Annexure 2: Overview of Global AI governance frameworks

Jurisdiction	Summary of Approach
Australia	Ongoing deliberations on a government whitepaper titled **"Safe and Responsible AI in Australia"**, proposing mandatory guardrails to regulate AI in high-risk settings and general-purpose AI models.
Brazil	Proposals for a new AI law (Bill No. 2,338/2023) that promotes secure, reliable AI systems, categorizing them by risk and imposing various compliance requirements.
Canada	Published the **draft Artificial Intelligence and Data Act (AIDA)** that focuses on responsible AI use, consumer protection, and fair competition. The law is still at the parliamentary review stage.
China	Technology-specific regulations aimed at specific issues, including algorithmic recommendations and generative AI. Various national standards for AI systems and 'Labeling Rules' have also been introduced to enhance the security and governance of generative AI.
European Union	Statutory framework in the form of the **Artificial Intelligence Act** that categorizes systems by risk levels, imposes stringent requirements on high-risk applications, and aims for transparency and accountability.
Japan	Adopted the law on **Promotion of AI-Related Technologies** in May 2025. It establishes an AI Strategy Center and implements non-binding guidelines to promote innovation and adoption. The framework emphasizes voluntary compliance and international cooperation.
Singapore	Voluntary, use-case based approach that emphasizes a sectoral approach based on governance frameworks. It has released a **draft Model AI Governance Framework for Generative AI** to address emerging risks and provide guidance for safety evaluations. It has developed practical testing methods such as Veritas and AI Verify, which allow organisations to evaluate fairness and transparency in real use cases.
United Kingdom	Context-based and cross sectoral framework that focuses on core principles (safety, transparency, fairness, accountability, contestability) that will be implemented by existing sectoral regulators.

United States of America	A pro-innovation approach that emphasises innovation, infrastructure development and international diplomacy to promote American leadership and global competitiveness. Voluntary commitments, such as the NIST AI Risk Management Framework, and some executive orders relating to AI governance are applicable.
South Korea	Adopted the **Basic Act on the Development of Artificial Intelligence and Establishment of Trust.** The Act adopts a risk-based approach focusing on high-impact AI systems and generative AI transparency requirements, with moderate enforcement through administrative fines.
New Zealand	Developed the **Algorithm Charter for Aotearoa New Zealand** in 2020 which applies specifically to public sector algorithmic decisions, establishing six commitments for fair, ethical, and transparent government algorithm use. The framework emphasizes human oversight and Māori data sovereignty considerations.
Israel	**"Artificial Intelligence Regulations and Ethics"** encourages "responsible AI innovation in the private sector" through a principled-based, sector-specific regulatory approach using 'soft' tools, such as non-binding ethical principles and voluntary standards.
South Africa	**National AI Policy Framework** establishes twelve strategic pillars for responsible AI development. The framework emphasizes human-centered AI, addressing socioeconomic disparities through talent development, digital infrastructure, and ethical governance.

Annexure 3: Overview of current laws in India relevant to AI systems (Illustrative)

Below is an illustrative list of statutes and regulations in India that may be applicable to the development, deployment and use of AI systems.

Information Technology Act, 2000 (IT Act):
The IT Act remains the backbone of India's digital regulation. Section 66D addresses cheating by personation using computer resources, applicable to AI-generated impersonations and deepfakes. Section 79, along with the 2021 Intermediary Guidelines, places due diligence obligations on online platforms, requiring active monitoring and takedown of unlawful AI-generated content, including misinformation and harmful deepfakes.

Bharatiya Nyaya Sanhita, 2023 (BNS):
In addition to the IT Act, certain harms/cybercrimes perpetuated by AI could also fall under the BNS. For instance, identity theft and cheating by personation are offences under Section 319(2) (cheating by personation), section 336(1) and 336(2) (forgery for the purpose of cheating), section 294 and 296 (selling/circulating/distributing obscene objects), and section 356(1) (causing harm to reputation/defamation).

Digital Personal Data Protection Act, 2023 (DPDP Act):
The DPDP Act introduces obligations of consent, purpose limitation, and data minimisation that have direct bearing on AI model training and deployment. It prohibits processing of personal data without consent, requires safeguards against misuse of sensitive data, and empowers the Data Protection Board to investigate harms caused by misuse of AI-driven profiling. These provisions create accountability pathways for AI developers and deployers handling personal data at scale.

Consumer Protection Act, 2019 (CPA):
The CPA protects consumers against unfair trade practices, misleading advertisements, and deficiency of service. Its provisions can be invoked where AI-enabled systems mis-sell financial products, misrepresent the capabilities of AI-driven health devices, or cause consumer harm through opaque algorithms in e-commerce. The Central Consumer Protection Authority is empowered to order corrective advertising or levy penalties on misleading AI claims, including advanced forms of dark patterns.

Sectoral legislations:
Sector-specific legislations such as the Telecommunications Act, 2023, under which rules are being notified in areas such as cybersecurity, critical infrastructure, and incident reporting also strengthen the implementation of AI governance principles.

AI-specific guidelines:
Sectoral regulators and technical bodies have been adapting their mandates to address AI-specific risks, issuing frameworks on cybersecurity, fairness, robustness, and ethical safeguards. These initiatives reflect the operational realities of each domain: financial stability in banking, integrity in securities markets, safety and reliability in telecom, and accountability in healthcare. Collectively, they demonstrate how India's oversight architecture is evolving in practice.

Reserve Bank of India (RBI):
RBI's regulatory architecture on technology risk has progressively expanded to cover AI. The *Cybersecurity Framework for Banks (2016)* established board-approved cyber policies, continuous monitoring, incident reporting, and resilience planning, all of which extend to AI-enabled services.

The Digital Lending Guidelines (2022) require transparency, consent, and accountability in automated decision-making, and are now expected to incorporate disclosure obligations for AI-driven credit scoring and fairness audits.

Building on these foundations, the *Framework for Responsible, Explainable and Ethical AI (FREE-AI) Committee Report (2025)* sets out detailed AI-specific measures: adoption of board-approved AI policies covering governance, lifecycle management, vendor oversight, and annual review; integration of AI-specific threats such as adversarial attacks and model poisoning into cybersecurity protocols; and the creation of a tiered incident reporting system for AI failures, including bias, explainability gaps, and unintended outcomes.

Securities and Exchange Board of India (SEBI):
SEBI's *Cybersecurity and Cyber Resilience Framework* requires market infrastructure institutions and intermediaries to maintain security operation centres, conduct vulnerability assessments, and submit compliance reports. AI-driven trading algorithms and surveillance systems fall under this framework, linking automation to accountability for market integrity. SEBI has also released a consultation paper on *"Guidelines for responsible usage of AI/ML In Indian Securities Markets"* in June, 2025.

Insurance Regulatory and Development Authority of India (IRDAI):
IRDAI mandates insurers and intermediaries to comply with its *Guidelines on Information and Cyber Security* for Insurers, with direct implications for AI-driven underwriting, claims management, and fraud detection.

Telecommunication Engineering Centre (TEC):
TEC has issued a Voluntary Standard for Fairness Assessment and Rating of AI Systems, covering bias detection and mitigation, and is developing a Standard for Assessing & Rating Robustness of AI Systems in Telecom Networks and Digital Infrastructure. TEC has also published a Draft Standard for the Schema and Taxonomy of an AI Incident Database in Telecommunications and Critical Digital Infrastructure. These standards provide structured pathways for trustworthy AI assessment focusing on fairness, robustness, and incident reporting in areas like critical infrastructure, network optimisation, and service quality management.

Indian Council of Medical Research (ICMR):
The *Ethical Guidelines for Application of AI in Biomedical Research and Healthcare* set expectations for safety, transparency, accountability, fairness, and human oversight. They require bias audits, independent ethics review, data quality checks, and delineation of responsibility between developers and healthcare providers.

CERT-In and NCIIPC (cross-sectoral cybersecurity):
Under the IT Act, 2000, *CERT-In Directions (2022)* mandate entities to report cybersecurity incidents within six hours, retain logs for 180 days, and enable audits. These requirements directly cover AI systems integrated into cloud platforms, fintech, or critical infrastructure. The *NCIIPC Rules (2014)* designate critical information infrastructure sectors and require mandatory safeguards, monitoring, and incident response, provisions highly relevant to AI deployment in energy, telecom, and transport.

Bureau of Indian Standards (BIS):
The BIS Technical committee LITD 30 develops standards in the area of artificial intelligence for India. This committee also contributes to the development of International Standards (for eg. ISO/IEC JTC 1/SC 42 “Artificial intelligence”). The list of standards published/under development by BIS are in Annexure 6.

Annexure 4: Applicability of existing laws in India to regulate AI harms (illustrative)

Nature of Harms	Applicable Statutory Law
Depiction of a child in a sexually explicit video that is AI-generated	• Information Technology Act, 2000 • Bharatiya Nyaya Sanhita, 2023 • Prevention of Children from Sexual Offences Act, 2012
Unauthorized impersonation using AI-generated deepfakes	• Bharatiya Nyaya Sanhita, 2023 • Information Technology Act, 2000
Discrimination in hiring decisions using AI recruitment tools	• Rights of Persons with Disabilities Act, 2016 • Transgender Persons (Protection of Rights) Act, 2019 • Code on Wages, 2019 • Scheduled Castes and the Scheduled Tribes (Prevention of Atrocities) Act, 1989
Use of an individual's personal data without consent to train AI models	• Digital Personal Data Protection Act, 2023 • Information Technology Act, 2000
Misleading ads about the reliability or performance of an AI service	• Consumer Protection Act, 2019
Use of copyright-protected material in AI-generated content without permission of the author or owner	• The Copyright Act, 1957
Use of AI/ML technologies in the securities market for the purpose of algorithmic trading and artificially affecting the market trends.	• SEBI Act, 1992 • Banking Regulation Act, 1949 • Sectoral Guidelines by SEBI and RBI