© IJARW | ISSN (O) - 2582-1008
April 2026 | Vol. 7 Issue. 10
www.ijarw.com

SYNTHETIC CONTENT, INTERMEDIARY RESPONSIBILITY, AND DATA PROTECTION: A STUDY OF INDIA'S EMERGING REGULATORY FRAMEWORK

Dr. Shilpa Khandelwal

Assistant Professor, Modi Law College, Kota, Rajasthan

ABSTRACT

The proliferation of synthetic content, driven by rapid advancements in artificial intelligence, has introduced complex regulatory challenges for legal systems worldwide. In India, the emergence of deepfakes, AI-generated media, and algorithmically curated information has intensified concerns surrounding intermediary liability, data protection, and accountability in digital ecosystems. This article critically examines India's evolving regulatory framework governing synthetic content, with particular focus on the Information Technology Rules, 2026, and the Digital Personal Data Protection Act, 2023. It explores how these frameworks attempt to translate legislative intent into practical compliance obligations for intermediaries, while simultaneously addressing the risks posed by emerging AI technologies.

The study highlights the operational challenges faced by intermediaries in implementing due diligence requirements, content moderation standards, and transparency obligations under the IT Rules, 2026. It further interrogates the governance architecture of the DPDP Act, 2023, arguing that while the Act emphasizes data protection and user consent, it may fall short in ensuring robust accountability for the inputs and training datasets that underpin AI systems. This disconnect raises critical concerns regarding regulatory blind spots, particularly in the context of synthetic content generation.

By conceptualizing a "governance trilemma" involving innovation, accountability, and user rights, the article identifies key regulatory gaps that hinder effective oversight of AI-driven content ecosystems. A comparative analysis of international AI governance models—including risk-based approaches and rights-centric frameworks—provides valuable insights into alternative regulatory strategies. These global perspectives underscore the need for adaptive, context-specific regulation that can respond to the dynamic nature of artificial intelligence technologies.

The article ultimately argues for the development of a bespoke regulatory framework tailored to India's socio-legal realities. Such a framework must integrate principles of transparency, accountability, and ethical AI use, while fostering innovation and safeguarding fundamental rights. By bridging existing regulatory gaps and aligning legal standards with technological advancements, India can establish a more coherent and future-ready approach to governing synthetic content and intermediary responsibility.

Keyword: Artificial Intelligence Regulation; Data Protection; Intermediary Liability; Regulatory Governance; Synthetic Content

1. INTRODUCTION

The emergence of generative AI has fundamentally altered the established paradigms of intermediary liability and data privacy in India. Historically, digital platforms have benefited from extensive protection under the "safe harbour" principle established in Section 79 of the IT Act, 2000.[1] The swift weaponization of algorithmically altered media termed 'deepfakes' has required a

[1] Information Technology Act, 2000, § 79.

legal shift from reactive content management to proactive algorithmic governance.

The new Information Technology (Intermediary Guidelines and Digital Media Ethics Code) Amendment Rules, 2026 (hereafter "2026 IT Rules")[2] explicitly codify this change by establishing the legal classification of "Synthetically Generated Information" (SGI). The 2026 IT Rules require intermediaries to conduct demanding technical verification, maintain unalterable metadata provenance, and adhere to ultra-short takedown windows of three hours to qualify for immunity, contingent upon rigorous, AI-specific due diligence. Safe harbour is no longer an inherent entitlement; it has become a very constrained privilege.

The Digital Personal Data Protection Act, 2023 (hereafter "DPDP Act")[3] concurrently regulates the foundational input phase of GenAI. The improper algorithmic acquisition of an individual's biometric data, facial features, or voice to produce SGI represents a significant violation of the consent framework outlined in Section 6 of the DPDP Act.[4] However, a significant regulatory conflict arises at the convergence of these two frameworks. The 2026 IT Rules impose stringent regulations on the output and dissemination of SGI, however the DPDP Act seemingly establishes a jurisprudential void concerning the entry of training data, especially due to the "publicly available data" exception outlined in Section 3(c)(ii).[5]

This article analyses the intricate legal relationship between the 2026 IT Rules and the DPDP Act. It asserts that India's existing regulatory strategy, which seeks to regulate core AI models via intermediary responsibility and a broad data protection law, results in a disjointed and uneven structure. This paper contends that the excessive burden imposed on digital intermediaries, when examining the statutory mechanics of SGI compliance and the jurisprudential development of informational privacy[6] , inadequately holds upstream foundational model developers accountable, thus requiring a tailored, comprehensive AI regulatory framework.

2.TRANSLATING LAW INTO PRACTICE: COMPLIANCE FRAMEWORKS UNDER THE IT RULES, 2026

The legal framework of the 2026 IT Rules profoundly transforms the operational landscape for digital platforms in India. Before these revisions, intermediaries functioned under a reactive "notice-and-takedown" framework. The exponential scale of generative AI makes retrospective moderation ineffective. Thus, the 2026 IT Rules establish a stringent, proactive due diligence framework, converting intermediaries from passive conduits into active gatekeepers of synthetic media.

2.2.Statutory Recognition of SGI

The primary accomplishment of the 2026 IT Rules is the official legal classification of deepfakes and AI-altered media. The regulations define SGI as any audio, visual, or audio-visual content that is intentionally or algorithmically produced, developed, edited, or altered by computer resources, in a way that renders the material seemingly legitimate.[7]

The legislation critically prevents the over-criminalization of innocuous internet conduct. It delineates particular exceptions for standard, bona fide technical modifications. Fundamental formatting, color correction, noise reduction, and the utilization of computational resources only for enhancing accessibility (such as text-to-speech applications) are not encompassed by the stringent limitations of SGI, as long as they do not substantially alter the foundational visual or auditory data.[8] This statutory distinction is vital for protecting legitimate academic and creative expression while targeting malicious generative outputs.

2.3.Proactive Due Diligence and the "Before Publish" Mandate

For Significant Social Media Intermediaries (SSMIs), the obligation to comply has transitioned

[2] Information Technology (Intermediary Guidelines and Digital Media Ethics Code) Amendment Rules, 2026, G.S.R. 120(E).
[3] Digital Personal Data Protection Act, 2023.
[4] *Id.* § 6.
[5] *Id.* § 3.
[6] *Puttaswamy v. Union of India*, (2017), MANU/SC/1044/2017.
[7] Information Technology (Intermediary Guidelines and Digital Media Ethics Code) Amendment Rules, 2026, G.S.R. 120(E).
[8] *Id.* Rule 2(ca).

from reactive removal to proactive technical verification. Under the revised framework, SSMIs cannot depend exclusively on the retrospective identification of deepfakes. They are now legally required to collect a user declaration at the time of upload, necessitating the user to reveal if the content is synthetically generated.[9]

Nonetheless, user self-reporting is notoriously unreliable. The 2026 IT Rules mandate platforms to implement "reasonable and proportionate technical measures," including automated detection technologies, to proactively verify the accuracy of user declarations prior to content publication.[10] This creates a complex technological mandate: platforms must build or procure enterprise-grade AI detection systems capable of identifying SGI at the point of ingestion.

2.4.Traceability, Labelling, and the Evaporation of Safe Harbour

To counteract the viral dissemination of misleading SGI, the new regulations establish a stringent traceability and labeling system. Intermediaries must guarantee that SGI is clearly and prominently branded to differentiate it from genuine material. Visual material necessitates distinct visual indicators, whereas synthetic audio must be accompanied by a clear audio disclosure.[11] Moreover, where technically practicable, platforms must incorporate permanent metadata or unique digital identities into the SGI. These digital fingerprints are engineered to endure cross-platform dissemination, allowing law enforcement to trace nefarious content to its source system.[12]

The most stringent aspect of the 2026 IT Rules is the reduction of content moderation timescales. The timeframe for adhering to government or court-mandated takedowns for illegal SGI has been significantly shortened from thirty-six hours to only three hours.[13] In cases of very egregious infractions, like as the unauthorized distribution of intimate synthetic photography or child sexual assault content, platforms have a mere two hours to revoke access.[14]

Noncompliance with these stringent timelines, or neglecting to properly mark SGI, leads to the instant loss of the safe harbour exemption conferred by Section 79 of the Information Technology Act, 2000.[15] Without this protection, platforms incur direct criminal culpability under the Bharatiya Nyaya Sanhita, 2023, and the Information Technology Act, thereby equating the middleman with the original producer of the harmful deepfake.

3.INPUT WITHOUT ACCOUNTABILITY? RETHINKING GOVERNANCE UNDER THE DPDP ACT, 2023

The 2026 IT Rules oversee the downstream output of generative AI, whereas the Digital Personal Data Protection Act, 2023, is ostensibly responsible for regulating the upstream input. This encompasses the computational assimilation of personal data utilized for training basic models. A thorough textual examination of the DPDP Act uncovers significant statutory constraints that diminish its efficacy as a comprehensive AI governance instrument.

3.1.The consent Paradigm and Biometric Ingestion

The foundation of the DPDP Act is its stringent, user-focused consent framework. Section 6 stipulates that consent for the processing of digital personal data must be voluntary, explicit, informed, unequivocal, and accompanied by a clear affirmative action.[16] In the context of generative AI, scraping an individual's biometric markers, facial topography, or voice data to train a voice-cloning or deepfake model without this heightened standard of consent constitutes a profound statutory violation.

Section 6(1) imposes a stringent principle of purpose limitation. A data fiduciary cannot arbitrarily repurpose user data provided for a specific service, such as social networking or e-commerce, for training proprietary AI models without acquiring new, explicit agreement. Additionally, Section 6(6) confers upon the Data Principal the unequivocal right to revoke consent, necessitating the prompt termination of data

[9] *Id.* Rule 4(2).
[10] *Id.*
[11] *Id.* Rule 3(1).
[12] *Id.*
[13] *Id.* Rule 3(2).
[14] *Id.* Rule 3(2)(b).
[15] Information Technology Act, 2000, § 79(2).
[16] Digital Personal Data Protection Act, 2023,§ 6(1).

processing.[17] For AI developers, technically unlearning or deleting specific user data from a pre-trained neural network presents a monumental, if not insurmountable, technological hurdle.

3.2.The section 3(c)(ii) Loophole: The Commodification of Public Data

Notwithstanding the strong safeguards instituted in Section 6, the DPDP Act includes a contentious exemption that effectively subsidizes the data acquisition processes of international AI companies. Section 3(c)(ii) explicitly excludes from the Act any personal data that a Data Principal voluntarily discloses to the public.[18]

This statutory exclusion is often construed by foundational model developers as a comprehensive permit for unregulated online scraping. AI models consistently assimilate billions of publicly available data points, encompassing public social media profiles, open-source blogs, and professional directories. The legislative ambiguity about this phrase engenders a perilous legal fiction. It assumes that a person's choice to publicly share a photograph or textual content implies ongoing consent for algorithmic processing and commercial AI training.

In contrast to the European Union's General Data Protection Regulation (GDPR), which maintains protections for publicly accessible personal data, the Indian framework entirely removes legislative safeguards for such data.[19] Where this exclusion applies, the rights ordinarily available to Data Principals evaporate, creating a massive regulatory vacuum precisely where AI companies source the bulk of their training material.

3.3.Jurisdictional Evaporation: Anonymization and Synthetic Outputs

The regulatory efficacy of the DPDP Act is completely nullified once the training data is transformed into a synthetic output. The Act's authority is exclusively linked to "personal data," which is defined as any information pertaining to an identified individual in regard to such data.[20]

Upon the anonymization of ingested data or the generation of entirely synthetic output by a fundamental model, the DPDP Act is no longer applicable. This results in a recurrent enforcement failure. The preliminary data extraction utilizes the Section 3(c)(ii) public data loophole, but the resultant synthetic product evades the legal definition of personal data. Thus, the fundamental mechanics of generative AI function predominantly beyond the confines of India's principal privacy regulation.

4.BRIDGING REGULATORY GAPS: THE GOVERNANCE TRILEMMA IN CONTEMPORARY LAW

The interaction between the 2026 IT Rules and the DPDP Act establishes a "governance trilemma" wherein the legislation must concurrently handle technological identification, privacy protections, and platform responsibility. The existing divided system results in considerable regulatory deficiencies that compromise legal certainty and constitutional safeguards.

4.1.The Input-Output Friction

A fundamental conflict occurs between the "input" exclusions of the DPDP Act and the "output" mandates of the IT Rules. According to Section 3(c)(ii) of the DPDP Act, personal data disclosed by a user is not subject to protection. Thus, an AI model can utilize a user's public images to train a basic model without infringing upon data protection legislation.

Once the model produces a synthetic output (SGI) utilizing the same data, the 2026 IT Rules impose stringent takedown and labeling obligations on the platform. This establishes a legal dilemma in which the law allows for the development of the capability while imposing severe penalties for its actualization. This segmentation disregards the fact that SGI is an inseparable process encompassing data ingestion and content generation.

4.2.Platform Liability versus Creator Accountability

The 2026 IT Rules perpetuate India's practice of governing the internet via the lens of intermediary responsibility. By threatening the revocation of safe harbor under Section 79 of the IT Act, the state effectively delegates the responsibility of

[17] *Id*. § 6(6).
[18] *Id*. § 3(c)(ii).
[19] Council Regulation 2016/679, art. 9(2)(e), 2016 O.J. (L 119) 1.
[20]Digital Personal Data Protection Act, 2023,§ 2(t).

policing AI to platforms. This imposes an inequitable burden on intermediaries such as social media platforms, which may lack the foundational AI technology.

Developers of foundational models, the entities constructing the generative engines, frequently evade direct liability under the IT Rules as they are not consistently categorized as intermediaries for the particular material disseminated. This establishes a liability disparity wherein the distributor is penalized for a three-hour takedown failure, while the author of the generative tool stays shielded from the repercussions of their system's design.

4.3.The Chilling Effect and Over-Censorship

The stringent three-hour removal requirement for SGI, coupled with the risk of criminal punishment, compels platforms to prioritize censorship over inquiry. Due to the imprecision of technological verification of AI-generated material, platforms are prone to censoring authentic satire, parody, or artistic expression that activates an automatic synthetic content alert. This excessive restriction jeopardizes the free expression principles outlined in Shreya Singhal, when the Supreme Court cautioned against establishing a system that compels intermediaries to adjudicate content.[21]

5.MAPPING AI GOVERNANCE: A COMPARATIVE JURISPRUDENTIAL STUDY OF DIVERGENT MODELS

The international regulatory approach to synthetic content is marked by a fundamental division in legal thought. India's 2026 IT Rules emphasize the governance of the digital "highway" via intermediate responsibilities, whereas other countries have shifted towards "upstream" regulation aimed at the AI systems themselves.

5.1.The EU AI Act: A Risk-Based Product Safety Model

The European Union's Artificial Intelligence Act constitutes the inaugural comprehensive, horizontal legislation on AI globally. In contrast to the Indian model, which largely views AI hazards as failures in content filtering, the EU framework employs a "product safety" approach.[22]

Article 50 of the EU AI Act mandates transparency by design. Providers of generative AI systems must ensure that outputs are labeled in a machine-readable format and identifiable as artificially created. Although India's 2026 IT Rules also require labeling, the EU Act assigns the technical responsibility to the Model Provider instead of solely the Deployer. Additionally, the EU has a stratified risk taxonomy, prohibiting "unacceptable" AI applications and imposing stringent compliance evaluations on "high-risk" systems prior to market entry. In contrast, India's approach is predominantly risk-agnostic, imposing an identical 3-hour takedown obligation on both small startups and giant conglomerates.

5.2.The UK Online Safety Act: The "Duty of Care" Philosophy

The United Kingdom has established a "duty of care" framework with the Online Safety Act 2023. This paradigm redirects attention from individual content removals to the overall resilience of a platform's risk management procedures.[23]

A notable distinction between the UK and Indian frameworks is to the regulation of non-consensual synthetic imaging. India employs the IT Rules and the Bharatiya Nyaya Sanhita to penalize dissemination, but the UK's Data (Use and Access) Act 2025 explicitly criminalizes the creation of sexually explicit deepfakes, irrespective of their distribution.[24] This "zero-tolerance" approach at the point of creation addresses the harm at its source, whereas the Indian 2026 IT Rules only trigger when the content is "uploaded" or "transmitted" on an intermediary's computer resource.

5.3.The "Agile" vs. "Comprehensive" Debate

India's approach is frequently described as "agile" or "techno-legal," leveraging existing statutes like the IT Act and the BNS rather than enacting a sweeping AI code.[25] This adaptability facilitates swift modifications, exemplified by the February 2026 revisions, however also engenders the

[21] *Shreya Singhal v. Union of India*, (2015), MANU/SC/0329/2015.
[22] Council Regulation 2024/1689, 2024 O.J.
[23] Online Safety Act 2023, c. 50 (UK).
[24] Data (Use and Access) Act 2025 (UK).
[25] *AI Governance, the India Way: Using Old Laws, Fixing New Gaps to Shape a Global South Template*, Econ. Times (Feb. 14, 2026), https://economictimes.indiatimes.com/news/company/corporate-trends/ai-governance-the-india-way.

"liability gap" addressed in Section IV. Comparative research indicates that although India's reactive strategy efficiently addresses immediate victim concerns, it lacks the structural accountability for AI developers present in the EU's complete framework.

6. MAKING THE CASE FOR A BESPOKE REGULATORY FRAMEWORK FOR ARTIFICIAL INTELLIGENCE

The above study indicates that India's existing regulatory framework is defined by "governance by proxy." The state has adapted intermediary responsibility regulations and broad data protection legislation to mitigate the hazards associated with synthetic content, rather than implementing a specific AI bill. This "agile" method facilitates swift notifications, such as the 2026 IT Rules, although it does not satisfy the constitutional criterion of proportionality and engenders a perilous accountability void.

6.1.The Failure of Intermediary-Centric Governance

The 2026 IT Rules effectively designate digital intermediaries as the principal enforcement mechanism for AI-related harms. The law imposes a three-hour removal window and requires technological verification for all uploads, thereby transferring the state's enforcement responsibility to private organizations. This presents jurisprudential issues. In contrast to conventional content moderation, which allows a platform to visibly detect explicit violations (such as copyright infringement or hate speech), the identification of SGI is a probabilistic endeavour.

Forcing platforms to act as quasi-judicial arbiters of "authenticity" under the threat of losing Section 79 immunity incentivizes over-censorship. This contradicts the mandate in *Shreya Singhal*[26], which warned that intermediaries should not be required to exercise independent judgment on the legality of content. A framework that regulates AI through the narrow lens of the "messenger" (the platform) ignores the "source" (the AI developer).

6.2.The Accountability Asymmetry

A significant deficiency in the existing system is the isolation of core model developers. The 2026 IT Rules focus on the dissemination of SGI, while the DPDP Act addresses the acquisition of training data; nevertheless, neither statute assigns liability to the developer for the model's design.

If a generative model is constructed without enough protections that facilitate the effortless production of non-consensual photography, the obligation under the 2026 IT Rules predominantly rests with the platform on which the image is disseminated. This establishes an asymmetrical accountability framework wherein the entity possessing the highest technical capability to avert harm at the "prompt level" encounters minimal legal liability. For optimal efficacy, Indian law should adopt a "distributed liability" framework wherein model developers, deployers, and intermediaries share accountability commensurate with their involvement in the content lifecycle.

6.3.The Inadequacy of the DPDP Act's Input Shield

The omission of publicly accessible data under Section 3(c)(ii) of the DPDP Act constitutes a substantial "input shield" for AI enterprises. The rule permits unrestricted scraping of public digital footprints, so establishing a "fair use" exception for AI training that far beyond the scope of the "fair dealing" protections outlined in Section 52 of Indian copyright law.[27]

The legal ambiguity about the "downstream reuse" of public data allows for an individual's digital identity to be assimilated, tokenized, and transformed into a deepfake without any infringement of the DPDP Act during the training phase. The existing approach is inherently reactive; it permits the privacy infringement to be integrated into the model and intervenes solely after the harm (the SGI) has occurred.

7. CONCLUSION

The announcement of the 2026 IT Rules signifies a definitive conclusion to the period of unregulated generative AI in India. By formalizing the SGI category and reducing compliance requirements to three hours, the State has indicated that safety now legally takes precedence over the conventional extensive safeguards of safe harbor. Nonetheless, as this article has illustrated,

[26] *Shreya Singhal v. Union of India*, (2015), MANU/SC/0329/2015.

[27] Copyright Act, 1957,§ 52.

the existing structure is a reactive amalgamation rather than a proactive code of behavior.

The trilemma of regulating synthetic content—balancing technology identification, person privacy under the DPDP Act, and platform accountability—cannot be addressed only through intermediary liability. The inherent conflict between the public data loophole of the DPDP Act and the output demands of the IT Rules fosters an environment where the development of detrimental AI capabilities frequently evades oversight, while their subsequent dissemination is severely punished. This imbalance transfers the entire responsibility of AI governance to digital messengers, rendering basic model developers entirely immune to the repercussions of their design decisions.

For India to establish itself as a global leader in AI jurisprudence, it is essential to migrate from proxy governance to a tailored AI statutory framework. This methodology must transcend the limited scope of Section 79 of the IT Act and implement a risk-based strategy that distinguishes between innocuous creative expression and significant synthetic damages. The legislative inaction regarding the downstream reuse of publicly accessible personal data must be rectified to maintain the significance of the DPDP Act's consent framework in an era of automated data scraping.

The primary objective of Indian AI legislation should be to cultivate an environment in which innovation is not compromised by informational privacy or constitutional due process. Until a comprehensive rule is implemented that mandates safety by design for AI developers, the responsibility for regulating synthetic content will continue to be inherently unequal, operationally challenging, and legally inadequate.