DSCI
PROMOTING DATA PROTECTION

Brief on

MeitY's Handbook on Data Empowerment and Protection Architecture for AI Training



Brief on MeitY's Handbook on Data Empowerment and Protection Architecture for AI Training

In October 2026, the Ministry of Electronics and Information Technology (MeitY) launched the Handbook on Data Empowerment and Protection Architecture for AI Training (DEPA). The DEPA enables purpose-bound data collaboration through a techno-legal framework that addresses privacy, confidentiality, intellectual property and institutional control. It combines contractual permissions, controlled computing environments, privacy safeguards and data provenance to ensure secure, transparent and accountable data sharing.

DEPA For AI Training

The DEPA proposes a Digital Public Infrastructure (DPI) for Artificial Intelligence (AI) that enables secure, privacy-preserving collaboration between institutions holding valuable datasets (referred to as Training Data Providers or TDPs) and entities seeking to use such datasets for AI development (referred to as Training Data Consumers or TDCs), including startups, enterprises and research institutions.

It addresses the limitations of existing data-sharing arrangements, where concerns around privacy, confidentiality, intellectual property and institutional control often prevent access to valuable datasets. Unlike DEPA for inference, which facilitates consent-based sharing of individual data for specific transactions, the present framework is designed to enable the use of aggregated, de-identified datasets for developing AI models.

The proposed architecture rests on four key technological components:

Data Exchange Platforms	Facilitates the discovery, registration, and exchange of data between participating institutions.
Electronic Contracts	Establishes machine-readable agreements governing data access, permitted purposes, usage restrictions and obligations of participating entities.
Secure Data Environments	Provides controlled computing environments where AI models can be trained without exposing or transferring raw datasets to data consumers.
Differential Privacy	Introduces calibrated statistical noise to protect individual-level information and reduce the risk of re-identification through model outputs.

The DEPA framework envisages Data Collaboratives (DCs) to aggregate and standardize cross-sectoral datasets within a federated, decentralized ecosystem, guided by the ADEPTS principles of Accountability, Democratization, Evolvability, Privacy, Transparency and Security, while ensuring data providers retain control over their data.

A Techno-Legal Approach to AI Governance

DEPA proposes a techno-legal approach to AI governance that embeds compliance into technical architecture and operational processes, combining enforceable legal agreements with technological safeguards to ensure that data sharing and AI training adhere to authorized purposes and applicable privacy, security and governance requirements:

Electronic Contracts

Automate the enforcement of data-sharing agreements, purpose limitations and usage restrictions, while creating an auditable record of obligations.

Secure Data Environments

Ensure that AI training takes place in controlled computing environments, preventing direct exposure of raw datasets.

Differential Privacy

Provides mathematical privacy guarantees by limiting the ability to infer information about individuals from aggregate outputs.

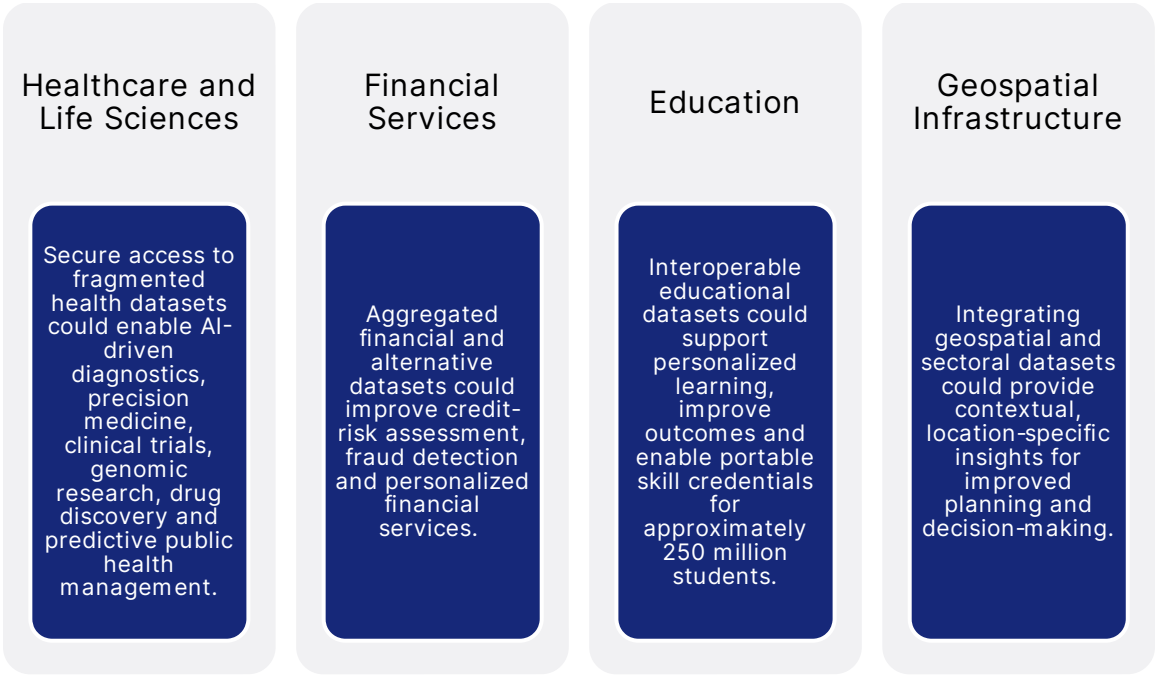
Zero-Trust Security

Requires continuous authentication and authorization of every interaction, supported by encryption, monitoring and audit trails.

The framework proposes a multi-stakeholder oversight body (comprising TDPs, TDCs, DCs, technology service providers, sectoral experts, and civil society representatives) responsible for participant registration, setting privacy and security standards, certifying infrastructure providers, monitoring compliance and resolving disputes, alongside independent AI model validation for bias, fairness, security, robustness, safety and explainability. It also envisages alignment with the Information Technology Act, 2000, the Digital Personal Data Protection Act, 2023, as well as international and sector-specific frameworks, including the General Data Protection Regulation and the AI Act of the European Union, and the Health Insurance Portability and Accountability Act of 1996, of the United States of America.

Unleashing Sectoral Transformation

DEPA training seeks to drive AI-led economic growth by enabling cross-sectoral data collaboration, with the potential to unlock approximately USD 500 billion in AI-driven economic value for India's GDP by 2030. It also envisages integration with national data platforms, including AI KOSH, API Setu, NAPIX, NDAP, GDI and OGD, to facilitate access to diverse datasets and enable the development of contextually relevant AI models. To effectively unleash sectoral transformation, the following sectoral opportunities have been identified:



A Global Vision for Responsible Innovation

The framework positions DEPA training as a potential global model for responsible AI governance, promoting ethical data sharing, privacy protection and democratized access while respecting national data sovereignty. Drawing on the UN AI Advisory Body's 2024 report titled "Governing AI for Humanity", DEPA is envisaged as a decentralized, protocol-based DPI that could be adapted and adopted by other countries. Such architecture would allow countries to develop sovereign AI models using sensitive, locally generated datasets without compromising control over their data, while enabling secure participation in cross-border research and development.