# Privacy and AI Training Models

**Mains**: *GSII – Governance*

## Why in news?

*The Artificial Intelligence (AI) is transforming governance, economic activities and public service delivery. However, using personal information to train AI models raises concerns about privacy, consent and individual autonomy.*

## How Can AI Training Models Compromise Privacy?

- **Personal data storing** – AI models learn patterns from large datasets containing text, images, videos and personal information.
    - During training, they may memorise specific personal details.
- **Data memorisation** – AI models may reproduce personal information through carefully designed queries.
- **Re-identification** – Anonymised datasets can be combined with publicly available information to identify individuals.
- **Sensitive information inference** – AI can infer health conditions, emotions, locations and behavioural patterns from seemingly harmless information.
- **Unauthorised profiling** – Extracted information can be used to construct detailed profiles without individuals' knowledge or consent.
    - Thus, anonymisation alone may not ensure privacy when AI can combine and analyse different information sources.

## What is the Existing Legal Framework?

- **Constitutional Protection** – The Supreme Court recognised privacy as a fundamental right under Article 21 in the 2017 Puttaswamy judgment.
    - Privacy protects individual dignity, autonomy and personal information against arbitrary interference.
- **Digital Personal Data Protection Act, 2023** – The DPDP Act establishes a general framework for regulating personal data, including consent, security and individual rights.
    - However, it does not specifically address AI training or the risks of information memorised and inferred by AI models.
    - Its substantive provisions are being implemented in phases, with major obligations becoming effective in 2027.
- **Information Technology Act, 2000** – The IT Act and its 2011 Rules provided safeguards for specified sensitive personal data.

- However, ordinary personal information, including browsing history and behavioural patterns, received limited protection.

## What are the Major Concerns associated?

- **Inadequate Protection of Sensitive Data** – The DPDP Act does not create a separate category for sensitive personal information.
  - Consequently, passwords, biometric information and health conditions receive no distinct category of protection, despite their potential for serious misuse.
- **Limited Consent and Objection Rights** – Indian users lack an effective mechanism to prevent their publicly available data from being used for AI training.
  - Even after the DPDP Act becomes fully operational, its treatment of publicly available information may leave important privacy gaps.
- **Weaknesses in the Right to Erasure** – The Act does not clearly establish whether individuals can demand deletion of information inferred from their data or memorised by trained AI models.
  - This creates uncertainty regarding individual control over personal information.
- **Research Exemptions** – The DPDP Act exempts certain processing activities undertaken for research, archival or statistical purposes.
  - AI training may be presented as research to avoid certain obligations, making clarification of the exemption essential.
- **Discrimination and Social Exclusion** – AI models trained on biased datasets may generate discriminatory outcomes.
  - Inaccurate profiling can affect access to employment, credit, healthcare and public services, particularly for vulnerable communities.

## What are the Challenges in Regulating AI Training?

- **Balancing Privacy and Innovation** – Excessive restrictions may discourage AI research and reduce technological competitiveness.
  - However, unrestricted data collection can undermine privacy and public trust.
- **Cross-Border Data Flows** – India is an important source of data for global AI training.
  - Weak safeguards may expose Indian citizens to privacy risks arising from foreign AI systems.
- **Technological Limitations** – AI models can retain information even after the original dataset is deleted.
  - Conventional data protection mechanisms may therefore be insufficient to address memorisation and re-identification.

## What is the Way Forward?

- **Strengthen the DPDP Act** – Extend data protection provisions to AI training, publicly available personal data and information inferred through AI analysis.
- **Introduce Sensitive Data Classification** – Create a separate category for sensitive information, with stricter requirements for collection, processing and disclosure.
- **Strengthen Individual Rights** – Provide a clear right to object to AI training and strengthen erasure rights, including information derived from personal data.

- **Promote Responsible AI** – Encourage data minimisation, security audits and accountability mechanisms to prevent unauthorised extraction.
- **Clarify Research Exemptions** – Define the scope of AI training under research exemptions and prevent these provisions from becoming a blanket mechanism for avoiding data protection obligations.
- **Strengthen Global Cooperation** – Participate in international discussions on responsible AI governance, privacy protection and cross-border data flows.

## What lies ahead?

- AI can improve productivity, governance and public service delivery, but its development should not come at the cost of constitutional rights.
- India must balance technological innovation with privacy, accountability and individual autonomy through a robust and forward-looking regulatory framework.

## Reference

The Hindu| How AI training models could compromise privacy

