TAKSHASHILA INSTITUTION

State of AI Governance, 2024

A report analysing AI governance measures across countries, companies, and multistakeholder gatherings

Bharath Reddy, Rijesh Panicker, Sridhar Krishna, Arindam Goswami, Anwesha Sen, Adya Madhavan

Takshashila Report 2025-10. Version 1.0, April 2025.

About the State of AI Governance Report

The rapid progression in artificial intelligence capabilities has far-reaching impacts. It has the potential to significantly boost economic productivity, disrupt labour markets, and alter the balance of power between countries. The widespread diffusion of this general-purpose technology presents significant governance challenges that are amplified by the ongoing geopolitical competition for AI dominance.

It is against this backdrop that the Takshashila Institution, an independent centre for research and education in public policy, presents its inaugural State of AI Governance Report. This report provides a systematic comparative analysis of AI governance approaches across different countries, revealing their strategic priorities. Further, the effectiveness of corporate self-regulation initiatives, and the progress of multistakeholder collaborative efforts are also analysed. It concludes by offering predictions in these areas in the coming year.

This annual report will track key developments, analyse trends and offer informed predictions for the AI governance environment. It is intended to provide policymakers, analysts, and interested citizens with insights that help navigate the evolving AI governance landscape. You can learn more about our work at **takshashila.org.in/**

Executive Summary

This report analyses AI governance in three contexts – countries, companies and multistakeholder gatherings.

Countries

- Being at the forefront of AI innovation, the US favours a pro-market regulatory environment while prioritising geopolitical considerations to maintain its competitive advantage.
- The EU has focused on creating a comprehensive regulatory framework that prioritises transparency, accountability, and the protection of individual rights.
- China's approach prioritises national security and favours heavy state control in enforcing regulations and driving innovation.
- India has opted for a light-touch regulatory environment while investing in developing Indigenous AI models for Indian use cases.

Companies

- Many companies are proactively establishing principles, guardrails and transparency and disclosure norms that guide how they build or use AI. These initiatives sometimes go beyond what is strictly expected by the regulatory environment in which they operate and are intended to build trust with users.
- However, reporting on AI governance efforts by companies is not standardised, and the details of specific initiatives vary considerably. It is also unclear to what extent these efforts involve meaningful external scrutiny. The report analyses the governance initiatives of a few companies operating at different parts of the AI value chain.

Multistakeholder Gatherings

- Various multistakeholder gatherings, such as the AI Summits and the Global Partnership on AI, have been formed to raise awareness and coordinate AI governance efforts among different countries.
- Although most of these groupings do not have 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). However, they serve as a platform to highlight important concerns and drive convergence in AI governance efforts.

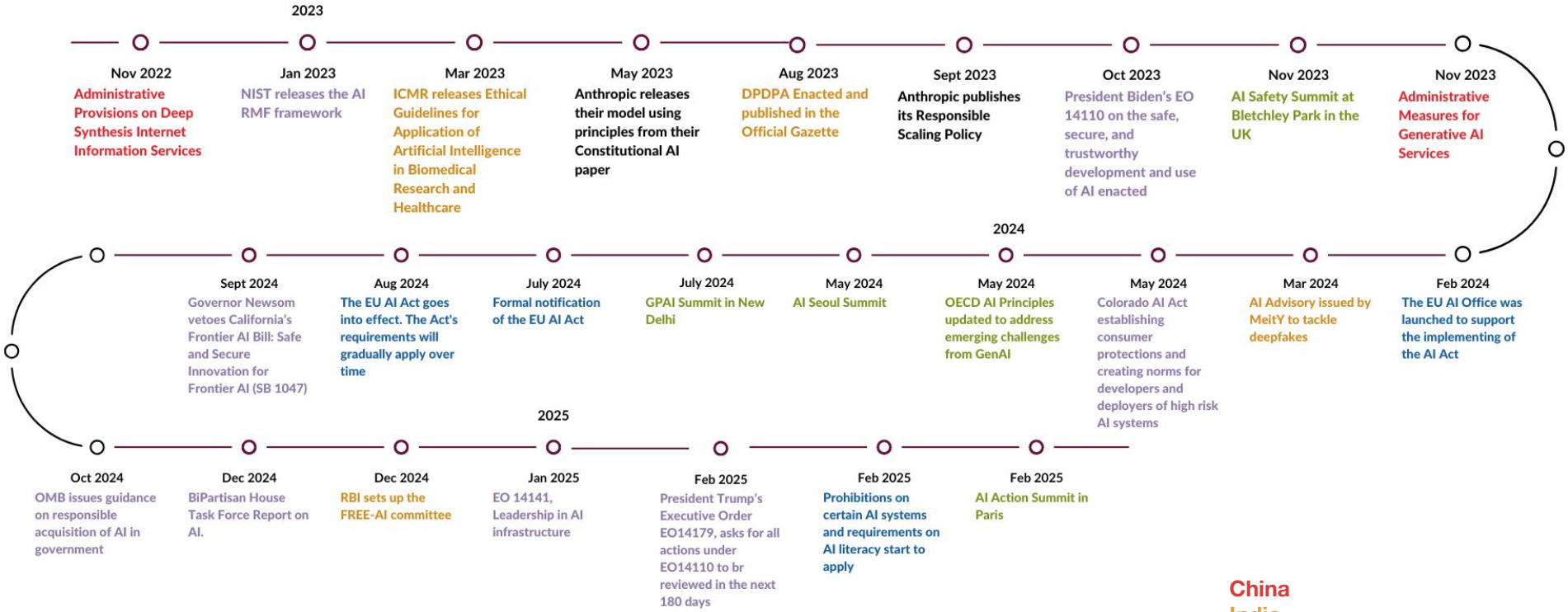
Finally, the report ends with some predictions on what we can expect in AI governance in the coming year.

Table of Contents

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

Timeline of AI Governance Events

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

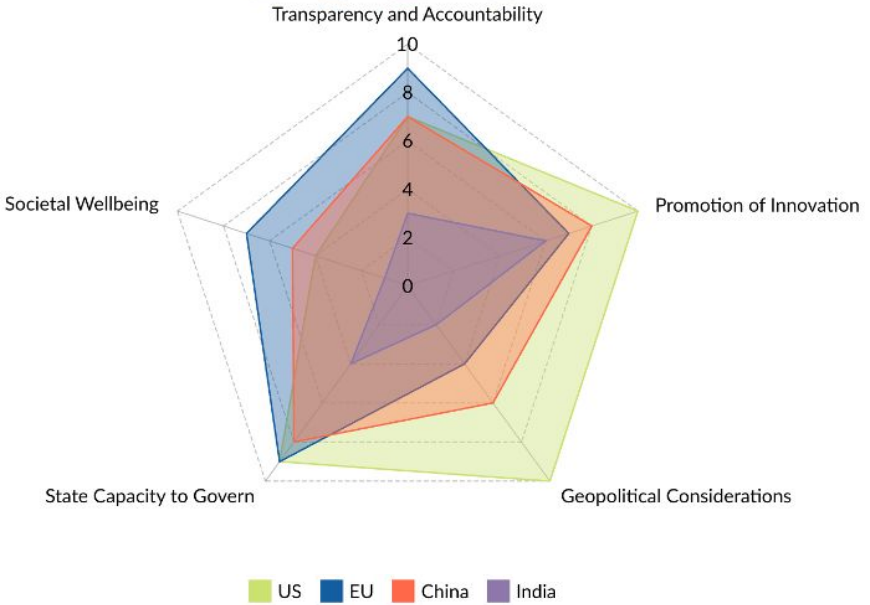


China
India
US
EU
Multistakeholder bodies
Companies

Analysis of AI Governance Measures Across Countries

Analysis of AI Governance Measures Across Countries

- 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.



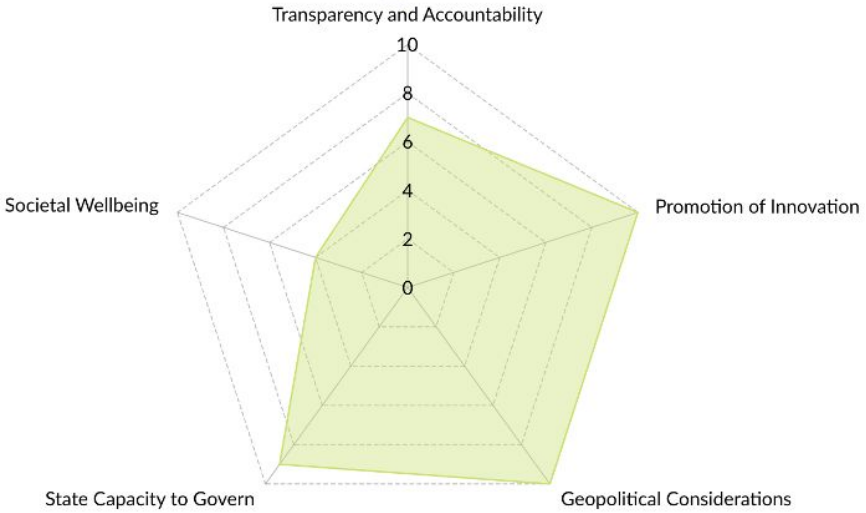
Analysis of AI Governance Measures Across Countries

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 U.S.

- The US regulatory approach utilizes existing regulatory capacity, supplementing it with AI-specific considerations.
- Federal AI regulation in the US involves executive orders setting goals, agency rulemaking enacting specifics (such as by the NIST, DOE, DOJ and HHS), and independent agency actions within their sectors.
- Several US states have implemented their own AI-specific regulations focusing on consumer protection, privacy, transparency, and algorithmic discrimination.
- President Trump's Executive Order on Removing Barriers to American Leadership in AI has revoked President Biden's EO 14110, that might lead to changes in federal AI regulations.
- Refer to **Appendix 1** for details about the specific governance measures.



Transparency & Accountability

- Compared to the EU, the US currently scores lower on transparency and accountability in AI governance partly due to the revocation of EO 14110 that had mandated disclosure for large foundational models and tasked NIST with developing safety testing standards.
- The OMB Act mandates risk management standards for AI use in identified high-risk government systems. The DOE is evaluating tools for identifying AI model risks in critical areas like nuclear and biological threats.

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

Analysis of AI Governance Measures in the U.S.

Promotion of Innovation

- The US strongly focuses on promoting AI innovation, with President Trump's executive order aiming to remove barriers to global AI dominance.
- Some initiatives from EO 14110 are expected to continue, such as NIST's NAIRR project and the OMB mandate for open-sourcing government AI models and data by default.
- The DOE will likely continue streamlining approvals for allied AI infrastructure like power and data centers.
- NIST's risk management framework and the US Patent Office's guidance on AI patentability will drive standardization and innovation.

Geopolitical Considerations

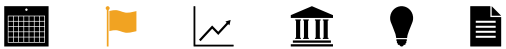
- Geopolitics is a significant focus in US AI regulation due to its leadership and competition with China.
- The US aims to maintain a competitive edge by controlling access to AI chips through the AI diffusion framework.
- A monitoring regime is in place for large AI model training runs on US Infrastructure as a Service (IaaS), including reporting by international customers.
- The US Treasury prohibits certain financial transactions involving high-risk AI systems.
- Security agencies are tasked with identifying and mitigating vulnerabilities across the AI supply chain.

Societal Wellbeing

- Several US states have enacted AI-specific regulations to prevent algorithmic discrimination, ensure disclosure, and label AI-generated content.
- State privacy laws have been updated to address the use of personal information and risks related to AI profiling and consent.
- Federal regulations from OMB and HHS establish standards for AI use in delivering public benefits and within government.

State Capacity to Govern

- The revocation of EO 14110 creates uncertainty around current US AI governance rules.
- Regulations focusing on building AI infrastructure, capital flows, and talent are expected to remain relatively stable.
- Regulations concerning AI governance capacity building, transparency, and auditability are likely to be reviewed and potentially rolled back.
- This analysis considers AI regulations at the federal and state levels as they relate to various governance dimensions.

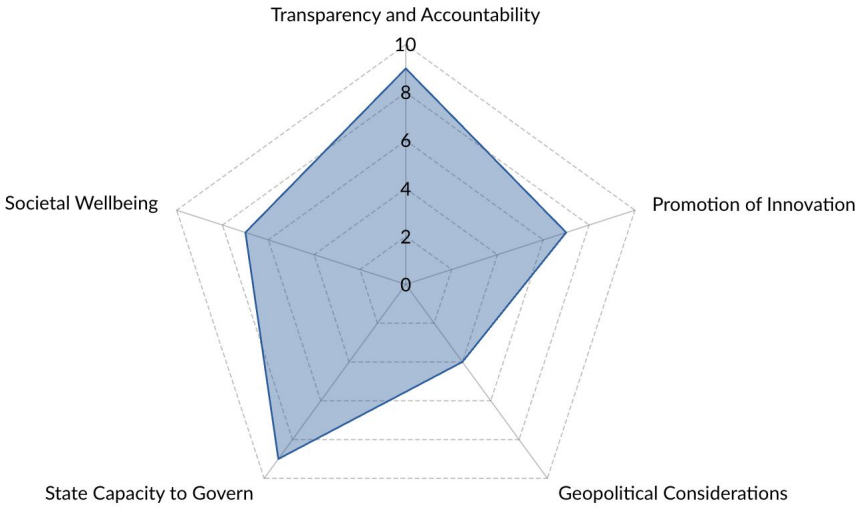


Analysis of AI Governance Measures in the E.U.

- The EU AI Act establishes comprehensive regulatory guidelines for AI systems operating within the European Union.
- It interacts 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.
- The analysis below primarily focuses on the requirements under the EU AI Act.
- Refer to **Appendix 2** for details about the specific governance measures.

Transparency & Accountability

- The EU has the most comprehensive measures for AI transparency and accountability, including risk tiering, evaluations, disclosure, licensing, penalties, and input controls.
- AI systems in the EU are classified into unacceptable, high, limited, and minimal risk categories, with corresponding compliance obligations.



Sources: EU AI Act, EU InvestAI Initiative, EU Chips Act, AGORA

Analysis of AI Governance Measures in the E.U.

Promotion of Innovation

- While compliance with the comprehensive regulations creates hurdles, there are provisions for promoting standards and regulatory sandboxes to facilitate adherence.
- The InvestAI initiative aims to mobilise $216 bn for open and collaborative development of complex AI models in Europe.
- The EU AI Act includes carve-outs to lessen the impact of penalties on SMEs and startups.

Geopolitical Considerations

- Strategic autonomy and technology sovereignty have gained prominence in EU AI policy discussions.
- The InvestAI initiative funds the open development of complex AI models, including AI gigafactories for data centers and the creation of datasets.
- The EU Chips Act involves significant investment to enhance competitiveness and resilience in the semiconductor industry.
- The EU AI Act includes exceptions for AI models released under free and open-source licenses.

Societal Wellbeing

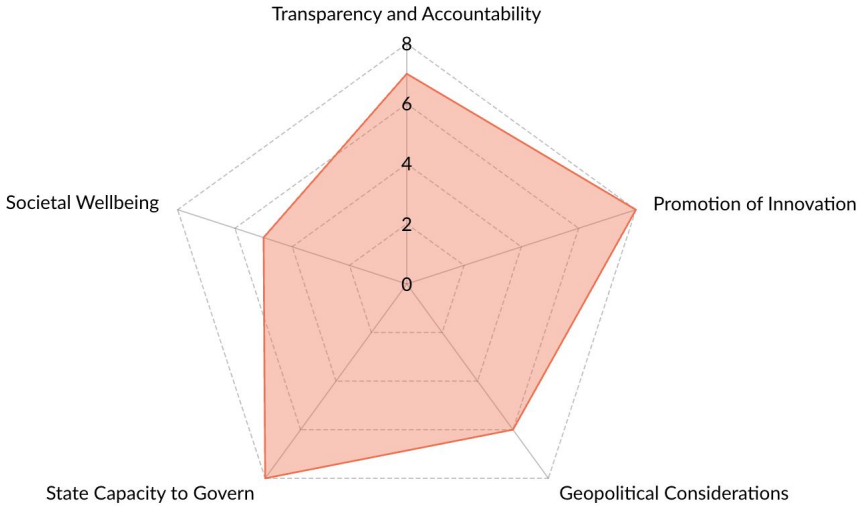
- Among the regions compared in this report, the EU has the most comprehensive measures to address societal wellbeing through various governance instruments such as tiering, evaluation and performance requirements, disclosure mandates, licensing and certifications, penalties and input controls
- Voluntary codes of conduct are encouraged for assessing and minimising the environmental impact of AI systems.

State Capacity to Govern

- The EU AI Act mandates the creation of governance institutions at both the EU and member state levels for implementation.
- This includes the establishment of the EU AI Office, advised by a scientific panel of independent experts.
- Member states are required to establish their own institutions to enforce compliance with the EU AI Act regulations.

Analysis of AI Governance Measures in China

- China's AI governance framework is a comprehensive, state-driven approach across multiple dimensions.
- The system prioritizes centralized control through proactive regulation requiring algorithm transparency, security assessments, ethics reviews, and content monitoring aligned with state objectives.
- China's governance model emphasizes public-private collaboration for AI innovation along with strategic state investments.
- China's governance strategy strongly reflects national security priorities through data sovereignty measures, indigenous computing development and military-civil fusion initiatives.
- Massive financial investments such as $140 bn from the Bank of China and $184 bn in government VC funding are announced for AI firms.
- Refer to **Appendix 3** for details about the specific governance measures.



Transparency & Accountability

- China's state-enforced AI regulations demonstrate a preference for proactive, centralised governance.
- Chinese regulations mandate algorithm filing, transparency, security assessments, and ethics reviews before AI deployment.
- Watermarking for AI-generated content and complaint redressal mechanisms are also required in China.

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

Promotion of Innovation

- China's governance mechanisms focus on state-driven public-private collaboration and strategic infrastructure development for innovation.
- There is a strong emphasis on AI talent cultivation and public-private research partnerships in China, including innovation hubs supported by AI-focused education guidelines and university-affiliated research centers.
- Guidelines reinforce an open AI ecosystem, promoting resource-sharing, interoperability, and common technical standards.
- Large-scale infrastructure projects ensure regional computing power distribution and provide tax incentives for AI infrastructure investments.

Geopolitical Considerations

- Policies emphasise technological self-sufficiency through measures requiring data sovereignty, AI security, indigenous computing infrastructure development and military-civil fusion.
- Strict data localization measures are imposed, requiring government approval for cross-border transfers of "important data."
- Mandatory algorithmic filing and real-time content monitoring ensure AI-generated content aligns with state narratives.
- Real-name verification requirements under cybersecurity laws address information control and foreign influence concerns.

Societal Wellbeing

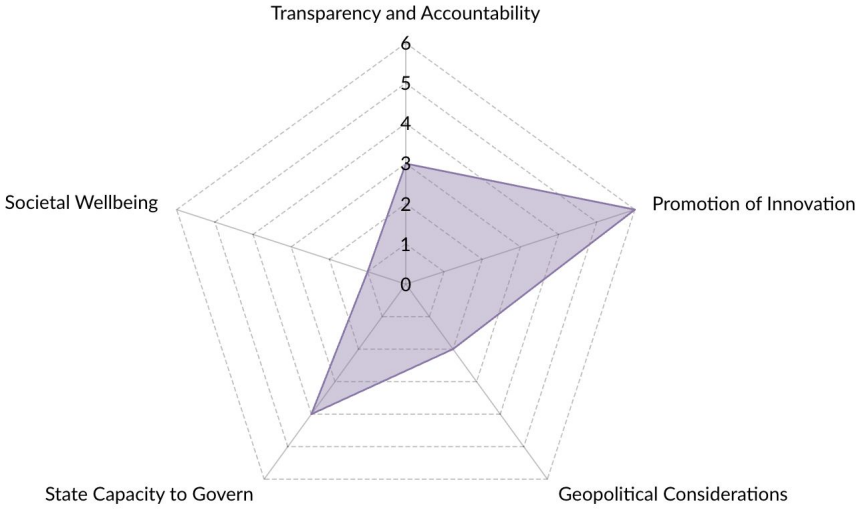
- A state-led approach focuses on AI ethics, data privacy, content moderation, and digital fairness.
- Explicit user consent for personal data usage and mandatory opt-out options for personalised recommendations are required.
- Algorithmic price discrimination is banned, and watermarking/labeling of AI-generated content is mandated.
- Explicit user consent is required for use of voices and images in synthetic media.

State Capacity to Govern

- China's capacity to govern AI development is demonstrated through massive financial investments and centralised infrastructure planning.
- The government directs significant capital into priority AI sectors at an unprecedented scale.
- Local governments have also dedicated billions to AI research and infrastructure, reflecting a decentralised but state-coordinated funding approach.
- Strong oversight is provided by government institutions like the Cyberspace Administration of China (CAC).
- The state has the capacity to fund and manage large-scale computing networks for AI development.

Analysis of AI Governance Measures in India

- India's AI governance is characterised by limited regulations, with sector-specific guidelines in finance and health.
- The government is promoting AI innovation through the IndiaAI Mission addressing priority use cases such as agriculture, education and healthcare.
- Some policy interventions aim to address vulnerabilities related to access to AI resources.
- There is a lack of comprehensive regulations in India for managing the broad risks associated with AI.The draft AI Governance Guidelines propose a Technical Secretariat to enhance state capacity to govern AI, but its implementation is uncertain.
- Refer to **Appendix 4** for details about the specific governance measures.



Transparency & Accountability

- India has limited cross-sector regulations. The finance and health sectors in India have more specific regulations and AI safety guidelines.The RBI has formed a panel to review AI regulation in finance, identify risks, and recommend a compliance framework. Similarly, ICMR Ethical Guidelines provide principles for AI development and deployment in healthcare.
- The AI Advisory (2024) recommends labeling deepfake and synthetic content, and draft guidelines suggest voluntary incident reporting.

Sources: AI Advisory (March 2024), 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, Report on AI Governance Guidelines Development

Analysis of AI Governance Measures in India

Promotion of Innovation

- The IndiaAI Mission has allocated over $1.2 bn to develop AI models, datasets, compute, and education.
- Efforts are underway to enable access to compute resources and datasets through portals like IndiaAI Compute and AIKosha.
- While aiming to encourage startups and MSMEs, the implementation of these innovation initiatives may face inefficiencies.

Geopolitical Considerations

- The US AI Diffusion Framework's export controls could potentially restrict India's access to computing resources and advanced AI models.
- India is attempting to mitigate these risks by promoting indigenous AI models and domestic computing clusters.

Societal Wellbeing

- Sector-specific guidelines from ICMR and RBI, along with the AI Advisory (2024), offer a voluntary ethical framework for societal well-being in AI.
- There are currently no overarching regulations in India to address the diverse risks posed by AI technologies.

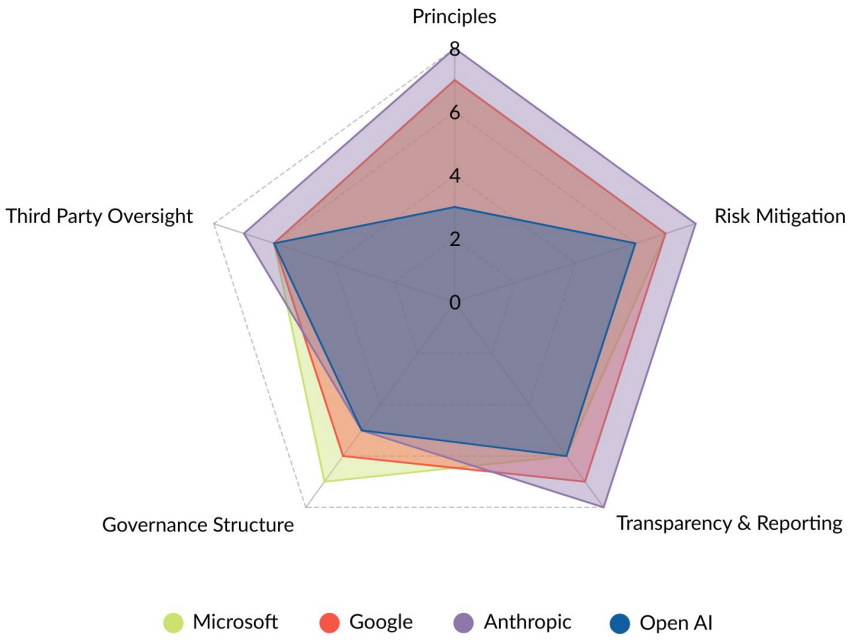
State Capacity to Govern

- MeitY's draft AI Governance Guidelines propose a Technical Secretariat to monitor and mitigate AI risks and harms in real-time, serving as a bridge between the industry and policymakers.
- As these are draft guidelines, it remains to be seen how state capacity will be enhanced to implement them.

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.
- However, there is currently a lack of standardization in how companies report on their AI governance efforts. The specific details of company AI governance initiatives vary considerably.
- The extent of meaningful external scrutiny of company AI governance efforts is unclear. It is also challenging to understand how companies are operationalizing their principles, measuring the effectiveness of their risk mitigation strategies, and ensuring accountability.
- 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 chart indicates the authors' scoring of the AI governance measures of these companies on the criteria selected.
- The following slides in this section provide the reasoning for the scoring.



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: In 2018, six ethical principles were adopted: accountability, transparency, fairness, reliability and safety, privacy and security, and inclusiveness. These principles have been translated into corporate policies, including a standard for engineering teams.

Google: First published in 2018 and updated recently, the principles are bold innovation, responsible development and deployment, and collaborative progress.

OpenAI: Committed to building safe and beneficial AGI. Specific principles are not detailed. Its usage policy prohibits uses such as biometric profiling, deceptive techniques, harm to children, etc.

Mistral: Mistral's responsible AI policy emphasises key principles such as neutrality, empowerment, building trust and minimising potential harm and misuse. Furthermore, its terms of use emphasise data privacy and security and restrict access to individuals under 13.

Anthropic: Anthropic emphasizes proportional safeguards that scale with the potential risks of AI systems. Their approach is inspired by biosafety standards, ensuring that safety measures align with the capabilities of their models.

Amazon: AWS has introduced an updated Responsible AI Policy in January, 2025, which supplements the AWS Acceptable Use Policy and AWS Service Terms. This policy outlines a set of prohibited actions, and includes commitments to developing safe, fair, and accurate AI and machine learning services.

Accenture: Accenture's four pillars of responsible AI implementation are: Organizational, Operational, Technical, and Reputational.

Deloitte: Uses a Trustworthy AI™ framework to help organizations develop ethical safeguards across seven key dimensions: transparent and explainable, fair and impartial, robust and reliable, respectful of privacy, safe and secure, responsible, and 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: AI risk management is part of the company's enterprise risk management system. They also prioritise development of open-access responsible AI tools to map, measure and manage risks.

Google: Work with external organizations to establish boundaries and reduce the risks of abuse. Google has developed a frontier safety framework to prepare for risks from frontier models. They have also published a generative AI toolkit and people + AI guidebook that guide AI development and deployment.

OpenAI: Investments in research to inform regulation – including techniques for assessing potentially dangerous capabilities. Their red-teaming and safety procedures are publicly disclosed.

Mistral: Includes measures that address data privacy and security. It prohibits use for illegal activities, hate and discrimination, misinformation, and professional advice. They have a zero-tolerance policy regarding child sexual abuse material.

Anthropic: They employ AI Safety Level (ASL) Standards to address catastrophic risks, such as misuse or unintended autonomous actions. These standards include rigorous testing and safeguards before deploying advanced models

Amazon: Prohibits use for disinformation, deception, violation of privacy, harm or abuse of minors. Also prohibits AI/ML services for circumvention of safety filters and use in a weapon without human intervention

Accenture: Ongoing testing of AI for human impact, fairness, explainability, transparency, accuracy, and safety. Employing use of state-of-the-art responsible AI tools and technologies to mitigate problems.

Deloitte: Using tools to test and monitor AI models.

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: Transparency reports are published with recent reports aligned with the NIST framework. Provide documentation on the intended use, limitations and risks of AI technologies to consumers.

Google: Transparency reports are published with recent reports aligned with the NIST framework. Gemini app, Google Cloud and Google Workspace are ISO/IEC 42001 certified. Engaging with governments, civil society, and academia on information sharing, establishing common standards, and promoting best practices for AI safety.

OpenAI: Transparency reports and system cards for different products detail the red-teaming and safety procedures.

Mistral: Publishes documentation for its AI models with details on their capabilities and limitations. Provides channels for reporting incidents, contact information for inquiries, and communities for users.

Anthropic: Anthropic commits to sharing updates on their governance practices and lessons learned. They aim to maintain public trust by openly communicating their safety measures and policies.

Amazon: Provides details on how AI systems are deployed, monitored and managed during their development and operations. Mentions openly sharing development choices, including data sources and algorithms.

Accenture: Public disclosure of AI systems' capabilities, limitations, and suitable uses, addressing both security and societal risks.

Deloitte: Recommends engaging stakeholders in AI governance by defining clear roles and responsibilities, documenting accountabilities, establishing expectations for AI ethics and trust, and empowering employees to voice concerns and act ethically.

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: Aether is an internal advisory body set up to focus on AI ethics and responsible AI practices within the company. In 2023, over 350 employees worked on responsible AI, developing best practices for building safe, secure, and transparent AI systems designed to benefit society.

Google: The AGI Safety Council and the Responsibility and Safety Council oversee alignment with AI principles. Google forced AI researcher Dr. Gebru to resign after she resisted orders to halt research indicating that speech technology, like Google's, could negatively impact marginalized groups. Others also resigned in response to how Dr. Gebru was treated.

OpenAI: Specific governance structures are not publicly available. Concrete governance practices are specifically tailored to highly capable foundation models. The company structure has also seen a lot of churn, changing from a non-profit to a "capped profit" structure to enable the pursuit of AGI. Former employees have also raised concerns about the safety practices at the company.

Mistral: Does not provide any details on the governance structure.

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.

Amazon: Does not provide any details on the governance structure.

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.

Deloitte: Evaluating roles and responsibilities and implementing change management and training.

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: Bug bounty programs are used to incentivise external discovery and reporting of issues and vulnerabilities. They are also building external red teaming capacity to enable third-party oversight, especially of sensitive capabilities of highly capable models.

Google: Google Cloud AI achieved a “mature” rating in a third-party evaluation. Encourages third-party discovery and reporting of issues and vulnerabilities. Works with industry peers and standards-setting bodies towards efforts such as developing technical frameworks to help users distinguish synthetic media.

OpenAI: Bounty systems to encourage responsible disclosure of weaknesses and vulnerabilities.

Mistral: Its policies do not explicitly seek third-party oversight, but its commitment to transparency indicates a willingness to be open about its AI practices which could extend to third-party oversight.

Anthropic: Anthropic collaborates with external experts for adversarial testing and red-teaming of their models. This ensures that their systems meet stringent safety and security standards.

Amazon: While AWS is committed to developing safe, fair and accurate AI/ML services and provides tools and guidance to assist in this, it does not talk about third-party oversight.

Accenture: No mention of third-party oversight.

Deloitte: Offers the Omnia Trustworthy AI module, which provides guidelines and guardrails for designing, developing, deploying, and operating ethical AI solutions. Does not mention using third party oversight over their own AI 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 Multistakeholder Groupings

Analysis of AI Governance Measures Across Multistakeholder Gatherings

- Various multistakeholder 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, multistakeholder 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).
- 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 Multistakeholder 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 Multistakeholder 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 multistakeholder 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 Multistakeholder 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

Predictions

Predictions

Confidence	Region	Prediction
High	Global	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	Investments in sovereign cloud infrastructure will increase, driven by geopolitical considerations.
High	China, EU	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	India	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.
Moderate	Global	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.

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

Predictions

Confidence	Region	Prediction
Moderate	US	US chip restrictions on China will not escalate further. This is because Deep Seek makes owning newer chips less relevant.
Moderate	US	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	EU	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	China	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	India	Regulatory focus will be on protecting the information ecosystem from content deemed harmful for the government or Indian society.

Predictions

Confidence	Region	Prediction
Moderate	Corporate	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.

Appendix

1. Governance Measures in the U.S.

Transparency and Accountability

- NIST to develop guidance, evaluations and benchmarks to develop consensus standards for AI safety and security in areas such as cybersecurity and biosecurity
- DOE to create AI evaluation tools and testbeds that evaluate AI model capabilities around risk to critical infrastructure and possible nuclear, biological, chemical threats and hazards
- OMB has passed rules for federal agencies which:
    - mandates agencies to adopt risk management regulations for the use of AI within their agencies, including incorporating viewpoints to mitigate bias from the use of AI models
    - mandates that all relevant agencies will identify high-risk AI systems - defined as those impacting civil rights, human rights, discrimination, bias, etc. and implement risk management rules for these systems
- EO 14110 mandates disclosures for any planned development activity and disclosure of model weights in dual-use models, which are large foundational models capable of being used for regular and military purposes. Models trained using greater than 10^26 integer or floating point operation, and models using biological sequencing data and more than 10^23 operations fall under this rule
- Several states have enacted AI-specific regulation or tightened existing privacy laws for AI, which provide for disclosures and transparency in the use of data and AI models, and accountability on developers and deployers for use of AI in specified use cases.

Promotion of Innovation

- NSF to use the NAIRR pilot and future programs to democratise access to AI resources, with a goal of promoting discovery and innovation in AI
- The White House and DOE are to streamline approval for the construction of AI-enabling infrastructure, as well as surrounding assets supporting the resilient operation of this infrastructure, such as clean energy generation, power transmission lines, and high-capacity fibre data links.
- OMB rules for advancing agency use of AI, stipulates that all AI models developed within U.S agencies must be released as open source and open weight unless they specifically violate IP , copyright, or any national security concerns

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 U.S.

Promotion of Innovation

- DOC to release inventorship guidance for AI-assisted inventions, requiring inventors or co-inventors on patents to be natural persons. AI-assisted inventions are accepted as long as natural persons have contributed, while pure AI system-led inventions are not patentable. This is designed to promote human innovation.
- President Trump's EO 14176 repealing EO14110 and asking for a review of all steps taken under the previous executive order is likely to accelerate innovation with its focus on reducing guardrails around safety and alignment
- DOE to launch pilot project to assess the performance and efficiency of federated AI and data sources for frontier AI models.

Geopolitical Considerations

- NSC staff, ODNI in coordination with others (DOD, DOJ etc) to identify and assess foreign intelligence threats to AI ecosystem and allied sectors like semiconductors. ODNI to identify vulnerable points in the supply chain for AI including allied infrastructure and talent
- The BIS released a framework for AI diffusion to manage the spread of AI while maintaining a competitive edge for the US. The focus is on 3-tier system to control access to AI chips, alongside controls on model weights for advanced AI models.
- The Outbound Investment Security Rule (US Treasury) implements EO14105, which requires disclosure of transactions involving national security tech (including AI) and prohibits transactions for AI in military, surveillance and intelligence use.
- Under EO14110, cloud deployers of AI compute are required to inform the US government of large model training runs by international customers. This extends to similar requirements for international resellers of compute. The limits for what constitutes a large model is also determined by the executive order.

State Capacity to Govern

- The White House and DOE to streamline approval for the construction of AI-enabling infrastructure, as well as surrounding assets supporting the resilient operation of this infrastructure, such as clean energy generation, power transmission lines, and high-capacity fibre data links.
- OMB Act: Relevant government agencies to assess current AI maturity and prepare plans for AI use, R&D with respect to AI and workforce development
- National AI Talent Surge implemented as part of EO14110 to increase AI capacity in federal government in order to implement AI in relevant agencies, and build capacity to regulate and administer federal R&D programmes.

1. Governance Measures in the U.S.

Societal Wellbeing

- Several states have enacted AI-specific regulation or tightened existing privacy laws for AI. Most regulations have focused on algorithmic discrimination and profiling using AI, privacy of personal information, and disclosure norms around AI-generated content, including use of AI in specific contexts (election advertising)
- OMB Act mandates that all relevant agencies will identify high-risk AI systems - defined as those impacting civil rights, human rights, discrimination, bias etc. - and implement risk management rules for these systems
- HHS to promote responsible AI for public benefits by providing guidance to enhance benefit access, efficiency and effectiveness of these systems, while mitigating risk and providing transparency.

State Regulations on AI (Enacted)

- Colorado AI Act SB24-205 – algorithmic discrimination in high risk systems
- Utah AI Policy Act
- California AB 2013, SB942, AB2355 – provides for disclosure of 1) data used in AI training, 2) AI content generation in systems, 3) Gen AI use in electoral advertisement
- Oregon SB619 (Consumer Privacy Law) - rules for profiling and automated decision-making
- Montana SB384 (Consumer Data Privacy Act) – personal information, profiling and automated decision making
- Tennessee ELVIS Act – prohibits use of AI to mimic a person's voice.

State Regulations on AI (Proposed)

- Montana SB212 – manual failsafe for critical infra AI systems
- Texas Responsible AI Governance Act (TRAIGA) – Risks from high risk AI systems

2. Governance Measures in the E.U.

Transparency and Accountability: Tiering

- 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, AGORA

2. Governance Measures in the E.U.

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.

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.

2. Governance Measures in the E.U.

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.

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).

2. Governance Measures in the E.U.

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.
- 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

- InvestAI to 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 datacentres (called AI gigafactories) that will be specialised in training the most complex, very large AI models
- 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.

2. Governance Measures in the E.U.

Promotion of Innovation: Funding

- 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.
  - 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: Funding

- 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.

2. Governance Measures in the E.U.

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.

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.
- Under the same provisions, platforms must provide users with options to turn off recommendation algorithms or manually adjust content preferences.
- Administrative Provisions on Deep Synthesis Internet Information Services (2023) provide for watermarking and labelling requirements for AI-generated content.
- User complaint mechanisms, requiring platforms to address user concerns about synthetically-generated content, are mandated under these same provisions.

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

3. Governance Measures in China

Transparency and Accountability

- Governance Principles for New Generation AI (2019) establish the ‘self-discipline’ principle - industry practitioners should self-regulate and establish codes of conduct.
- 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
- AI Standardization Guidelines (2018) provide for data annotation and quality standards to improve training data.
- Data Security Law (2021) mandates data security risk assessments for data processors.
- Personal Information Protection Law (2021) provides for the algorithmic transparency requirement - automated decision-making must be explainable.

Promotion of Innovation

- Governance Principles for New Generation AI (2019) encourages ‘open and shared’ principle - encourages open AI platforms and resource sharing.
- New Generation AI Development Plan 2017-2030 encourages public-private research partnerships and innovation centres for AI.
- It also encourages AI talent cultivation programs with specialized educational initiatives. China’s Ministry of Education (MOE) issued guidelines to introduce AI-focused courses in primary, secondary, and higher education. China established AI research centres in universities to cultivate talent, like, for instance, the Baidu AI Research Institute.
- AI Standardization Guidelines (2018) encourage interoperability standards for AI systems across different platforms, as also data annotation and quality standards to improve training data.
- National Computing Network Coordination Plan (2022) aims to have regional computing centres linking eastern and western regions of China.
- Computing Power Hub Plan (2022) provides for tax incentives for computing infrastructure investments. There is also a nationwide AI computing power sharing platform.

3. Governance Measures in China

Geopolitical Considerations

- 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 generation of illegal content.
- The Administrative Measures for Generative AI Services (2023) specifies algorithm filing and security assessment requirements before service launch.
- The Ethical Norms for New Generation AI (2021) specify the ‘controllability’ principle - AI systems must remain under meaningful human control.
- New Generation AI Development Plan 2017-2030 focuses on public-private research partnerships (In 2017, the Ministry of Science and Technology (MOST) designated key firms to lead AI research in specialized areas, like Baidu in autonomous driving, Alibaba Cloud for smart cities) and innovation centres for AI (like Beijing AI Innovation Centre, Shanghai AI Lab, etc.).
- There is increased emphasis on military-civil fusion in AI development - coordinated development between defence and civilian sectors. For instance, the Chinese Academy of Sciences (CAS) and defence contractors partner with firms like Huawei and Hikvision to develop AI for dual-use applications (civilian and military).
- Personal Information Protection Law (2021) mandates cross-border data transfer restrictions requiring security assessments.
- Data Security Law (2021) has put in place a data classification system with heightened protection for "important data". The classification framework categorizes data into general, important, and core data, with escalating levels of regulatory oversight. Important data collected within China must be stored domestically with exceptions requiring government approval before cross-border transfer. These data localisation norms also exist under the Cybersecurity Law (2017).
- Under the Cybersecurity Law (2017), critical information infrastructure has enhanced security requirements There is a national security review mechanism for data activities that might affect national security.
- Companies violating important data protection rules face fines up to $1.5 million, and business license revocation for severe violations.
- There are real-name verification requirements using approved IDs for internet services under China's Cybersecurity Law (2017).
- National Computing Network Coordination Plan (2022) aims to establish regional computing centres linking eastern and western regions of China.

3. Governance Measures in China

Geopolitical Considerations

- National Computing Network Coordination Plan (2022) provides support for indigenous computing technology with preferential policies. For instance, state-owned enterprises (SOEs) like China Mobile and China Telecom have shifted to homegrown cloud solutions (e.g., Alibaba Cloud, Baidu Cloud).
- Computing Power Hub Plan (2022) aims to establish a nationwide AI computing power sharing platform, and provides preferential land and energy policies for computing facilities in approved zones. For instance, the plan designates 8 national computing power hubs, including Beijing, Shanghai, Guangzhou, Chengdu, and Inner Mongolia, to create a distributed AI computing network.

Societal Wellbeing

- 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.
- Under the Internet Information Service Algorithmic Recommendation Management Provisions (2022), there must be provision of user opt-out option - platforms must provide users options to decline personalized recommendations.
- Under the Internet Information Service Algorithmic Recommendation Management Provisions (2022), there are anti-monopoly and fair competition requirements, for instance, to prevent algorithmic price discrimination.
- Administrative Provisions on Deep Synthesis Internet Information Services (2023) provide for watermarking and labelling requirements for AI-generated content. These also provide a consent mechanism for using a person's voice or image in deep synthesis.
- User complaint mechanisms, requiring platforms to address user concerns about synthetically-generated content, are mandated under these same provisions.
- Governance Principles for New Generation AI (2019) specify the ‘harmony and friendliness’ principle, i.e. AI should promote economic development and social progress.
- Ethical Norms for New Generation AI (2021) specify the ‘human autonomy’ principle - AI systems should enhance human autonomy rather than weaken it.

3. Governance Measures in China

Societal Wellbeing

- These norms also specify the fairness and justice principle - AI should avoid unfair bias against groups - and the controllability principle - AI systems must remain under meaningful human control.
- Opinions on Strengthening the Ethical Governance of Science and Technology (2022) mandate ethics review requirements for high-risk AI research projects. There is also an ethical education mandate for researchers in emerging technologies.
- AI Standardization Guidelines (2018) provide for AI safety testing frameworks for evaluating system reliability.
- National Computing Network Coordination Plan (2022) has green computing requirements for energy efficiency in data centres.
- Personal Information Protection Law (2021) mandates ‘algorithmic transparency’ requirement - automated decision-making must be explainable. It also mandates the ‘data minimization’ principle - only collect necessary personal information.
- There are real-name verification requirements using approved IDs for internet services under China's Cybersecurity Law (2017).

State Capacity to Govern

- National Level AI funding estimates:
  - Bank of China Investment: The Bank of China plans to invest at least 1 trillion Yuan (approximately $140 billion USD) in the AI sector over the next five years. This includes financial support for companies across the entire AI industry chain, with stocks and bonds amounting to no less than 300 billion Yuan.
  - Government Guidance Funds: Between 2000 and 2023, Chinese government VC funds invested in over 9,623 unique AI firms through more than 20,000 transactions, totalling $184 billion. These funds are crucial for supporting AI development in less-developed regions.
  - AI Chip Development – SMIC, Huawei, and Biren Technology received billions in state subsidies to develop AI processors after U.S. sanctions on NVIDIA and Intel chips.
  - New AI Investment Fund: China recently established a new AI investment fund with an initial capital of 60 billion Yuan (approximately $8.2 billion USD). This fund aims to enhance China's AI capabilities amidst tightened US trade restrictions.

3. Governance Measures in China

State Capacity to Govern

- Local Government AI Funds:
  - Beijing AI Fund: $2 billion (2018) – For AI research, smart city projects, and AI applications.
  - Tianjin AI Fund: $16 billion (2018) – AI infrastructure and industrial automation.
- Infrastructure support:
  - Under the National Computing Network Coordination Plan (2022), there are plans to establish regional computing centres linking eastern and western regions of China.
  - Under the Computing Power Hub Plan (2022), there are tax incentives for computing infrastructure investments. There are also plans to establish a nationwide AI computing power sharing platform.
- Oversight mechanisms: There are strong oversight mechanisms through government institutions like the Cyberspace Administration of China (CAC).

4. Governance Measures in India

Transparency and Accountability, Social Wellbeing

- RBI has set up an 8-member panel to review AI regulation and supervise its adoption in the financial sector, focusing on identifying risks and recommending frameworks for evaluation, mitigation, monitoring, and compliance for financial institutions.
- ICMR Ethical Guidelines outline principles for the development, validation, and deployment of AI in healthcare.
- AI Advisory (2024) requires labelling deep fake or synthetically-modified content and mandates platforms to ensure their AI models don't allow content violations under the IT Rules or IT Act, 2000.
- The Digital Personal Data Protection Act is not yet implemented, but it excludes publicly available data and allows an exemption for personal data used for "research, archiving, or statistical purposes." The act might have implications for
- MeitY's AI Governance Guidelines aims to create a Technical Secretariat to mitigate AI risks and act as a bridge between industry and policymakers for real-time intervention; voluntary incident reporting and transparency guidelines.

Promotion of Innovation

- $230 mn for Innovation Centres, to develop and deploy indigenous LMMs and domain-specific models for critical sectors.
- $533 mn for providing compute capacity (18,693 GPUs) to support start-ups, researchers, students, and academics over five years.
- $23 mn for the AIKosha Datasets Platform, aimed at making public sector datasets AI-ready and providing real-time access through APIs.
- $80 mn of IndiaAI mission budget is allocated for the Application Development Initiative.

Sources: AI Advisory (March 2024), 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, Report on AI Governance Guidelines Development

Acronyms

Acronyms

AI	Artificial Intelligence	GPAI	Global Partnership on AI
CAC	Cyberspace Administration of China	HHS	Health and Human Services
DMA	Digital Markets Act	IaaS	Infrastructure as a Service
DOE	Department of Energy	ICMR	Indian Council of Medical Research
DOJ	Department of Justice	MeitY	Ministry of Electronics and Information Technology
DPDPA	Digital Personal Data Protection Act	MSMEs	Micro, Small and Medium Enterprises
DSA	Digital Services Act	NAIRR	National Artificial Intelligence Research Resource
EO	Executive Order	NIST	National Institute of Standards and Technology
EU	European Union	OECD	Organization for Economic Co-operation and Development
GDPR	General Data Protection Regulation	OMB	Office of Management and Budget

Acronyms

RBI	Reserve Bank of India
SMEs	Small and Medium-sized Enterprises
UN	United Nations
UNESCO	United Nations Educational, Scientific and Cultural Organization
US	United States of America
VC	Venture Capital

Acknowledgements and Disclosure

The authors would like to thank their colleague, Pranay Kotasthane, for his 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, Perplexity, and Grammarly for assistance in analysis and copy-editing.