



UPSC & STATE PCS CURRENT AFFAIRS · UJIYARI.COM

DAILY CURRENT AFFAIRS

Sarvam AI and India's Sovereign AI Push — IndiaAI Mission and Multilingual Models

11 February 2026

SUBJECTS COVERED

SCIENCE & TECH

ECONOMY

CURATED & WRITTEN BY

Bharat Choudhary

UPSC Educator & Content Creator •

[linkedin.com/in/epicbharat](https://www.linkedin.com/in/epicbharat)

Free UPSC & State PCS Resources

ujiyari.com

Sarvam AI and India's Sovereign AI Push — IndiaAI Mission and Multilingual Models

11 February 2026

WHY IN NEWS

Sarvam AI's new Vision model set benchmark records on Indian-language OCR tasks, outperforming Google Gemini 3 Pro and DeepSeek OCR v2 — underscoring India's accelerating progress in building sovereign AI systems under the Rs 10,300-crore IndiaAI Mission.

WHAT IS SARVAM AI?

Sarvam AI is an Indian AI startup focused on building foundation models tuned for Indian languages and cultural contexts. It is one of the flagship beneficiaries of the **IndiaAI Mission** — the Government of India's national initiative to build indigenous AI infrastructure and capabilities.

Unlike general-purpose international AI models (GPT-4, Gemini, Claude) trained predominantly on English-language internet data, Sarvam AI's mission is to build models that genuinely understand India's 22 scheduled languages, regional contexts, and domain-specific requirements (healthcare, agriculture, governance).

KEY TECHNICAL ACHIEVEMENTS — FEBRUARY 2026

Sarvam Vision (Multimodal Model)

A **3-billion-parameter vision-language model** trained on 22 Indian languages

Benchmark performance:

olmOCR-Bench: 84.3% (surpassing Google Gemini 3 Pro and DeepSeek OCR v2) — the first time an Indian model has outperformed major international models on a recognised AI benchmark

OmniDocBench v1.5: 93.28%

The model is designed for document understanding tasks critical to India's digital transformation: reading handwritten forms, processing Aadhaar/PAN documents in regional scripts, digitising historical archives

Bulbul V3 (Text-to-Speech)

Supports **35 professional voice profiles** across **11 Indian languages**

Roadmap: Expand to all 22 scheduled languages (listed in the Eighth Schedule of the Constitution)

Applications: Automated government service announcements in regional languages, IVR systems for rural citizens, accessibility tools for the visually impaired

Upcoming LLM Family

Three variants in development:

Sarvam-Large — frontier-class general-purpose model

Sarvam-Small — efficient deployment on limited compute

Sarvam-Edge — a 70-billion-parameter model optimised for edge deployment (running on devices without cloud connectivity)

INDIAAI MISSION — THE POLICY BACKBONE

The **IndiaAI Mission** is a **Rs 10,300-crore** initiative approved by the Government in March 2024, running through 2029. It operates across seven pillars:

IndiaAI Compute Capacity: Procuring 10,000+ GPUs for government-accessible AI compute; distributed across public institutions

IndiaAI Innovation Centre: Building foundation models for public interest applications

IndiaAI Datasets Platform: Creating shared national datasets (health records, land records, satellite imagery) for AI training

IndiaAI Application Development Initiative: Funding AI solutions in agriculture, health, education, and governance

IndiaAI FutureSkills: Training 1 million+ AI-skilled professionals

IndiaAI Startup Financing: Providing risk capital for deep-tech AI startups

Safe and Trusted AI: Developing India's AI safety and ethics framework

Key institutions:

MeitY (Ministry of Electronics and IT): Nodal ministry

NASSCOM, IITs, IISc: Technical partners for skilling and research

Digital India Corporation (DIC): Programme management

WHY SOVEREIGN AI MATTERS FOR INDIA

Strategic dependency risk: India's current AI adoption is almost entirely dependent on models built and controlled by US companies (OpenAI, Google, Meta) and Chinese alternatives. This creates:

Data sovereignty concerns: Sensitive user data (medical records, financial transactions) processed on foreign servers under foreign legal jurisdictions

Alignment risk: Models trained primarily on Western/English data may be poorly calibrated for Indian cultural norms, legal frameworks, and language nuances

Geopolitical leverage: A country dependent on foreign AI infrastructure for critical services (tax processing, defence analytics, healthcare diagnostics) faces potential supply disruptions in conflict scenarios

Digital Public Infrastructure (DPI) angle: India’s success with DPI (Aadhaar, UPI, CoWIN) demonstrates that public-good infrastructure built on open standards and domestic capabilities can achieve global-scale impact. Sovereign AI is the next frontier of this DPI philosophy.

Language imperative: Over 500 million Indians are not functionally proficient in English. Without AI models that genuinely work in Hindi, Tamil, Telugu, Kannada, Odia, Marathi, and other Indian languages, AI’s productivity benefits will deepen India’s existing digital divide rather than bridging it.

CHALLENGES FOR INDIA’S SOVEREIGN AI PROJECT

Compute gap: India currently has limited high-end AI GPU infrastructure. The IndiaAI Mission’s 10,000-GPU target, while significant domestically, compares unfavourably with US hyperscalers (Microsoft, Google, Amazon) which deploy hundreds of thousands of H100-class GPUs. The global semiconductor supply chain — dominated by TSMC (Taiwan) and constrained by US export controls — limits India’s ability to rapidly scale compute.

Talent pipeline: India produces large numbers of software engineers but relatively few AI researchers with experience in large-scale model training, architecture innovation, or AI safety. Most AI PhDs educated in India move to US industry (Google DeepMind, OpenAI, Meta AI). The “brain circulation” problem remains unresolved.

Data quality: Building high-quality multilingual training datasets for 22 Indian languages requires massive curation investment. Much available text on the internet for regional Indian languages is of poor quality (transliterated, inconsistent script, sparse).

Evaluation standards: India lacks indigenous AI benchmarking frameworks analogous to MMLU or BIG-Bench (US) or CEVAL (China). Relying on international benchmarks may not capture what “good performance” means for Indian language and cultural contexts.

UPSC RELEVANCE

IndiaAI Mission (Rs 10,300 crore, 7 pillars), Sarvam AI (3B parameter vision model, 22 languages, olmOCR-Bench 84.3%), Bulbul V3 (35 voices, 11 languages), MeitY, Digital India Corporation (DIC), Eighth Schedule (22 scheduled languages), Digital Public Infrastructure (DPI).

*Government policies — IndiaAI Mission; digital governance; AI regulation. **GS-3:** Science and technology — artificial intelligence; Indigenous R&D; Atmanirbhar Bharat in technology; data sovereignty.*

FACTS CORNER — KNOWLEDGEPEDIA

INDIAAI MISSION:

Budget: **Rs 10,300 crore** (approved March 2024, valid through 2029)

Nodal Ministry: **MeitY** (Ministry of Electronics and Information Technology)

Compute target: **10,000+ GPUs** for public institutions

Skilling target: **1 million+** AI-skilled professionals (IndiaAI FutureSkills pillar)

Seven pillars: Compute Capacity, Innovation Centre, Datasets Platform, Application Development, FutureSkills, Startup Financing, Safe and Trusted AI

SARVAM AI — KEY DATA:

Vision model: **3 billion parameters**; trained on **22 Indian languages**

olmOCR-Bench score: **84.3%** (beats Google Gemini 3 Pro, DeepSeek OCR v2)

OmniDocBench v1.5 score: **93.28%**

Bulbul V3 TTS: **35 voices, 11 Indian languages**; target: all 22 scheduled languages

Upcoming: Sarvam-Large, Sarvam-Small, **Sarvam-Edge** (70B parameter edge model)

INDIA'S AI COMPETITIVE POSITION:

India ranks **45th** on Network Readiness Index 2025 (Portulans Institute); up 4 positions

India ranks 2nd among lower-middle-income economies (after Vietnam)

Global top 3 NRI: USA (79.13), Finland (75.82), Singapore (75.46)

CONSTITUTIONAL & LEGAL CONTEXT:

Eighth Schedule of Constitution: Lists 22 scheduled languages of India

Digital Personal Data Protection Act (DPDPA), 2023: Governs personal data processing — including by AI systems

OTHER RELEVANT FACTS:

DPI = Aadhaar + UPI + CoWIN; India's DPI framework cited as model globally at G20 and World Bank

AI compute globally dominated by TSMC (chip fabrication), NVIDIA (GPU design)

US BIS (Bureau of Industry and Security) export controls on advanced AI chips: restrict direct supply to several countries

China's comparable initiative: ERNIE Bot (Baidu), Tongyi Qianwen (Alibaba), Kimi (Moonshot AI)

Global AI investment 2024: ~USD 200 billion (Stanford HAI report); USA leads with 67% share

Sources: Drishti IAS, Next IAS

CURATED & WRITTEN BY

Bharat Choudhary

UPSC Educator & Content Creator

 linkedin.com/in/epicbharat

Published on ujjari.com · Free UPSC & State PCS Current Affairs