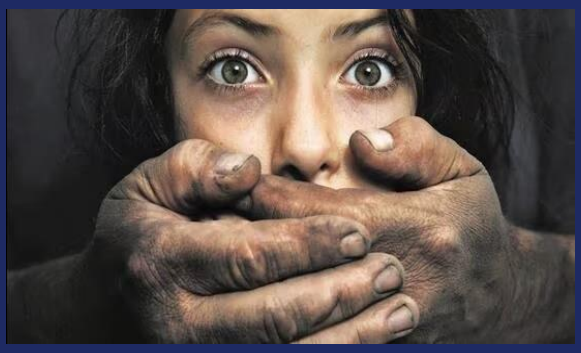

POLICY BRIEF | JULY 2026

Strengthening Platform Accountability to Prevent the Promotion of Child Sexual Abuse Material on Social Media in India

A review of India's legal and institutional response to CSAM on digital platforms, and a prioritized agenda for government, platforms, and civil society.

900 million+ Users

India is one of the world's largest online markets, with over 900M internet users and 500M+ active social media users — a huge share of them minors.

1.9 million Reports

India logged 1.9M CSAM reports to the NCMEC CyberTipline in 2025 — the most of any country besides the United States.

2–3 Hour Window

The 2026 Amendment Rules compress CSAM takedown timelines to 2–3 hours, far outpacing most platforms' and regulators' current response capacity.

1. Executive Summary

India is one of the largest and fastest-growing online markets globally, with more than 900 million internet users and over 500 million active social media users.

High smartphone penetration among minors, low digital literacy in many households, and the scale of encrypted messaging including grooming, live-streamed abuse, and increasingly AI-generated "synthetic" CSAM have made Indian platforms a significant vector for the dissemination of Child Sexual Abuse Material (CSAM).

India's regulatory framework has developed rapidly, the IT (Intermediary Guidelines) Rules, 2021, followed by the 2026 Amendment Rules, have shortened takedown timelines, introduced new POCSO reporting requirements, and set binding guidelines on synthetic and AI-generated content for the first time.

Nevertheless, enforcement gaps remain uneven technology adoption across platforms, encrypted services outside effective detection, slow cross-border evidence sharing with NCMEC, and far lower compliance capacity among small and mid-sized platforms than large ones.

2. Policy Brief Objectives

- Close detection and reporting gaps across all platform sizes and content types, from encrypted to synthetic media.
- Decrease the delay between content appearing and takedown or law-enforcement notification without incentivizing overly broad or reflexive removal.
- Improve India's own detection and verification capabilities so intervention does not rely largely on referrals from US intermediaries.
- Build compliance capacity at small and mid-sized firms, not just a handful of large ones.
- Strengthen survivor-centred reporting, takedown-request, and data-deletion processes.

3. The Problem

Multiple, overlapping forms of CSAM circulate on Indian platforms:

- **Re-circulation of known material** — previously identified images and videos, re-uploaded and shared, usually in closed messaging groups.
- **Grooming and live-streamed abuse** — solicitation of minors via games, social apps, and direct messages, sometimes escalating to live-streamed abuse that is harder to detect than static images.
- **Synthetic and AI-generated CSAM** — "nudify" tools and generative AI producing sexually explicit images of real or composite children, with no existing hash (unique digital fingerprints) to match against and legal frameworks not designed with synthetic imagery in mind.
- **Peer-to-peer and encrypted distribution** — WhatsApp and other end-to-end encrypted (E2EE) services, where server-side content scanning is technically and legally limited.

4. Why Current Safeguards Fall Short

An investigation by BBC Eye, released July 3, 2026, found that Instagram had approved and promoted paid advertisements in India containing explicit content and links to Telegram channels advertising CSAM for as little as ₹99 around 30 CSAM-promoting ads and 20 adult-content ads, with one flagged ad left live for 24 hours despite review.

MeitY responded within days, IT Minister Ashwini Vaishnaw directed officials to summon Meta, ordered Instagram to block the ads, and required a compliance explanation within seven days, still slower than the statutory 2–3-hour takedown window.

Separately, Telegram told the BBC it had removed more than 274,000 CSAM-linked groups and channels in 2026 through its own moderation, while India logged 1.9 million CSAM reports to NCMEC in 2025, the most of any country besides the US.

5. Key structural gaps

- **Inequitable technology detection.** Large platforms rely on PhotoDNA, CSAI Match, and ML classifiers reporting to NCMEC, but smaller apps, gaming chats, and content-sharing tools without hash-matching have become unintentional safe havens and the July 2026 Instagram episode shows even large platforms run weaker moderation pipelines for ads than for organic content.
- **The encryption gap.** CSAM circulation occurs substantially on E2EE channels not vulnerable to perceptual hashing; India's framework lacks anything like the EU's temporary e-Privacy derogation for CSAM detection on encrypted services.
- **Cross-border reporting friction.** Most hash-match reporting runs through NCMEC forwarding CyberTipline reports to India's NCRB, functional, but with jurisdictional and timing delays, and without direct Indian access to global hash databases (IWF, NCMEC, Project Arachnid).
- **A compliance-capacity gap at smaller platforms.** Differentiated IT Rules requirements still demand the most of SSMIs (5M+ users), while mid-size and regional platforms - where abuse often migrates once major platforms tighten controls lack the incentive and technical capacity to comply.
- **Under-reporting and missing NCRB data.** Reported cases remain a small share of actual incidence, due to low awareness, stigma, and platforms reporting to NCMEC in the US without always relaying reports to Indian authorities.
- **Synthetic content is outpacing detection.** The 2026 Amendment Rules require removal of synthetic CSAM and provenance labelling, but detection still relies on tools less effective against never-before-seen AI-generated image

6. Legal and Institutional Framework

Instrument	Relevant provisions for CSAM
IT Act, 2000	Section 67B criminalizes publishing/transmitting material depicting children in sexually explicit acts; Section 79 provides safe harbour for intermediaries conditional on due diligence.
POCSO Act, 2012	Section 15 penalizes storage/possession of CSAM for commercial or non-reporting purposes; Section 19 imposes mandatory reporting obligations on any person including platforms aware of an offence against a child.
IT Intermediary Guidelines Rules, 2021	Establishes due-diligence obligations, grievance redressal mechanisms, and a differentiated compliance regime for Significant Social Media Intermediaries (SSMIs).
IT Intermediary Guidelines Amendment Rules, 2026	Compresses the general takedown window from 36 to 3 hours, and to 2 hours for CSAM and non-consensual intimate imagery; mandates labelling and provenance metadata for AI-generated content; requires automated blocking of synthetic CSAM; ties platform knowledge to mandatory POCSO reporting; shortens grievance redressal from 15 to 7 days; conditions safe harbour on compliance.
Bharatiya Nyaya Sanhita (BNS) / Bharatiya Nagarik Suraksha Sanhita (BNSS), 2023	Replaced the IPC/CrPC; contain provisions relevant to sexual offences and platform reporting obligations referenced in the 2026 Rules.
Digital Personal Data Protection Act, 2023	Imposes fiduciary duties on platforms regarding minors' data, including verifiable parental consent, indirectly relevant to age-verification and child-safety-by-design obligations.
National Crime Records Bureau (NCRB) / National Cyber Crime Reporting Portal (NCRP)	Domestic reporting channel; receives CyberTipline referrals relayed from NCMEC as well as direct public complaints.

7. Recent judicial and enforcement developments (2024–2026)

Just Rights for Children Alliance v. S. Harish (Supreme Court, September 2024) resolved a split among the Madhya Pradesh, Bombay, Madras, and Kerala High Courts on whether mere possession of CSAM without intent to transmit or commercial purpose was punishable. The Court held that accessing CSAM online even without downloading or storing it constitutes constructive possession under POCSO Section 15; that failure to delete, destroy, or report is a distinct offence from transmission; and that the term "child pornography" should be replaced with "Child Sexual Exploitative and Abuse Material" (CSEAM) in judicial usage. Critically, the Court ruled that reporting CSAM to NCMEC does not replace intermediaries' independent statutory duty to report to Indian authorities under POCSO Section 19, a dual-reporting requirement reiterated in the 2026 Amendment Rules, and one commentator note poses real operational challenges for platforms not built for parallel bulk reporting.

MeitY notice on "username" contact features (July 2026): in the same week as the Instagram ad episode, MeitY wrote to WhatsApp, Telegram, and Signal questioning username-based contact features that let users connect without sharing phone numbers, citing fraud, impersonation, and traceability concerns relevant to CSAM investigations. Meta agreed to postpone the WhatsApp rollout pending further consultation.

8. Policy Recommendations

A coordinated agenda across government, platforms, and civil society is needed to close detection, reporting, and enforcement gaps at every platform size.

A. For Government / Regulators *(MeitY, NCRB, Ministry of Women and Child Development)*

- Establish a national CSAM hash-matching repository interoperable with NCMEC, IWF, and Project Arachnid, hosted with NCRB/I4C enabling Indian law enforcement to search and submit hashes directly rather than only through foreign intermediaries.
- Set tiered compliance obligations by platform size and risk, with a staged path to hash-matching and classifier tools (e.g., subsidised access to PhotoDNA, Google's Content Safety API, Thorn's Safer) rather than a cliff-edge SSMI/non-SSMI split.
- Provide legal certainty on E2EE detection, clarifying how content-blocking versus client-side detection can apply to encrypted messaging, consistent with the Puttaswamy privacy jurisprudence.
- Require standardized, independently verifiable transparency reporting, CSAM detected, response timelines, and NCRB referral outcomes, so self-reported figures (like Telegram's 274,000 group takedowns) can be checked.
- Expand ad-review systems specifically, as a separate and verifiable transparency item, given the July 2026 Instagram case showed weaker moderation on ad pipelines than organic content.
- Operationalize the Supreme Court's dual-reporting standard from S. Harish (2024), with practical technical steps so NCMEC reporting and independent POCSO Section 19 reporting both function in practice.
- Fast-track training and tooling for state police cybercrime units, since the 2–3-hour takedown window is unworkable without matching state-level investigative capacity.

Scope any advisory-based safe-harbour conditions in the pending 2026 draft amendment carefully, so CSAM-specific rapid-response provisions are not diluted.

B. For Platforms

- Adopt hash-matching and classifier tools (PhotoDNA, CSAI Match, PDQ, or equivalent) across all user-generated content surfaces — not only flagship apps, but smaller regional apps, gaming, and chat features.
- Build India-focused, adequately resourced trust & safety and law-enforcement liaison teams able to operate within the 2–3-hour statutory window.
- Deploy classifier-based detection for unknown/synthetic CSAM, not only hash-matching against known material, given the growing volume of AI-generated content with no pre-existing hash.
- Offer trauma-informed, survivor-accessible reporting flows in Indian languages, including simplified in-app NCRP integration.
- Strengthen safeguards for accounts likely used by minors — age-assurance and protective design — aligned with DPDPA verifiable parental consent requirements.

C. For Civil Society, Researchers, and Industry Coalitions

- Increase India-specific engagement in global bodies such as WeProtect Global Alliance and the Technology Coalition, so India's threat landscape — vernacular content, regional platforms, low-bandwidth contexts — shapes global practice, not only US/EU use cases.
- Conduct independent audits of platform compliance claims, given the reliance on self-reported transparency data.
- Expand public digital-literacy and parental-awareness efforts, recognizing these are necessary but insufficient without platform-level detection.

9. Risks and Trade-offs to Manage

- **Over-removal and due process.** Rapid takedown schedules and conditional safe harbours create strong disincentives to over-remove; CSAM-specific fast-track processes should stay analytically distinct from broader content-moderation and misinformation measures.
- **Privacy and encryption.** Any move toward detection on encrypted services must account for India's constitutional right to privacy, assessing proportionality and efficacy rather than pursuing detection automatically.
- **False positives and hash-database integrity.** Detection technology is not error-proof; human review and appeals processes must continue even as automation scales.
- **Inconsistent global cooperation.** Reliance on NCMEC and international hash databases serves internationally known material well, but under-serves CSAM originating in India or in vernacular languages not indexed internationally.

10. Conclusion

India's regulatory regime has advanced substantially through the IT Rules of 2021 and 2026, the Supreme Court's 2024 POCSO clarification in S. Harish, and early adoption of synthetic-CSAM rules ahead of many other countries. Yet the events of early July 2026 — MeitY's summons to Meta over CSAM-linked Instagram ads, and Telegram's disclosure of 274,000 CSAM-related group takedowns in a year — show the remaining gap is less about rules than about uniform enforcement and verified compliance.

Closing that gap requires detection across all content surfaces including advertising, an independent domestic hash-matching and investigation capability, closing the dual-reporting implementation gap the Supreme Court identified, and confronting the harder problem of detection on encrypted channels within India's constitutional framework.

References

- World Health Organization (WHO) and international child-safety guidance.
- Ministry of Electronics and Information Technology (MeitY), Government of India.
- IT (Intermediary Guidelines and Digital Media Ethics Code) Rules, 2021 and Amendment Rules, 2026.
- Protection of Children from Sexual Offences (POCSO) Act, 2012.
- Just Rights for Children Alliance v. S. Harish, Supreme Court of India (September 2024).
- National Crime Records Bureau (NCRB) / National Cyber Crime Reporting Portal (NCRP).
- BBC Eye investigation into CSAM-linked advertising on Instagram (July 2026).
- NCMEC CyberTipline annual reporting data.

Subhanshu Jaiswal – Assistant Manager - MLE

Policy Brief | July 2026