

To,
The Oversight Board

Date: 26 February, 2026

IFF/2026/004

Subject: Submission by IFF on the Case of Non-Consensual AI Sexualized Impersonation

1. About us

The Internet Freedom Foundation (“IFF”) is a registered charitable trust based in New Delhi, India that advances constitutional freedoms for every Indian in a digital society. We work across a wide spectrum of issues, with expertise in free speech, electronic surveillance, data protection, and net neutrality. We aim to champion privacy protections, digital security, and individual freedoms in the digital age.

2. About the case

We are writing to submit public comments as solicited by the Oversight Board (“Board”) dated February 12, 2026 where the Board announced a new case for consideration wherein it would be addressing non-consensual Artificial Intelligence (“AI”) sexualized impersonation (2026-021-IG-RA). The case included an AI-generated non-consensual sexualized video an Instagram user posted a short video of a woman in a form-fitting dress in which the woman is adjusting her dress and moving her body, with her underwear visible in a few frames. The case was appealed by a user who stated that the video in question was impersonating one of their friends, without their consent thereby damaging their friend’s reputation.

3. Our Reservations about the Oversight Board

At the outset, we would like to clearly state that while we are choosing to engage with the Board as we are committed to working towards a free, open and rights respecting internet especially for vulnerable communities. Our engagement comes with reservations against the structure and the function of the Board which are laid out here in the spirit of transparency. These criticisms have been echoed by academics and civil society organizations in the past include the limited jurisdiction of the Board, concerns regarding the non-binding nature of the policy recommendations of the Board, the caveats of



“technical and operational feasibility”¹ regarding the application of the Board’s past decisions to “identical content in parallel context”, to name a few.² However, despite these concerns, and recognising the significance of the Board’s decision and widespread use of META platforms in India, we hope to engage with the Board through the medium of public comments in a good faith manner.

4. Significance for India and Indian Users

Before engaging with the primary issues related to the case, it is necessary to highlight the significance of Meta’s Policies and the Board’s decisions, particularly in the Indian context. As per Statista, an online platform that provides market insights, India has the “largest Facebook audience worldwide, with over 403 million users” as of October 2025.³ Further, India also had the “largest Instagram audience worldwide, with 480.55 million users”, as of October 2025.⁴ Indian users are also one of the most active users of Threads and Meta’s platforms exercise huge influence in the domain of digital advertising as well⁵. These statistics clearly indicate that Meta and its platforms occupy a dominant position in India, wield great influence, and have the capacity to deeply impact Indian users. Therefore, decisions of the Board, especially with respect to sensitive issues such as AI-driven harms, hold a unique interest for Indian users, legislators, policy-makers, and civil society.

5. Background on Non-consensual Intimate Imagery and AI-driven Harms

AI-driven harms, particularly, the creation of non-consensual intimate imagery (NCII) and its subsequent dissemination using online social media platforms disproportionately affects women and children. These harms cannot be undermined because they occur in

¹ Article 4, Implementation, Oversight Board Charter, available at: <https://www.oversightboard.com/wp-content/uploads/2025/07/Oversight-Board-Charter-June-2025.pdf> (Last accessed: February 23, 2026).

² David Wong and Luciano Floridi, Meta’s Oversight Board: A Review and Critical Assessment, October 24, 2022, available at: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4255817 (Last accessed: February 23, 2026).

³ Leading countries based on Facebook audience size as of October 2025, available at: <https://www.statista.com/statistics/268136/top-15-countries-based-on-number-of-facebook-users/#:~:text=Countries%20with%20the%20most%20Facebook%20users%202025&text=As%20of%20October%202025%2C%20India,United%20States%2C%20Indonesia%20and%20Brazil> (Last accessed: February 24, 2026).

⁴ Leading countries based on Instagram audience size as of October 2025, available at: https://www.statista.com/statistics/578364/countries-with-most-instagram-users/?srsltid=AfmBOop5gNef92oJWp7C_BMp41dDfNK2EeAOE4_8mDN9HjWIQKNb6YHC (Last accessed: February 24, 2026).

⁵ Suneera Tandon, Meta’s short videos are reshaping how India shops, Live Mint, June 12, 2025, available at: <https://www.livemint.com/companies/meta-india-retail-report-instagram-reels-shopping-india-facebook-product-discovery-digital-advertising-trends-india-11749701482644.html> (Last accessed: February 24, 2026).



virtual spaces as they extensively infringe upon the privacy, dignity, sexual autonomy, and freedom of expression of victims, and cause significant psychological and emotional distress.⁶ NCII also has an overarching chilling effect and negatively impacts women's visibility and participation in digital spaces.⁷ This in turn discourages their involvement in political, social, and economic life, thereby deepening the existing gendered digital divide. A lack of serious action in combating AI also perpetuates collective harms such as the normalization of NCII and contributes towards a culture where gender-based violence against women and children is condoned.⁸

Harms perpetuated by Generative AI (GenAI) including the creation of non-consensual sexually explicit images have already been flagged by regulators of several jurisdictions.⁹ Safeguards against such harms can be found in numerous international human rights instruments as well, such as the Universal Declaration of Human Rights, 1948 ("UDHR")¹⁰, the International Covenant on Civil and Political Rights, 1966 ("ICCPR")¹¹, and the International Covenant on Economic, Social and Cultural Rights ("ICESCR")¹². For instance, Article 12 of the UDHR protects the right to privacy and arbitrary interference with an individual's "privacy, family, home or correspondence", and also protects against any attacks upon an individual's honour and reputation. Article 12(1) of the ICESCR recognizes the "right of everyone to the enjoyment of the highest attainable standard of physical and mental health".

⁶ Avanti Deshpande and Jhanvi Anam, Grok runs amok: Understanding the repercussions of AI-driven Sexual Abuse, Internet Freedom Foundation, available at: <https://internetfreedom.in/grok-runs-amok-understanding-the-repercussions-of-ai-driven-sexual-abuse/> (Last accessed: February 24, 2026).

⁷ *ibid.*

⁸ Oversight Board Public Comment, Explicit AI Image Cases, Digital Speech Lab, UCL and Sussex Law and Technology Group, University of Sussex, available at: <https://osbcontent.s3.eu-west-1.amazonaws.com/PC-27065.pdf> (Last accessed: February 24, 2026).

⁹ Grok obscene AI content: govt. gives X time till January 7 to submit report, The Hindu, January 7, 2026, <https://www.thehindu.com/news/national/grok-obscene-ai-content-govt-gives-x-time-till-january-7-to-submit-report/article70478927.ece> (Last accessed: February 23, 2026); Dan Milmo, Helena Horton and Kiran Stacey, UK media regulator investigating Elon Musk's X after outcry over sexualised AI images, The Guardian, January 12, 2026, <https://www.theguardian.com/technology/2026/jan/12/ofcom-investigating-x-outcry-sexualised-ai-images-grok-elon-musk> (Last accessed: February 23, 2026).

¹⁰ Universal Declaration of Human Rights, 1948, <https://www.un.org/en/about-us/universal-declaration-of-human-rights>

¹¹ International Covenant on Civil and Political Rights, 1966, <https://www.ohchr.org/en/instruments-mechanisms/instruments/international-covenant-civil-and-political-rights>.

¹² International Covenant on Economic, Social and Cultural Rights, <https://www.ohchr.org/en/instruments-mechanisms/instruments/international-covenant-economic-social-and-cultural-rights>.



Further, the right to freedom of opinion and expression is protected under Article 19 of the ICCPR and Article 19 of the UDHR. The Report of the Special Rapporteur on violence against women, its causes and consequences in 2018 acknowledged the central role that intermediaries play in providing digital spaces for interaction and their specific human rights responsibilities.¹³ The recommendations made by the Special Rapporteur included *inter alia*, the adoption of a transparent complaint mechanisms for cases of online and technology-facilitated violence against women and girls, easily accessible and transparent policies and procedures for reporting and requesting the removal of harmful content, and a comprehensive content moderation policy and human rights safeguards against arbitrary censorship, and transparent reviews and appeal processes.¹⁴ The Special Rapporteur also recommended that intermediaries should provide terms of service and reporting tools in local languages and that reporting tools should be “accessible, user-friendly and easy to find”.¹⁵

6. Previous Case involving ‘Explicit AI Images of Female Public Figures’

The Board has previously adjudicated upon a case involving ‘Explicit AI Images of Female Public Figures’ in 2024 where the Board deliberated two cases that involved female public figures from India and United States. The Board recognized that restrictions on such content were “legitimate to protect individuals from the creation and dissemination of sexual images made without their consent”. In both cases, the Board found that the images violated Meta’s rule that prohibits “derogatory sexualized photoshop” under Meta’s Bullying and Harassment Policy and overturned Meta’s decision to leave the post up in the first case (in India) and the upheld Meta’s decision to take down the post in the second case (in the United States).

The Board further recommended that Meta, a) move the prohibition on “derogatory sexualized photoshop” into the Adult Sexual Exploitation Community Standard; b) change the word “derogatory” in the prohibition on “derogatory sexualized photoshop” to “non-consensual”; and c) replace the the word “photoshop” in the prohibition on “derogatory sexualized photoshop” with a more generalized term for manipulated media; and d) that Meta harmonize its policies on non-consensual content by adding a new signal for lack of consent in the Adult Sexual Exploitation policy: context that content is AI-generated or manipulated.

¹³ Report of the Special Rapporteur on violence against women, its causes and consequences on online violence against women and girls from a human rights perspective, Human Rights, Council, A/HRC/38/47, 18 June 2018, available at: <https://digitallibrary.un.org/record/1641160?ln=en&v=pdf> (Last accessed: February 25, 2026).

¹⁴ *ibid.*

¹⁵ Report of the Special Rapporteur (n10).



This case thus highlights the Board’s previous experience in AI driven non-consensual intimate images (“NCII”) and indicates the Board’s acknowledgement that changes in Meta’s policies are needed to appropriately address the harms being perpetuated by images/content generated by AI. The Board also recognized that due to the severity of the harms, “removing the content was the only effective way to protect the people impacted”.

7. India’s Legal Vacuum on Non-consensual Intimate Imagery (NCII)

At present, there is no specific legal provision that applies to the creation and dissemination of non-consensual intimate imagery (NCII). As a result of this legal vacuum, the existing framework governing intermediary liability has been applied to tackle the issue, in the context of seeking accountability from platforms.

Under the criminal law framework, while there is no specific law against a NCII violation, several provisions of the Indian Penal Code, 1860 (IPC) (sections 292, 354C, 499, 509) (now the and the Information and Technology Act, 2000 (IT Act) (sections 66E, 67A)¹⁶ are applicable to the dissemination of NCII. Section 66E of the IT Act criminalises the intentional capture, publishing, or transmission of an image of a person’s private area without their consent. Further, rule 3(2)(b) of the Information Technology (Intermediary Guidelines and Digital Media Ethics Code) Rules, 2021 (IT Rules) which have recently been amended on February 10, 2026¹⁷, provides for the grievance redressal mechanism for an intermediary upon receipt of a complaint, but fails to acknowledge the aspect of lack of consent. The 2026 Amendment Rules also add Section 2(1)(wa) to the 2021 IT Rules, to define “synthetically generated information”.

In accordance with the earlier version of these rules, the broad scope of the definition still includes all audio, visual, or audio-visual information which is “artificially or algorithmically created, generated, modified or altered using a computer resource” in a manner that such information appears to be “real, authentic or true”. The definition captures all audio, visual, or audio-visual information that appears to portray individuals or events in a manner that is “likely to be perceived as indistinguishable from a natural person or real-world event”.¹⁸ While the amendment includes a proviso which exempts certain innocuous usage of AI such as routine, good-faith editing etc, at present, the

¹⁶ The Information and Technology Act, 2000, available at: https://www.indiacode.nic.in/bitstream/123456789/13116/1/it_act_2000_updated.pdf.

¹⁷ The Information Technology (Intermediary Guidelines and Digital Media Ethics Code) Rules, 2021,

<https://drive.google.com/file/d/1p0zkFFA92qRmaC7GeEKLFikj0YKdSIOX/view?usp=sharing>.

¹⁸ Indumugi C. and Jhanvi Anam, IT Intermediary Amendment Rules, 2026 contradict their purpose, Internet Freedom Foundation, February 11, 2026, available at: <https://internetfreedom.in/it-intermediary-amendment-rules-2026-contradict-their-purpose/>



definition of synthetically generated information will now include deepfakes, AI-generated or AI-altered images and videos, voice cloning, and other forms of realistic synthetic audio-visual content.¹⁹ The core due diligence obligations of online intermediaries will now include a) ensuring that there is no generation or publication of high-risk unlawful synthetically generated information using their softwares; b) the mandatory labelling of all synthetically generated information; and c) ensuring that the label, permanent metadata, or unique identifier cannot be modified, suppressed, or removed. A detailed explainer of the amendment and its implications published by IFF is available [here](#). IFF had also made a detailed submission on the draft version of the amendment to the IT Rules where it had detailed its reservations about these changes.²⁰

Further, the issue of NCII content has garnered attention from the judiciary as well. The Hon'ble High Court of Delhi, in a case involving NCII, titled, *Mrs X v. Union of India and Ors.*²¹, ordered online intermediaries to proactively take down NCII content and admonished platforms for their lackadaisical approach due to which the victim had been forced to approach the court each time the content in question was re-uploaded online.

The absence of a specific provision in this regard effectively leaves victims with little choice but to pursue the usual criminal process to seek recourse for NCII. Thus, options for redressal would typically entail filing a First Information Report (FIR) and/or filing a complaint with the National Cyber Crime Reporting Portal. Both these systems do not guarantee swift takedown of the content in question which is typically a victim's priority at this point. Therefore, due to such a legal vacuum, it is even more important that the redressal mechanism provided by the online platform is easily accessible, victim-friendly and rights respecting, and effective in taking down the content with adequate safeguards and ensuring that it is not disseminated further.

8. Need for Change in Meta's Adult Nudity and Sexual Activity Policy

At present, Meta's Adult Nudity and Sexual Activity Policy ("ANSA Policy") is narrowly framed and circumspect in nature. Due to its framing, the ANSA Policy covers a limited remit of *inter alia*, photorealistic/digital images of genitalia, explicit and implicit sexual

¹⁹ *ibid.*

²⁰ Apar Gupta and Naman Kumar, Withdraw the Draft Synthetic Information IT Rules, 2025, Internet Freedom Foundation, November 15, 2025, available at: <https://internetfreedom.in/withdraw-the-draft-synthetic-information-it-rules-2025/> (Last accessed: February 25, 2026).

²¹ W.P. (CrI). 1505/2021, available at: <https://indiankanoon.org/doc/105980506/>. IFF's analysis: Radhika Roy, Delhi HC issues directions to search engines and other authorities for dealing with the dissemination of Non-Consensual Intimate Images (NCII), Internet Freedom Foundation, May 2, 2023, available at: <https://internetfreedom.in/delhi-hc-issues-directions-to-search-engines-and-other-authorities-for-dealing-with-the-dissemination-of-non-consensual-intimate-images-ncii/> (Last accessed: February 25, 2026).



activity or stimulation under the heads of Adult Nudity and Adult Sexual Activity, respectively.

Briefly put, we submit that the Policy's three major shortfalls as it stands at present are its failure to:

- a) Understand and capture the harms of non-consensual images that are sexualised without containing sexually explicit images of *inter alia* genitalia, breasts, buttocks, etc.
- b) Anchoring in consent of the individual(s) depicted in the image(s)/video(s) in question.
- c) Recognise exceptions for journalism, parody, political speech and satire of public officials and persons.

A definition that is solely reliant on "nudity" and "sexual activity" effectively excludes a large portion of images/videos that while not depicting explicitly sexual genitalia may be of a highly sexualised and be an impersonation, as in the present instance, or be derogatory/degrading, violent or otherwise dehumanizing and humiliating in nature. A failure to account for such sexualised images which while not containing anything that can be deemed to be sexually explicit but carries sexual connotations is precisely why the present case was not considered to be a violation of the ANSA Policy.

The reasoning provided by Meta's subject matter expert reviewers for their decision was that the video in question did not contain any "visual representations of a sexual encounter" or "language that facilitates or encourages sexual encounter", and was therefore not removed. A change to the ANSA Policy that broadens it to include non-consensual sexualised content, irrespective of whether it is AI generated or not, would better serve the interests of victims who are facing issues with ensuring that manipulated content uploaded without their consent is permanently taken down from online platforms. Further, any takedown must be accompanied by an evaluation of its impact on free expression in protected categories of speech such as journalism, parody, political speech and satire of public officials and persons.

9. Our Recommendations

With this background, IFF would like to set out certain recommendations to Meta's ANSA Policy to ensure a more victim-friendly, trauma-informed process involving NCII content. Our recommendations are as follows:

- a) We recommend that Meta update its existing ANSA Policy to include sexualised images uploaded without consent apart from merely sexually explicit images of



genitalia as covered under the Adult Nudity and Adult Sexual Activity sub-heads of the ANSA Policy. This must be through the use of precise language that respects free expression.

- b) We recommend that Meta create a separate category under the reporting mechanism for NCII content. This can be done by creating a new category for reporting that is anchored in consent, and can be used by victims to report content irrespective of whether the content in question is AI generated/AI manipulated or not. This would distinguish NCII from being reported under the current category of, “pornography” as was done in the present case where two users reported the content for pornography, but the content continued to remain on Instagram.
- c) We recommend that any change to the ANSA Policy be rooted in a trauma-informed and victim-centric response. In order to facilitate this, we recommend: a) a graduated, time-bound response from Meta’s side once it receives a report of NCII content; b) temporary removal of the content till a human review can take place to ensure immediate relief for the victim; c) prioritisation of a human review of the content in question; d) usage of technology such as hash matching technology to ensure that the reported content in question can be taken down swiftly if it is re-uploaded, disseminated, shared again on any of Meta’s platforms.
- d) Given the highly sensitive nature of NCII content and the lasting repercussions it can have on victims, we recommend that there exist a presumption in favour of the victim leading to immediate take down of content upon it being reported, subject to human review to ensure safeguards with respect to rights guaranteeing the freedom of speech and expression, as they exist in various jurisdictions.
- e) We are also cognizant of ensuring the freedom of expression in online spaces at a time where the world is witnessing democratic backsliding in several countries. Therefore, any changes to the ANSA Policy must contain safeguards to ensure that such a Policy is not weaponised against political expression, parody, satire, and journalism. Further, we recommend that such takedown be accompanied by an ‘Evaluation’ which provides the platform’s reasoning and explanation for taking down the content in question.
- f) Finally, while it is important for such takedowns to be swift to alleviate further harm to the victims, it is also necessary to provide recourse to contest such removal. Therefore, we recommend Meta to provide an easily accessible process of Appeal to contest the takedown of content. Such mechanisms are necessary



in addressing concerns related to censorship and safeguarding the freedom of speech and expression while also ensuring that victims receive prompt support.

We welcome the Oversight's Board interest in addressing this important issue and reiterate our commitment to combatting technology-facilitated gender-based violence (TGBV) and ensuring a safer internet for everyone. We are happy to provide further support based on our work and experience in India.

Kind regards,

A handwritten signature in blue ink that reads "Apar Gupta".

Apar Gupta,
Founder Director
Internet Freedom Foundation