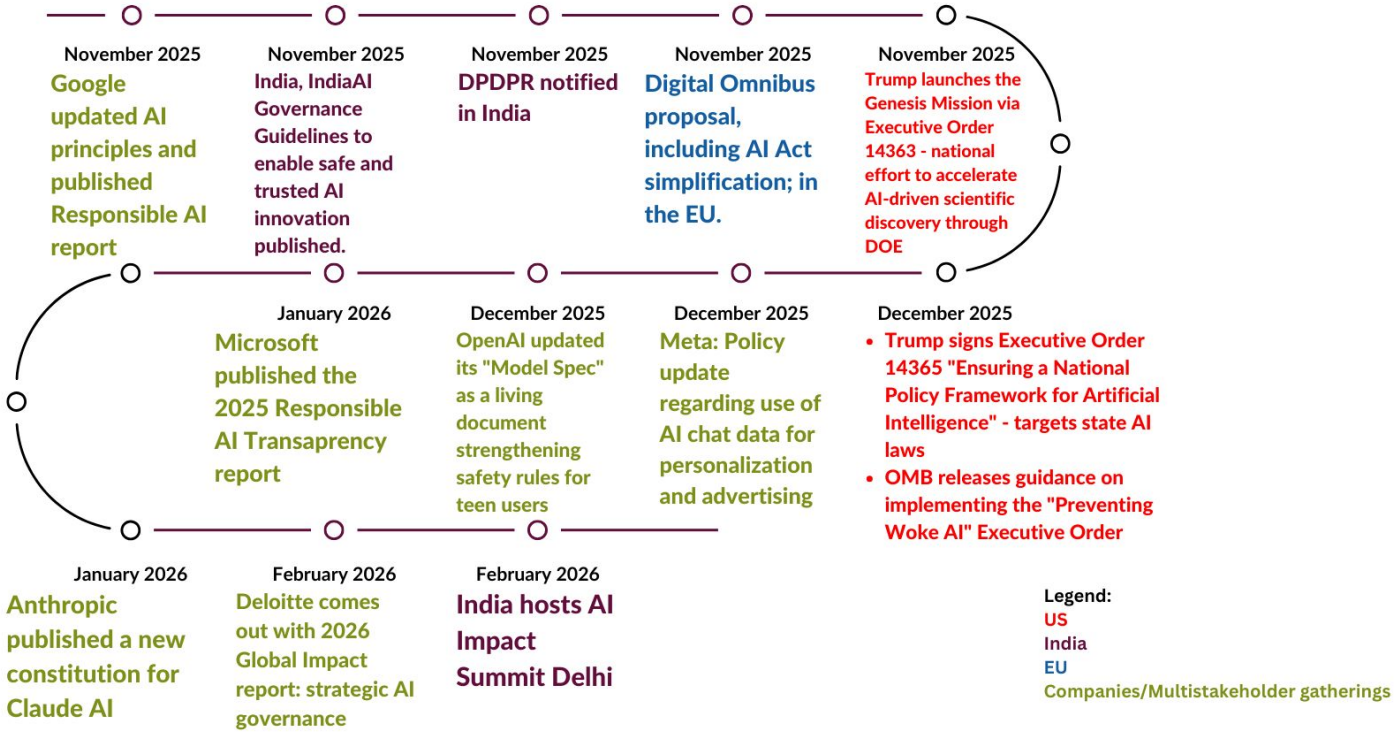
Confidence reflects how likely the authors believe their prediction will be accurate in the coming year.

Timeline of AI Governance Events

A timeline of significant AI governance events across countries, companies, and multi-stakeholder gatherings is presented below. The timeline focuses only on AI Governance events and does not list milestones related to advancements in the technology.


January 2025
Biden administration publishes the AI Diffusion Rule
January 2025
EO 14141, Leadership in AI infrastructure; in the US
January 2025
Trump takes office and immediately revokes Biden's EO 14148
January 2025
Trump issues Executive Order 14179 "Removing Barriers to American Leadership in Artificial Intelligence" - revokes Biden's EO 14110 on AI safety
January 2025
President Trump's Executive Order EO14179, asks for all actions under EO14110 to be reviewed in the next 180 days
February 2025
• Prohibitions on certain AI systems and requirements on AI literacy start to apply; in the EU.
February 2025
Substantive provisions of the EU AI Act take effect including prohibitions on "unacceptable risk" practices
February 2025
AI Action Summit Paris
April 2025
The comprehensive AI Continent Action Plan published by the EU.
May 2025
Trump Administration announces rescission of Biden's AI Diffusion Rule (before it takes effect)
May 2025
President Trump signs the TAKE IT DOWN Act into law - first major federal AI legislation
June 2025
In India, SEBI Consultation Paper on "Guidelines for Responsible Usage of AI/ML in Indian Securities Markets"
Legend:
US
India
EU
Companies/Multistakeholder gatherings


July 2025
In the US, One Big Beautiful Bill Act signed into law (without the AI moratorium provision that was removed from Senate version)
July 2025
Over 40 of Europe's most influential companies signed an open letter calling on the European Commission to pause or delay the EU AI Act's most stringent requirements for at least two years.
July 2025
In the EU, the General-Purpose AI Code of Practice is published
July 2025
White House releases "Winning the Race: America's AI Action Plan"
Trump signs Executive Order 14318 "Accelerating Federal Permitting of Data Center Infrastructure"
July 2025
Trump signs Executive Order 14319 "Preventing Woke AI in the Federal Government"
Trump signs Executive Order on "Promoting the Export of the American AI Technology Stack"
August 2025
In the EU, rules on notified bodies, GPAI models, governance, confidentiality, and penalties start to apply
August 2025
In the EU, AI Act obligations for providers of GPAI models enter into application
August 2025
In India, stakeholder consultations held for the Draft National Data Centre Policy
October 2025
In the EU:
Commission adopts Apply AI Strategy
Launch of AI Act Service Desk
October 2025
Amazon announces the Responsible AI lens helping responsible AI by design
October 2025
CDSCO's Draft Guidance on Medical Device Software; in India
October 2025
In India, IT Rules draft Amendment 2025 tackling synthetically generated information published
October 2025
OpenAI releases update to Usage policies
Legend:
US
India
EU
Companies/Multistakeholder gatherings

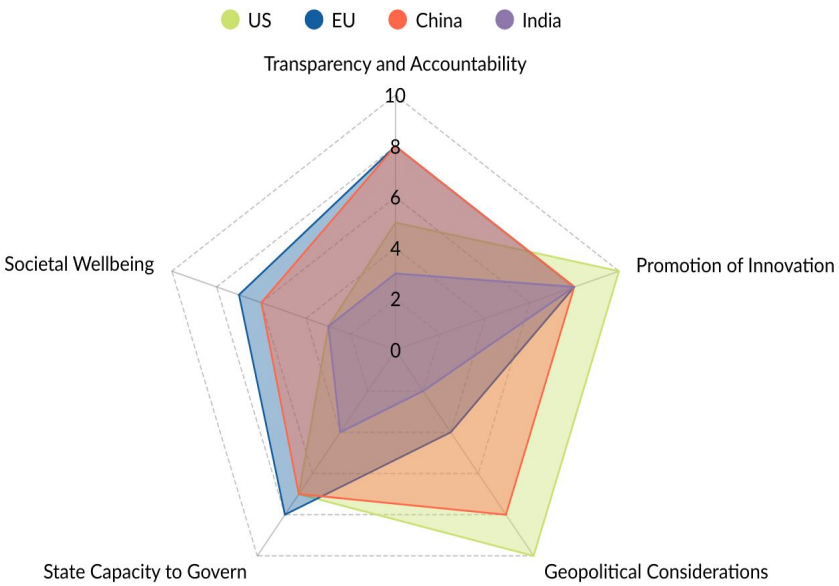

November 2025
Google updated AI principles and published Responsible AI report
November 2025
India, IndiaAI Governance Guidelines to enable safe and trusted AI innovation published.
November 2025
DPDPR notified in India
November 2025
Digital Omnibus proposal, including AI Act simplification; in the EU.
November 2025
Trump launches the Genesis Mission via Executive Order 14363 - national effort to accelerate AI-driven scientific discovery through DOE
December 2025
• Trump signs Executive Order 14365 "Ensuring a National Policy Framework for Artificial Intelligence" - targets state AI laws
• OMB releases guidance on implementing the "Preventing Woke AI" Executive Order
December 2025
Meta: Policy update regarding use of AI chat data for personalization and advertising
December 2025
OpenAI updated its "Model Spec" as a living document strengthening safety rules for teen users
January 2026
Microsoft published the 2025 Responsible AI Transaprency report
January 2026
Anthropic published a new constitution for Claude AI
February 2026
Deloitte comes out with 2026 Global Impact report: strategic AI governance
February 2026
India hosts AI Impact Summit Delhi
Legend:
US
India
EU
Companies/Multistakeholder gatherings

Analysis of AI Governance Measures Across Countries

Analysis of AI Governance Measures Across Countries

Summary

The chart compares the US, EU, China, and India in five areas: transparency and accountability, promotion of innovation, geopolitical considerations, societal well-being, and state capacity to govern. The geopolitical innovation race and the characteristics of AI governance have led many countries to promote domestic innovation to address technological sovereignty and supply-chain resilience to different extents. While the EU has been one of the front-runners in taking the lead on prioritising safety and societal well-being, industry and member state pushback have led to some realignment. Overall, all countries are leaning towards a lighter-touch regulation and scaling up regulatory capacity. Addressing societal well-being through use cases in healthcare and public-service delivery also sees significant interest.



- AI governance measures often address multiple objectives. These include ensuring transparency and accountability, promoting innovation, addressing geopolitical considerations, enabling state capacity to implement the measures, and promoting societal well-being.
- The authors have analysed and compared the country-specific AI governance measures across these different criteria. There is some subjectivity in this comparative analysis, but the authors feel it is a useful representation of how countries are pursuing these different AI governance priorities.
- The United States of America, the European Union, China, and India are selected as countries/regions for comparison. These have been chosen for their significant role in influencing the path of innovation, governance or adoption of AI.
- The chart shows the author's scoring of the AI governance measures in the different countries on the selected criteria on a scale of 0-10. The following slides in this section provide the reasoning for the scoring for each state/region.

Analysis of AI Governance Measures Across Countries

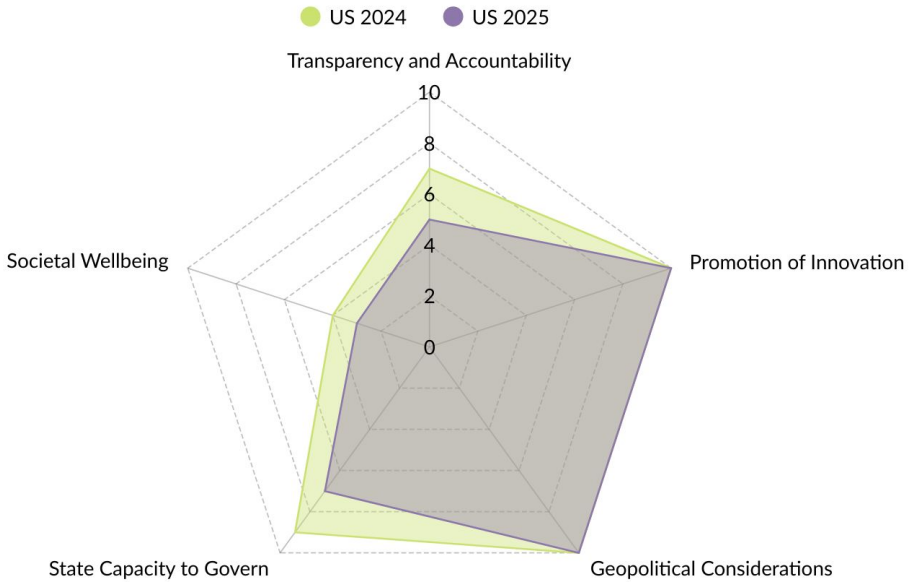
A description of the different criteria is provided below.

- **Transparency and Accountability:** Assesses the extent to which governance frameworks attempt to promote transparency and accountability. This includes examining governance instruments such as evaluation and disclosure requirements, licensing requirements, penalties for non-compliance, and grievance redressal mechanisms.
- **Promotion of Innovation:** Evaluates how regulations foster innovation by creating an enabling environment. This includes examining the quantum of funding for AI infrastructure, restrictions on market participation, education and skilling initiatives and maturity of the R&D ecosystem.
- **Geopolitical Considerations:** This assesses the extent to which policy decisions address geopolitical priorities. It includes assessing whether a state can secure access to the building blocks of AI and deny access to other countries. Relevant policy measures include export controls, investments in domestic infrastructure, promotion of open-source technologies, and policies that reduce vulnerabilities in the value chain.
- **Societal Wellbeing:** Assesses how regulations address broader societal concerns, such as protecting individuals from risks and harms from the adoption of AI in various sectors, and reducing environmental costs associated with AI.
- **State Capacity to Govern:** An estimation of the financial resources, institutional frameworks, and skilled human capital being created to enforce compliance with AI regulations effectively.

Analysis of AI Governance Measures in the US

Trends in 2025

- Innovation First: The Trump administration has significantly changed course on safety and oversight prioritising an innovation first policy. This includes relaxing safety testing and disclosure requirements, bias and discrimination oversight, equity and civil rights considerations, and international cooperation frameworks.
- Geopolitics Dominates: AI governance has become a tool to achieve economic and military dominance. Restrictions on key inputs in the AI value chain continue to be used as diplomatic leverage rather than pure security measures.
- Rise of State-Level Regulation: With federal deregulation, states have taken the lead on AI governance. Focus areas include deep fakes, non consensual intimate imagery, algorithmic discrimination, transparency and high-risk systems.
- Federal-State Conflict Escalates: Actions like the Executive Order 14365 represent an unprecedented federal challenge to state authority. This sets up major legal battles between federal government and states.
- Narrow Bipartisan Consensus on Limited Issues: Despite deep partisan divides, bipartisan consensus exists on narrow issues like child safety online, deepfakes, infrastructure investment, opposition to adversary access to advanced AI, etc.



Sources: US Executive Orders and Acts, IAPP - US State AI Legislation Tracker, BCLP Law - US State AI Legislation Snapshot, NIST Risk Framework, AGORA

Analysis of AI Governance Measures in the US

- "Woke AI" as a Political Framework: The Trump Administration has politicised AI safety concepts, conflating technical bias mitigation with political ideology. This represents a unique approach where AI governance principles themselves become political battlegrounds.
- Refer to **Appendix 1** for details about the specific governance measures.

Transparency & Accountability

- **Rollbacks**:
  - The revocation of Biden's EO 14110 eliminated requirements for disclosure of large foundational model training activities and model weights for dual-use models (those trained using >10^26 FLOPs)
  - NIST's mandate to develop comprehensive safety testing standards was discontinued.
  - The AI Diffusion Rule's reporting requirements for large model training runs were rescinded before implementation.
- **Remaining Measures:**
  - The OMB Act continues to mandate risk management standards for AI use in identified high-risk government systems
  - State-level regulations (California, Colorado, Texas, Utah) maintain various disclosure and transparency requirements, though these face federal challenge
  - The DOE retains evaluation capabilities for identifying AI model risks in critical areas like nuclear and biological threats

- **New Concerns:**
  - The "Preventing Woke AI" Executive Order mandates federal procurement prioritise "truth-seeking" and "ideological neutrality" principles, which critics argue could politicise AI safety assessments
  - NIST AI Risk Management Framework is under revision to remove references to DEI, fairness, and bias considerations
  - Federal oversight mechanisms have been significantly weakened

Sources: US Executive Orders and Acts, IAPP - US State AI Legislation Tracker, BCLP Law - US State AI Legislation Snapshot, NIST Risk Framework, AGORA

Analysis of AI Governance Measures in the US

Promotion of Innovation

The Trump Administration has made innovation acceleration its primary AI governance priority:

Major Initiatives:

- AI Action Plan (July 2025): Comprehensive framework with over 90 federal policy actions across three pillars: Accelerating Innovation, Building AI Infrastructure, and Leading in International Diplomacy and Security
- Genesis Mission (November 2025): $Multi-billion initiative leveraging DOE's 17 National Laboratories to build integrated AI discovery platform, with goal to double American science/engineering productivity within a decade
- Infrastructure Development: Executive Order 14318 streamlines federal permitting for AI data centres (>$500M investment or >100MW load), designates qualifying projects on federal lands, provides preferential land and energy policies
- AI Technology Stack Export Program: Commerce and State Departments partnering with industry to deliver secure, full-stack AI export packages (hardware, models, software, applications, standards) to allies

Deregulation Efforts:

- Elimination of Biden-era safety and reporting requirements
- Focus on removing "bureaucratic barriers" to AI development
- Attempted (but failed) 10-year moratorium on new state AI regulations through One Big Beautiful Bill Act
- OMB mandate for open-sourcing government AI models and data by default continues

Funding & Support:

- Continued support for NIST's NAIRR (National AI Research Resource) project
- DOE continues streamlining approvals for AI infrastructure including power and data centres
- Bank of China-style funding initiatives being explored for strategic AI sectors
- Tax incentives and federal financing tools (loans, equity investments, technical assistance) for qualifying AI projects

Sources: AI Action Plan, One Big Beautiful Bill Act, Genesis Mission, US Executive Orders and Acts, IAPP - US State AI Legislation Tracker, BCLP Law - US State AI Legislation Snapshot, NIST Risk Framework, AGORA

Analysis of AI Governance Measures in the US

Geopolitical Considerations

Geopolitics has become the dominant consideration in US AI governance:

Export Controls - Major Pivot:

- AI Diffusion Rule (January 2025): Biden Administration attempted to create three-tier global framework:
  - Tier 1: US + 18 allies (unrestricted access)
  - Tier 2: ~150 countries (limited quotas via Data Center Validated End User program)
  - Tier 3: China, Russia, adversaries (prohibited)
- Rescission (May 2025): Trump Administration rescinded rule before implementation, citing concerns it would "stifle American innovation" and "undermine diplomatic relations"
- Current Status: Ad-hoc export licensing continues; focus shifted to leveraging chip access as diplomatic/trade tool rather than security control

Sources: Rescission of AI Diffusion Rule

Strategic Measures:

- Monitoring regime continues for large AI model training runs on US Infrastructure as a Service (IaaS)
- Treasury Outbound Investment Security Rule prohibits transactions involving high-risk AI systems for military/surveillance use. The rule represents a notable shift in US economic security policy. Historically, the US scrutinised inbound foreign investment through CFIUS, but this program flips the lens — monitoring and restricting where US money flows outward.
- The Comprehensive Outbound Investment National Security (COINS) Act of 2025, which was part of the FY 2026 National Defense Authorization Act, became law on December 18, 2025. It significantly expands the program in several ways.
- Security agencies tasked with identifying AI supply chain vulnerabilities
- "America First" approach to AI leadership explicitly prioritises US dominance

International Strategy:

- AI Action Plan emphasises exporting "American AI Technology Stack" to align allies with US standards
- Focus on preventing technology transfer to adversaries while enabling diffusion to partners
- De-emphasis on multilateral cooperation compared to Biden approach

Analysis of AI Governance Measures in the US

Societal Wellbeing

Societal wellbeing protections have been significantly reduced at the federal level:

State-Level Activity:

- Most societal wellbeing regulations now occur at state level
- Over 70 AI-related laws passed in at least 27 states in 2025
- California: 13 new AI laws including SB 53 (Transparency in Frontier AI Act), companion chatbot regulations (SB 243), AI transparency requirements
- Colorado: AI Act requiring reasonable care against algorithmic discrimination (effective date delayed to June 30, 2026)
- Texas: Responsible AI Governance Act (TRAIGA) effective January 1, 2026
- Utah: Amended AI Policy Act with safe harbour provisions

Federal Measures - Limited:

- TAKE IT DOWN Act (May 2025): Only major federal AI legislation passed; criminalises non-consensual intimate imagery (NCII) including deepfakes
- Platforms must remove flagged NCII within 48 hours
- Criminal penalties: up to 2 years imprisonment (adults), harsher for minors
- FTC enforcement with civil penalties up to $51,000 per violation
- Platform compliance deadline: May 19, 2026

Rollbacks:

- Biden EO 14110 provisions addressing algorithmic discrimination, bias, and civil rights protections eliminated
- Federal emphasis on equity and civil rights in AI removed
- NIST AI RMF being revised to remove DEI, fairness, and bias mitigation guidance

Federal-State Conflict:

- December 2025 Executive Order 14365 directs DOJ to establish AI Litigation Task Force to challenge "onerous" state AI laws
- Attorney General tasked with challenging state laws on grounds of:
  - Unconstitutional burdens on interstate commerce
  - Preemption by federal statutes
  - Violation of free speech (First Amendment)
- Federal funding (BEAD program, discretionary grants) conditioned on states not enforcing conflicting AI laws
- FCC directed to consider federal reporting/disclosure standards that would preempt state laws

Sources: TAKE IT DOWN Act

Analysis of AI Governance Measures in the US

State Capacity to Govern

State capacity remains uncertain following major policy reversals:

Federal Level:

- Uncertainty: Revocation of EO 14110 creates regulatory vacuum
- Weakened Agencies: FTC Commissioners terminated, raising enforcement concerns
- Limited Mandates: OMB Act requires agencies to assess AI maturity and prepare AI use plans, but implementation unclear

Infrastructure Investments:

- Strong: Genesis Mission represents major federal commitment ($multi-billion, though exact figures not disclosed)
- DOE National Labs: Leveraging existing infrastructure of 17 National Laboratories
- Computing Power: Focus on building world's most powerful scientific AI platform
- Private Partnerships: 24 organisations signed MOUs for Genesis Mission collaboration

Regulatory Capacity:

- Reduced: National AI Talent Surge from Biden EO discontinued
- Limited: NIST continues AI standards work but with reduced scope
- Fragmented: State-level capacity building creates 50 different regulatory regimes

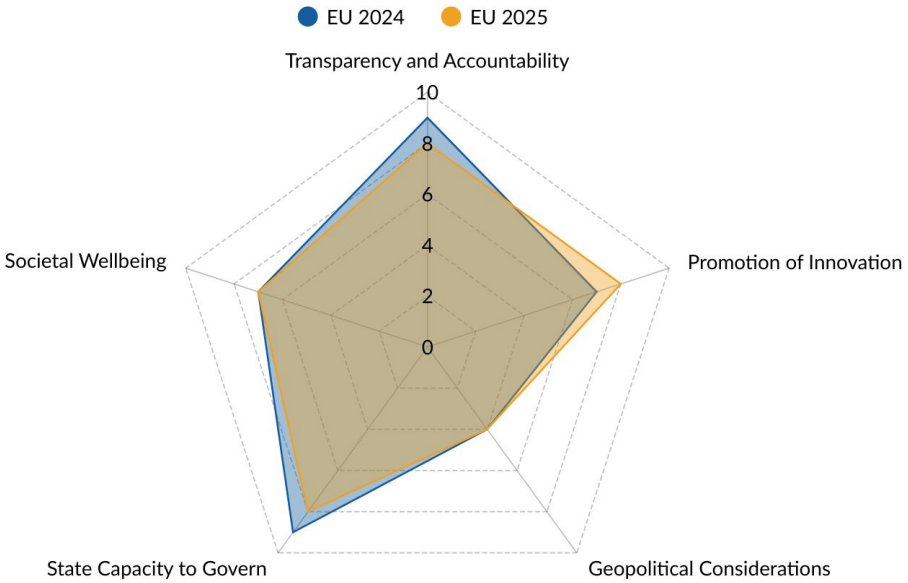
Outlook:

- Regulations on AI infrastructure, capital flows, and talent relatively stable
- Transparency, auditability, and governance capacity measures under review or eliminated
- Shift from federal oversight to market-driven governance

Analysis of AI Governance Measures in the EU

Trends in 2025

- Building on the regulatory framework established by the EU AI Act, many other initiatives have been undertaken to advance Europe as an AI power. These include the AI Continent Action Plan, Data Union Strategy, and Apply AI strategy, which commit resources and outline strategies to promote European innovation and technological sovereignty.
- There has been progress in establishing codes of practice, voluntary commitments, and regulatory bodies specified by the Act. However, the implementation of the national-level authorities by member states is not uniform. Only three states have designated both notifying and market surveillance authorities, while others are yet to do so or have made partial progress.
- Companies like Meta and Apple have decided not to release some features or models, citing regulatory uncertainties. Pressure from industry and some member states to pause or delay enforcement of the EU AI Act's compliance requirements led to the Digital Omnibus proposal to simplify and rationalise the Act's requirements.
- They interact with other regulations like the GDPR, DMA, DSA, Chips Act, and Cyber Resilience Act, that collectively influence the operations entities involved in the manufacture, deployment, import or distribution of AI systems.
- Refer to **Appendix 2** for details about the specific governance measures.



Sources: EU AI Act, EU InvestAI Initiative, EU Chips Act, AI Continent Action Plan, AI Factories, AI Pact, Apply AI Strategy, Artificial Intelligence in Health, AI Act Service Desk, AI Act Single Information Platform, Guidelines on Prohibited AI Practices, Digital Omnibus on AI, European AI Office, European Data Union Strategy, The General-Purpose AI Code of Practice, AGORA

Analysis of AI Governance Measures in the EU

Transparency & Accountability

- The EU has the most comprehensive measures for AI transparency and accountability, including risk tiering, evaluations, disclosure, licensing, penalties, and input controls.
- General-Purpose AI Code of Practice: A voluntary tool offering practical solutions for providers to comply with the AI Act.
- The AI Pact: A voluntary initiative where participants create a collaborative community, sharing their experiences and best practices. Companies also pledge to implement transparency measures ahead of the legal deadline.
- Compliance Checker & AI Act Explorer: Digital tools available on the Single Information Platform that help stakeholders determine their legal obligations and browse the legislation to understand compliance requirements

Promotion of Innovation

- AI Factories and Gigafactories: A major initiative to create "dynamic ecosystems" that bring together computing power, data, and talent to train cutting-edge AI models and applications. The EU plans to establish at least 15 AI Factories and up to 5 AI Gigafactories (which are four times more powerful and akin to a CERN for AI).
- Apply AI Strategy: This strategy aims to boost AI adoption across 10 key industrial sectors (such as healthcare, robotics, and manufacturing) and the public sector. It is designed to enhance the competitiveness of strategic sectors and strengthen the EU's technological sovereignty.
- Data Union Strategy: This initiative seeks to increase the availability of data for AI development, simplify EU data rules and strengthen the EU's position on international data flows.
- The proposed Digital Omnibus amendments seek to reduce compliance burdens by extending timelines for key requirements, eliminating certain obligations, simplifying compliance for smaller enterprises, and more.

Geopolitical Considerations

- Technological Sovereignty: The Apply AI Strategy explicitly promotes a "buy European" approach, particularly for public sector procurement, with a focus on open-source.
- Chips and Infrastructure: Investments in AI factories with the goal of tripling data centre capacity over five to seven years. The Chips Act aims to reduce reliance on non-EU technology by boosting local design and production of AI semiconductors.
- The Data Union Strategy prioritises EU data sovereignty, ensuring that international data flows are secure and that unjustified data localisation measures are countered.

Analysis of AI Governance Measures in the EU

- The AI Office is tasked with promoting the EU's approach to trustworthy AI globally and fostering international cooperation on AI governance.

Societal Wellbeing

- AI for Health and Societal Good: Initiatives like the European Cancer Imaging Initiative and 1+ Million Genomes Initiative use AI to improve disease prevention and treatment outcomes.
- AI Omnibus requires Member States and the Commission to actively foster AI literacy among staff and the public to ensure people can use AI safely and effectively.

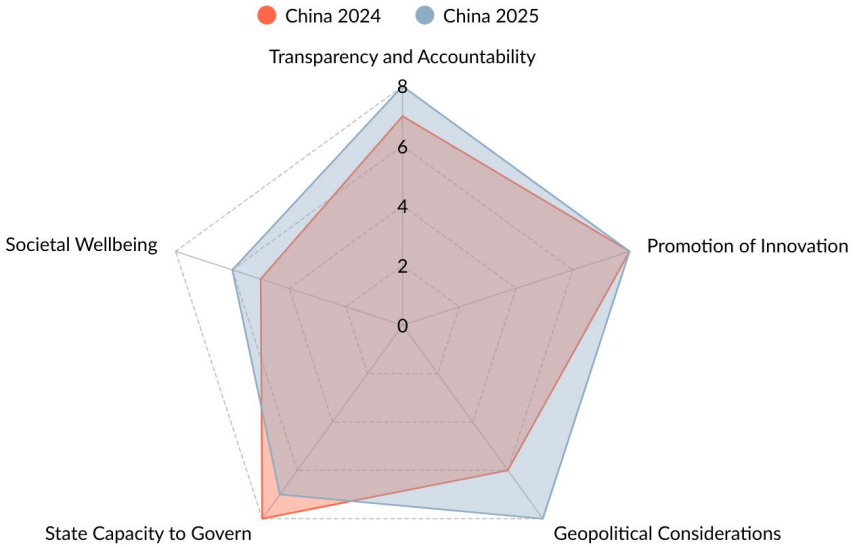
State Capacity to Govern

- The Digital Omnibus proposes centralising oversight of complex AI systems within the AI Office to avoid fragmentation and ensure consistent enforcement across Member States.
- AI Act Service Desk: A team of experts and an online tool allowing stakeholders to submit questions and receive guidance, enhancing the administration's capacity to support implementation.
- Implementation among member states of market surveillance and national authorities is not uniform.

Analysis of AI Governance Measures in China

Trends in 2025

- China's AI governance has evolved in a layered and incremental manner, with the 2023 Generative AI Measures continuing to serve as the backbone of the regulatory framework. While these Measures apply primarily to generative AI services provided to the public and therefore exclude internal enterprise use, R&D, and services directed solely at overseas clients, developments in 2024–2025 have deepened compliance expectations. These include new labeling requirements, data governance standards, and model security specifications.
- Push for Diffusion: While competition with the US continued to taint AI's strategic outlook, 2025 saw a massive push for the technology's diffusion across the economy - from education to government services, to R&D, manufacturing and several industries
- Geopolitical Considerations: 2025 saw US export controls retracted, expanded and diluted, and Chinese policy pushed for domestic capability building while differing from time to time on whether or not to import US chips. Hardware supply chain security came to the fore when China summoned Nvidia's representatives over reports of tracking devices in their chips, adding to further mistrust.
- State-led investments in AI: A national trillion yuan guidance fund for deep-tech was launched, semiconductor big fund 3 continued investments, and land subsidies and local government push across the AI stack continued



Sources: Administrative Measures for Generative AI Services, Internet Information Service Algorithm Recommendation Management Regulations, Administrative Provisions on Deep Synthesis Internet Information Services, Governance Principles for New Generation AI, Ethical Norms for New Generation AI, Opinions on Strengthening the Ethical Governance of Science and Technology, New Generation AI Development Plan, AI Standardization Guidelines, National Computing Network Coordination Plan, Computing Power Hub Plan, Personal Information Protection Law, Data Security Law, Cybersecurity Law

Analysis of AI Governance Measures in China

- AI companions surged in popularity in China with 2025 initiating a regulatory framework on humanised AI services
- Global AI Governance Plan and International Norm Setting: Calling for “global solidarity”, China continued to seek an active role in international AI governance, whether in standards, environmental management, or data sharing.
- Refer to **Appendix 3** for details about the specific governance measures.

Transparency & Accountability

- The 2022 regulation requiring all public-facing algorithmic recommendation service providers to report to a national filing system with an "algorithm self-assessment report" and other disclosure requirements on the data and model used. This continues to be ongoing.
- New measures for explicit and implicit metadata labeling of AI-Generated synthetic content. However, enforcement remains uneven.
- China’s central legislature approves major amendments to the Cybersecurity Law, adding dedicated AI governance provisions and strengthening penalties for data/security violations — scheduled to take effect 1 Jan 2026.

Promotion of Innovation

- Several policy initiatives aimed at building a data economy - annotating and labelling of data that China produces, building data companies, and promoting trading of this data through data exchanges
- AI+ initiative launched - AI for R&D, manufacturing, government services, etc. Integration of AI across sectors. Ambitious, unattainable targets set for local and provincial governments
- Embodied AI, supported by the 15th five year plan, is a key focus. The Ministry of Industry and Information Technology (MIIT) specifically named humanoid robots in its list of work priorities for 2025. And throughout the second half of 2025, the Chinese Institute of Electronics has been working on standards for the humanoid robots industry

Geopolitical Considerations

- Strong push - subsidies, land/ tax, etc - to incentivise domestic chip production. Domestic companies incentivised to buy local alternatives, and local chip equipment. Ban/ restrictions on foreign chips.
- Local governments actively incentivise data center set up - subsidised land, lower electricity rates, local compute mandates etc

Analysis of AI Governance Measures in China

- National Venture Capital Guidance Fund, of a trillion yuan, launched to invest in long-term (20 years), early and hard, technologies
- Antitrust and competition supervision (SAMR) used to check big tech power, market concentration used as a tool of strategic retaliation
- The Global AI Governance Action Plan, announced in July 2025, covers 13 commitments spanning innovation, ethical standards, security, and the creation of a World Artificial Intelligence Cooperation Organization (WAICO).
- China's Big Fund 3 launched in 2024 continued deployments into semiconductor chokepoint areas

Societal Wellbeing

- February 2025, CAC's list explicitly lists rectifying AI misuse (scams, impersonation, fraud) as a priority
- Ethical committees and review boards through the lifecycle of AI development being piloted under the Administrative Measures for the Ethical Management of AI Science and Technology
- Regulation of AI that interacts with citizens in a humanlike way - Measures for the Administration of Humanized Interactive Services Based on Artificial Intelligence (notified December 2025)
- Guidelines for the incorporation of Gen AI based learning published. AI skilling and exposure has been made mandatory for children from the age of six.

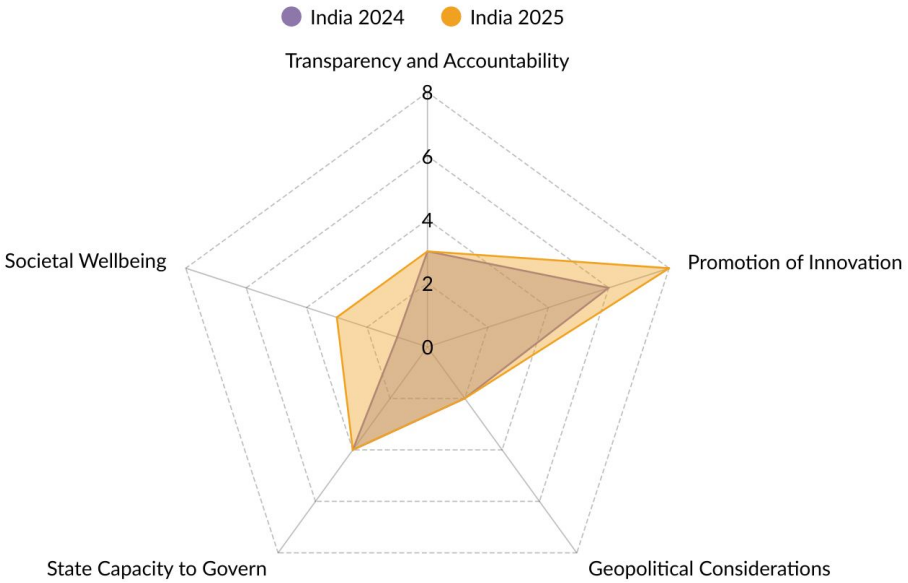
State Capacity to Govern

- China's top-down executive system means that the central government has high directive capacity to mobilise massive capital, human resources and bureaucracy machinery towards achieving its goals, and the political ability to absorb the costs incurred.
- The state has strong enforcement leverage over large platforms and state-funded projects with administrative penalties, procurement controls and cybersecurity reviews reinforcing compliance.
- However, enforcement is weak in some areas vs others - labelling of synthetic data (weak) vs filing mechanism for recommendation systems (high).
- The regulators' technical capacity to review datasets, algorithms filed, classify risk through the lifecycle of AI deployment especially for large-scale generative outputs and frontier models is still weak.

Analysis of AI Governance Measures in India

Trends in 2025

- Innovation and Self-Regulation: Continued prioritisation of innovation through self-regulation and voluntary disclosures. Instead of a single comprehensive AI law, sectoral regulators in areas such as finance and healthcare aim to address sector-specific high-risk use cases.
- De-risking Supply Chains: Leveraging its strength in semiconductor design, India is set to join Pax Silica, a trusted semiconductor supply chain initiative. The design and production-linked incentives under the semiconductor mission are focused on building long-term resilience in other parts of the semiconductor supply chain, such as fabrication, assembly and testing.
- Incentives for Sovereign Data Centres: 20-year tax holidays for data centres setting up shop in India aim to build data centre capacity in India. Several states are also providing incentives for data centres to increase domestic capacity.
- GPU Clusters and Priority Use Cases: Domestic GPU clusters from empanelled vendors subsidise compute for priority use cases such as indigenous models and applications across agriculture, healthcare and education. The focus is on multilingual models and applications that can bridge state capacity limitations in delivering public services. Digital public infrastructure is expected to transform access to public services at scale and is pitched as India's unique advantage.



Sources: India AI Governance Guidelines, IndiaAI Mission, DPDPA, ICMR Ethical Guidelines for Application of AI in Biomedical Research and Healthcare, RBI's framework for responsible and ethical enablement: Towards ethical AI in finance, CDSCO's Draft guidance document on Medical Device Software, SEBI's guidelines for responsible usage of AI/ML In Indian Securities Markets

Analysis of AI Governance Measures in India

- Progress in Establishing Institutional Oversight: Establishment of the AI Governance Group (AIGG) and the AI Safety Institute (AISI) to monitor standards and provide effective testing and safety protocols.
- The Digital Personal Data Protection Rules were notified. However, protections for personal data use in AI research are limited.
- Refer to Appendix 4 for details about the specific governance measures.

Transparency & Accountability

- India's AI regulations emphasise transparency and accountability through voluntary guidelines and existing laws, but enforcement remains light-touch and self-regulatory, prioritising innovation over mandates. Harms from outcomes are regulated through existing regulations.
- IT Rules draft Amendment 2025 (now notified) tackling synthetically generated information introduces mandatory labeling.
- SEBI issued guidelines for reporting AI/ML use by market participants, enhancing transparency in financial markets.

Promotion of Innovation

- The IndiaAI Mission has allocated over USD 1.2 bn to develop AI models, datasets, compute, and education.
- Portals like IndiaAI Compute Portal and expanded AIKosh aims to democratise access for startups/MSMEs, yielding 350+ BHASHINI models and hackathon wins, though scaling efficiencies lag due to fragmented procurement.

Geopolitical Considerations

- US-led export controls continue to limit India's access to cutting-edge chips and AI models, heightening supply chain vulnerabilities.
- IndiaAI Mission fast-tracked sovereign capabilities in domestic GPU clusters (38,000 subsidised units), open foundational models, and fab investments such as those for Tata-PSMC.
- Strategic partnerships and G20 advocacy leveraged for balanced norms, alongside state data centre incentives for drastically increasing capacity.

Analysis of AI Governance Measures in India

Societal Wellbeing

- Sector-specific guidelines from ICMR, RBI, and CDSCO (such as Class C/D medical devices) provide voluntary ethical frameworks for wellbeing in healthcare and finance.
- India AI Governance Guidelines introduce risk-based principles like fairness and safety, but no comprehensive law exists to uniformly address AI risks such as bias or misinformation.

State Capacity to Govern

- Finalised India AI Governance Guidelines establish the AI Governance Group (AIGG) and propose an AI Safety Institute (AISI) for real-time monitoring, testing, and standards, linking industry with policymakers.
- Capacity builds via IndiaAI Mission's training programmes, BIS standards, and sectoral regulators, though implementation relies on voluntary adoption and lacks dedicated enforcement funding.

Analysis of AI Governance Measures Across Companies

Analysis of AI Governance Measures Across Companies

- Companies are proactively adopting AI governance measures. These measures include developing AI principles, implementing risk mitigation strategies, enhancing transparency and establishing governance structures.
- Companies are tailoring their AI governance plans in compliance with regulatory requirements in the US and EU. There is an increased acknowledgement of the risks of misuse, limitations, and disruptions to the wider labour market from the widespread diffusion of generative AI.
- The tension between speed of innovation and robustness of safety mechanisms more often than not prioritises speed in a bid to be the first to market.
- Microsoft, Google, OpenAI, Mistral, Anthropic, Amazon, Accenture, and Deloitte are selected for the comparative analysis. These companies operate across different stages of the AI value chain, including big technology platforms, AI model developers, and technology services firms.
- The authors have not provided a comparative chart scoring the AI governance measures of these companies as there is much variance between the approaches of different companies that is hard to measure.
- The following slides in this section outline some of the AI governance initiatives by the different companies.

Sources: Microsoft's AI policy Open AI's policy OpenAI's Bounty Programme, Google's AI policy, Anthropic's Responsible Scaling Policy, Anthropic's Constitutional AI

Analysis of AI Governance Measures Across Companies

Principles: The extent to which the organisation's responsible AI policy is articulated and identifies the principles it seeks to adhere to. A comparison of the articulation of principles on responsible AI development and use by different companies is provided below.

Microsoft: Six ethical principles established in 2018: Fairness, Reliability & Safety, Privacy & Security, Inclusiveness, Transparency, and Accountability.
It has matured from these to Responsible AI Standard V2 providing actionable engineering requirements for its teams.

Google: Seven AI Principles (2018): Socially beneficial, avoid unfair bias, safety, accountability, privacy, scientific excellence, and usage limits. These continue.

OpenAI: Mission-led focus on building safe, beneficial AGI; values include Democratic Values, Safety, Responsibility, and Accountability.In 2025, it shifted to a "democratic vision" that is focused on freedom, individual rights and innovation.

Mistral: Mistral's responsible AI policy emphasises core values: Neutrality, empowering people through robust controls, and building trust via transparency.

Anthropic: An 80 page "Claude's Constitution" released in Jan 2026: Shifting from rules to reason-based alignment with a 4-tier hierarchy: safety, ethics, compliance, and helpfulness.

Amazon: Guided by eight priorities: Fairness, Explainability, Privacy/Security, Safety, Controllability, Veracity/Robustness, Governance, and Transparency.

Accenture: Committed to responsible design that prioritises ethics, transparency, accountability, and inclusivity.

Deloitte: Trustworthy AI™ framework based on seven dimensions: transparent/explainable, fair/impartial, robust/reliable, private, safe/secure, and responsible/accountable.

Sources: Microsoft's AI policy Open AI's policy OpenAI's Bounty Programme, Google's AI policy, Mistral's AI policy, Anthropic's Responsible Scaling Policy, Anthropic's Constitutional AI, Amazon's AI policy, Accenture's AI policy, Deloitte's AI policy

Analysis of AI Governance Measures Across Companies

Risk Mitigation: The extent to which policy identifies the potential risks and lists actions to mitigate against those risks. The risk mitigation efforts by different companies are listed below

Microsoft: The company implemented a "sensitive use review" program that conducted rigorous reviews of 600+ nove AI use cases by 2023 and further increased these in 2025.

Google: Guided by the Secure AI Framework and Frontier Safety Framework. Conducts red-teaming for election and national security concerns.

OpenAI: Employs internal and external red-teaming for CBRN (biological/ nuclear) and cyber risks prior to release. Maintains a dedicated Preparedness team.

Mistral: Implements Usage Policies against illegal acts and CSAM. Currently developing measures for EU AI Act compliance for non-high-risk systems ahead of 2026 deadlines.

Anthropic: Distinguishes between hardcoded prohibitions (e.g. bioweapons) and soft-coded defaults. Conducted 2025 pre-deployment testing with the UK AI Safety Institute.

Amazon: Offers 70+ responsible AI tools like Bedrock Guardrails. AWS Responsible AI Policy updated in Jan 2025, mandates risk evaluation for "consequential decisions".

Accenture: Deploys a four-pillar approach: Organisational, Operational, Technical, and Reputational. Uses an Algorithmic Assessment toolkit to help clients mitigate bias in financial and government models.

Deloitte: Uses a cross-functional approach to identify high-risk exposure areas across the AI life cycle. Bolsters trust using validating technology and monitoring solutions.

Sources: Microsoft's AI policy Open AI's policy OpenAI's Bounty Programme, Google's AI policy, Mistral's AI policy, Anthropic's Responsible Scaling Policy, Anthropic's Constitutional AI, Amazon's AI policy, Accenture's AI policy, Deloitte's AI policy

Analysis of AI Governance Measures Across Companies

Transparency and Reporting: The extent to which governance frameworks attempt to promote transparency and accountability. The measures by companies that promote transparency and accountability are listed below.

Microsoft: Publishes an annual AI Transparency Report. Released 33+ Transparency Notes for specific services since 2019.
It has implemented Automated Transparency logs that generate audit trails of model decisions, data inputs and user overrides to meet EU AI act requirements.
To distinguish between human led and AI generated content, they now automatically attach C2PA metadata to images and videos generated by models.

Google: Published AI Responsibility Reports annually since 2019. The latest report was published in Feb 2025. Uses Model Cards to document intended purpose and performance.

OpenAI: Commits to publicly reporting model capabilities, limitations, and safety evaluations for all significant releases.

Mistral: Publicly updates Terms of Service and Usage Policy (latest Feb 2025). Terms require users to disclose AI generation.

Anthropic: Published full constitution under "creative commons(2026) to set a precedent for industry disclosure.

Amazon: Participates in the US AI Safety Institute. Provides information on intended uses and policy compliance upon request.

Accenture: Collaborated with World Economic Forum on a 2024 playbook for turning governance principles into practice.

Deloitte: Recognized for its Omnia Trustworthy AI auditing module that provide ethical guardrails for its clients.

Sources: Microsoft's AI policy Open AI's policy OpenAI's Bounty Programme, Google's AI policy, Mistral's AI policy, Anthropic's Responsible Scaling Policy, Anthropic's Constitutional AI, Amazon's AI policy, Accenture's AI policy, Deloitte's AI policy

Analysis of AI Governance Measures Across Companies

Governance Structure: An estimation of the financial resources, institutional frameworks, and skilled human capital made available to enforce compliance with AI regulations effectively. A comparison of the governance structure of different companies is provided below.

Microsoft: Nearly 350 employees specialised in responsible AI by 2025. Includes Aether (advisory committee) and the Office of Responsible AI. Governance is now deeply prescriptive and integrated engine that is embedded directly into the Azure AI platform rather than a separate audit process.

Google: Governance process covers the full lifecycle: development, application deployment, and post-launch monitoring.

OpenAI: Formal evaluation process led by Product Policy and National Security teams. Expanded board oversight in March 2024. In Oct 2025, it proposed a "Classified Stargate" for US security.

Mistral: French limited joint-stock corporation with specific admin account features for workspace management.

Anthropic: Their governance framework is iterative and adaptable, incorporating lessons from high-consequence industries. It includes internal evaluations and external inputs to refine their policies. Signed the EU General-Purpose AI Code of Practice in July 2025, facilitating presumption of conformity for regulated sectors.

Amazon: Cross-functional expert collaboration across security, privacy, science, engineering, public policy, and legal teams. Calls it "governance by design".

Accenture: Establishment of transparent governance structures across domains with defined roles, expectations, and accountability. Creation cross-domain ethics committees. Has establishment a Chief Responsible AI officer role. This continues.

Deloitte: Deloitte AI Institute coordinates ecosystem dialogue. Documentation covers roles, responsibilities, and accountabilities throughout the AI lifecycle.

Sources: Microsoft's AI policy Open AI's policy, OpenAI's Bounty Programme, OpenAI's Structure, OpenAI Safety Concerns, Google's AI policy, Google AI Researcher Fired, Mistral's AI policy, Anthropic's Responsible Scaling Policy, Anthropic's Constitutional AI, Amazon's AI policy, Accenture's AI policy, Deloitte's AI policy

Analysis of AI Governance Measures Across Companies

Third-Party Oversight: The willingness to subject itself to third-party oversight. Examines policies that encourage third-party oversight to identify and report risks they might have overlooked.

Microsoft: Worked with NewsGuard to mitigate deep fake risks in text-to-image tools. Signed voluntary White House commitments in July 2023. Microsoft is a founding member of the Frontier Model Forum, to facilitate industry wide discussions on AI safety and responsibility.

Google: Collaborates with NGOs, industry partners, and experts at every stage. Founder of Frontier Model Forum. Contributes to the National AI research resource pilot.

OpenAI: Incentivises discovery through bug bounty systems. Member of the Frontier Model Forum for information sharing.

Mistral: Standardised DPA and SCCs for business customers. Relies on Partner Infrastructure (Azure, GCP) with shared responsibilities.

Anthropic: Allows external exploratory exploration of model capabilities by research nonprofits like METR.

Amazon: Member of the Frontier Model Forum and Partnership on AI. Facilitates third-party vulnerability reporting. Joined NIST AISIC

Accenture: No mention of third-party oversight.

Deloitte: Member of the NIST AISIC.

Sources: Microsoft's AI policy Open AI's policy OpenAI's Bounty Programme, Google's AI policy, Mistral's AI policy, Anthropic's Responsible Scaling Policy, Anthropic's Constitutional AI, Amazon's AI policy, Accenture's AI policy, Deloitte's AI policy

Analysis of AI Governance Measures Across Multi-Stakeholder Gatherings

Analysis of AI Governance Measures Across Multi-Stakeholder Gatherings

- Various multi-stakeholder gatherings, including the AI Summits and the Global Partnership on AI, have been established to raise awareness and coordinate international AI governance efforts.
- While state-level efforts have tended to focus on innovation and geopolitics, multi-stakeholder gatherings highlight broader societal concerns arising from the rapid development of advanced AI.
- Most gatherings do not have legally binding commitments or backing from all members (for instance, the US and EU refusing refused to sign the declaration on inclusive and sustainable AI at the AI Action Summit in February 2025).
- While organisations involving different countries are not making significant progress, public private partnerships such as NIST AISIC seems to be gaining credibility and interest.
- Knowledge sharing between companies on AI safety is gaining momentum with organisations like AI Frontier Forum taking the lead.
- Achieving alignment or convergence on AI regulations through these platforms can simplify compliance for multinational technology companies.
- The analysis in this section focuses on the membership composition, guiding principles, and recent developments in these gatherings.

Analysis of AI Governance Measures Across Multi-Stakeholder Gatherings

The Organization for Economic Co-operation and Development

Membership:

- OECD has 38 member countries committed to democracy, collaborating on addressing global policy changes, and is not an AI-specific body.

Principles and areas of focus:

- OECD promotes inclusive growth, human-centric values, transparency and explainability, robustness and accountability of AI systems.
- The OECD AI Principles are the first intergovernmental standard on AI.

Global Partnership on AI

Membership:

- GPAI has 44 member countries, including the US, EU, UK, Japan, and India.

Principles and areas of focus:

- GPAI promotes the responsible development of AI grounded in human rights, inclusion, diversity, innovation, and economic growth
- Areas of focus include responsible AI, data governance, the future of work, and innovation and commercialisation.

Developments:

- As of 2024, GPAI and the OECD formally joined forces to combine their work on AI and implement human-centric, safe, secure, and trustworthy AI. The two bodies are committed to implementing the OECD Recommendation on Artificial Intelligence.

Sources: OECD AI Principles, GPAI

Analysis of AI Governance Measures Across Multi-Stakeholder Gatherings

AI Governance Alliance

Membership:

- The AI governance alliance is a global initiative launched by the World Economic Forum. The alliance has over 603 members from more than 500 organisations globally.

Principles and areas of focus:

- The principles of the AI Governance Alliance include responsible and ethical AI, inclusivity, transparency, international collaboration and multi-stakeholder engagement.
- The areas of focus include safe systems and technologies, responsible applications and transformation, resilient governance and regulation.

AI Summits

Membership:

- The AI summits are a series of international conferences addressing the challenges and opportunities presented by AI. Participants include heads of state and major companies such as Meta and DeepMind.

Principles and areas of focus:

- Each AI summit has set its own agenda, but some common principles are ethical AI development, safety and security, transparency and accountability, and international collaboration.
- The AI Summits have been held thrice since their inception. The first summit, focussing on AI safety, was held at Bletchley Park in the UK in 2023. The second summit was held in Seoul, South Korea, in 2024.
- The third event, the AI Action Summit was held in Paris in February 2025 and was attended by representatives from more than 100 countries. While 58 countries, including France, China and India, signed a joint declaration, the US and UK refused to sign the declaration on inclusive and sustainable AI.
- The agenda has evolved from existential risks and global cooperation at Bletchley, to risk management frameworks and company commitments at Seoul to an action oriented focus on public interest, sustainability and global governance at Paris.

Sources: AI Action Summit, AI Seoul Summit, Bletchley Declaration, AI Governance Alliance

Analysis of AI Governance Measures Across Multi-Stakeholder Gatherings

United Nations

Membership:

- The UN is an international organisation committed to global peace and security, with 193 member states, including almost all internationally-recognised sovereign states. The safe development of AI is one of their many areas of work.

Principles and areas of focus:

- Some of their core principles include doing no harm. AI applications should have a clear purpose, fairness and non-discrimination, safety and security to prevent misuse and harm, responsibility and accountability.
- The UN Secretary-General is convening a multi-stakeholder High-level Advisory Body on AI to study and provide recommendations for the international governance of AI.
- Other efforts include convening global dialogues, developing standards and building capacity.

United Nations Educational, Scientific and Cultural Organization

Membership:

- UNESCO is a specialised agency of the UN with 194 member states and 12 associate member states

Principles and areas of focus:

- The UNESCO general conference adopted the recommendation on the ethics of artificial intelligence – the first global standard on AI ethics principles aligned with the UN's principles on AI.
- Areas of focus include developing an AI Readiness Assessment Methodology, facilitating policy dialogues and capacity building initiatives.

Sources:UN Ethical AI Principles, UN AI Advisory Body, UNESCO Recommendations on Ethics of AI

Analysis of AI Governance Measures Across Multi-Stakeholder Gatherings

AI Frontier Forum

Membership:

- The current members of the forum are Amazon, Anthropic, Google, Meta, Microsoft and Open AI. It is also known as Frontier Model Forum

Principles and areas of focus:

- Promote AI safety
- Facilitate information sharing between companies and governments
- They have initiated a small AI safety fund of $10M

NIST AISIC

Membership:

- Set up by the National institute of Standards and Technology (NIST) as the AI Safety Institute Consortium (AISIC). A US government initiative.
- Has over 280 members including US based researchers, academic teams, AI creators and civil society organizations. Includes companies like NVIDIA, Amazon, Microsoft.

Principles and areas of focus:

- Risk Management for generative AI
- Synthetic content development
- Capability evaluation
- Red teaming
- Safety & Security

Appendix

1. Governance Measures in the US

Transparency and Accountability

Eliminated Measures:

- Disclosure requirements for large foundational models (10^26 FLOPs) - REMOVED with EO 14110 revocation
- NIST safety testing standards development - Discontinued
- Algorithm filing and transparency provisions - No federal requirement
- Mandatory impact assessments for high-risk systems - Limited to OMB ACT federal use only

Remaining Federal Measures:

- OMB Act: Risk management standards for AI in federal government high-risk systems affecting civil rights
- DOE Evaluation: Tools for identifying AI model risks in nuclear/biological threat domains
- IaaS Monitoring: Reporting by international customers for large model training runs (though enforcement unclear post-AI Diffusion Rule rescission)

State-Level Measures (Examples):

- Colorado AI Act: Documentation requirements, impact assessments, appeal processes for adverse decisions
- California SB 53: Quarterly compliance reports, incident reporting for frontier models
- Texas TRAIGA: Disclosure requirements for AI use in healthcare, biometric systems
- New York RAISE Act: Transparency and risk safeguards for frontier models (pending governor signature)

Promotion of Innovation

AI Action Plan (July 2025) - Key Components:

Pillar I: Accelerating Innovation

- Increased federal investment in AI R&D
- Public-private partnerships for AI development
- Support for commercialisation of cutting-edge AI
- Promotion of open-source and open-weight AI
- Interoperability standards development
- NIST Directive: Revise AI Risk Management Framework to remove DEI references
- Patent Guidance: Maintain inventorship requirements for natural persons
- AI Evaluation Ecosystem: Build infrastructure for assessing AI reliability in regulated industries

Sources:Trump EO on AI, Biden EO on AI, IAPP - US State AI Legislation Tracker, BCLP Law - US State AI Legislation Snapshot, NIST Risk Framework, AGORA, AI Diffusion Rule

1. Governance Measures in the US

Promotion of Innovation

Pillar II: Building American AI Infrastructure

- **Data Centre Acceleration (EO 14318):**
  - Qualifying Projects: >$500M investment or >100MW electric load
  - Expedited NEPA reviews and categorical exclusions
  - Federal land designation for AI facilities
  - Preferential permitting and siting
  - Federal financing tools: loans, loan guarantees, grants, tax incentives, offtake agreements
- **Chip Manufacturing:**
  - Revitalising domestic semiconductor production
  - CHIPS Act implementation continues
  - Reducing dependence on foreign supply chains
- **Power and Energy:**
  - Streamlined approval for power generation and transmission
  - Focus on reliable, affordable energy for AI infrastructure
  - Nuclear, fusion, and grid modernisation prioritisation

Pillar III: International Diplomacy and Security

- Export of "American AI Technology Stack" to allies
- Alignment of partner country standards with US frameworks
- Trade policy integration with AI leadership
- Competitive positioning against China

Genesis Mission (November 2025):

- **Leadership:** Department of Energy (Under Secretary Darío Gil)
- **Scope:** DOE's 17 National Laboratories + industry + academia
- **Platform:** Integrated AI discovery platform connecting:
  - World's most powerful supercomputers
  - Advanced scientific instruments
  - Next-generation quantum systems
  - Vast scientific datasets
- **Goals:**
  - Double American science/engineering productivity within decade
  - Create world's most powerful scientific instrument
  - Focus areas: US energy dominance, discovery science, national security
- **Priority Domains:**
  - Advanced manufacturing
  - Biotechnology
  - Critical materials
  - Nuclear energy
  - Quantum information science
  - Semiconductors and microelectronics
- **Partnerships:** 24 organisations signed initial MOUs; ongoing RFIs for additional collaborators
- **Reporting:** Annual progress reports on operational status, scientific advances, partnership outcomes

Sources: Trump EO on AI, Biden EO on AI, IAPP - US State AI Legislation Tracker, BCLP Law - US State AI Legislation Snapshot, NIST Risk Framework, AGORA

1. Governance Measures in the US

Geopolitical Considerations

Export Control Evolution:

AI Diffusion Framework (January 2025 - Rescinded May 2025):

- Created three-tier global system for AI chip exports
- Tier 1 (US + 18 allies): Unrestricted
- Tier 2 (~150 countries): Quota-based through DC VEU program
  - Universal VEU authorisation: 75% compute in Tier 1 countries, max 7% per Tier 2 country
  - National VEU authorisation: ~100,000 H100-equivalents per country (2025), scaling to 320,000 (2027)
  - Country caps: 50,000 H100-equivalents via standard licenses
- Tier 3 (China, Russia, etc.): Prohibited
- Status: Rescinded before May 15, 2025 implementation date

Current Export Control Status:

- Traditional licensing requirements continue
- Ad-hoc approach rather than systematic framework
- Focus on using chip access as trade/diplomatic leverage
- Huawei Guidance (May 2025): Warned about Chinese chip use globally (later softened after China objections)
- EDA Software: Licensing restrictions removed as part of US-China trade negotiations (June 2025)

Strategic Initiatives:

- Outbound Investment Security Rule: US Treasury prohibits transactions involving AI for military/surveillance in adversary countries
- Supply Chain Vulnerability Assessments: NSC, ODNI, DOD, DOJ identifying threats to AI ecosystem
- "America First" Doctrine: Explicit prioritisation of US dominance over international cooperation
- Technology Stack Exports: Leveraging AI as diplomatic tool with allies

Sources: Trump EO on AI, Biden EO on AI, IAPP - US State AI Legislation Tracker, BCLP Law - US State AI Legislation Snapshot, NIST Risk Framework, AGORA

1. Governance Measures in the US

Societal Wellbeing

Federal Level - Limited:

TAKE IT DOWN Act (May 19, 2025):

- **Criminal Provisions:**
  - Prohibition: Knowingly publishing or threatening to publish NCII (including AI-generated)
  - Penalties: Up to 2 years imprisonment (adults), increased penalties for minors
  - No consent defence: Prior consent to create ≠ consent to publish
  - Scope: Both authentic and "digital forgery" (AI-generated) intimate imagery

- **Platform Requirements (Effective May 19, 2026):**
  - 48-hour takedown requirement upon valid notice
  - Must remove known identical copies
  - Clear, accessible reporting process required
  - FTC enforcement: civil penalties up to ~$51,000 per violation

- **Exceptions:**
  - Good faith disclosure for law enforcement, medical treatment
  - Access/connection providers (Communications Decency Act defense)

Federal Procurement:

- OMB Act: High-risk AI systems in federal use must address civil rights impacts
- HHS guidance: AI in public benefits delivery must ensure transparency, mitigate risk

State-Level Activity (Primary Locus):

Deepfakes and NCII:

- Laws addressing deepfakes in political communications
- Deepfake-related bills enacted
- Many states criminalised AI-generated child sexual abuse material (CSAM)
- Mix of disclosure requirements (preferred due to First Amendment) and prohibitions

Algorithmic Discrimination:

- Colorado AI Act: Reasonable care duty for developers/deployers to protect against discrimination
  - Impact assessments required
  - Appeal processes for adverse decisions
  - Effective June 30, 2026 (delayed from February 1, 2026)
- Illinois HB 3773: Prohibits employer use of AI discriminating against protected classes
- Connecticut SB 1295: Amendments to existing data privacy law's automated decision-making provisions
- New Jersey: Ongoing rulemaking on automated decision-making systems

Healthcare AI:

- Illinois: Prohibits AI from making therapeutic decisions, interacting with clients, creating treatment plans
- Indiana: Requires disclosure to patients if AI used in healthcare decisions
- Texas SB 1188: Permits AI for diagnostic purposes with practitioner review and patient disclosure

1. Governance Measures in the US

Societal Wellbeing

Privacy and Consumer Protection:

- State privacy laws (CCPA, CPPA, others) updated for AI-specific provisions
- California CPPA Regulations: Cybersecurity audits, risk assessments for ADMT (proposed May 2025)
- Oregon SB 619: Rules for profiling and automated decision-making
- Montana SB 384, Utah amendments: Personal information protection in AI contexts

Content Labeling and Disclosure:

- California AB 2013: Disclosure of data used in AI training
- California SB 942: AI content generation disclosure requirements
- California AB 2355: GenAI use in electoral advertisements
- California SB 243: Companion chatbot disclosures, self-harm prevention protocols
- Tennessee ELVIS Act: Prohibits AI mimicking person's voice

State Capacity to Govern

Federal Level:

Reduced Capacity:

- National AI Talent Surge (from Biden EO 14110) - DISCONTINUED
- NIST AI standards development - SCOPE REDUCED, focus shifted to "neutral" frameworks
- Federal agency AI maturity assessments - LIMITED implementation under OMB Act
- FTC enforcement capacity - WEAKENED by Commissioner terminations

New Infrastructure:

- Genesis Mission: Massive federal commitment through DOE
    - Under Secretary for Science as Mission Director
    - Assistant to President for Science and Technology (APST) provides general leadership
    - National Science and Technology Council (NSTC) coordination
    - 17 National Laboratories mobilised
    - Standardised partnership frameworks for industry/academia collaboration

1. Governance Measures in the US

State Capacity to Govern

Federal Level:

New Infrastructure (continued):

- **DOE Platform Development:**
  - Integrated supercomputing resources
  - AI model training infrastructure
  - Robotic laboratories with AI-directed experimentation
  - Secure, unified platform for federal scientific data

- **Coordination Mechanisms:**
  - Special Advisor for AI and Crypto (David Sacks)
  - Assistant to President for Science and Technology (Michael Kratsios)
  - Economic Diplomacy Action Group (EDAG) for AI exports
  - AI Litigation Task Force (DOJ) - for challenging state laws

Agency Roles:

- NIST: Standards development, AI Risk Management Framework (under revision), evaluation ecosystem
- DOE: Genesis Mission leadership, national laboratory coordination, energy infrastructure
- Commerce: State law evaluation, BEAD funding decisions, export controls
- FTC: TAKE IT DOWN Act enforcement, guidance on "deceptive AI models"
- FCC: Directed to consider federal reporting/disclosure standards for AI models
- OMB: Federal AI procurement guidance, "Preventing Woke AI" implementation

State Capacity to Govern

State-Level Capacity Building:

- 34 states studying AI governance
- 24 states created dedicated AI study groups
- 10 states delegated to standing committees
- State attorneys general developing enforcement expertise
- Likely coalition of states to challenge federal preemption efforts

Funding:

- Federal: Genesis Mission (exact figures undisclosed, estimated multi-billion)
- Federal: CHIPS Act implementation continues (~$50B authorised)
- Federal: Infrastructure investments via EO 14318 (scale dependent on qualifying projects)
- State: Varies significantly; California leading state-level investment

Challenges:

- Policy instability creating compliance uncertainty
- Federal-state conflict consuming resources
- Weakened regulatory agencies at federal level
- Fragmented approach across 50 state jurisdictions
- International partners unsure of US policy direction
  -

2. Governance Measures in the EU

Transparency and Accountability

- The EU AI Act classifies AI systems into tiers of unacceptable, high, limited, and minimal risk categories. Unacceptable-risk systems are banned, high-risk systems are subject to higher compliance requirements, limited-risk systems are subject to transparency obligations, and minimal-risk systems are unregulated.

Tiering: High-risk AI Systems:

- Biometric identification and categorisation of natural persons
- Management and operation of critical infrastructure
- Education and vocational training
- Employment, worker management, and access to self-employment
- Access to essential private services and public services and benefits
- Law enforcement
- Migration, asylum and border control management
- Administration of justice and democratic processes

Tiering: Prohibited AI Practices

- Using subliminal techniques or manipulative methods to distort behaviour, impair decision-making, and cause significant harm.
- Exploiting vulnerabilities of individuals or groups (e.g., based on age, disability, or socioeconomic status).
- Evaluating individuals or groups based on social behaviour or personal characteristics that leads to detrimental treatment.
- Criminal risk prediction that is based on profiling or personality traits, except when supporting objective human assessments.
- Creating facial recognition databases through untargeted scraping of images from the internet or CCTV footage.
- Inferring emotions in workplaces or educational institutions, except for medical or safety purposes.
- Biometric categorisation systems to infer sensitive attributes (e.g., race, political opinions, religious beliefs, sexual orientation), except for law enforcement.
- Real-time remote biometric identification systems for law enforcement in public spaces, with some exceptions.

Sources: EU AI Act, EU InvestAI Initiative, EU Chips Act, AI Continent Action Plan, AI Factories, AI Pact, Apply AI Strategy, Artificial Intelligence in Health, AI Act Service Desk, AI Act Single Information Platform, Guidelines on Prohibited AI Practices, Digital Omnibus on AI, European AI Office, European Data Union Strategy, The General-Purpose AI Code of Practice, AGORA

2. Governance Measures in the EU

Transparency and Accountability: Tiering

- General-purpose AI systems trained on large amounts of data using self-supervision at scale are subject to additional requirements, including copyright directives.
- General-purpose AI systems using beyond 10^25 FLOPs in training or evaluated as having a high impact are classified as those with systemic risks. These have additional evaluation and disclosure requirements.

Transparency and Accountability: Evaluation and Performance Requirements

- High-risk AI systems must
    - have a comprehensive risk management system throughout the system's lifecycle. Post-market monitoring systems must also be established proportionate to the nature of the risks involved. They must also meet applicable cybersecurity requirements.
    - incorporate appropriate human oversight to prevent or minimise the risks to health, safety or fundamental rights.
    - implement a quality management system to ensure regulatory compliance.
    - implement technical solutions to safeguard against attacks trying to manipulate the training data set (data poisoning), or pre-trained components used in training (model poisoning), inputs designed to cause the AI model to make a mistake (adversarial examples or model evasion), confidentiality attacks or model flaws
- Models with systemic risks are required to assess and mitigate risks using standardised protocols and tools and meet adequate levels of cybersecurity protection.
- Data sets used for training, validation and testing must meet quality criteria and data governance and management practices.
- General-Purpose AI Code of Practice
    - It provides a practical compliance toolkit for providers of GPAI models.
    - Published on 10th July, 2025 and signed by most leading AI providers. Meta and Apple are significant holdouts.
    - It has three chapters on transparency, copyright and safety and security.
    - While it is voluntary, it offers a way to demonstrate compliance with their obligations under the AI Act. The safety and security section is only applicable for providers of advanced models with systemic risks.

2. Governance Measures in the EU

Transparency and Accountability: Licensing and certification

- High-risk AI systems must ensure compliance with certification requirements by drawing up an EU declaration of conformity, affixing the CE marking to indicate regulatory conformity, and fulfilling registration obligations in the EU database.
- Developers and deployers must take corrective actions, provide necessary information, demonstrate conformity upon request by national authorities, and ensure that accessibility requirements meet EU directives.
- General-purpose AI systems using beyond 10^25 FLOPs in training or evaluated as having a high impact are classified as those with systemic risks. These have additional evaluation and disclosure requirements.

Transparency and Accountability: Disclosure

- Providers of general purpose AI systems must keep technical documentation of the model, including its training and testing process and the results of its evaluations. Information and documentation of the capabilities and limitations must be shared with stakeholders in the value chain.
- Providers of general purpose AI systems with systemic risks must report information about serious incidents and possible corrective measures to address them.
- Providers of AI systems must ensure that natural persons are informed that they are interacting with an AI system.
- Outputs of AI systems generating synthetic audio, image, video or text content must be marked in a machine-readable format and detectable as artificially generated or manipulated.
- In addition to the above, providers or deployers of High-risk AI systems must
    - have technical documentation that demonstrates that the system complies with the requirements
    - have record-keeping to ensure traceability of the functioning of the system
    - ensure communication of its use when deployed in the workplace.

Transparency and Accountability: Fines and Penalties

- Authorities shall lay down the rules on penalties and other enforcement measures, which may include warnings and non-monetary measures applicable to infringements of this Regulation.
- Administrative fines can go up to €15,000,000 or 3% of worldwide annual turnover, whichever is higher. Violations of prohibited uses can be fined up to €35,000,000 or 7% of worldwide annual turnover, whichever is higher.

2. Governance Measures in the EU

Transparency and Accountability: Input Controls

- High-risk AI systems are
  - required to be trained and tested on data reflecting their intended geographical, behavioural, contextual, or functional settings.
  - allowed to process special categories of personal data for the purpose of ensuring bias detection and correction concerning high-risk AI systems.
  - allowed to transfer data collected during real-world testing to third countries only provided that appropriate and applicable safeguards are implemented.
- The processing of personal and non-personal data must also comply with the General Data Protection Regulation (GDPR).

Transparency and Accountability: Other Measures

- AI Act Service Desk: To ensure rules are followed, the Commission launched a Service Desk with a team of experts and an online tool allowing stakeholders to submit questions and receive guidance, enhancing the administration's capacity to support implementation. These help stakeholders determine their legal obligations and ensure that products entering the market comply with safety standards.
- The AI Pact: It is a voluntary initiative that encourages organisations to plan ahead for AI Act implementation. Participants can create a collaborative community that shares best practices and pledges to implement transparency measures ahead of the legal deadline.
- Compliance Checker & AI Act Explorer: Digital tools available on the Single Information Platform that help stakeholders determine their legal obligations and browse the legislation to understand compliance requirements

Promotion of Innovation: Governance

- Establishment of AI regulatory sandboxes that provide for a controlled environment that facilitates development, testing and validation before deployment. This aims to improve compliance, share best practices, contribute to regulatory learning, and foster innovation. Startups are to be given priority access to AI regulatory sandboxes.
- Establishment of the European AI Office that heads regulatory efforts, an AI Board with representation from all states, a multi-disciplinary advisory forum, and a scientific panel of independent experts and relevant national authorities.
- The AI Act calls for standard development covering requirements for high-risk applications and compliance and disclosure requirements. This aims to provide legal certainty, competitiveness, and growth of the Union market and strengthen global cooperation on standardisation.

2. Governance Measures in the EU

- Models released under free and open-source licenses are exempt from certain evaluation, disclosure and compliance requirements.

Promotion of Innovation: Fines and Penalties

- In the case of SMEs, including start-ups, fines can go up to the lower of the percentages or amounts specified.

Promotion of Innovation: Funding

- AI Factories and Gigafactories: A major initiative to create "dynamic ecosystems" that bring together computing power, data, and talent to train cutting-edge AI models and applications.
- They foster collaboration across Europe, linking supercomputing centres, universities, small and medium-sized enterprises (SMEs), industry, and financial actors. AI Factories serve as hubs driving advancements in AI applications across various sectors such as health, manufacturing, climate, finance, space, and more.
- Through 2025-2026, at least 15 AI Factories and several Antennas (associated with AI-optimised supercomputers in existing AI Factories) are expected to be operational, enabling the pan-EU AI ecosystem and promoting growth by prioritising access for AI startups and SMEs. In this context, at least 9 new AI-optimised supercomputers will be procured and deployed across the EU. This will more than triple the current EuroHPC AI computing capacity.
- InvestAI Initiative aims to mobilise €200 billion for investment in AI, including a new European fund of €20 billion for AI gigafactories (CERN for AI).
- InvestAI also includes funding for Common European Data Spaces, a single market for data that will make more data available for access and reuse. This includes a number of strategic fields, such as health, agriculture, manufacturing, energy, mobility, finance, public administration, and skills.

- Under the InvestAI initiative, the following activities are planned:
  - mobilise € 200 billion towards an open, collaborative development of the most complex AI models and to make Europe an AI continent. € 20 billion of the InvestAI fund is towards building data centres (called AI gigafactories) that will be specialised in training the most complex, very large AI models
  - funding for Common European Data Spaces, a single market for data that will make more data available for access and reuse. This includes a number of strategic fields, such as health, agriculture, manufacturing, energy, mobility, finance, public administration, and skills.

2. Governance Measures in the EU

- strengthen the EU's generative AI talent pool through education, training, skilling and reskilling activities.
- support public and private investments in AI startups and scale-ups through venture capital or equity support.
- ‘GenAI4EU' initiative, which aims to support the development of novel use cases and emerging applications in Europe's 14 industrial ecosystems and public sector.

Geopolitical Considerations

- Models released under free and open-source licenses are exempt from certain evaluation, disclosure and compliance requirements.
- The InvestAI initiatives on building AI gigafactories and data spaces also address geopolitical considerations.
- The EU Chips Act aims to strengthen competitiveness and resilience in semiconductor technologies and applications. Over €43 billion of policy-driven investment will support the Chips Act until 2030, which will be broadly matched by long-term private investment.
- Apply AI Strategy: This strategy aims to boost AI adoption across 10 key industrial sectors (such as healthcare, robotics, and manufacturing) and the public sector. It is designed to enhance the competitiveness of strategic sectors and strengthen the EU’s technological sovereignty.
- It aims to boost AI adoption and innovation across Europe, particularly among Small and Medium-sized Enterprises (SMEs). The Strategy encourages an AI-first policy where AI is considered as a potential solution whenever organisations make strategic or policy decisions, taking into careful consideration the benefits and the risks of the technology. The Apply AI also promotes a 'buy European' approach, particularly for the public sector, with a focus on open source AI solutions.
- Data Sovereignty: The Data Union Strategy includes a priority to safeguard the EU’s data sovereignty by countering unjustified data localisation or leakage and ensuring international data flows are secure and consistent with EU values.

2. Governance Measures in the EU

Societal Wellbeing: Evaluation, disclosure and bans

- Disclosure obligations for manipulated content that might constitute a deepfake.
- Voluntary codes of conduct for assessing and minimising the impact of AI systems on environmental sustainability, such as energy-efficient programming and techniques for the efficient design, training and use of AI.
- Bias detection and evaluation is mandated for high-risk systems where it can impact health and safety and have a negative impact on fundamental rights or discrimination under the law.
- Prohibition of detrimental or unfavourable treatment of certain natural persons or groups of persons that are disproportionate to their social behaviour or unrelated to the contexts in which the data was originally generated.
- Ban on emotion recognition systems in the workplace or in educational institutions.
- AI for Health and Societal Good: Initiatives like the European Cancer Imaging Initiative and 1+ Million Genomes Initiative use AI to improve disease prevention and treatment outcomes.
- Bias Detection and Correction: The Digital Omnibus proposes a legal basis allowing providers to process special categories of personal data (e.g., ethnicity) specifically to detect and correct bias in AI systems, protecting citizens from discrimination.
- AI for Societal Good Unit: A dedicated unit within the European AI Office focuses on ensuring AI benefits society.
- AI Literacy: A proposed amendment requires Member States and the Commission to actively foster AI literacy among staff and the public to ensure people can use AI safely and effectively.

State Capacity to Govern

- European AI Office: A new body within the Commission serving as the centre of AI expertise. The Digital Omnibus proposes centralising oversight of complex AI systems (like those built on general-purpose models) within the AI Office to avoid fragmentation and ensure consistent enforcement across Member States.
- AI Act Service Desk: A team of experts and an online tool allowing stakeholders to submit questions and receive guidance, enhancing the administration's capacity to support implementation.
- The AI Act’s governance will be steered by three advisory bodies: the European Artificial Intelligence Board, composed of representatives from the EU Member States, the Scientific Panel, composed of independent experts in the field of AI, and the Advisory Forum, representing a diverse selection of stakeholders, both commercial and non-commercial.

3. Governance Measures in China

Transparency and Accountability

- The Administrative Measures for Generative AI Services (2023) specifies algorithm filing and security assessment requirements before service launch
- Internet Information Service Algorithmic Recommendation Management Provisions (2022) mandates algorithm transparency provisions - key parameters and operation principles must be explainable.
- Administrative Provisions on Deep Synthesis Internet Information Services (2023) provide for watermarking and labelling requirements for AI-generated content.
- Amendments to the 2017 cybersecurity law elevated AI governance from regulation to legislation level, with heavier penalties for violations - both monetary and administrative, personal data protection, and legal obligations to ensure cybersecurity compliance across the supply chain.
- Opinions on Strengthening the Ethical Governance of Science and Technology (2022) mandate ethics review requirements for high-risk AI research projects. These also mandate proper whistleblower protection mechanisms for reporting ethical violations
- Provisions on the Management of Algorithmic Recommendations in Internet Information Services 2021 requires algorithm filing and maintenance of a national registry database
- Measures for Labeling of AI-Generated Synthetic Content 2026 requires the explicit and implicit labelling of data
- Decision of the Standing Committee of the National People's Congress on Amending the Cybersecurity Law of the People's Republic of China amended the 2017 law to include higher penalties and stricter compliance

Sources: Administrative Measures for Generative AI Services, Internet Information Service Algorithm Recommendation Management Regulations, Administrative Provisions on Deep Synthesis Internet Information Services, Governance Principles for New Generation AI, Ethical Norms for New Generation AI, Opinions on Strengthening the Ethical Governance of Science and Technology, New Generation AI Development Plan, AI Standardization Guidelines, National Computing Network Coordination Plan, Computing Power Hub Plan, Personal Information Protection Law, Data Security Law, Cybersecurity Law, Amendments to the Cybersecurity Law

3. Governance Measures in China

Promotion of Innovation

- Opinions of the State Council on Deepening the Implementation of the "Artificial Intelligence+" Action was a key milestone and pushed for integration of AI across multiple industries
- Opinions of the National Data Administration and other departments on promoting the development and utilization of enterprise data resources, Notice on Issuing the "Pilot Program for Full-Process Management of Data Assets", Implementation Opinions of the National Development and Reform Commission and Other Departments on Promoting the High-Quality Development of the Data Labelling, Guidelines for the Construction of National Data Infrastructure - all attempt to build China's data economy
- The Chinese Institute of Electronics held a symposium on humanoid robot standardization—government, industry, academia, research, and application stakeholders jointly discussed urgently needed standards to promote the high-quality development of the humanoid robot industry.
- Embodied AI emerged as a key area of government focus. The Government Work Report 2025 explicitly mentioned embodied AI for the first time, placing it alongside longstanding tech aspirations like quantum and 6G. The Ministry of Industry and Information Technology (MIIT) specifically named humanoid robots in its list of work priorities for 2025.

Geopolitical Considerations

- National Computing Network Coordination Plan (2022) aims to establish regional computing centres linking eastern and western regions of China.
- Global AI Governance Action Plan (July 2025) Announced at the World AI Conference calls for international cooperation, sovereignty-respecting governance, and global standards. It covers infrastructure, open ecosystems, data, and safety principles.
- National Computing Network Coordination Plan (2022) provides support for indigenous computing technology with preferential policies. 2025 saw domestic big tech encouraged to use homegrown compute and cloud services.

Societal Wellbeing

- Interim Measures for the Administration of Humanized Interactive Services Based on Artificial Intelligence tried to control emotional dependence on human-like chat interfaces
- Guidelines to promote the use of AI in education from six years of age
- Content filtering and censorship requirements under the Administrative Measures for Generative AI Services (2023) - Providers must implement real-time monitoring of AI-generated content and prevent the generation of illegal content.
- Under the Administrative Measures for Generative AI Services (2023), there are data privacy protections - services must obtain explicit consent before using personal information.

4. Governance Measures in India

Transparency and Accountability, Social Wellbeing

- SEBI mandates reporting for AI/ML use by market participants, enhancing transparency in financial markets
- CDSCO Directive on AI/ML Software: Issued in 2025, it requires manufacturing/import licenses, Class C classification for diagnostic AI (moderate-high risk), Indian patient validation, quality management systems, and post-market surveillance for adverse events.
- IndiaAI Governance Guidelines to enable safe and trusted AI innovation
- IT Rules Amendment 2025 regulating synthetically generated information
- Digital Personal Data Protection Rules implemented. There is an 18-month implementation period overall. The Act does not apply to publicly available data. Additionally, there are exemptions for personal data being used for "research, archiving or statistical purposes".

Sources: India AI Governance Guidelines, IndiaAI Mission, DPDPA, ICMR Ethical Guidelines for Application of AI in Biomedical Research and Healthcare, RBI's framework for responsible and ethical enablement: Towards ethical AI in finance, CDSCO's Draft guidance document on Medical Device Software, SEBI's guidelines for responsible usage of AI/ML In Indian Securities Markets

4. Governance Measures in India

Promotion of Innovation

- Financed 30+ AI apps via hackathons for sectors like healthcare, agriculture, and governance, with subsidised compute access
- 6.6%, or ₹689.05 crore of IndiaAI's budget is allocated for the Application Development Initiative.
- IndiaAI initiative to build foundational AI models trained on Indian datasets such as Bhashini, BharatGen.
- 19%, or ₹1,971.37 crore of IndiaAI's budget is allocated for the innovation centre. This IAIC will develop and deploy of indigenous Large Multimodal Models (LMMs) and domain-specific foundational models in critical sectors.
- 44%, or ₹4,563.36 crore, of the ₹10,371.92 crore ($1.25 billion) approved by the cabinet for the IndiaAI Mission in March is earmarked for providing compute capacity of more than 38,000 GPUs for start-ups, researchers, students and academics over a period of five years.
- Stakeholder consultations held for the Draft National Data Centre Policy (not yet released). Policy expected in 2026. Focus on tax holidays for data centre developers, streamlined national and state policies, provisions/incentives for green power use and sustainability, provisions for edge data centres.
- IndiaAI has established 27 state-of-the-art AI & Data Labs across NIELIT Centres in Tier 2 and Tier 3 cities.
- In partnership with National Institute of Electronics & Information Technology, IndiaAI offers four NCVET-recognised foundational certification courses on data annotation, data curation, AI, and data science
- 1.9%, or ₹199.55 crore of IndiaAI's budget, has been allocated for the Datasets Platform through which the government will make public sector datasets AI-ready and give real time access to data through APIs wherever

Sources: IndiaAI Mission, Industry Consultation Meeting of National Data Centre Policy 2025

Acknowledgements and Disclosure

The authors would like to thank their colleagues, Rijesh Panicker and Pranay Kotasthane, for their valuable feedback and suggestions.

The authors also extend their appreciation to the creators of AGORA, an exploration and analysis tool for AI-relevant laws, regulations, standards, and other governance documents by the Emerging Technology Observatory.

The authors acknowledge the use of generative AI tools such as NotebookLM, Claude and Grammarly for assistance in analysis and copy-editing.