# Artificial Intelligence: Impact and Governance

Satya S Sahu, Shrikrishna Upadhyaya, Bharath Reddy, Sridhar Krishna, Bharat Sharma & Saurabh Todi

Takshashila Internal Conference Compendium 2024-02
Version 1.0, February 2024

This document is a compendium of four Working Papers presented at the February 2024 Internal Conference organized by the Takshashila Institution on the theme of AI and AI Governance. The papers featured in this document cover four key aspects – a comparative analysis of AI regulation in India and four other jurisdictions, the impact of AI on jobs and policy solutions, global governance framework for AI, and leveraging AI to augment state capacity in India.

## Table of Contents

**TAKSHASHILA**
**INSTITUTION**

# Executive Note

## Sarthak Pradhan

The Takshashila Institution organised a conference on 14th February 2024 to understand Artificial Intelligence's impact and governance. Takshashila's in-house scholars, Satya Shoova Sahu, Shrikrishna Upadhyaya, Sridhar Krishna, Bharat Sharma, Saurabh Todi and Bharath Reddy, presented papers covering diverse themes related to Artificial Intelligence. Shambhavi Naik chaired the conference. This document is a compendium of the four working papers presented at the conference.

In the first paper, Satya and Shrikrishna comprehensively analyse India's AI regulatory framework, emphasising the diverse policies, frameworks, and guidelines that contribute to its structure. They also analyse AI regulations in various jurisdictions, such as the US, UK, EU, and China, and identify convergent and divergent trends. The analysis

3

**Authors**

Satya S Sahu is a Research Analyst with the High-Technology Geopolitics Programme at the Takshashila Institution.

Shrikrishna Upadhyaya is a consultant and knowledge manager at Takshashila Institution.

Bharath Reddy is a Programme Manager (Tech & Policy) and is also part of the High-Technology Geopolitics Programme at the Takshashila Institution.

Sridhar Krishna is a Senior Scholar at the Takshashila Institution, where he led the 20 Million Jobs Project and is now part of the AI Governance research group.

Bharat Sharma is a Quad Studies Research Analyst with the Indo-Pacific Studies Programme at the Takshashila Institution.

Saurabh Todi is a Research Analyst with the High-Technology Geopolitics Programme at the Takshashila Institution

reveals how varying priorities, strategies, and stages of legislative development shape the landscape.

In his paper, Sridhar Krishna assesses the legitimacy of concerns about Artificial Intelligence displacing jobs on a large scale. After reviewing relevant literature on the subject, he advocates for proactive measures, suggesting that preparing for the future involves staying ahead of the AI wave. He emphasises the importance of focusing on tasks that AI cannot perform and adapting to new endeavours when AI achieves significant breakthroughs.

Bharat Sharma and Saurabh Todi analyse the current global governance landscape concerning key technologies and identify valuable lessons applicable to the governance of artificial intelligence. Their analysis indicates the need for a nimble and recommendatory approach towards AI, allowing room for regulations to evolve into more direct, targeted, and binding measures as the technology matures.

TAKSHASHILA
INSTITUTION

In the concluding paper, Bharath Reddy explores how Artificial Intelligence can augment state capacity. The paper breaks down government processes and assesses them based on transaction volume and discretion. The paper identifies specific areas where AI systems can effectively augment state capacity through this analysis.

We welcome comments to build on and add to the ideas in this document. If you have any feedback, please get in touch with us at research@takshashila.org.in.

TAKSHASHILA
INSTITUTION

# Comparison of Artificial Intelligence (AI) Regulations in India, US, UK, European Union and China

*How is AI Currently Regulated in India? How Does it Compare with Other Prominent Economies?*

## Satya S Sahu & Shrikrishna Upadhyaya

## Introduction

The rapid advancement of Artificial Intelligence (AI) technologies has precipitated a global conversation on the need for comprehensive regulatory frameworks. This discourse underscores the critical balance between fostering innovation and addressing ethical, legal, and

societal concerns posed by AI. As AI technologies permeate various sectors — ranging from healthcare to finance — their regulation becomes imperative to ensure ethical use, data privacy, and accountability. This paper seeks to analyse the regulatory landscape of AI, with a specific focus on India. It extends the comparison to four major jurisdictions: the United States, the United Kingdom, the European Union, and China. Investigating these jurisdictions is vital for understanding the global regulatory stance on AI, providing insights into convergent and divergent trends.

The scope of this study is delineated by its focus on the overview of existing regulatory frameworks without delving too deeply into the potential applicability of existing civil and criminal legislation for AI regulation in India. The authors aim to map out the current regulatory terrain rather than propose specific legislative adaptations. "Regulation" in the context of this paper relates primarily to policy measures and legislations aiming to mitigate risks or harm that arise in the course of the AI lifecycle or use cases of AI technology. It does not include policy and governance measures that aim to build capability,

TAKSHASHILA
INSTITUTION

upskilling, develop sovereign infrastructure for leveraging AI, the adoption of AI technologies, or foster the underlying ecosystem.

This focus also recognises the critical role of regulation in addressing AI's immediate and tangible challenges rather than the broader and more diffused aspects of governance, encompassing the wider societal, ethical, economic, and organisational considerations of AI integration into society. The latter aspects are covered by both precise and ad hoc initiatives, like the Ministry of Electronics and Information Technology (MEITY) AI Committee reports,[1] and IndiaAI,[2] aiming to chart a roadmap for India's AI ambitions and challenges before it.

Structured in three sections, the paper first provides an overview of India's regulatory framework for AI, followed by comparative analyses

---

[1] Ministry of Electronics and Information Technology, Government of India. "Artificial Intelligence Committees Reports." Accessed February 12, 2024. https://www.meity.gov.in/artificial-intelligence-committees-reports.
[2] INDIAai. "Home Pillar." Accessed February 12, 2024. https://indiaai.gov.in/.

TAKSHASHILA
INSTITUTION

of AI regulations in the aforementioned jurisdictions. The final section discusses converging and diverging trends, offering insights into the global regulatory discourse on AI. This structure aims to furnish a nuanced understanding of current regulatory approaches, facilitating a discussion on the direction of global AI regulation, offering insights into the effectiveness of different regulatory strategies, and the potential for international regulatory coherence.

## AI Regulation in India

Over the past few years, India has launched various initiatives and set forth guidelines for creating and implementing AI technologies. However, as of February 2024, India has not yet introduced specific laws dedicated solely to regulating AI technologies or their applications, including Deepfakes[3] or AI-enabled dissemination of

---

[3] Sample, Ian. "What Are Deepfakes – and How Can You Spot Them?" The Guardian, January 13, 2020. https://www.theguardian.com/technology/2020/jan/13/what-are-deepfakes-and-how-can-you-spot-them.

misinformation. The current, actionable regulatory framework comprises mainly broader IT and data protection policies that indirectly cover AI technologies.[4] Therefore, India's approach to AI regulation focuses on creating ethical guidelines, tentative policy frameworks, and international collaboration. This is also borne out by the Union government's commitment to a light-touch regulatory environment, seeking to foster innovation in AI despite recognising the risks posed.[56] This approach would potentially leverage non-regulatory

---

[4] Desk, TOI News. "As European Union Passes World's First Law to Curb Artificial Intelligence, India Set on Different Approach." Times Of India, December 9, 2023. https://timesofindia.indiatimes.com/world/europe/as-european-union-passes-worlds-first-law-to-curb-ai-india-set-on-different-approach-artificial-intelligence-pm-modi/articleshow/105859769.cms?from=mdr.

[5] Das, Shouvik. "'AI Regulations Will Ideally Be Light Touch, Though Harm Is Concerning.'" Mint, December 7, 2023. https://www.livemint.com/news/india/ai-regulations-will-ideally-be-light-touch-though-harm-is-concerning-11701961915125.html.

[6] Desk, TOI News. "As European Union Passes World's First Law to Curb Artificial Intelligence, India Set on Different Approach." Times Of India, December 9, 2023. https://timesofindia.indiatimes.com/world/europe/as-european-union-passes-worlds-first-law-to-curb-ai-india-set-on-different-approach-artificial-intelligence-pm-modi/articleshow/105859769.cms?from=mdr.

TAKSHASHILA
INSTITUTION

instruments like "*market mechanisms, such as principles-based accreditation*"[7] instead.

That said, the Digital India Bill is speculated to introduce specific provisions regulating AI companies, models, and algorithms, focusing on addressing questions of bias and countering deepfakes and misinformation.[8] In the interim period before the law is tabled, the MEITY will likely introduce amendments to the IT Rules, 2021 to address the same issues.[9]

---

[7] Desk, TOI News. "As European Union Passes World's First Law to Curb Artificial Intelligence, India Set on Different Approach." Times Of India, December 9, 2023. https://timesofindia.indiatimes.com/world/europe/as-european-union-passes-worlds-first-law-to-curb-ai-india-set-on-different-approach-artificial-intelligence-pm-modi/articleshow/105859769.cms?from=mdr.

[8] "Government May Amend IT Act 2021 to Add These New Rules: Why and What They Are to Be." Times Of India, January 4, 2024. https://timesofindia.indiatimes.com/gadgets-news/government-may-amend-it-act-2021-to-add-these-new-rules-why-and-what-they-are-to-be/articleshow/106528324.cms.

[9] Aryan, Aashish. "Govt May Amend IT Act to Add New Rules for AI, GenAI Models." Economic Times, January 4, 2024. https://economictimes.indiatimes.com/tech/technology/govt-may-amend-it-act-to-add-new-rules-for-ai-genai-models/articleshow/106524019.cms?from=mdr.

TAKSHASHILA
INSTITUTION

## Non-Binding Guidelines, Principles and Frameworks

The National Institution for Transforming India (NITI Aayog, the government's apex think tank) has been pivotal, releasing the National Strategy for Artificial Intelligence #AIForAll,[10] exploring ethical considerations for AI deployment in 2018 and outlining a strategic vision for AI in sectors such as healthcare, agriculture, education, smart cities and infrastructure, and smart mobility and transformation.[11]

In February 2021, NITI Aayog published Part 1 — Principles for Responsible AI,[12] a paper detailing ethical considerations for AI

---

[10] INDIAai Web Portal. Accessed February 12, 2024. https://indiaai.gov.in/research-reports/national-strategy-for-artificial-intelligence/.

[11] Kapoor, Rahul, Shokoh H. Yaghoubi, and Theresa T. Kalathil. "AI Regulation in India Current State and Future Perspectives," January 26, 2024. https://www.morganlewis.com/blogs/sourcingatmorganlewis/2024/01/ai-regulation-in-india-current-state-and-future-perspectives.

[12] INDIAai. "Responsible AI: Part 1 - Principles for Responsible AI." Accessed February 12, 2024. https://indiaai.gov.in/research-reports/responsible-ai-part-1-principles-for-responsible-ai.

TAKSHASHILA
INSTITUTION

deployment in India, categorising them into ethical challenges, system and societal considerations. System considerations include decision-making principles, ensuring beneficiaries' inclusion, and accountability for AI decisions, while societal considerations examine automation's effects on job creation and employment. The paper identifies seven broad principles for responsible management of AI systems: 1) safety and reliability; 2) inclusivity and non-discrimination; 3) equality; 4) privacy and security; 5) transparency; 6) accountability; and 7) protection and reinforcement of positive human values.[13]

August 2021 saw the release of Part 2 — Operationalising Principles for Responsible AI[14], aiming at practically applying these principles. This segment outlines necessary actions for government and private

---

[13] Njie, Sheriff. "The Key Policy Frameworks Governing AI in India." Access Partnership, October 2, 2023. https://accesspartnership.com/the-key-policy-frameworks-governing-ai-in-india/.

[14] INDIAai. "Responsible AI: Part 2 - Operationalizing Principles for Responsible AI." Accessed February 12, 2024. https://indiaai.gov.in/research-reports/responsible-ai-part-2-operationalizing-principles-for-responsible-ai.

TAKSHASHILA
INSTITUTION

sector collaboration with research institutions, encompassing "*regulatory and policy interventions, enhancing capacity, promoting ethics by design, and adhering to AI standards.*"[15] The push to the private sector to self-regulate while developing responsible AI is quite evident in this paper.

Other key initiatives include India's involvement in the Global Partnership on Artificial Intelligence (GPAI) and the development of standards in line with international norms through the Bureau of Indian Standards (BIS).

India's participation in the Global Partnership on Artificial Intelligence (GPAI) is significant for its AI regulatory landscape primarily because it offers a direct pathway to integrating international standards and best

---

[15] Kapoor, Rahul, Shokoh H. Yaghoubi, and Theresa T. Kalathil. "AI Regulation in India Current State and Future Perspectives," January 26, 2024. https://www.morganlewis.com/blogs/sourcingatmorganlewis/2024/01/ai-regulation-in-india-current-state-and-future-perspectives.

practices into its domestic AI regulations. By adopting the 2023 Ministerial Declaration, India, alongside other GPAI members, *"reaffirmed their commitment to the responsible stewardship of trustworthy AI and values reflected in the OECD Recommendation on AI."*[16][17]

Therefore, any potential domestic regulation of AI would align with the OECD AI Principles' focus on promoting innovative, trustworthy AI and respecting human rights and democratic values. They emphasise inclusive growth, sustainable development, well-being, human-centred values, fairness, transparency, explainability, robustness, security, safety, and accountability.

---

[16] Barik, Soumyarendra. "Innovation or Safety? How Delhi Declaration Found Middle Path between Promotion and Regulation of AI." The Indian Express, December 14, 2023. https://indianexpress.com/article/explained/explained-sci-tech/delhi-declaration-gpai-regulation-ai-explained-9067865/.

[17] N.d. https://one.oecd.org/document/GPAI(2023)2/en/pdf.

TAKSHASHILA
INSTITUTION

The Bureau of Indian Standards (BIS) is working on draft Indian standards for AI that align with International Standards Organisation standards, focusing on risk management, among other areas.[18] These efforts are aimed at creating a robust framework for AI technologies in India, ensuring they meet global benchmarks for quality, safety, and reliability.[19] Draft standards corresponding to ISO/IEC 38507 and ISO/IEC TR 24368 provide principles for ethical use and regulatory compliance and frameworks for identifying, assessing, and mitigating risks associated with deploying and using AI systems. This includes addressing data privacy concerns, ensuring transparency in AI decision-making processes, and establishing accountability mechanisms for AI outcomes.[20]

---

[18]      "Standards."      Accessed      February      12,      2024. https://www.services.bis.gov.in/php/BIS_2.0/dgdashboard/Published_Standards_new/standards?com mttid=Mzg2&commttname=TElURCAzMA%3D%3D&aspect=&doe=&from=2022-07-21&to=2023-07-21.
[19] Njie, Sheriff. "The Key Policy Frameworks Governing AI in India." Access Partnership, October 2, 2023. https://accesspartnership.com/the-key-policy-frameworks-governing-ai-in-india/.
[20]      "Standards."      Accessed      February      12,      2024. https://www.services.bis.gov.in/php/BIS_2.0/dgdashboard/Published_Standards_new/standards?com mttid=Mzg2&commttname=TElURCAzMA%3D%3D&aspect=&doe=&from=2022-07-21&to=2023-07-21.

TAKSHASHILA
INSTITUTION

Following a brief consultation period, the Telecom Regulatory Authority of India (TRAI) published its suggestions titled 'Leveraging Artificial Intelligence and Big Data in the Telecommunication Sector' on July 20, 2023. It recommended establishing a common regulatory framework for AI across sectors, advocating for an independent statutory body to oversee AI governance, which includes categorising AI use cases by risk and imposing legally binding obligations on high-risk applications.[21]

The proposed framework suggests forming an independent statutory body designated as the "Artificial Intelligence and Data Authority of India" alongside a multistakeholder advisory group to assist it. Furthermore, it introduces a risk-based classification for AI applications, mandating that those with higher risk adhere to strict legal requirements. The statutory body would also be tasked with

---

[21] N.d. https://trai.gov.in/sites/default/files/PR_No.62of2023.pdf.

formulating guidelines on AI governance and ethical standards for uses considered lower risk, applicable to both governmental and private entities.[22]

## Legislations

For AI and AI use cases in India, the regulatory approach includes considering existing laws that can indirectly impact AI development and deployment, such as the Information Technology Act, 2000 (IT Act) and its subsidiary laws, for cybersecurity and cyber-offences, while the Digital Personal Data Protection Act can govern privacy concerns related to AI applications. However, these laws do not specifically address the unique risks and challenges AI technologies pose.

The IT Act includes provisions for the authentication of electronic records, digital signatures, and cybercrime offences. It addresses

---

[22] N.d. https://trai.gov.in/sites/default/files/PR_No.62of2023.pdf.

issues related to unauthorised access, data theft, and privacy violations. Though not AI-specific, its provisions on electronic content and cyber offences could be interpreted to include AI-generated content like deepfakes or to combat misinformation spread through digital platforms.[23]

Section 66D of the IT Act deals with punishment for cheating by personation using a computer resource or a communication device. It is particularly relevant for cases involving deepfakes used for impersonation or fraud.[24] At the same time, Sections 67, 67A and 67B can be applied to deepfakes involving obscene, sexually explicit content or Child Sexual Abuse Material (CSAM).[25] Similarly, Sections

---

[23] TNN. "Follow IT Rules: Govt to Social Media Platforms on Deepfakes." Times Of India, December 27, 2023. https://timesofindia.indiatimes.com/india/follow-it-rules-govt-to-social-media-platforms-on-deepfakes/articleshow/106302330.cms?from=mdr.

[24] "India Code: Section Details." Accessed February 12, 2024. https://www.indiacode.nic.in/show-data?actid=AC_CEN_45_76_00001_200021_1517807324077&orderno=80.

[25] "Cyber Crimes." Accessed February 12, 2024. https://pib.gov.in/PressReleasePage.aspx?PRID=1602398.

TAKSHASHILA
INSTITUTION

66E and 66C can be read with Sections 420 and 468 of the Indian Penal Code, 1860 (IPC), which criminalise AI use cases that involve cheating, identity theft, and forgery. Sections like 153A, 295A, and 505(1) of the IPC can be used against the spread of AI-generated misinformation that incites hatred, communal tensions, public mischief, etc. Finally, criminal defamation under Sections 499 and 500 of the IPC can also be invoked in either case.[26]

In a recent prominent deepfake case, the Delhi Police filed a case against the suspects in the matter under sections 465 (forgery) and 469 (harming reputation) of the IPC and sections 66C (identity theft) and 66E (privacy violation) of the IT Act.[27] The fact that AI may have

---

[26] Deshkar, Ankita. "Fake News and Misinformation: Sections of the Law That Police Can Invoke." The Indian Express, September 14, 2023. https://indianexpress.com/article/explained/everyday-explainers/fake-news-misinformation-ipc-sections-explained-8940115/.
[27] Desk, India Today News. "Rashmika Mandanna Deepfake Case: Delhi Police Tracks down 4 Who Uploaded Video." India Today, December 20, 2023. https://www.indiatoday.in/india/story/rashmika-mandanna-deepfake-case-delhi-police-action-bollywood-deepfake-videos-2478222-2023-12-20.

been involved in the creation of the deepfake is of no apparent importance in the initiation of legal recourse.

The Digital Personal Data Protection Act, 2023 (DPDPA) provides a foundational framework for regulating artificial intelligence (AI) use cases, particularly concerning data protection and privacy. Although it does not directly address AI, its data privacy, consent, and security provisions can significantly influence how AI systems are developed, deployed, and managed in India.

The definitions of a "Data Principal", "automated system", "processing", and "artificial juristic person" have not yet been tested vis-a-vis AI systems yet. However, the definitions are broad enough to encompass a diverse range of AI systems and, thus, be the crux of questions surrounding legal accountability.[28] However, it would be

---

[28] Abhivardhan. "The Digital Personal Data Protection Act & Shaping AI Regulation in India." Indic Pacific (blog), August 13, 2023. https://www.indicpacific.com/post/the-digital-personal-data-protection-act-shaping-ai-regulation-in-india.

TAKSHASHILA
INSTITUTION

necessary to harmonise liabilities that can possibly be incurred by AI systems under the DPDPA with the proposed Digital India Bill's purported risk-based coverage to prevent the creation of an onerous regulatory framework.
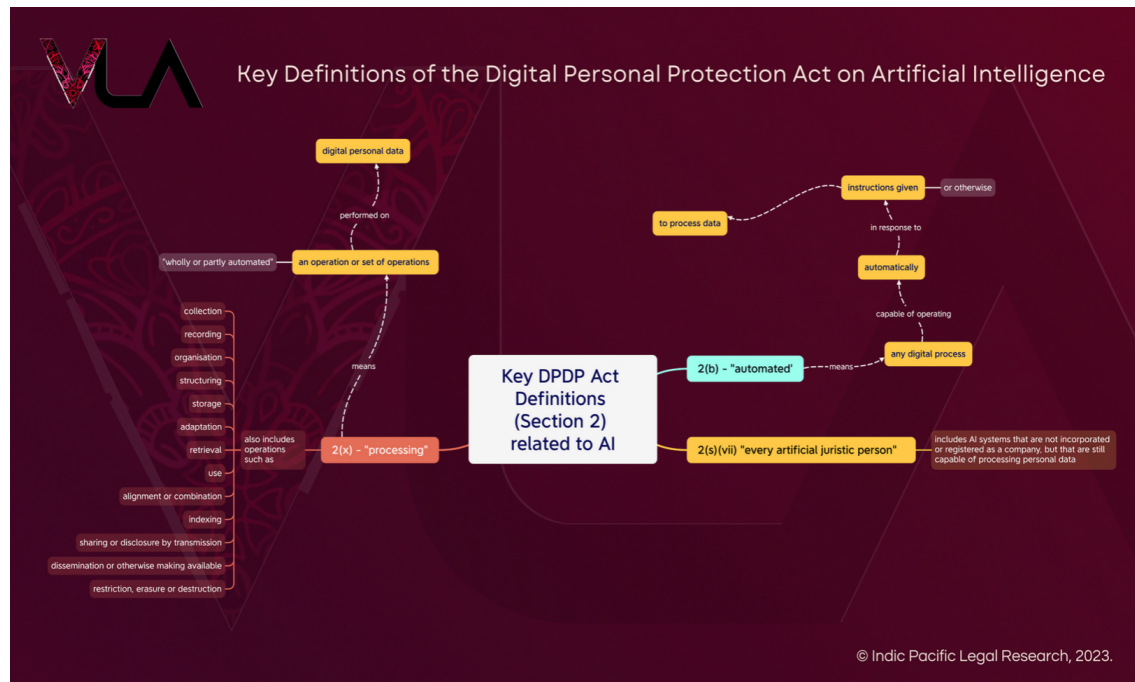
Figure 1: Key Definitions of the DPDP Act on AI (Indic Pacific Legal Research)[29]

The DPDP Act emphasises the protection of personal data through consent, transparency, and security measures, which are essential for AI applications that process vast amounts of personal information. The DPDPA ensures that AI systems respect user privacy and autonomy by requiring explicit consent for data collection and processing. Moreover, its focus on transparency obligates AI developers and operators to disclose how they collect, use, and protect personal data, fostering trust among users.

Another critical aspect of the DPDP Act relevant to AI regulation is the emphasis on data security. AI systems, being data-intensive, must adhere to stringent security protocols to prevent data breaches and unauthorised access. This requirement protects personal information and enhances the reliability and credibility of deployed AI systems. In

---

[29] Ibid.

identifying Significant Data Fiduciaries (SDFs), the Government will consider several criteria, notably the amount of data managed by the fiduciary and the potential threats to the privacy and rights of individuals. Key among these considerations is the potential repercussions on India's national integrity and sovereignty, along with concerns regarding the safeguarding of electoral democracy,[30] both of which are also concerns surrounding AI use cases.

The Act's provisions for penalties in cases of non-compliance underline the seriousness with which data protection is treated, thereby encouraging AI developers and companies to prioritise privacy and security in their operations. The requirement for data protection impact assessments (DPIAs) underscores the proactive obligation of SDFs to identify and mitigate data protection risks in AI applications.[31]

---

[30] LLP, Hunton Andrews Kurth. "India Passes Digital Personal Data Protection Act." Privacy & Information Security Law Blog, August 22, 2023. https://www.huntonprivacyblog.com/2023/08/22/india-passes-digital-personal-data-protection-act/.

[31] "India DPDP Act 2023." EY. Last modified August 2023. https://assets.ey.com/content/dam/ey-sites/ey-com/en_in/topics/cybersecurity/2023/08/ey-india-dpdp-act-2023.pdf?download.

TAKSHASHILA
INSTITUTION

## India's Complexity Problem

On January 30, 2024, the Economic Advisory Council to the Prime Minister (EAC-PM) unveiled a groundbreaking approach titled 'A Complex Adaptive System Framework to Regulate Artificial Intelligence.'[32] In stark contrast to the Union government's existing stance on a light-touch regulatory approach to AI, this framework employs a Complex Adaptive System (CAS) methodology for comprehensively regulating AI across various dimensions, including algorithms, training data sets, models, and applications. It outlines five key principles aimed at fostering a balanced and effective regulatory environment for AI:

---

[32] "Economic Advisory Council to the Prime Minister. 'A Complex Adaptive System Framework to Regulate Artificial Intelligence.' Last modified January 2024. https://eacpm.gov.in/wp-content/uploads/2024/01/EACPM_AI_WP-1.pdf."

1) *"Establishing guardrails and partitions — implement clear boundary conditions to limit undesirable AI behaviours, including creating 'partition walls' between distinct systems and within deep learning AI models to prevent systemic failures.*

2) *Mandating manual 'overrides' and 'authorization chokepoints' — critical infrastructure should include human control mechanisms at key stages to intervene when necessary, emphasising the need for specialised skills and dedicated attention without limiting the automation of systems. Manual overrides empower humans to intervene when AI systems behave erratically or create pathways to cross-pollinate partitions. Meanwhile, multi-factor authentication authorization protocols provide robust checks before executing high-risk actions, requiring consensus from multiple credentialed humans.*

3) *Ensuring transparency and explainability — open licensing of core algorithms for external audits, AI factsheets, and continuous monitoring of AI systems is crucial for accountability. There*

*should be periodic mandatory audits for transparency and explainability.*

4) *Defining clear lines of AI accountability — includes, among other things, establishing predefined liability protocols to ensure that entities or individuals are held accountable for AI-related malfunctions or unintended outcomes;*

5) *Creating a specialist regulator — this would also entail having a national registry of algorithms as compliance and a repository of national algorithms for innovations in AI.*"[33]

While innovative, as an actionable framework, it faces significant challenges due to its complexity and potential implementation difficulties. It runs the risk of stifling innovation through over-

---

[33] DataGuidance. "India: EAC-PM Publishes Complex Adaptive System Framework to Regulate AI," February 6, 2024. https://www.dataguidance.com/news/india-eac-pm-publishes-complex-adaptive-system.

regulation and manual intervention points, not to mention the logistical and legal hurdles in ensuring low-level explainability of AI systems, the oversimplification of AI accountability in a field marked by collaborative development and evolving capabilities, and concerns about the feasibility and effectiveness of a national registry of algorithms in keeping pace with rapid AI advancements.

## India's Regulatory Landscape at a Glance

| Policy/ Framework/ Guideline | Objective | Salient Features | Application to AI |
|---|---|---|---|
| IT Act, 2000, the IPC, 1860, and subsidiary legislation | To provide legal recognition for electronic transactions and address cybercrimes. | Legal framework for digital commerce and cybersecurity.<br><br>Includes provisions for digital verification, | Indirectly regulates AI through provisions on electronic content, cyber offences, and data protection. |

TAKSHASHILA
INSTITUTION

| Policy/ Framework/ Guideline | Objective | Salient Features | Application to AI |
|---|---|---|---|
| | | data protection, and privacy.<br><br>Broad definitions can be invoked to encompass a wide range of offending conduct. | |
| Digital Personal Data Protection Act, 2023 | To protect individuals' privacy in relation to the processing of their personal digital data. | Proposes a framework for protecting personal data, including provisions for data processing consent, data subject rights (access, correction, deletion), and data fiduciary obligations. | Indirectly impacts AI technologies that process personal data.<br><br>Sets a regulatory framework that could dictate how AI systems must handle personal data, including the |

| Policy/ Framework/ Guideline | Objective | Salient Features | Application to AI |
|---|---|---|---|
| | | Introduces penalties for non-compliance. | development of AI models that might process or generate personal information, thus indirectly affecting applications like deepfakes.<br><br>AI developers must design AI systems with data protection by design and default, ensuring compliance with consent mechanisms and data subject rights. |
| National Strategy for | To promote AI in India for | Outlines a vision for leveraging AI in key | While not a regulatory document, it |

TAKSHASHILA
INSTITUTION

| Policy/ Framework/ Guideline | Objective | Salient Features | Application to AI |
|---|---|---|---|
| Artificial Intelligence by NITI Aayog | economic growth, social development, and inclusive technology advancement. | sectors: healthcare, agriculture, education, smart cities, and infrastructure.<br><br>Advocates for ethical AI development, focusing on building AI research capabilities and fostering public-private partnerships. | influences regulatory intentions by promoting ethical AI use and suggesting the development of sector-specific guidelines and frameworks.<br><br>Impacts how AI technologies are developed and deployed, emphasising inclusivity, ethical use, and responsible AI practices. |

| Policy/ Framework/ Guideline | Objective | Salient Features | Application to AI |
|---|---|---|---|
| Advisory Guidelines and Recommendations (NITI Aayog, MEITY, TRAI, PM-EAC) | To recommend best practices for the ethical development and deployment of AI technologies. | Issued by various bodies like MeitY and NITI Aayog, these guidelines focus on ethical principles such as fairness, accountability, and transparency in AI.<br><br>Encourages stakeholder engagement and responsible AI practices. | Provides a foundation for responsible AI development and usage. While these guidelines are advisory, they shape the development of future AI-specific regulations and standards by outlining best practices for ethical AI, addressing issues like bias, privacy, and transparency.<br><br>These guidelines indirectly influence AI applications, including |

| Policy/ Framework/ Guideline | Objective | Salient Features | Application to AI |
|---|---|---|---|
| | | | content generation in the private sector. |
| Global Partnership on Artificial Intelligence (GPAI) | To support and guide the responsible development and use of AI based on human rights, inclusion, diversity, innovation, and economic growth. | An international and multi-stakeholder initiative that brings together experts from academia, industry, and government.<br><br>Focuses on bridging the gap between theory and practice in AI, promoting responsible AI research and policies.<br><br>Works on projects under themes like | While GPAI is not a regulatory body, India's membership in GPAI indicates a commitment to international collaboration on ethical AI standards and practices.<br><br>GPAI's work influences global AI policy discussions and can inform domestic policy development, encouraging |

TAKSHASHILA
INSTITUTION

| Policy/ Framework/ Guideline | Objective | Salient Features | Application to AI |
|---|---|---|---|
|  |  | responsible AI, data governance, the future of work, and innovation. | alignment with international best practices for responsible AI use. |

# How is AI Regulated in Other Jurisdictions

## United States

The US President Joe Biden issued an Executive Order on the Safe, Secure, and Trustworthy Development and Use of Artificial

TAKSHASHILA
INSTITUTION

Intelligence on October 30, 2023 ("Executive Order").[34] It recognises AI's promise and perils, emphasising the need for responsible use to mitigate substantial risks like discrimination, disinformation, job losses and threats to national security. It calls for a coordinated effort involving government, private sector, academia, and civil society.

The Executive Order lists eight guiding principles: ensuring AI systems are safe and secure; promoting responsible AI innovation, competition and collaboration while guarding against risks; supporting the American workforce in adapting to AI changes; ensuring American values in administration, including equity and civil rights, through AI; protecting interests of American users of AI products through consumer protection; protecting Americans' privacy and civil liberties, including their data privacy; managing risks from Federal

---

[34] The White House. "Executive Order on the Safe, Secure, and Trustworthy Development and Use of Artificial Intelligence." Accessed February 12, 2024. https://www.whitehouse.gov/briefing-room/presidential-actions/2023/10/30/executive-order-on-the-safe-secure-and-trustworthy-development-and-use-of-artificial-intelligence/.

Government's use of AI and improving its state capacity to govern AI through hiring AI professionals; and promoting American AI leadership globally by engaging internationally to promote trustworthy AI development and use.

The Executive Order provides a list of definitions, including AI, AI models, AI systems, critical infrastructure, dual-use foundation models, generative AI, machine learning and synthetic content. It contains a mix of regulatory obligations and policy measures for each principle. For every regulatory requirement or government action under each principle, it identifies the relevant nodal authority for carrying out the functions and timelines.

To ensure the safety and security of AI technology, the Order proposes the following:

1. Developing guidelines, standards and best practices for promoting consensus industry standards for the safe deployment of AI systems and regular audits

TAKSHASHILA
INSTITUTION

2. The energy department will develop tools to evaluate AI capabilities to generate outputs that may represent nuclear, nonproliferation, biological, chemical, critical infrastructure, and energy-security threats or hazards

3. In furtherance of national defence and protection of critical infrastructure objectives under the Defence Protection Act, companies developing dual-use foundational models are bound to report to the Federal Government with prescribed information and records

4. To address malicious cyber-enabled activities, all US Infrastructure as a Service (IAAS) providers are required to comply with additional record-keeping and identity verification obligations with respect to foreign transactions

5. For AI in critical infrastructure and cybersecurity, concerned departments will issue best practices and risk management framework guidelines for companies to follow, and the Artificial Intelligence Safety and Security Board will also be set up as an advisory body to the Federal Government

6. To reduce risks at the intersection of AI and Chemical, Biological, Radiological and Nuclear (CBRN) weapons, measures like expert consultation and a sponsored study of synthetic biology threats by premier institutions will be undertaken and biosecurity requirements for biosecurity standards for the nucleic acid synthesis industry will be imposed

7. To reduce the risks of synthetic content, capabilities for identifying and labelling synthetic content produced by AI systems and to establish the authenticity and provenance of digital content, both synthetic and not synthetic, produced by the Federal Government or on its behalf will be fostered

8. For promoting the safe release and preventing malicious use of government data, guidelines for performing security reviews, including reviews to identify and manage the potential security risks of releasing Federal data that could aid in the development of CBRN weapons and development of autonomous offensive cyber capabilities will be released

9. The NSA shall develop a National Security Memorandum on AI on the governance of AI used as a component of a national security system or for military and intelligence purposes

To promote competition in AI, the FTC and other agencies must address risks arising from *"concentrated control of key inputs, taking steps to stop unlawful collusion and prevent dominant firms from disadvantaging competitors, and working to provide new opportunities for small businesses and entrepreneurs"*.[35]

To strengthen AI and civil rights in the criminal justice system, the Attorney General shall enforce federal laws to protect, prevent and address civil liberties violations and discrimination related to AI, including algorithmic discrimination. Federal agencies are also required to ensure there is no denial or discrimination in civil rights related to government benefits and programmes due to the use of

---

[35] Ibid.

TAKSHASHILA
INSTITUTION

automated or algorithmic systems. The scope of the Order includes protecting civil rights and preventing discrimination in hiring by federal contractors, housing markets and consumer financial markets.

Regulatory action by relevant authorities is encouraged to protect American consumers from fraud, discrimination, and threats to privacy, especially in the healthcare and human services sector (including research and drug development). Similarly, various agencies have been tasked with developing safety standards for using AI in automated transportation systems.

For privacy risks exacerbated by AI, steps will be taken to identify and evaluate commercially available information with federal agencies and revise guidelines for implementing privacy under relevant statutes.

A new council will be established for AI management to coordinate the use of AI across the federal government. Certain state capacity-building measures, including attracting and retaining AI talent within the government, were also announced.

TAKSHASHILA
INSTITUTION

A White House Artificial Intelligence Council (White House AI Council) was also established to *"coordinate the activities of agencies across the Federal Government to ensure the effective formulation, development, communication, industry engagement related to, and timely implementation of AI-related policies, including policies set forth in this order."*[36]

Other than the above measures, various other policy measures are proposed under the Executive Order for American workers protection, attracting AI to the US and enabling US leadership in the international AI ecosystem. Notably, most of the regulations are directed at the Federal Government and its agencies. This inward-looking focus of the Order reflects the importance of identifying and mitigating AI risks within the government and its agencies.

---

[36] Ibid.

TAKSHASHILA
INSTITUTION

## United Kingdom

The United Kingdom (UK) government published an AI Regulation White Paper in March 2023[37] and called for consultation. On February 6, 2024, the government gave its responses to the earlier consultation paper titled 'A pro-innovation approach to AI regulation: government response'.[38] The response of the UK government gives insight into the current thinking on AI regulation in the country and implications for businesses and individuals in the future.

---

[37] Department of Science, Innovation & Technology, Government of the United Kingdom. "A pro-innovation approach to AI regulation." Accessed February 12, 2024. https://assets.publishing.service.gov.uk/media/64cb71a547915a00142a91c4/a-pro-innovation-approach-to-ai-regulation-amended-web-ready.pdf.

[38] Department of Science, Innovation & Technology, Government of the United Kingdom. "Consultation outcome: A pro-innovation approach to AI regulation: government response." Accessed February 12, 2024. https://www.gov.uk/government/consultations/ai-regulation-a-pro-innovation-approach-policy-proposals/outcome/a-pro-innovation-approach-to-ai-regulation-government-response#introduction.

The government has stated that the existing approach involving "*combining cross-sectoral principles and a context-specific framework, international leadership and collaboration, and voluntary measures on developers*"[39] will continue for now. Given the rapid and uncertain developments in the field of AI, the document notes that legislative action will be needed once a greater understanding of AI risks is achieved. It also warns that responding prematurely will not only hinder technological progress but also inhibit quick adaptation to emerging risks.

The UK has opted for a less centralised approach to AI regulation than the EU, placing greater responsibility on existing sector-specific regulators. The government has put forth a pro-innovation stance, emphasising the need for regulations that do not unnecessarily hinder technological advancement.

---

[39] Ibid.

TAKSHASHILA
INSTITUTION

The regulatory approach towards AI systems categorises them based on their capabilities and scope of application, aiming to effectively address risks posed by the most powerful models. These categories include

1. **Highly Capable General-Purpose AI**: These are foundation models designed to perform a wide array of tasks, ranging from beginner to expert levels. Some may exhibit superhuman performance across various tasks.
2. **Highly Capable Narrow AI**: Unlike their general-purpose counterparts, these foundation models specialise in a limited set of tasks within specific domains, such as biology. They match or surpass the most advanced models in their specialised field, often demonstrating superhuman abilities in their focused tasks or domains.
3. **Agentic AI or AI Agents**: This category refers to AI technologies capable of independently completing complex, multi-step tasks over extended periods.

TAKSHASHILA
INSTITUTION

This classification helps in tailoring regulatory measures that are proportionate to the risks and capabilities of different AI systems, ensuring that oversight is effective without stifling innovation.

The UK has proposed a framework for AI regulation that will be pro-innovation, proportionate, trustworthy, adaptable, clear and collaborative. The elements of this approach are:

1. Defining AI based on its unique characteristics to support regulator coordination, i.e., a definition of AI based on adaptivity and autonomy of the application
2. Adopting a context-specific approach i.e. regulating the use and not the technology itself. Risk levels or rules will not be assigned to the entire sector. Instead, regulation will be based on the outcomes of use of AI in different contexts.
3. The UK framework is based on five cross-sectoral principles. Instead of rigid legislative requirements, the principles will be issued on a non-statutory basis and implemented by existing regulators. Initially, this will be at the discretion of the regulators

TAKSHASHILA
INSTITUTION

to enable them to prioritise. Later, a statutory duty to regard these principles will be imposed on the regulators. These principles are:

- Safety, security and robustness.
- Appropriate transparency and explainability.
- Fairness.
- Accountability and governance.
- Contestability and redress.

4. New central functions to support this framework: A new central function within government "to monitor and assess risks across the whole economy and support regulator coordination and clarity" will be set up. This body will consist of a cross-disciplinary team to look at AI risk assessment and come up with risk management frameworks similar to those in the US. It will also host regulatory sandboxes and carry out education and awareness functions.

The UK Competition and Markets Authority, Information Commissioner's Office, Office of Gas and Electricity Markets (Ofgem), Civil Aviation Authority (CAA), and Medicines and Healthcare products Regulatory Agency have already taken various steps to implement the principles outlined above in their regulatory efforts. The regulators have been tasked with releasing their respective strategic approach to AI by April. Additionally, the government plans to bolster the capability of existing regulators through a GBP 10 million funding pipeline and identify and mitigate regulatory gaps.

## European Union

The EU member states have agreed on the final text of the upcoming AI regulation. The regulations will be voted on in the European Parliament in March or April and ratified thereafter. This paper relies on the pre-final text of the regulations published by the media. Once enacted, these regulations will be the most comprehensive AI regulations worldwide.

The Regulation of the European Parliament and of the Council Laying Down Harmonised Rules on Artificial Intelligence (Artificial Intelligence Act) and Amending Certain Union Legislative Acts ("EU AI Regulations") is being brought in with the objective of "*improv[ing] the functioning of the internal market and promoting the uptake of human centric and trustworthy artificial intelligence, while ensuring a high level of protection of health, safety, fundamental rights enshrined in the Charter, including democracy, the rule of law and environmental protection against harmful effects of artificial intelligence systems in the Union and supporting innovation.*"[40]

The EU AI Regulations adopt a risk-based approach and classify AI applications into three categories: unacceptable risk, high-risk and

---

[40] European Commission. "Regulation Of The European Parliament And Of The Council Laying Down Harmonised Rules On Artificial Intelligence (Artificial Intelligence Act) And Amending Certain Union Legislative Acts" Accessed February 12, 2024. https://eur-lex.europa.eu/resource.html?uri=cellar:e0649735-a372-11eb-9585-01aa75ed71a1.0001.02/DOC_1&format=PDF.

TAKSHASHILA
INSTITUTION

non-high-risk. The Act bans certain AI practices deemed a threat to fundamental rights and falls within the unacceptable risk category. The Act imposes various compliance requirements on high-risk AI applications, including general purpose AI (GPAI). Unlike the US AI regulations, the EU law on AI is applicable to all actors and proposes significant compliance obligations on businesses.

The Regulations have wide application and cover providers of AI systems in the EU market, whether based in the EU or outside, deployers of AI systems, providers and deployers located outside the EU but whose output is used in the EU, importers and distributors, product manufacturers and affected persons. It exempts AI systems put in place for military, defence, and national security purposes, research, testing and development activities in non-live environments and individuals using AI in personal, non-professional activity. All AI systems released under free and open-source licences are exempt unless they fall under the high-risk category. The Regulations also provide detailed definitions.

The following are listed as prohibited AI practices under the Regulations:

1. Subliminal techniques beyond a person's consciousness or purposefully manipulative or deceptive techniques to distort behaviour and impair decision-making
2. Systems that exploit any of the vulnerabilities of a person or a specific group of persons due to their age, disability or a specific social or economic situation to distort behaviour and impair decision-making
3. Biometric categorisation systems that categorise natural persons individually based on their biometric data to deduce or infer their race, political opinions, trade union membership, religious or philosophical beliefs, sex life or sexual orientation.
4. Systems for the evaluation or classification of natural persons or groups thereof over a certain period of time based on their social behaviour or known, inferred or predicted personal or personality characteristics, with the social score leading to specified detrimental or unfavourable treatment

TAKSHASHILA
INSTITUTION

5. Use of 'real-time' remote biometric identification systems in publicly accessible spaces for the purpose of law enforcement unless it falls under specified exceptions like terrorist attacks, search for victims, localisation and identification of criminals, etc., subject to prior authorisation

The classification rules for high-risk AI systems mandate that any AI system shall be considered high-risk if the AI system is intended to be used as a safety component of a product or the AI system is itself a product covered by specific EU harmonisation legislations and the product is required to undergo a third-party conformity assessment. Some designated high-risk AI system examples include critical infrastructure, biometric identification, access to essential public and private services, education and vocational training, employment, worker management and access to self-employment, medical devices, such as diagnostic tools or treatment planning systems, law enforcement, migration, asylum and border control management and administration of justice and democratic processes. Any GPAI is also generally considered a high-risk AI system.

The Regulations also provide that "*AI systems shall not be considered as high risk if they do not pose a significant risk of harm, to the health, safety or fundamental rights of natural persons, including by not materially influencing the outcome of decision making*",[41] particularly if these are used to make narrow procedural decisions or improve results of human activities. A provider can avail exemptions from the EU Commission upon satisfying specific criteria.

Much of the Regulations is devoted to compliance requirements for high-risk AI systems. These include establishing risk management systems, data governance and management practices for testing and validation of data, technical documentation, record keeping including automated logs, transparency and information sharing requirements towards deployers, human oversight through appropriate human-machine interface tools commensurate with the AI autonomy and risk,

---

[41] Ibid.

TAKSHASHILA
INSTITUTION

and appropriate levels of accuracy, robustness, and cybersecurity in operations.

These compliances have to be met by the providers of high-risk AI systems who need to put in place quality management systems, record keeping, corrective action, and cooperation with authorities' mechanisms. Certain quality assurance and legal compliance obligations also fall upon importers and distributors of AI systems. Deployers of AI systems must also comply with technical and human oversight requirements.

If the deployers are government bodies, public institutions, and financial and insurance companies, they must also perform a fundamental rights impact assessment for high-risk AI systems before deployment, which provides for intended use, time and frequency of use, affected natural persons, human oversight, specific harms that may accrue and measures for risk mitigation. The Regulations also provide for standards, common specifications, conformity assessment, certificates, and registration processes for high-risk AI systems.

TAKSHASHILA
INSTITUTION

The transparency obligations for providers and users of AI systems and GPAI models provide for concerned natural persons to be informed that they are interacting with an AI system; synthetic audio, image, video or text content, being outputs of the AI system, are marked in a machine-readable format and detectable as artificially generated or manipulated; and image, audio or video content constituting a deep fake, shall disclose that the content has been artificially generated or manipulated.

A GPAI will be classified as GPAI with systemic risk if it meets specific criteria and has high-impact capabilities evaluated on the basis of appropriate technical tools and methodologies. There are immediate notifications and certain technical obligations imposed on the providers of GPAI with systemic risk.

For the purposes of governance, the Regulations provide for the establishment and structure of the European Artificial Intelligence Board at the Union level, as well as a system of notified bodies at the level of a member state. The new AI Regulator is proposed to oversee

advanced AI models and promote safety standards. Apart from this, the Regulations also establish regulatory sandboxes for innovation and measures for developing SMEs.

The Regulations also provide for a list of remedies for dealing with contraventions, including the right of any natural or legal person to make a complaint to the relevant national authority if they have grounds to consider that there has been an infringement of its provisions. To deal with non-compliance, the Regulations propose a regime of penalties. Member states can set rules on penalties and other enforcement measures, including warnings and non-monetary measures.

There is a graded system of penalties. For non-compliance with prohibited AI practices, administrative fines of up to 35,000,000 Euros or 7% of the company's global turnover, whichever is higher, can be levied. For violations by high-risk AI system providers in relation to their special obligations, penalties can go up to fines of up to the higher of 15,000,000 Euros or 3% of the global turnover of the company. For

TAKSHASHILA
INSTITUTION

supplying incorrect or misleading information, fines of up to 7,500,000 Euros or up to 1 % of the company's global turnover, whichever is higher, are applicable. A separate penalty structure for Union and government bodies is also provided.

The Regulations are expected to be enforceable two years after its official publication. A glide path for implementation over the next few years up till 2030 is provided under the Regulations.

## China

The People's Republic of China will likely bring comprehensive AI legislation later in 2024. So far, it has enforced piecemeal regulations tackling specific subsets of the AI universe, such as algorithm

TAKSHASHILA
INSTITUTION

recommendation technology, deep synthesis technologies (Deepfakes), and management of generative AI.[42]

Administrative Provisions on Algorithm Recommendation for Internet Information Services, 2022 ("Algorithmic Recommendation Regulation") regulates all algorithm recommendation technologies that provide online or internet information services in China, including generative and synthetic algorithms.

It imposes the following obligations on the service providers:

1. On or before the launch of any AI-powered service or product, the service providers must file the relevant algorithm with the relevant government authority. This applies to AI-powered products, services, or applications that have "*public opinion*

---

[42] Latham & Watkins. "China's New AI Regulations." Latham & Watkins Client Alert Commentary, August 16, 2023. Accessed February 12, 2024. https://www.lw.com/admin/upload/SiteAttachments/Chinas-New-AI-Regulations.pdf.

*attributes or social mobilization capabilities*", which are defined to mean "*forums, blogs, microblogs, chat rooms, communication groups, public accounts, short videos, webcasts*" or other "*internet information services that provide channels for the public to express their opinions.*"[43]

2. They must conduct a security assessment as per prescribed regulations and submit a report to the local city-level cyberspace administration.

3. They must abide by the laws and regulations, respect social morality and ethics, and uphold principles of impartiality, fairness, openness, transparency, scientific rationality and honesty.

4. Refrain from indulging in manipulation using AI recommendation services, including fake accounts, false lists and search results, blocking information, falsifying likes, comments, shares, etc.

---

[43] Ibid.

TAKSHASHILA
INSTITUTION

5. Refrain from creating monopolies or engaging in anti-competitive behaviour, including imposing unreasonable restrictions on other internet service providers.
6. Adhere to mainstream value orientation, promote socialist values and design services in a manner that "actively spread positive energy", and ensure that the homepage/ first screen is aligned with these values.
7. Follow prohibitions on the creation and dissemination of fake news and other fake content; take measures to prevent and stop the dissemination of false, illegal, or harmful information and have screening mechanisms to spot fake and illegal content
8. Protect user rights such as the right not to be targeted based on personal information, the right to opt out, the right to delete personal information and the right against differentiated treatment.
9. Disclose principles, purposes and operational mechanisms of deployment.
10.     Protection to minors, elderly and workers against user addiction and over-consumption

TAKSHASHILA
INSTITUTION

11.    Put in place user complaints and redress mechanisms like online portals and feedback mechanisms

For violations of the Algorithmic Recommendation Regulation, penalties like a fine of CNY10,000 to CNY100,000, suspension of information updates, warnings and public denouncements have been prescribed.

The Provisional Provisions on Management of Generative Artificial Intelligence Services, 2023 ("Generative AI Regulation"), apply to all generative AI technologies used to provide goods and services in China. These include "*models and related technologies that have the ability to generate texts, pictures, sounds, videos, and other content.*"[44] It imposes the following obligations on service providers, which include technical providers who extend access through APIs:

---

[44] Ibid.

TAKSHASHILA
INSTITUTION

1. On or before launch, generative or synthetic algorithms must be filed with relevant government agencies if they have "*public opinion attributes or social mobilisation capabilities*".[45]
2. They must conduct a security assessment as per prescribed regulations and submit a report to the local city-level cyberspace administration.
3. Refrain from creating monopolies or unfair competition by taking advantage of the algorithms, data, and platforms
4. Pre-training, optimising training and data processing shall comply with the law and ensure that data is from lawful sources, data does not infringe on third-party IP rights, consent of users is obtained, etc.
5. Comply with data labelling requirements to ensure identification of raw data used to train AI models and extend cooperation to regulators in explaining training data.
6. Protect users' input data and usage records.

---

[45] Ibid.

TAKSHASHILA
INSTITUTION

7. Generative AI must reflect socialist core values and not generate content which harms national security or propagates extremism, ethnic hatred, terrorism, pornography etc.
8. Take measures to enhance generated content's accuracy, authenticity, and reliability and identify and take down illegal content.
9. Labelling requirements for all information generated by AI and deep synthesis methods.
10.     Take measures to protect minors from becoming addicted to generative AI services.

For violation of any of the obligations contained in the Generative AI Regulation, penalties, as prescribed under other cybersecurity, data security, and personal information laws, will be applicable.

The Provisions on Management of Deep Synthesis in Internet Information Service, 2023 ("Deep Synthesis Regulation") are aimed at Deepfakes and apply to any use of deep synthesis technologies to provide internet information services in China. The use cases include

TAKSHASHILA
INSTITUTION

generating or editing text and voice content, non-speech content, biometric features, 3D reconstruction and digital simulation. Certain obligations are imposed on service providers, technical supporters and users of such services.

Similar to the Algorithmic Recommendation Regulation and Generative AI Regulation, the Deep Synthesis Regulation imposes several obligations upon service providers relating to filing of algorithms, security assessments, compliance with laws and social morality, maintaining mainstream values orientation and adherence to correct political direction, prohibition on fake news and false information generation and dissemination, screening of illegal and other harmful information and corrective measures, user complaints redressal, etc.

In addition, the Deep Synthesis Regulation also contains provisions on procedural review and management measures by online app distribution platforms, verification of users' real identities using phone numbers, ID cards, etc. before extending information dissemination

services, notification and consent of subjects whose biometrics (like faces and voice) is edited, explicit and prominent labeling of deepfake content as synthetic, and measures for dispelling rumours.

While there are no explicit penalties under the Deep Synthesis Regulation, it instead refers to other laws under which defaulting service providers can be penalised. There are no penalties for users so far. For instance, the cyberspace department may suspend information updates or new user registration against defaulting service providers.

Finally, Chinese state-owned research institutes have also released negative lists of AI technologies of areas and products that AI companies should avoid unless they have explicit government approval.[46]

---

[46] Yang, Zeyi. "Four Things to Know about China's New AI Rules in 2024." MIT Technology Review, January 17, 2024. https://www.technologyreview.com/2024/01/17/1086704/china-ai-regulation-changes-2024/.

# Trends in Global AI Regulation

## Convergent Trends

1. **Risk-Based Regulation**: There's a consensus on adopting a risk-based approach to AI regulation, where AI systems are categorised based on their potential harm to society and individuals. The EU explicitly adopts this with its classification into unacceptable risk, high-risk, and non-high-risk categories. At the same time, the US, UK, and China implicitly follow a similar approach through their targeted regulations on specific AI applications. The US has not adopted a formal risk-based classification but has emphasised the need to balance the benefits and harms of AI in different sectoral contexts. China has not proposed a clear risk-based classification but has highlighted

the importance of ensuring the safety, reliability, and controllability of AI.

However, the conception and the extent of risks associated with AI can differ drastically between jurisdictions. For example, China's strict regulation of generative AI,[47] is not reflected in its approach to AI-enabled facial and voice recognition/surveillance technology.[48]

2. **Global Collaboration on AI Governance and Standards:** The push for international collaboration is evident, with entities like the GPAI involving multiple countries, including India, in setting global AI governance standards, which consequently inform domestic policy.

---

[47] Xiang, Nina. "China's AI Sector Has No Time for End-of-the-World Worries." Nikkei Asia, December 18, 2023. https://asia.nikkei.com/Opinion/China-s-AI-sector-has-no-time-for-end-of-the-world-worries.

[48] Press, Associated. "China Warns of Artificial Intelligence Risks, Calls for Increased National Security Measures." PBS NewsHour, May 31, 2023. https://www.pbs.org/newshour/world/china-warns-of-artificial-intelligence-risks-calls-for-increased-national-security-measures.

TAKSHASHILA
INSTITUTION

3. **Privacy and Data Protection Concerns**: Data protection remains a critical aspect of AI regulation, with comprehensive privacy laws like the GDPR in the EU influencing global discussions. The US's state-specific regulations, such as California's CPPA, reflect a fragmented but constant focus on privacy in AI applications.[49]

## Divergent Trends

1. **Scope and Centralisation of Regulatory Frameworks**: The US adopts an inward-looking approach focused on guiding principles for federal agencies, emphasising sector-specific regulations, while the UK opts for a decentralised, sector-specific regulatory role to avoid stifling innovation. Meanwhile, the European Union proposes comprehensive, centralised regulations applicable across all AI actors, with stringent

---

[49] "California Privacy Protection Agency (CPPA)." Accessed February 12, 2024. https://cppa.ca.gov/.

requirements for high-risk applications. China's AI regulations predominantly address specific applications and provide detailed guidelines for companies, including restrictions on AI-generated content. This focus contrasts with the broader, more general regulatory discussions in the EU, UK, and US.

The manner in which different jurisdictions prioritise their governance measures also reflects the divergent assumptions regardin their state capacity. While the US approaches the question granularly fostering talent and upskilling its population vis a vis AI technologies, the UK takes a more hands off approach, letting market forces take the driver's seat for now. The EU's comprehensive legislative framework and China's strict penal provisions showcase a contrasting position.

2. **Legislative Speed and Specificity**: The UK's cautious, principles-based approach to AI regulation contrasts with the EU's swift *ex-ante* legislative actions. Meanwhile, the US has not proposed federal AI legislation, relying instead on *ad hoc* periodic

executive orders and state-specific measures. China's comprehensive AI legislation is likely to come out later this year.

3. **Innovation vs. Precaution**: The U.S. and U.K. prioritise fostering AI innovation with policies designed to promote leadership and adaptability in AI development. EU, while supportive of innovation, adopts a more precautionary stance with its stringent compliance obligations for high-risk AI systems, suggesting a balance between innovation and risk management. China emphasises controlling and promoting socialist values in its AI development, focusing on national interests alongside innovation.

## Conclusion

The examination of AI regulation across the abovementioned jurisdictions reveals a landscape shaped by varying priorities, strategies, and stages of legislative development. The primary

TAKSHASHILA
INSTITUTION

balancing concerns seem to be innovation facilitation and risk mitigation, with the exception of China's focus on preserving national and socialistic values as it regulates generative AI. India's regulatory posture is marked by an absence of specific AI laws but a potential for comprehensive coverage under the anticipated Digital India Act. However, India's participation in the GPAI and existing rhetoric from the Union government indicates that future legislation would align with global trends towards ethical, secure, and responsible AI development.

Comparatively, the U.S. and UK's emphasis on fostering innovation, the EU's comprehensive risk-based regulatory framework, and China's focused control over AI applications illustrate the global regulatory spectrum from light-touch principles to stringent legislative mandates.

TAKSHASHILA
INSTITUTION

# AI & Jobs

## Sridhar Krishna

## Executive Summary

The world economy was stagnant in real terms for centuries before the industrial revolution towards the latter part of the 18th century. The industrial revolutions of the past have accelerated growth, and contributed towards poverty alleviation across the world. Yet during times of change, they disrupted jobs and often ways of life. Many protested but the economy grew leaps and bounds.

With the current industrial revolution underway, many of us shudder to imagine a world where artificial intelligence (AI) can potentially replace millions of jobs, and render a majority of humankind jobless. Purely by past evidence, one could shrug off that fear and have faith in the world's ability to find new things to do.

Earlier revolutions often replaced routine tasks with machines that performed them faster and better; now, the threat is to well-paid and highly-skilled jobs. We wonder — and worry — whether the entire human race may be rendered irrelevant.

This document touches upon history, tries to understand the science, and shares ideas about the future. There are more questions than answers, of course, but it is an attempt to understand the fear and find out which of the fears are unfounded and which ones are not.

## The Fears

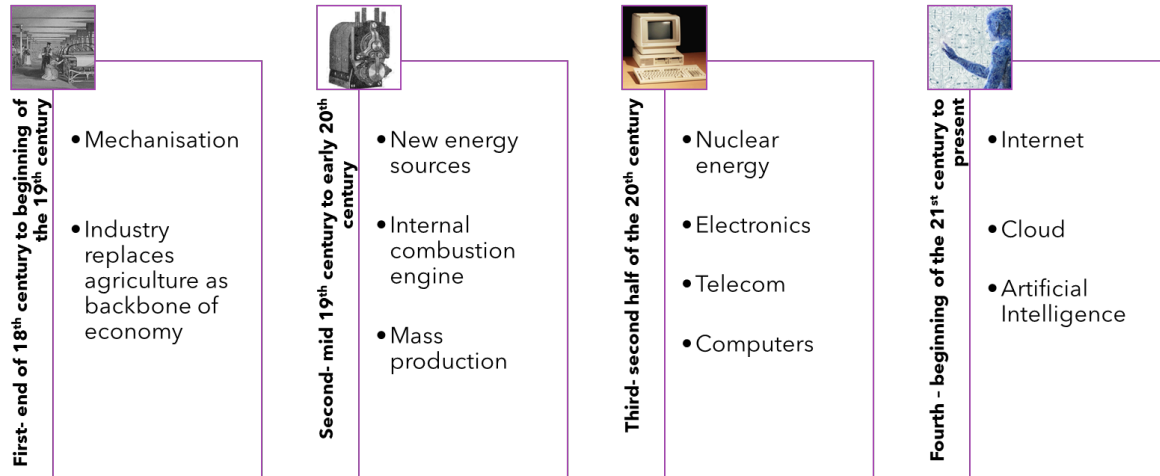Disruptive technology has often instilled a fear of  impact on employment.

British economist John Maynard Keynes spoke about technological unemployment almost a century ago, in 1930. He felt that, by 2030,

the world would have solved the problem of ensuring there is enough for everyone to meet their basic needs through the power of compounding and with scientific progress, but that there may not be enough for everyone to do in terms of work. He articulates this in an essay titled "Economic Possibilities for our Grandchildren." (Keynes 1930)

Oxford University economist Daniel Susskind in 'A World without Work', — which he wrote during the Covid-19 pandemic — comes to a similar conclusion. (Susskind 2021) We shall talk about that in a bit. Each industrial revolution since the first one in the late 18th century has had a significant impact on jobs, disrupted economies, dislodged people, and caused heartburn and anxiety to many while opening new opportunities and a shift in the balance of power.

## The Facts

Here are the four industrial revolutions and how they impacted jobs.

TAKSHASHILA
INSTITUTION

**First- end of 18th century to beginning of the 19th century**

- Mechanisation
- Industry replaces agriculture as backbone of economy

**Second- mid 19th century to early 20th century**

- New energy sources
- Internal combustion engine
- Mass production

**Third- second half of the 20th century**

- Nuclear energy
- Electronics
- Telecom
- Computers

**Fourth - beginning of the 21st century to present**

- Internet
- Cloud
- Artificial Intelligence

The industrial revolutions brought with them unprecedented growth, and while there were fears of job losses, there was significant growth in employment and real wages. The apprehensions of Luddites over the ages have been proven to be unfounded.

Each industrial revolution impacted jobs in different ways.

1. **First industrial revolution:** A bias in favour of unskilled workers over skilled workers
2. **Second Industrial Revolution:** A tilt in bias towards formal education but continued hiring of unskilled workers.
3. **Third Industrial Revolution:** A strong bias towards formal education but rising inequality within countries.
4. **The Fourth Industrial revolution:** This seems to threaten even well-paying highly skilled jobs
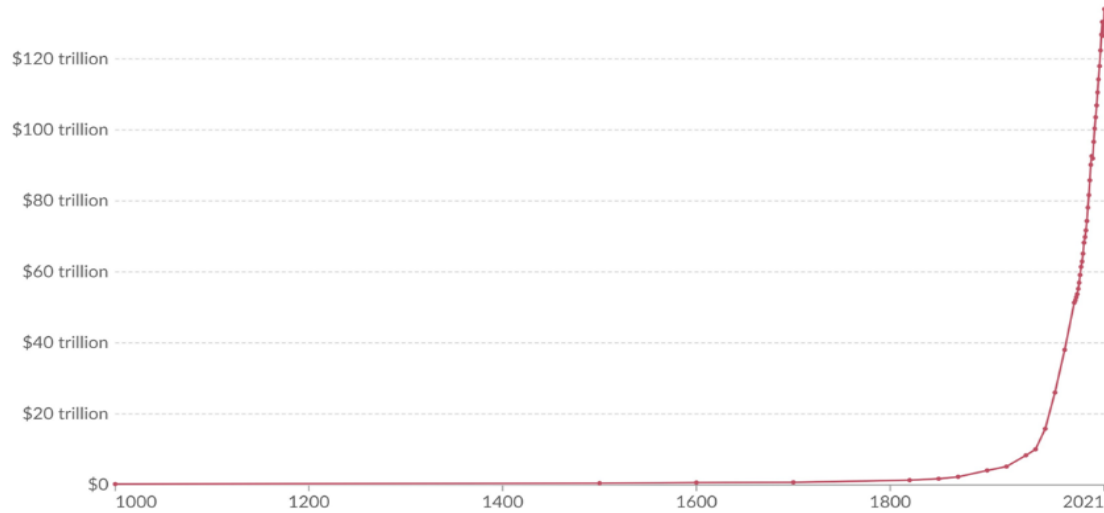
The above chart shows how the industrial revolutions have helped the economy grow. Abject poverty has declined over time, even if inequality may have increased. See the charts below.

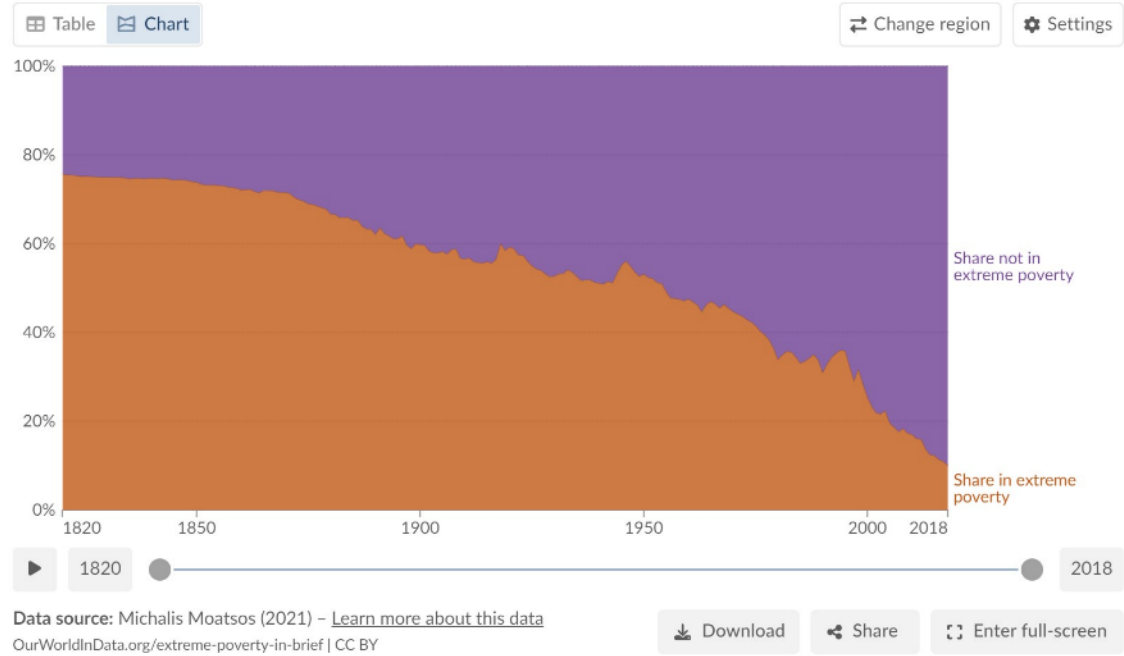**Figure 2.3** *Global income inequality: Gini index, 1820-2020*
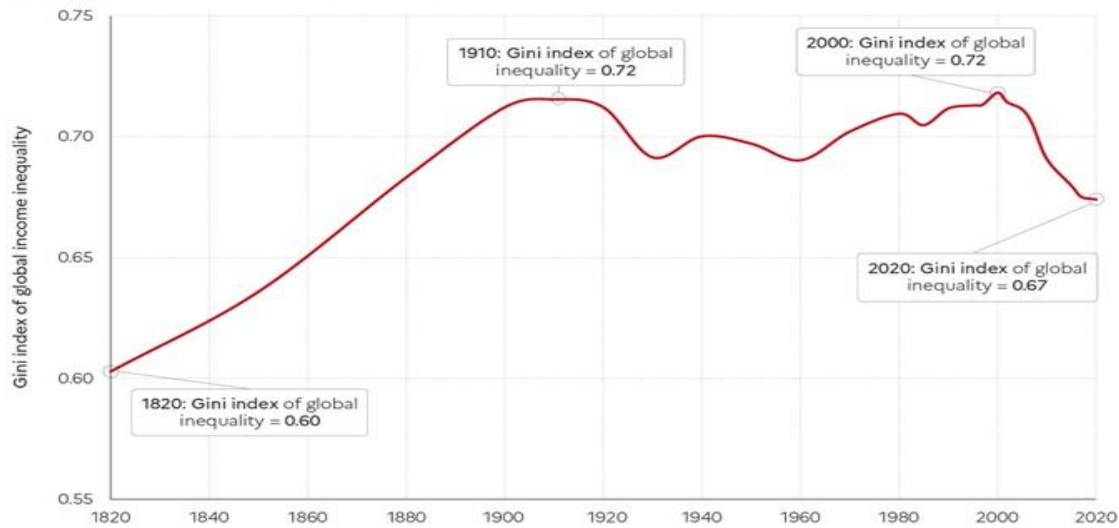


**Interpretation:** *Global inequality, as measured by the global Gini coefficient, rose from about 0.6 in 1820 to about 0.7 in 1910, and then stabilized around 0.7 between 1910 and 2020. It is too early to say whether the decline in the global Gini coefficient observed since 2000 will continue. Income is measured per capita after pension and unemployment insurance transfers and before income and wealth taxes.* **Sources and series:** *wir2022.wid.world/methodology and Chancel and Piketty (2021).*

TAKSHASHILA
INSTITUTION

### Share of population living in extreme poverty, World

This data follows a "cost of basic needs" approach: it represents the share of the population unable to meet basic needs (including minimal nutrition and adequately heated shelter) according to prices of locally available goods and services.

**Data source:** Michalis Moatsos (2021) – Learn more about this data
OurWorldInData.org/extreme-poverty-in-brief | CC BY

TAKSHASHILA
INSTITUTION

To sum up,

- In each industrial revolution, there were Luddites who resisted modern technology and protested that jobs were being threatened.
- However, the economy grew and there were more net new jobs.
- What is different about AI and the fourth industrial revolution? There is no reason to believe that even this AI revolution will not create new jobs, even if we do not know what those jobs will look like in the future.

## AI and the Fourth Industrial Revolution

Here are some recent milestones in AI.

TAKSHASHILA
INSTITUTION
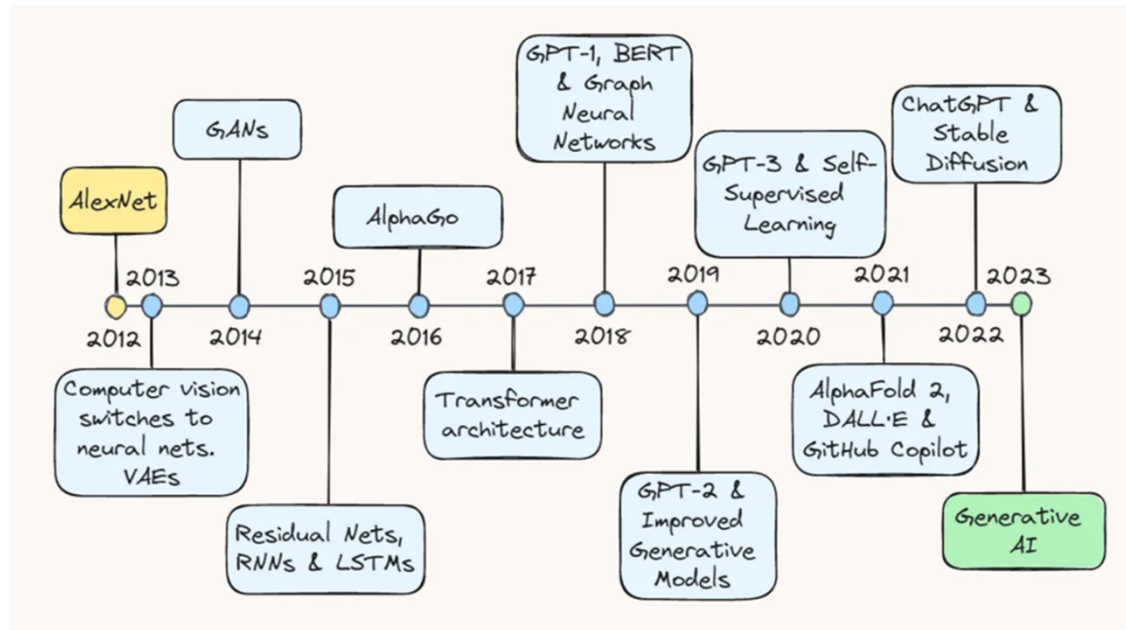
Source: (Dorfer 2023)

In the last decade and more, we have studied, with great depth, how the human brain works and developed artificial neural networks. AlexNet is a convolutional neural network capable of classifying images at near-human levels. With Variational Autoencoders,

programmers could generate new content and create images that were near ones in the real world. GANs are Generative Adversarial Networks where machine learning models create two neural networks that compete with each other and help increase the accuracy of the new content created.

Deep learning is an interesting space. When IBM created Deep Blue, it was an expert system that learned chess by observing experts. Deep learning models such as AlphaGo create their data and learn how to play the game based on just an input of the rules of the game. It produced moves no human to date had played. With ChatGPT and generative AI, near human-like interactions between humans and machines have become possible. These models are trained on hundreds of billions of parameters and vast amounts of data to be able to answer questions that not even the most well-informed and intelligent human can. These developments are the source of the fear. Will there be anything left for humans to do?

TAKSHASHILA
INSTITUTION

In a 2023 article titled, "The age of AI has begun",, Microsoft co-founder Bill Gates said AI can accomplish a lot and feels there will be a host of companies working on AI as well as ways to improve the technology itself. All these could lead to new jobs that do not exist today, but not much is said about jobs that could be lost.

The significant difference between this and the earlier industrial revolution is that this time around, many highly paid and highly skilled jobs could be lost to AI. In a conference paper titled "Artificial Intelligence and High-Skilled Work: Evidence from Analysts", Jillian Grennan and Roni Michaely (Grennan-Michaely 2021) provide evidence to show how AI could significantly impact equity analysts.

## WEF – The Future of Jobs Report

The World Economic Forum has been publishing a bi-annual report on the Future of Jobs since 2016. It has been tracking the labour market impact of the Fourth Industrial Revolution, the scale of the

TAKSHASHILA
INSTITUTION

occupational disruption and growth and recommendations for job transitions from declining to emerging economies. (WEF 2023)

The maturing of AI, along with the economic and geopolitical disruptions and the rising pressures on the environment, are expected to have a significant impact. The fourth edition of 'Future of Jobs Report' therefore broadens its scope beyond technological change also to consider and address the labour-market impact of a multitude of concurrent trends, including the green and energy transitions, macroeconomic factors, and geoeconomic and supply-chain shifts. Having said that, we shall retain our focus on artificial intelligence and its labour market impact.

We would like to discuss the findings of this report and review recommendations to policymakers so that these transitions lead to better jobs and opportunities.

Key Findings (WEF 2023)

TAKSHASHILA
INSTITUTION

1. Economic, health and geopolitical trends have created divergent outcomes for labour markets globally in 2023. Tight labour market conditions in high-income countries and higher unemployment in low to middle-income countries compared to pre-COVID times.
2. Technology adoption will remain a key driver of business transformation in the next five years.
3. The largest job creation effects come from environmental, technology and economic trends.
4. The largest job destruction effects also come from environmental, technological and economic trends.
5. Within technology adoption, big data, cloud computing and AI feature highly on the likelihood of adoption
6. The impact of most technologies on jobs is expected to be a net positive over the next five years. This is the good news
7. Employers anticipate a structural labour market churn of 23% of jobs (jobs requiring reskilling or upskilling)  in the next five years.

8. The human-machine frontier has shifted, with businesses introducing automation into their operations at a slower pace than previously anticipated.
   a. Thirty-four per cent by machines and 66 per cent by humans (only a 1 per cent increase since 2020)
   b. The 2020 report predicted that 47 per cent of tasks would be done by machines by 2025.
   c. Moderated now to expect 42 per cent by machines by 2027.

**The following jobs will see a notable change:**

TAKSHASHILA
INSTITUTION

| Jobs that will see growth | Jobs that will see a decline |
|---|---|
| AI and Machine Learning (ML) specialists | Bank tellers |
| Sustainability specialists | Postal clerks |
| Sustainability specialists | Data entry clerks |
| Business intelligence analysts | Cashiers |
| Information security analysts | Ticket clerks |
| Renewable energy engineers | |
| Solar energy installation | |
| System engineers | |

The Indian IT industry may be in a position to acquire these skills and take advantage of the developments in AI. India also needs to grow and be mindful of its carbon footprint. This could create conditions for increased demand and competitiveness in sustainability and renewable energy.

TAKSHASHILA
INSTITUTION

**The most important skills in this environment are**

1. Analytical thinking
2. Creative thinking
3. Curiosity and lifelong learning

Acquiring these analytical and creative skills is a start but those who exhibit curiosity and a willingness and ability to seek knowledge and skills that are relevant can hope for a lifetime of economic success.

The report seems to indicate that the adoption of AI may significantly lag the capabilities of AI. They refer to the fact that between 2020 and 2023, there has been insignificant growth in the use of machines in production. If 33 per cent of the work was done by machines in 2020, only 34 per cent of the work will be done by machines in 2023. They seem to use this information to suggest that people can upskill themselves in time to deal with future AI-caused disruptions.

TAKSHASHILA
INSTITUTION

They are overly optimistic about reskilling and upskilling efforts. The lack of a systematic and market-driven approach to upskilling may be the biggest hurdle to dealing with job losses caused by AI.



FIGURE 3.4 | **Largest job growth, millions**
Top roles ordered by largest net job growth, calculated based on ILO Occupation Employment statistics and growth reported by organizations surveyed
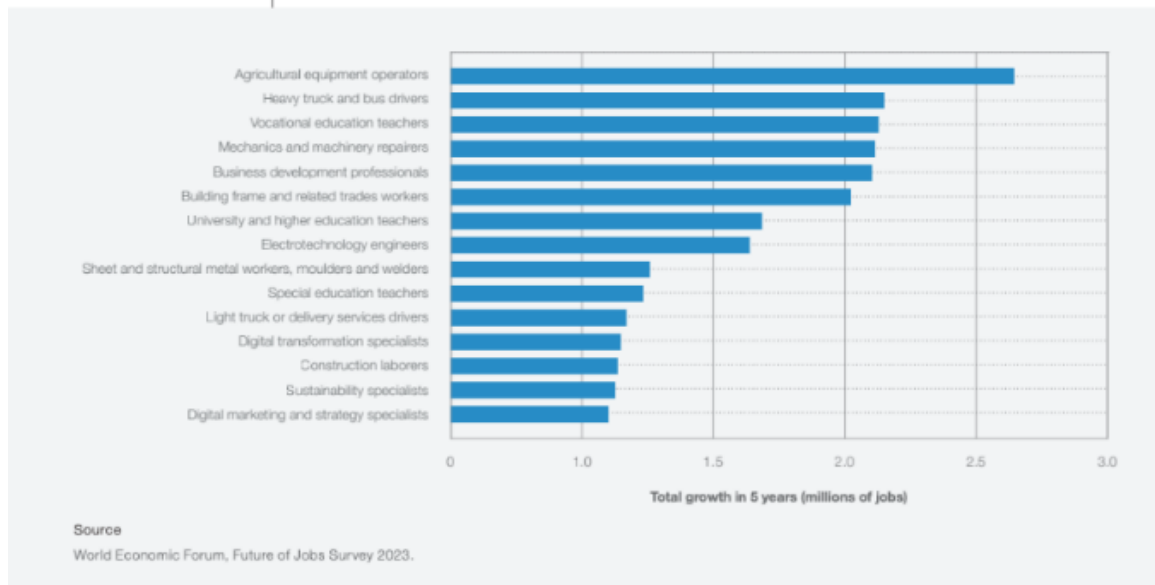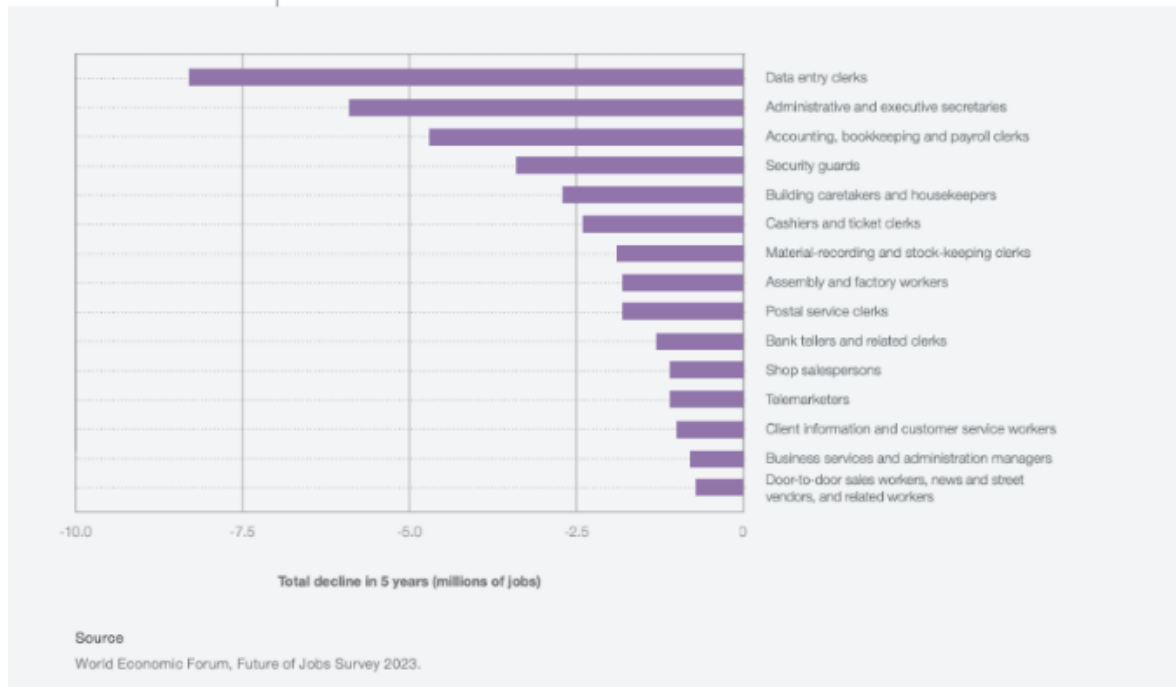
Source
World Economic Forum, Future of Jobs Survey 2023.

FIGURE 3.5 | **Largest job decline, millions**

Top roles ordered by the largest net jobs reduction, calculated based on ILO Occupation Employment statistics and growth reported by organizations surveyed



| | | | | |
|---|---|---|---|---|
| Data entry clerks | | | | |
| Administrative and executive secretaries | | | | |
| Accounting, bookkeeping and payroll clerks | | | | |
| Security guards | | | | |
| Building caretakers and housekeepers | | | | |
| Cashiers and ticket clerks | | | | |
| Material-recording and stock-keeping clerks | | | | |
| Assembly and factory workers | | | | |
| Postal service clerks | | | | |
| Bank tellers and related clerks | | | | |
| Shop salespersons | | | | |
| Telemarketers | | | | |
| Client information and customer service workers | | | | |
| Business services and administration managers | | | | |
| Door-to-door sales workers, news and street vendors, and related workers | | | | |

-10.0      -7.5      -5.0      -2.5      0

Total decline in 5 years (millions of jobs)

Source

World Economic Forum, Future of Jobs Survey 2023.

## IMF on AI and jobs

The International Monetary Fund (IMF) published a report in January 2024 titled 'Gen-AI: Artificial Intelligence and the Future of Work" (Mauro Cazzaniga 2024).

The key findings of this report are:

- AI will impact 40 per cent of all jobs
- Developed economies will see a greater impact affecting 60-70 per cent of jobs, both positive and negative
- Unlike earlier revolutions, high-paying and high-skilled jobs may also see an impact
- The less-developed agrarian economies will see a lower negative impact, but will also not get access to the benefits.
- AI challenges the belief that technology affects mainly middle and, in some cases, low-skill jobs. Its advanced algorithms can

TAKSHASHILA
INSTITUTION

now augment or replace high-skill roles previously thought immune to automation.

The IMF puts jobs into three buckets.

a) Low exposure: Gen AI does not impact these jobs
b) High exposure, high complementarity: Gen AI impacts these jobs, but they require human supervision
c) High exposure, low complementarity: AI can do this work without human supervision.

TAKSHASHILA
INSTITUTION

|  | High Exposure | |
| --- | --- | --- |
| | • AI can complete the task unsupervised<br>• Reduces need for human labour<br>• Lowers wages<br><br>Example: telemarketing employees | • AI helps improve human productivity<br>• Increases wages for incumbent workers<br>• Workers without AI related skills risk losing employment<br><br>Example: Surgeons, lawyers, judges |
| | • No change<br><br><br>Example: dishwashers, performers | • No material change |

Low Exposure ------------High Exposure

Low Complimetarity ------------------------------------------ High Complimentarity

The report acknowledges that societal acceptance of AI may determine the extent to which it is adopted in some professions due to cultural, ethical, or operational concerns.

The report also suggests this for workers to remain relevant and employed.

**TAKSHASHILA**
INSTITUTION

I.   Move from low complementarity roles in high-exposure industries to high complementarity roles. A case in point could be a telemarketing executive learning to be a prompt engineer. Easier for someone young and college-educated than for others.

II.  Move from high-exposure industries to low-exposure ones. This could also mean a lower wage.

This report seems to suggest that many industries will see low exposure to AI including areas like agriculture. There are already some farms where everything from sowing seeds to harvesting is done by AI-augmented robots. The impact of AI will be felt across every industry over time. The sequence in which they impact will depend on how easily AI can help and also whether the cost of implementing AI will be commercially viable.

The other important thing to keep in mind is that as automation lowers cost, demand will increase in many industries. The increased demand will reduce the adverse impact on jobs.

The findings of this report are useful if we want to get some predictions for the next few years but not if we are trying to prepare for the future. Preparing for the future will mean trying to stay ahead of the AI wave, do today what AI cannot do and move to something else when AI seems to be making a breakthrough.

TAKSHASHILA
INSTITUTION

## Conclusion

Autor Levy Murnane (ALM) hypothesis suggests that routine processes are most vulnerable to digital substitution. However, it is obvious to anyone following technology that even non-routine cognitive and manual tasks can be substituted by AI. (Khatwani-Raghuram-Mishra 2023)

The progress of science will ensure that eventually, all jobs that are done today can be done by AI. The issue is not "what jobs" but when will all jobs be automated. Society could slow things down because we may not be comfortable with some jobs being done by AI. Deep Blue showed that AI could defeat world champions in chess over two decades ago, but even today, we only watch humans playing chess, and its popularity is only growing.

TAKSHASHILA
INSTITUTION

The larger question thereafter could be whether humanity will survive the lack of jobs, even assuming some form of social safety net will exist when the economy sees hypergrowth fuelled by AI.

Do people need work to find meaning in their lives? The answer, according to Susskind (Susskind 2021), lies in looking at the lives of the very affluent and the leisure classes. Bertrand Russell in his essay "In Praise of Idleness" (Russell 1932) writes, "Without the leisure class, mankind would never have emerged from barbarism". The point is that leisure is necessary to cultivate the arts, discover the sciences, author books, and invent philosophies.

Teaching children how to use leisure well may be as important in the future as knowing how to learn. It is likely that humankind will find many new things to do. Sixty per cent of the jobs that exist today did not exist in 1940 (Autor 2022). Dell Technologies, in a report co-authored with the Institute for the Future, suggests that 85 per cent of the jobs that will exist in 2030 have not been invented yet. (IFTF n.d.)

TAKSHASHILA
INSTITUTION

The future is unknown, but we need to keep track of what is happening around us and respond through new learning and move forward to a place where we can be gainfully employed and live a meaningful life.
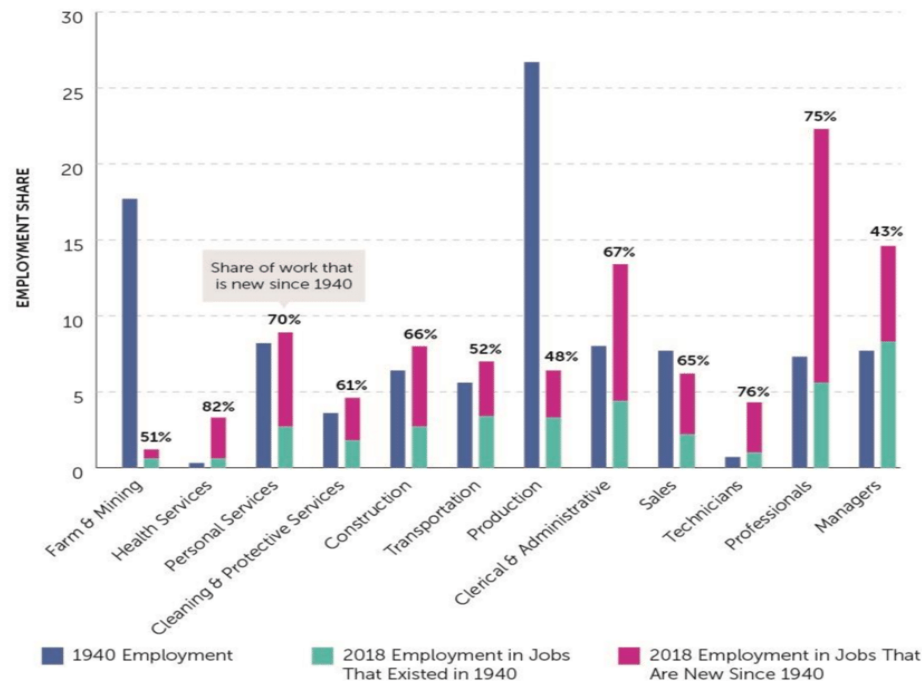


Source: (Autor 2022)

# References

Autor, David. 2022. "The Labor Market Impacts of Technological Change: From Unbridled Enthusiasm to Qualified Optimism to Vast Uncertainty." *SSRN Electronic Journal.*

Dorfer, Thomas A. 2023. *Ten years of AI in review.* 23 May. https://towardsdatascience.com/ten-years-of-ai-in-review-85decdb2a540.

Gates, Bill. 2023. *The Age of AI has begun.* https://www.gatesnotes.com/The-Age-of-AI-Has-Begun.

George Gray Molina, María Montoya-Aguirre, and Eduardo Ortiz-Juarez. 2022. *ADDRESSING THE COST-OF-LIVING crisis in developing countries.* UNDP.

TAKSHASHILA
INSTITUTION

Grennan-Michaely. 2021. *Artificial Intelligence and High-Skilled Work: evidence from analysts.* https://www.jbs.cam.ac.uk/wp-content/uploads/2021/05/ccaf-2021-conference-paper-grennan-michaely.pdf.

IFTF. n.d. *Realizing 2030 - a divided vision of the future.* Dell Technologies.

Keynes, JM. 1930. *Economic possibilities for our grandchildren.* http://www.econ.yale.edu/smith/econ116a/keynes1.pdf.

Khatwani-Raghuram-Mishra. 2023. "Automation, technological employment, and society - a review." *AIP Publishing.*

Mauro Cazzaniga, Florence Jaumotte, Longji Li, Giovanni Melina, Augustus J. Panton, Carlo Pizzinelli, Emma Rockall, and Marina M. Tavares. 2024. *Gen-AI: Artificial Intelligence and the future of work.* IMF.

TAKSHASHILA
INSTITUTION

Russell, Betrand. 1932. "In praise of idleness." *Harpers*, October.

Susskind, Daniel. 2021. *A world without work.* Penguin.

WEF. 2023. *The Future of Jobs Report.* World Economic Forum. https://www.weforum.org/publications/the-future-of-jobs-report-2023/.

TAKSHASHILA
INSTITUTION

# Global Governance Framework for Artificial Intelligence

## Bharat Sharma & Saurabh Todi

### What is Artificial Intelligence (AI)?

Artificial Intelligence (AI) is an emerging technology that empowers machines to simulate human intelligence, enabling them to perform tasks that typically require human intelligence, such as learning, reasoning, problem-solving, perception, and language understanding. AI systems analyse and interpret data, adapt to changing environments, and make decisions based on their understanding. AI technologies are being deployed in various fields, including healthcare, finance, education, autonomous vehicles, and more. AI can be broadly

[categorised](#) into two main types: Narrow AI (or Weak AI) and General AI (or Strong AI).

- Narrow AI (Weak AI) is designed to perform a specific or narrow set of tasks. It excels in well-defined and focused activities but lacks the ability to generalise its intelligence to other domains. This includes virtual personal assistants, facial recognition software or speech-to-text applications.
- General AI (Strong AI) refers to machines able to understand, learn, and apply knowledge across a broad range of tasks, similar to human intelligence. As of now, true General AI remains a theoretical concept and has not been realised. However, significant research is ongoing in this domain.

Several AI tools have been developed to aid productivity in various domains, while research and development are ongoing to improve these capabilities further. Some of the domains where AI tools have been employed include

- Healthcare: disease diagnosis, personalised treatment plans, drug discovery, and medical image analysis, enhancing efficiency and accuracy in healthcare services.
- Finance: fraud detection, credit scoring, algorithmic trading, and financial analysis to improve decision-making and risk management.
- Education: personalised learning experiences, adaptive assessments, and intelligent tutoring systems to cater to individual student needs.
- Autonomous Vehicles: AI plays a crucial role in self-driving cars, enabling them to perceive their environment, make real-time decisions, and navigate safely.
- Customer Service: Virtual chatbots and AI-driven customer support systems
- Manufacturing and Robotics: process optimisation, predictive maintenance, quality control, and the operation of robotic systems in assembly lines.

**TAKSHASHILA**
**INSTITUTION**

- Natural Language Processing: AI tools power language translation, sentiment analysis, and chatbots, facilitating communication between humans and machines.
- Cybersecurity: AI can help identify and mitigate cyber threats by analysing patterns, detecting anomalies, and enhancing security measures.
- Agriculture: crop monitoring, pest control, precision farming, and yield prediction, contributing to increased efficiency and sustainability.

As AI continues to advance, it holds the potential to revolutionise industries, streamline processes, and contribute to solving complex problems. However, ethical considerations and responsible development are crucial to ensuring AI's beneficial and safe integration into society.

TAKSHASHILA
INSTITUTION

## Need for Global Governance of AI

Given the depth and breadth of applications powered by AI tools, the economic, political and cultural impact of AI on human society is going to be immense. In some ways, the impact that AI is expected to have on human society and industrial productivity will be similar to what the world experienced due to the widespread adoption of Information and communications technology, specifically the internet. Therefore, evaluating the existing global governance structures for various technologies for potential models for regulating artificial intelligence (AI) technologies and applications is important.

AI governance frameworks will have a critical role in setting the stage for international collaboration within AI technologies and adopting a global approach to shaping AI norms, regulations, and institutional frameworks. A truly "global" AI governance system is [critical to managing risks](#) that AI presents, including disinformation,

TAKSHASHILA
INSTITUTION

discrimination and bias in algorithmic systems, and [societal polarisation](#).

## Current Global Governance Landscape

The history of institutional governance concerning key technologies — for instance, nuclear energy, the internet, and particular climate change technologies — might offer models for regulating AI. These aspects include creating a legal-institutional framework, setting up a standards-setting body, representing all countries, establishing norms around technology governance, and representing industry, academia, civil society, and technical experts.

For prevailing models of global governance, we looked at several technologies that were or are considered critical to the global system, especially during the period when their governing institutions were first established. For the purpose of this analysis, we looked at the following technologies:

1. Nuclear Energy
2. Climate Change Technologies
3. Telecommunications
4. Civil Aviation
5. Biological Weapons
6. Chemical Weapons
7. Biological Resources
8. Financial System
9. Data Privacy
10.      Internet
11.      Postal System

We looked at various aspects of such organisations:

1. Is there a global representative body for this technology?
2. Is there an underlying international treaty or organisation from which it derives its existence?

TAKSHASHILA
INSTITUTION

3. Is the organisation global or regional? If the former, is it part of the UN system?
4. Does it have a separate technical body to support state parties in their decision-making?
5. Are the recommendations of the body binding on state parties?
6. What is the level of representation of civil society, industry and academia during discussions in these bodies?

Examining these aspects offers insights into how similar processes may be utilised for developing global governance and regulatory frameworks for AI.

TAKSHASHILA
INSTITUTION

| Technology | Organisation | Underlying Authority/Treaty/Agreement | Type | Part of the UN System | Separate Technical Body | Recommendations | Industry (I) | Academia (A) | Civil Society (C) |
|---|---|---|---|---|---|---|---|---|---|
| Nuclear Energy | International Atomic Energy Agency (IAEA) | International Atomic Energy Agency (IAEA) | Multilateral | Yes | Yes | Binding | None | None | None |
| Climate Change Technologies | United Nations Framework Convention on Climate Change | The United Nations Framework Convention on Climate Change | Multilateral | Yes | Yes | Optional | Consultative | Consultative | Consultative |
| Telecommunications | International Telecommunications Union | International Telecommunications Union | Multilateral | Yes | Yes | Optional | Consultative | Consultative | Consultative |
| Civil Aviation | International Civil Aviation Organization (ICAO) | Convention on International Civil Aviation | Multilateral | Yes | Yes | Optional | Consultative | Consultative | Consultative |
| Biological Weapons | Biological Weapons Convention | Biological Weapons Convention | Multilateral | No | No | Binding | None | None | None |
| Chemical Weapons | Organisation for the Prohibition of Chemical Weapons | Chemical Weapons Convention | Multilateral | No | Yes | Binding | None | None | None |
| Biological Resources | Convention on Biological Diversity | Conference of the Parties (COP) | Multilateral | Yes | Yes | Binding | Consultative | Consultative | Consultative |
| Financial System | Bank for International Settlements | | Multilateral | No | Yes | Optional | None | None | None |
| Data Privacy | General Data Protection Regulation | Charter of Fundamental Rights of the European Union | Regional | No | Yes | Binding | Consultative | Consultative | Consultative |
| Internet | Internet Corporation for Assigned Names and Numbers (ICANN) | | Multilateral | No | Yes | Binding | Consultative | Consultative | Consultative |
| Internet | Internet Assigned Numbers Authority (IANA) | | Multilateral | No | Yes | Optional | None | None | None |
| Postal System | Universal Postal Union | United Nations Economic and Social Council | Multilateral | Yes | Yes | Optional | None | None | None |

TAKSHASHILA
INSTITUTION

## Notable Global Governance Bodies

Among the technologies examined above, focusing on two specifically will be instructive. First is the International Atomic Energy Agency (IAEA), which regulates nuclear technology, especially nuclear power, and second is the United Nations Framework Convention on Climate Change (UNFCCC). This global treaty is the foundation for international cooperation on climate action, for which the Conference of Parties (CoP) serves as the supreme decision-making body.

## The International Atomic Energy Agency

The year 1951 saw the first instance of electricity generated through nuclear fission by the EBR-1 reactor in Idaho, United States. The Soviet Union followed suit in 1954. The International Atomic Energy Agency (IAEA) was established in response to rapid development in the nuclear energy sector and growing international concern towards nuclear weapons, especially amid rising tensions between the

foremost nuclear powers. Its history can be traced back to U.S. President Dwight D. Eisenhower's "Atoms for Peace" address to the United Nations General Assembly on December 8, 1953.

The IAEA is an independent international organisation and operates within the United Nations (UN) framework. It was established in 1957 and is headquartered in Vienna, Austria. The IAEA's primary objective is to promote the peaceful use of nuclear energy while preventing the spread of nuclear weapons and facilitating international cooperation in the field of atomic energy. It also fulfils the Treaty on the Non-Proliferation of Nuclear Weapons (NPT) mandate. The structure of the IAEA is as follows:

1. The General Conference is the highest decision-making body of the IAEA and consists of representatives from all member states. It meets annually and makes decisions on matters related to the agency's work. Each member state has one vote in the General Conference.

2. The Board of Governors oversees the IAEA's activities between General Conference sessions. It comprises 35 member states elected by the General Conference, representing different regions. The Board meets several times a year to provide policy guidance, approve the budget, and decide on various issues.
3. The Secretariat is headed by the Director General, appointed by the Board of Governors and confirmed by the General Conference. The Director General is responsible for the day-to-day operations of the IAEA. The Secretariat comprises various departments and divisions, each responsible for specific areas of the agency's work, such as nuclear safety, security, technical cooperation, and safeguards.

The IAEA carries out several functions, such as implementing safeguards to prevent the spread of nuclear weapons, providing technical assistance and expertise to its member states to promote the peaceful use of nuclear energy, establishing international standards and guidelines for nuclear safety and providing expert advice to member states to enhance the safety of nuclear installations and

activities. It also strengthens global nuclear security by assisting member states in implementing measures to protect nuclear and radioactive materials from theft, sabotage, or unauthorised access. In addition, it supports the peaceful applications of nuclear technology in areas such as agriculture, medicine, industry, and water resource management. To effectively carry out many of these functions, the IAEA has been [significantly supported](#) by academia and civil society, especially with regard to promoting nuclear education and safety.

Over the years, the IAEA has played a crucial role in advancing international cooperation in the peaceful use of nuclear energy and in promoting nuclear safety and security. However, the current structure has associated weaknesses, too, which hamper its effective functioning.

The foremost challenge is the lack of enforcement authority within the IAEA. The IAEA's inspections and safeguards system relies heavily on cooperation from member states. It cannot directly enforce compliance and has no power to punish violations of safeguards

agreements. Furthermore, there have been instances of states influencing or pressuring the IAEA to downplay or turn a blind eye to violations. The IAEA has also struggled with budget limitations, impacting its ability to update safeguards technology, retain expert staff, and conduct sufficient inspections worldwide.

## United Nations Framework Convention on Climate Change

While questions of ethics and regulation beset global AI governance, questions regarding the inequitable deployment of AI technologies are equally important. Research by the International Monetary Fund highlights that AI risks widening the gap between developed and developing countries, creating a "great divergence". AI technologies possess the potential to radically meet challenges in agriculture, education, and healthcare domains, particularly in those countries with low state capacity. In this vein, these technologies have been linked to values like human flourishing. AI governance models should

TAKSHASHILA
INSTITUTION

emphasise a sustainable adoption of AI technologies, particularly noting that these run the risk of increasing misinformation, governance models, and building surveillance states.

In this vein, how the UNFCCC regulates the development and transfer of climate change technologies is fruitful to look at. Governance of these technologies under the treaty may provide insights into how sensitive technologies like AI may be governed. The Intergovernmental Panel on Climate Change (IPCC) defines climate technologies as "any piece of equipment, technique, practical knowledge or skills for performing a particular activity that can be used to face climate change".

The technology provisions of the UNFCCC flow from the Paris Agreement's climate commitments for developed countries and, crucially, their obligations towards developing countries. Article 10 of the Paris Agreement requires that "[p]arties share a long-term vision on the importance of fully realising technology development and transfer to improve resilience to climate change and reduce

greenhouse gas emissions". Its conceptual precedent is the 1992 Convention on Climate Change, which placed responsibility on developed countries to lead developing countries in climate-change technological development.

In this regard, the UNFCCC established the 'Technology Mechanism' ("Mechanism") to facilitate the implementation of actions for technology development and transfer to support mitigation and adaptation activities in developing countries. The UNFCCC defines 'technology transfer' as "a broad set of processes covering the flows of know-how, experience and equipment for mitigating and adapting to climate change amongst different stakeholders such as governments, private sector entities, financial institutions, non-governmental organisations (NGOs) and research/education institutions".

A more practical definition of technology "development and transfer" can be conceptualised as cooperation between developed and developing countries on technology use through direct support of

TAKSHASHILA
INSTITUTION

initiatives and projects that allow the diffusion of a specific technology. Other definitions include supporting a more enabling regulatory environment in developing countries.

From an AI governance perspective, the Mechanism's institutional structure is interesting. It works through its policy and implementation arms, the Technology Executive Committee (TEC) and the Climate Technology Centre and Network (CTCN), respectively.
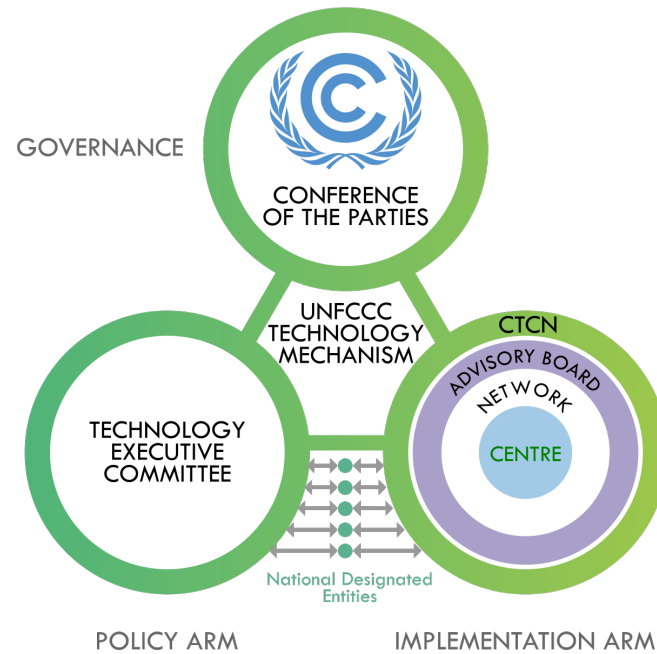
TAKSHASHILA
INSTITUTION

Figure 1: The UNFCCC Technology Mechanism
*(Source: The UNFCCC)*

The TEC works on providing policy recommendations to the CoP, focused on supporting countries' efforts to enhance technology

development and transfer, in particular focusing on "increasing and improving the transfer of, and access to, environmentally sound technologies (ESTs) and know-how". The TEC works closely with the CTCN to address technology development and transfer issues and is answerable directly to the CoP. It comprises 22 technical experts, with representatives from both developing and developing countries,50 and includes environmental NGOs and industry organisations.

---

50 For instance, three members to be included from Africa, Asia and the Pacific, and Latin America and the Caribbean; one member from a small island developing State; and one member from at least developed country Party.
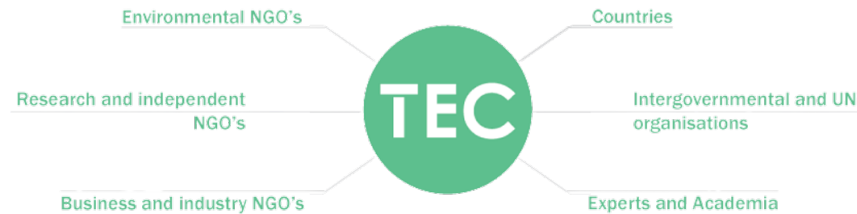
TAKSHASHILA
INSTITUTION

Figure 2: Composition of the Technology Executive Committee
*(Source: The UNFCCC)*

The CTCN — the implementation arm of the Technology Mechanism — works on helping implement and enhance the development and transfer of technologies to developing countries through access to information and knowledge. The CTCN consists of the Climate Technology Centre, hosted by the United Nations Environment Programme, and the Climate Technology Network. It follows a demand-driven mechanism: the Centre works to manage and respond to the needs of developing countries to accelerate ESTs and is the coordinating entity in the CTCN. It utilises the Climate Technology

Network of organisations that comprise academia, the private sector, and public and research institutions. The CTCN is answerable to an advisory board accountable to the CoP.

The UNFCCC serves as a critical platform for international collaboration, providing a framework for ongoing discussions, negotiations, and actions to address climate change's complex and interconnected challenges. The annual COP meetings are vital for countries to unite, share information, and make decisions shaping the global climate change response. Academia, civil society, and industry representation in the CTCN are important lessons for AI governance, as is the provision for a Technology Mechanism that addresses an inequitable diffusion of "green" technologies.

## Lessons for Governance of Artificial Intelligence

Given AI's vast applications, several countries have attempted to regulate AI based on their understanding of its risks and the opportunities it provides. The regulatory landscape for artificial

intelligence (AI) is continuously evolving. Various countries and regions have taken steps to establish guidelines, frameworks, and regulations to address the ethical, privacy, and safety concerns associated with AI technologies. For example, the Federal Trade Commission (FTC) and the National Institute of Standards and Technology (NIST) of the US have explored AI-related regulations, and the UK has established the Centre for Data Ethics and Innovation to provide guidance on the ethical use of AI and data-driven technologies. Singapore has developed a Model AI Governance Framework to guide organisations in deploying AI responsibly, and so has China. Although there's no specific law regulating AI in India, there have been a few initiatives and guidelines for the responsible development and deployment of AI technologies, such as the National Strategy for Artificial Intelligence, Principles for Responsible AI, and Operationalizing Principles for Responsible AI.

At the multilateral level, the Organisation for Economic Co-operation and Development (OECD) has developed the OECD AI Principles, which are a set of principles emphasising the responsible use of AI, including

transparency, accountability, and inclusiveness. Similar efforts [have been made](#) by the United Nations and the World Economic Forum. The EU has gone further by proposing an Artificial Intelligence Act, which, when passed, will be the world's first AI legislation.

The governance structure of nuclear technologies through the IAEA is not well suited for AI technology due to fundamental differences in the nature of these technologies. Nuclear technology is a highly expensive, specialised and clearly differentiated technology with its use cases (weapons, energy, and medical/food tech). Its users are also very limited, with violators relatively easier to detect. However, AI is a much more expansive domain with several players and many more applicable domains. The dual-use nature of AI technologies makes them harder to monitor than nuclear technologies. This makes a stringent, detailed, and very clearly defined regulatory framework at a global level an unlikely option for AI. However, elements from the UNFCCC framework may be more suitable for AI regulation, at least from the perspective of where the technology and its application currently stand.

TAKSHASHILA
INSTITUTION

Broad principles for such an attempt should be:

1. As with several international organisations, a foundational ethical framework would be necessary to establish a body tasked with defining the regulatory scope for AI. UNESCO's Recommendation on the Ethics of Artificial Intelligence could be a starting point for such an endeavour.
2. Given the evolving nature of the technology and the significant economic and security consequences of AI applications, any such principles derived from the ethical framework must be broad enough to leave room for individual interpretation by state parties.
3. A Technical Committee should be set up under the framework to help recommend broad directives based on evolving AI technology and its applications. It should have representation from civil society, academia, industry, and technical experts voted on by state parties while maintaining regional representation.

TAKSHASHILA
INSTITUTION

4. The conference of parties involving all signatories to the framework would meet periodically to review and update it as per the technical committee's recommendation.

## Conclusion

Given the rapidly evolving nature of AI, a rigid global governance framework would be counterproductive and risk not having a broad based acceptance due to the prospects of technological, economic and military benefits involved. At this rate of AI technology development, a nimble and recommendatory approach to global governance should be preferred, with the possibility of regulations becoming more direct, targeted, and binding as the technology matures. In due course of time, the international community would have to tackle complex issues such as the dual use of AI and the use of AI in warfare. As these issues are sensitive, building trust between countries first is better.

TAKSHASHILA
INSTITUTION

# Augmenting State Capacity with AI: A Working Paper

## Bharath Reddy

## Executive Summary

Despite having capable elite institutions, India's local governments often struggle to effectively implement policies and provide essential services due to understaffing, lack of resources, social cleavages, and the challenges of a precocious democracy. Artificial intelligence (AI) has the potential to augment state capacity by improving governance and public service delivery in India.

This paper proposes a framework for identifying areas where AI can enhance state capacity. By breaking down government processes into

tasks and analysing them based on transaction volume and discretion, the framework pinpoints tasks that would benefit most from AI adoption. In addition, those that have higher stakes for individuals carry higher risks and need robust checks and balances.

However, successful AI adoption requires essential prerequisites, including representative datasets, appropriate infrastructure, strict adherence to data protection principles, and effective risk management frameworks. In conclusion, this paper also highlights the need for further research on lessons learned from previous e-governance applications, identifying areas where AI should not be integrated, and mitigating the effects of the digital divide.

## Introduction

Artificial intelligence (AI) typically refers to machines that can perform tasks commonly associated with human intelligence. The understanding of the term in the popular imagination is heavily

TAKSHASHILA
INSTITUTION

influenced by science fiction — recursive self-improving systems that surpass human intelligence or comprehension. However, the reality of AI technology today is quite different from these portrayals.

Currently, AI technologies excel at specific, narrow tasks rather than broad, general intelligence. The term artificial intelligence seems like a misnomer when it is used to refer to such systems. While their "intelligence" might be debatable, they far exceed human abilities at performing repetitive tasks, pattern recognition, and probabilistic reasoning. Such tasks are valuable in various governmental functions, and AI can augment the state's ability to deliver more efficient and effective governance.

State capacity refers to the ability of a government to effectively implement policies, provide public goods and services, and enforce laws throughout its territory. High state capacity is often associated with better governance outcomes, such as higher economic growth, lower corruption, and more effective provision of public goods like healthcare, education, and infrastructure.

This paper examines the limitations of India's state capacity and the characteristics of complex problems for a state. It proposes a framework for identifying areas where state capacity can be augmented with AI. Finally, it identifies the necessary prerequisites and enabling requirements for the successful adoption of AI.

## The State of India's State Capacity

Lant Pritchett has described India as a "flailing state", where the head remains sound and functional while the arms and limbs are not connected to the head. In other words, the elite institutions (at the Union and sometimes state level) are capable of formulating policies, but the field-level agents of the state might actively subvert those policies by acting in their own interests.

This is also borne out in Devesh Kapur's work, who observes that the state performs poorly in providing basic public services like education, healthcare, water, and sanitation. But it has effectively managed

complex tasks like conducting massive elections, implementing large government programmes,, and maintaining macroeconomic stability. The performance is better for episodic activities with automatic exits like elections rather than ongoing regulation and service delivery. The state also fares better on goods and services where social hierarchies and norms matter less.
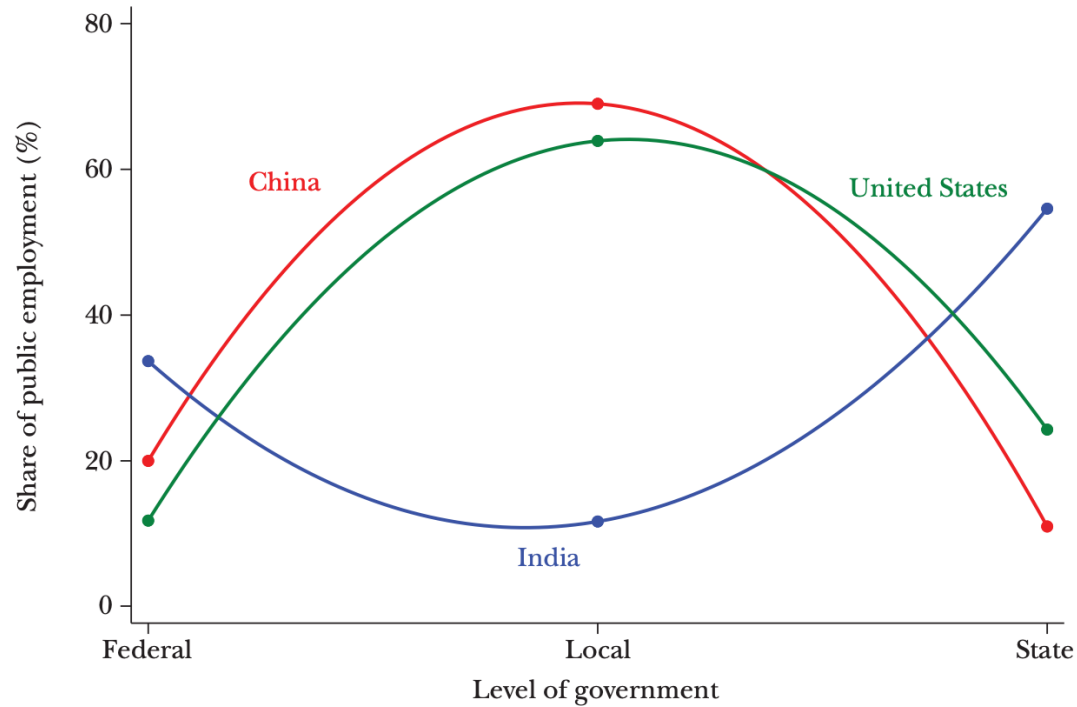
TAKSHASHILA
INSTITUTION

*Figure 1: Structure of Employment across Three Tiers of Government: China, India, and the United States. Source - Devesh Kapur, 2020.*

TAKSHASHILA
INSTITUTION

[Devesh Kapur](#) identified three reasons for the woefully poor public service delivery. Firstly, local governments suffer from severe understaffing and lack of financial resources, a situation not seen in other federal countries. This significantly limits their ability to provide services effectively. Secondly, persistent social cleavages such as caste and gender discrimination continue to affect service delivery, especially in areas where these norms have a greater influence on outcomes.

Lastly, compared to other countries, Kapur claims India is a precocious democracy given its low level of income and literacy and high levels of ethnic and religious conflict. This leads to a reduced ability to raise revenues and redistributive pressures that limit the effectiveness of public service delivery.

AI systems can potentially address some of these challenges to a significant extent. By strengthening state capabilities in service delivery at the local level, AI could help alleviate some of the shortcomings of the flailing state. Taking the metaphor forward, it

strengthens the arms and limbs of the state while maintaining control at the head. Additionally, although AI systems introduce fairness-related concerns of their own, they could accelerate progress in overcoming obstacles posed by social and cultural norms in public service delivery.

## Complex Problems for a State

Vijay Kelkar and Ajay Shah propose that states face their toughest challenges when dealing with processes characterised by a high number of transactions, the necessity for discretion, significant stakes for individuals, and a high level of secrecy.

A high number of transactions implies a greater administrative burden, necessitating more capacity to manage the workload effectively. Increased discretion often leads to ambiguity in accountability and creates opportunities for state agents to engage in rent-seeking behaviour. When individuals have much at stake, they are more likely

to invest considerable time and resources to influence outcomes in their favour. Lastly, more secrecy within processes weakens feedback mechanisms, limiting opportunities for constructive criticism.

For example, systems such as the criminal justice system, judiciary, tax administration, and financial regulation rank high on all these dimensions, presenting formidable challenges for state governance.

Information technology (IT) solutions and as an extension, AI solutions, could reduce the complexity of such challenges in areas that have a high transaction volume and low discretion, making it easier to overcome state capacity limitations and deliver better governance and public services.

## Framework for Augmenting State Capacity with AI

The framework proposed below is applicable for integrating information technology solutions more broadly, in which AI can be

seen as a sub-field. While conventional IT solutions enable data analysis and decision-making through explicitly programmed rules, AI systems self-learn rules using statistical techniques. AI systems can be prone to errors due to limitations in the algorithms or the data they are trained on. This framework provides a pathway to benefit from AI adoption while minimising the risks involved.

Autor's widely accepted approach facilitates the identification of tasks suitable for IT solutions (including AI systems). This approach involves breaking down job processes into a series of tasks. Routine and codifiable tasks are deemed more suitable for IT solutions, while non-routine tasks requiring abstract reasoning and interpersonal communication are typically more challenging for machines to replace.

Kelkar and Shah's criteria for complex problems can be applied to analyse this task breakdown approach, identifying tasks that would benefit most from AI adoption. Analysing tasks based on transaction volume and discretion helps pinpoint areas that stand to gain the most

from integrating AI. Further, those with significant stakes for individuals need more robust checks and balances, such as human oversight, continuous monitoring and evaluation, and risk mitigation strategies.

The [recent State of the Judiciary report](#) released by the Supreme Court of India has highlighted several tasks that align with this framework. These tasks include translating judgments/orders into various languages, case management, AI-supported roster management, dictating orders and judgments, natural language processing and judicial knowledge management.

In contrast, tasks requiring greater discretion have higher risks. Algorithms for predicting recidivism are often utilised to assess a criminal defendant's likelihood of reoffending. These predictions play a role in pretrial, parole, and sentencing decisions. Research indicates that such algorithms are no more accurate than predictions made by individuals lacking judicial expertise. Likewise, predictive and preventive policing systems, like the one [recently deployed in](#)

[Bengaluru](#), have [demonstrated risks](#). These include the perpetuation of existing biases, the potential for self-fulfilling prophecies arising from proactive policing in neighbourhoods labelled as "high-risk," and the erosion of due process due to excessive dependence on these imperfect systems.

## Prerequisites for AI adoption

While the previously described framework identifies tasks ripe for AI integration, they require many prerequisites to be used effectively. These are listed below.

### Data

Fairness issues exist where certain groups are underrepresented in the training data. This can lead to higher error rates and biased outcomes for minorities, women, the elderly, and vulnerable populations. Given India's incredible diversity, the government must create publicly

TAKSHASHILA
INSTITUTION

accessible open datasets representative of population diversity. Such data will be invaluable not just for government functions but also have significant positive externalities for research and commercial applications. Bhashini is one such initiative under the Indian government that attempts to capture the diversity of Indian languages and provide open-source databases and tools for real-time translation.

## Infrastructure

Implementing AI for different tasks necessitates the appropriate hardware, software, and communication networks. These resources must be accessible where AI systems are deployed, which might bring challenges in remote areas. Therefore, acquiring and maintaining these essentials requires substantial investment. The political economy of using AI to perform such tasks might also be challenging, considering that such systems will replace human tasks in some scenarios.

TAKSHASHILA
INSTITUTION

## Data Protection

[Significant exemptions are carved out](#) for the state's use of personal data for national security, public order, and crime prevention. These exemptions may lead to data collection, processing, and retention beyond what is necessary, violating the fundamental right to privacy.

Strict adherence to privacy by design, data minimisation, purpose limitation, and accountability is crucial. However, the procedures followed during the rollout of systems like [Digi Yatra highlight concerns](#). In this case, a private company registered under Section 8 of the Companies Act is gathering sensitive biometric data from airport passengers, often without their consent. Furthermore, as a private company, there are no established mechanisms for accountability or transparency.

## Risk Management

Ensuring responsible AI adoption requires frameworks that effectively manage risks linked with AI systems. Embracing AI risk management frameworks, like the one crafted by the US's National Institute of Standards and Technology (NIST), offers a systematic approach to monitoring and managing risks throughout the AI system's lifecycle. The framework is mandatory for US Federal agencies and widely embraced in the private sector.

## Future work

This working paper proposes a framework for identifying areas where AI systems can augment state capacity. Future work includes analysing what has been learned from previous adoption of e-governance applications that can prove helpful while adopting AI. There is also a need to identify a negative list of areas where AI should not be integrated, irrespective of the gains that they might provide. These might include areas where they might undermine or erode trust in democratic institutions. Finally, the adoption of AI might exacerbate

TAKSHASHILA
INSTITUTION

the already existing digital divide. Educated individuals with access to technology might be able to take advantage of these tools whereas the historically marginalised might find it harder, making it a problem which also needs to be studied and mitigated.

TAKSHASHILA
INSTITUTION

TAKSHASHILA
INSTITUTION

The Takshashila Institution is an independent centre for research and education in public policy. It is a non-partisan, non-profit organisation that advocates the values of freedom, openness, tolerance, pluralism, and responsible citizenship. It seeks to transform India through better public policies, bridging the governance gap by developing better public servants, civil society leaders, professionals, and informed citizens.

Takshashila creates change by connecting good pe, to  Takshashila creates change by connecting good people, to good ideas and good networks. It produces independent policy research in a number of areas of governance, it grooms civic leaders through its online education programmes and engages in public discourse through its publications and digital media.