

AI Content Labelling Rules 2026: New IT Norms on Deepfakes and Takedown Timelines

February 15, 2026 • vajiramandravi.com



AI Content Labelling Latest News

- The Ministry of Electronics and Information Technology (MeitY) has amended the Information Technology (IT) Rules, 2021 to mandate labelling of AI-generated content by users and social media platforms.
- The amendment also significantly reduces the content takedown timeline — from the earlier 24–36 hours to just 2–3 hours — for all types of online content, not limited to AI-generated material.
- The revised rules will come into effect from February 20.

AI-Generated Content Under the 2026 IT Rules

- **Mandatory Labelling of Synthetic Content**
 - The Information Technology (Intermediary Guidelines and Digital Media Ethics Code) Amendment Rules, 2026 require social media platforms to prominently label “**synthetically generated**” or **AI-generated images and videos**.

- Platforms with over five million users must obtain a user declaration for AI-generated content and conduct technical verification before publishing it.
- According to MeitY, the measure aims to combat deepfakes, misinformation, privacy violations, and threats to national integrity.
- The goal is to ensure users are clearly informed when content is AI-generated or inauthentic.
- **Narrowed Definition and Exemptions**
 - While the earlier draft had a broader definition of “**Synthetically Generated Information**” (SGI), the final rules provide exemptions.
 - Automatically retouched smartphone photos and film special effects are excluded from mandatory labelling.
- **Prohibited Synthetic Content**
 - The rules strictly prohibit certain types of AI-generated material, including:
 - Child sexual exploitation and abuse content
 - Forged documents
 - Information on developing explosives
 - Deepfakes falsely impersonating real individuals
 - These provisions strengthen safeguards against harmful or unlawful AI use.

Detecting AI-Generated Content: Platform Responsibilities and Technical Standards

- The government has directed large social media platforms to **deploy** “*reasonable and appropriate technical measures*” to detect and prevent unlawful synthetically generated information (SGI), while ensuring proper labelling, provenance tracking, and identifier requirements for permissible AI-generated content.
 - According to IT Ministry officials, major platforms already possess sophisticated AI tools capable of detecting synthetic content.
 - The new rules formalise and mandate the consistent use of these detection systems rather than introducing entirely new obligations.
 - In addition, several AI companies and digital platforms are part of the **Coalition for Content Provenance and Authenticity (C2PA)**, which has developed technical standards to embed invisible digital markers in AI-generated content.
 - These markers can help other platforms identify such content if automated detection systems fail.
- While the amended rules refer to “**provenance/identifier requirements**,” the government has clarified that it does not intend to endorse any single technological framework.
 - Provenance and identifier requirements define the necessary, documented record of a resource’s origin, ownership, and history, ensuring authenticity and trust.
- Instead, it seeks to institutionalise the broader objective of reliable AI content detection and traceability through collaborative standards.

Tightened Takedown Timelines Under Amended IT Rules

- The amended IT Rules significantly shorten the timelines for removing unlawful content.

- **Government- and Court-Ordered Takedowns:** Platforms must now comply within 2–3 hours, instead of the earlier 24–36 hours.
- **User Complaints (General Categories):** For issues such as defamation and misinformation, response time has been reduced from two weeks to one week.
- **Sensitive Content Reports (Rule 3(2)(b)):** The deadline has been cut from 72 hours to 36 hours.
- The government justified these changes by arguing that longer timelines previously allowed harmful content to cause significant damage before removal, making stricter response windows necessary.

Stricter User Notifications and Liability Warnings

- The amended IT Rules introduce tighter obligations for platforms to inform and caution users.
- **More Frequent Policy Reminders:** Platforms must now notify users of their terms and conditions every **three months**, instead of once a year. These notifications must clearly explain consequences of non-compliance and reporting obligations.

Explicit Warnings on AI Misuse

- Users must be warned that posting harmful deepfakes or other illegal AI-generated content may lead to:
 - Immediate content removal
 - Suspension or termination of accounts
 - Disclosure of identity to law enforcement agencies
- These changes aim to strengthen accountability and deter misuse of AI-generated content online.

Source: TH

Source: <https://vajiramandravi.com/current-affairs/ai-content-labelling-rules-2026-new-it-norms-on-deepfakes-and-takedown-timelines/>

© Vajiram & Ravi. For personal use only.