ujiyari

UPSC & STATE PCS CURRENT AFFAIRS · UJIYARI.COM

EDITORIAL ANALYSIS

Regulating the Synthetic World — India's Deepfake Rules and the AI Governance Gap

THE HINDU

12 February 2026

SUBJECTS COVERED

SCIENCE & TECH POLITY

GS PAPERS

GS2 GS3

CURATED & WRITTEN BY

Bharat Choudhary

UPSC Educator & Content Creator •
linkedin.com/in/epicbharat

ujiyari

Regulating the Synthetic World — India's Deepfake Rules and the AI Governance Gap

The Hindu 12 February 2026 GS2 GS3

TH The Hindu | MAINS RELEVANCE: GS Paper 2 GS Paper 3 |

INTERVIEW ANGLE

"How should India balance freedom of expression with the harms caused by AI-generated synthetic media? Is the IT Amendment Rules 2026 approach adequate?"

WHY IN NEWS

MeitY's notification of the IT (Intermediary Guidelines) Amendment Rules 2026, mandating AI content labelling and a 2-hour takedown window for deepfakes, has renewed debate on whether India's regulatory architecture is adequate to govern synthetic media at scale.

WHY THIS MATTERS

The proliferation of AI-generated synthetic media — deepfakes, voice clones, AI-written disinformation — represents a qualitatively new challenge for democratic governance. Unlike traditional misinformation, deepfakes weaponise the face and voice of real individuals, making falsehoods viscerally convincing.

India's particular vulnerability is acute: a 2024 Home Affairs report noted that synthetic media abuse is concentrated in three domains — **electoral manipulation**, **non-consensual intimate imagery (NCII)** targeting women, and **financial fraud** (voice-cloned CEOs authorising wire transfers). Each of these directly threatens democratic participation, gender equality, and economic security.

WHAT THE RULES DO — AND WHY IT IS NOT ENOUGH

The IT Amendment Rules 2026 represent meaningful progress. The 2-hour takedown window for deepfakes is one of the **fastest globally** — even faster than the EU's Digital Services Act (DSA) timelines for most illegal content. The mandatory labelling regime creates a shared accountability structure between platforms and users.

But the framework has four structural weaknesses.

First, enforcement is detection-dependent. The rules require platforms to label and act on AI-generated content — but AI detection accuracy ranges from **60–85%** (IBM Research, 2024). Platforms cannot reliably distinguish deepfakes from real content at scale. The regulation assumes a technical capability that does not yet exist with sufficient reliability.

Second, the liability shift is incomplete. By tying safe harbour loss to platforms that "knowingly permit" violative content, the rules create incentives for wilful blindness. Platforms that invest minimally in detection can argue they lacked knowledge. A stronger standard — "knew or ought reasonably to have known" — would close this gap, aligning with EU DSA Article 6 obligations.

Third, the labelling exemption is too narrow. Only "minor automatic smartphone camera touch-ups" are exempted. This means documentary filmmakers using AI colour restoration, musicians using AI-mastered audio, and artists using AI-assisted image editing would all face labelling obligations — a disproportionate burden on legitimate creative industries.

Fourth, decentralised enforcement risks inconsistency. Authorising state officers to issue takedown orders without a centralised coordination mechanism opens the door to politically motivated takedown demands — a pattern already documented under Section 69A IT Act orders.

THE DEEPER QUESTION: INDEPENDENT AI REGULATOR

India's digital governance architecture suffers from a fundamental design flaw: there is no **independent AI regulator** equivalent to SEBI (securities), TRAI (telecom), or RBI (finance). The IT Amendment Rules 2026 are executive rules issued under MeitY — they can be changed, suspended, or selectively enforced by the same ministry that oversees both regulation and industry promotion.

Regulatory capture risk: MeitY is simultaneously India's AI promoter (INDIAAI Mission, IndiaAI portal) and its AI regulator. This conflict is analogous to the pre-SEBI era when the Finance Ministry both promoted capital markets and regulated them — a arrangement that enabled the Harshad Mehta scam.

The **EU AI Act (2024)** provides a comparative model. It creates a **European AI Office** with independent enforcement powers, establishes a risk-based tiered framework (unacceptable risk → prohibited; high-risk → strict compliance; limited/minimal risk → lighter obligations), and applies proportionality standards. Deepfakes fall under "high risk" when used in political contexts.

India's IT Ministry discussion paper (2023) considered but ultimately rejected a standalone AI law in favour of amending existing sectoral laws (IT Act, IPC, POCSO, etc.). This piecemeal approach risks leaving gaps precisely where harms are most novel.

FREEDOM OF EXPRESSION DIMENSION

Article 19(1)(a) protects freedom of speech and expression. Mandatory labelling of all AI content faces three potential legal challenges:

Over-breadth: Labelling AI-assisted journalism (background removal, noise reduction) as "synthetic" misleads users about editorial integrity without corresponding public interest

Chilling effects: Creators may self-censor rather than navigate complex compliance requirements

Proportionality: The Supreme Court in *Shreya Singhal v. Union of India* (2015) — which struck down Section 66A IT Act — established that digital speech restrictions must be proportionate and narrowly tailored

However, protecting individuals from non-consensual deepfakes is a **compelling state interest** under Article 21 (right to privacy and dignity — *K.S. Puttaswamy v. Union of India*, 2017). The rights balance tilts toward regulation when identity theft causes concrete harm.

WHAT ADEQUATE REGULATION REQUIRES

Technically: Platforms should be required to disclose their detection accuracy rates (similar to how pharmaceutical companies disclose drug efficacy) and invest in provenance technology — embedding cryptographic watermarks at AI model output level (as developed by C2PA — Coalition for Content Provenance and Authenticity).

Institutionally: India needs either an independent AI regulator or a high-powered standing committee (like the Financial Stability and Development Council model) with cross-ministry coordination on AI harms.

Legally: A standalone Synthetic Media Disclosure Act — criminalising non-consensual deepfake creation with civil and criminal liability, on the model of the UK Online Safety Act 2023 — would provide clearer deterrence than rules under the IT Act.

Internationally: India should actively engage in the **Global Partnership on AI (GPAI)** and push for binding international norms on deepfake attribution, analogous to the Nuclear Non-Proliferation Treaty's inspection regime but for AI-generated content provenance.

UPSC RELEVANCE

> *IT (Intermediary Guidelines) Rules 2021, IT Amendment Rules 2026, Section 79 IT Act (safe harbour), MeitY, EU AI Act 2024, C2PA, EU Digital Services Act (DSA), Global Partnership on AI (GPAI), Article 19(1)(a), Article 21, Shreya Singhal v. UoI 2015, K.S. Puttaswamy v. UoI 2017, SEBI, TRAI.*
>
> *Regulatory bodies and their independence; digital governance; right to privacy; freedom of expression vs. national security; comparison with global AI regulatory models.* **GS-3:** *Artificial intelligence governance; deepfakes and disinformation; cybersecurity; technology policy.*

📌 FACTS CORNER — KNOWLEDGEPEDIA

IT AMENDMENT RULES 2026:

Effective: **February 20, 2026**; Ministry: **MeitY**

Deepfake/NCII takedown: **2 hours**; Court/govt-ordered: **3 hours**

Previous timeline (2021 Rules): 24–36 hours

Safe harbour lost if platform "knowingly permits, promotes, or fails to act"

Legal basis: Section 79, IT Act 2000

DEEPFAKE DETECTION GAP:

Current AI detection accuracy: **60–85%** (IBM Research 2024)

Detection gap is a key implementation challenge for platform compliance

GLOBAL COMPARISONS:

EU AI Act (2024): Risk-based tiered framework; European AI Office for enforcement; deepfakes in political contexts = high risk

UK Online Safety Act 2023: Criminalises non-consensual deepfake creation

China Deepfake Regulations (2022): Real-name registration + mandatory watermarking

USA: No federal AI law; California AB 602 (2019) and AB 730 (2019) — state-level only

CONSTITUTIONAL PROVISIONS:

Article **19(1)(a):** Freedom of speech and expression

Article **19(2):** Reasonable restrictions on free speech

Article **21:** Right to privacy (expanded in K.S. Puttaswamy v. UoI 2017)

Shreya Singhal v. UoI (2015): Struck down Section 66A — digital restrictions must be proportionate

CONTENT PROVENANCE TECHNOLOGY:

C2PA (Coalition for Content Provenance and Authenticity): Industry consortium embedding cryptographic watermarks in AI-generated content; members include Adobe, Microsoft, Sony

INDIA AI GOVERNANCE BODIES:

INDIAAI Mission: ₹10,300 crore; compute + datasets + startups + research

GPAI (Global Partnership on AI): India joined in 2020; hosted GPAI summit 2023

No standalone AI law yet; piecemeal sectoral approach adopted

Sources: The Hindu, MeitY Notification, Next IAS

CURATED & WRITTEN BY

Bharat Choudhary

UPSC Educator & Content Creator

linkedin.com/in/epicbharat

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