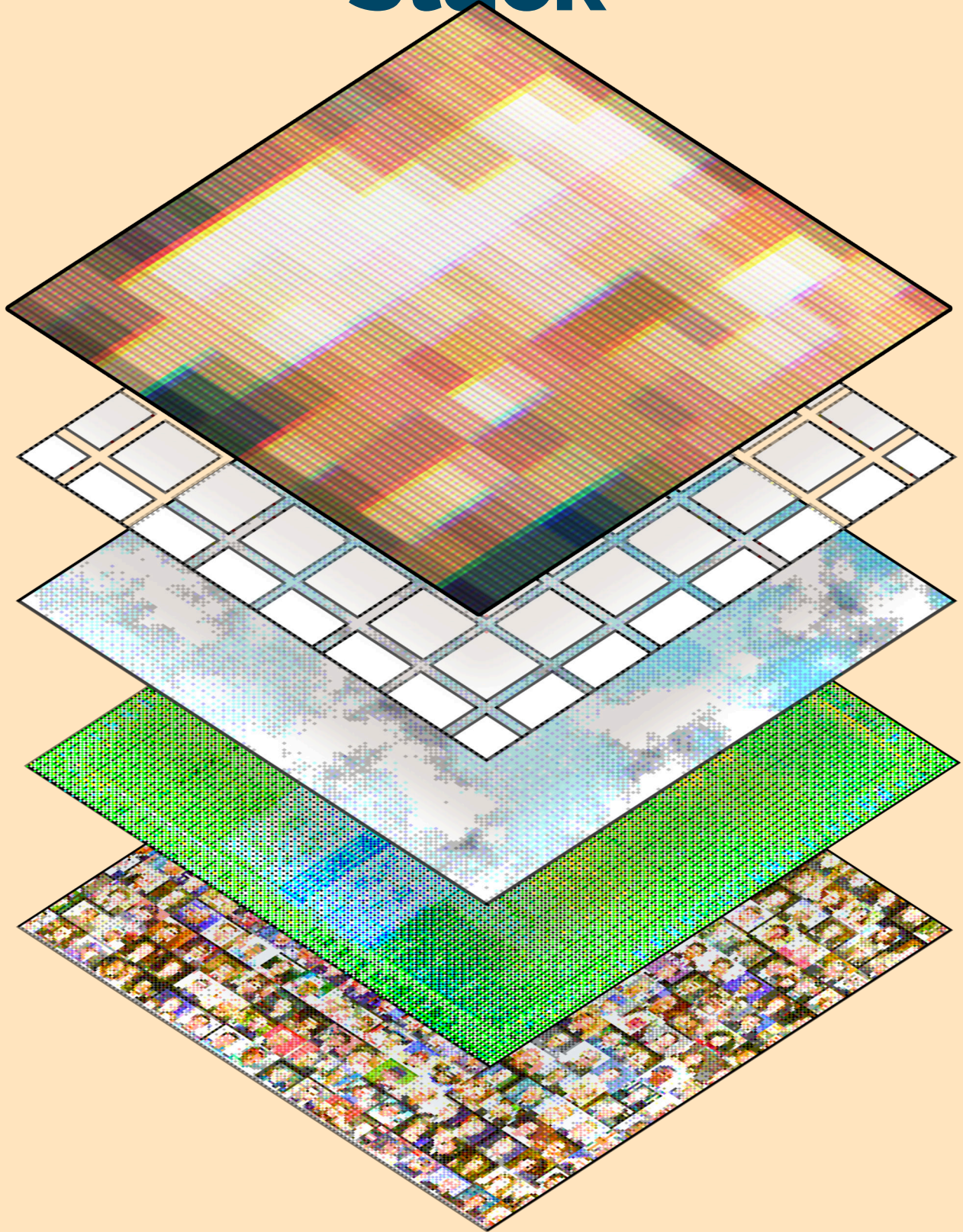


Deconstruct the Deepfake Stack



An Action Plan To Tackle
Non-Consensual Synthetic Imagery In India

Deconstruct the Deepfake Stack: An Action Plan to Tackle NCSI in India

Date:

12 October 2026

Authors

Tarunima Prabhakar, Kaustubha Kalidindi & Adrija Bose

License

The text and visuals in this document are licensed under CC BY SA 4.0. The report may be cited as "Prabhakar, Kalidindi, Bose. Deconstructing the Stack: An Action Plan for NCSI. Tattle Civic Technologies. 12 October 2026"

This report builds on the investigation “**Desi Faces, Foreign Servers: Inside the AI Economy Undressing Women**” conducted by Decode at Boom and Tattle Civic Technologies on the infrastructure that enables people to use AI to manipulate images of women without their consent. The journalistic investigation described the phenomenon and the real world harm from the use of AI. This policy report focuses on the interventions that can be taken to prevent the use of AI for creating and distributing non-consensual synthetic imagery.

The following people played a key role in supporting the investigation:

- Aatman Vaidya
- Denny George
- Poorvi Gupta
- Swair Shah
- Yogesh Girikumar

We gratefully acknowledge the valuable insights provided by the following individuals, whose expertise and feedback greatly strengthened this work: Vijay Keswani, Hamza Farooqui, Kundan Mishra, Krishnesh Bapat, Srishti Joshi, Soumya S, Rattanmeek Kaur, Vishakha Deshwal, Divya Chandra, Hera Rizwan, Aishani Jaiswal, Rhydhi Gupta, Sejal Patel, Shruti Dhapola, Pamposh Raina, Abhilash Mallick, Vaishnavi Sundararajan, Vasudev Devadasan, Akshit, Ashlesh Biradar, Arjun Bhagoji, Prateek Waghre.

This report is supported by the
Pulitzer Center AI Accountability Grant.

Design

Yatharth ([Strangerobot.Design](#))

Glossary	5
Introduction	7
NCSI's Multi-layered Infrastructure	10
Salient Findings from the Investigation	13
Legal Foundation For Action Against NCSI in India	16
1. Image as Extension of Personhood	17
2. Image as Intellectual property	18
3. Image as Societal Amorality	18
Interventions	19
Strategy 1: Clarify Applicability of Laws to Model Training	19
Strategy 3: Prevent Generation of NCSI	24
Strategy 4: Prevent Monetisation	27
Discussion	30

Actual knowledge: As per the Supreme Court's judgment in *Shreya Singhal v Union of India* (2015), an intermediary must comply with a takedown request pursuant to receipt of a court order or a government notification. This constitutes 'actual knowledge'.

API: An application programming interface (API) enables different software applications to communicate with each other. In the context of this report, an API enables a software developer to build an application using a synthetic image generation model hosted by another entity.

Blocking: Under 69A of the IT Act, the Central Government may issue an order to an agency of the Government or an intermediary to block public access to any 'information' generated, transmitted, received, stored, or hosted in any computer resource.

CSAM: Child sexual abuse material. This content is globally banned, and under Indian law, this content cannot be hosted, stored or disseminated in any form.

Datasets: A dataset is a collection of data points organised in a structured manner that enables computer algorithms to read, analyse and make inferences.

Deepfakes: A catchall term used to describe images, photos or videos digitally altered or generated using AI to create a falsified depiction of an individual.

Domain Name System (DNS): A server which stores DNS records and converts domain names into IP addresses.

Extra-territorial jurisdiction: Power under law which allows courts to adjudicate on matters that may have taken place outside of India, or have been committed by persons outside India.

Graphical User Interface: An easy-to-use modality through which non-technical users can upload images to generate synthetic content (ex: apps, websites, bots).

IT Act: Information Technology Act, 2000. This law governs electronic commerce, intermediaries, and sets penalties for cyber-offences.

IT Rules, 2021: Information Technology (Intermediary Guidelines and Digital Media Ethics Code), 2021 which sets out due diligence obligations on intermediaries and publishers.

IT Rules (Proposed Amendment), March 2026: The proposed amendments to the IT Rules, 2021.

Machine Learning (ML) Model: Machine learning models are computational systems that learn patterns from large datasets to predict or create content. In the context of this report, we are concerned with models that produce images or videos as their outputs. Some base media models are fine-tuned to create specific sexual scenes.

Model Access Provider: Entity through which more technical users can access models and build on top of AI models. Providers might also support model downloads, gamification and monetisation features.

Non-consensual Intimate Imagery (NCII): Images of a person depicted in an explicit fashion including through falsification such as morphing, editing etc.

Poisoning: Dataset poisoning is an activity where one tampers with the dataset used to train an AI model to influence the performance of the model.

Right to be Forgotten: A right recognised through judicial precedent as part of Art. 21 of the Constitution. Courts have permitted the delisting of search results or masking of names on a case-to-case basis on the grounds of protecting this right.

Sahyog Portal: A portal set up by the Government of India, which automates takedown notices to intermediaries under s.79 of the IT Act.

Synthetically Generated Information (SGI): Audio, visual, or audio-visual information, which has been artificially/algorithmically to depict information or individuals in a manner that leads one to believe that they are indistinguishable from real events and persons. Formatting, photo editing, PDFs etc. are not included.

Takedown: Refers to the removal of content by an intermediary hosting content to retain safe harbour under s.79 of the IT Act.

Decode at Boom and Tattle Civic Tech partnered on an investigation to understand the infrastructure that enables the use of AI for creating non consensual content.¹ The story published on Decode in June 2026 followed the case of an engineer who created a deepfake influencer 'Babydoll Archi' using his ex-girlfriend's face, ran multiple accounts, and made money from them.² Unpacking the story revealed layers of an infrastructure that allowed an ex-boyfriend to use a woman's images and create a persona. The woman receded from public life. The case was closed, but the machinery still chugs away as the survivor and law enforcement grapple with understanding and responding to this harm.

The case epitomises a worrying trend of using advancements in AI to depict women, without their consent, in scenes that have never occurred. While some of the manipulations are directed to shame and harass those targeted, in many cases people are unaware that their AI morphs are being used and monetised. In India, frontline data shows AI-generated sexual imagery is a common and growing form of harassment. Of the digitally manipulated content reported to Rati Foundation's Meri Trustline since 2022, most involved AI, and 92 percent of the AI cases targeted women and girls.³ The challenges in responding to non consensual intimate imagery (NCII) – complex and opaque reporting requirements, onerous platform-level reporting for cross-platform spread, tracking cross-border action, and institutional delays⁴ – are compounded by AI-enabled manipulations. Even if the content is fake, those targeted by the videos and images experience a loss of autonomy⁵, and a feeling of 'having done something dirty'.⁶ Experiences like these drive individuals to recede from online spaces, and cause considerable harm, including job loss, reputational damage, and social isolation.⁷

In 2023, the spread of deepfake content online received particular attention and furore when a 'faceswap' video of a prominent

1 Desi Faces, Foreign Servers: Inside the AI Economy Undressing Women," Decode, June 3, 2026, <https://www.decodeinternet.in/life/investigation/desi-faces-foreign-servers-inside-the-ai-economy-undressing-women-31691>

2 Bose, Desi Faces, Foreign Servers.

3 Pillai, Siddharth, Tarunima Prabhakar, and Kaustubha Kalidindi. Make It Real: Mapping AI-Facilitated Gendered Harm.(Rati Foundation and Tattle Civic Tech, 2025).<https://tattle.co.in/make-it-real-report.pdf>.

4 Qiwei, Li, Katelyn Kennon, Nicole Bedera, Asia A. Eaton, Eric Gilbert, and Sarita Schoenebeck. 'Platforms as Crime Scene, Judge, and Jury: How Victim-Survivors of Non-Consensual Intimate Imagery Report Abuse Online'. Preprint, arXiv, 15 December 2025. <https://doi.org/10.48550/arXiv.2512.13500>.

5 "Experiencing Technology-Facilitated Gender-Based Violence in India". Equality Now and Breakthrough Trust, November 5, 2025. <https://equalitynow.org/resource/reports/experiencing-technology-facilitated-gender-based-violence-in-india/>

6 Pillai, Prabhakar and Kalidindi, Make It Real

7 Mrs X v. Union of India, 2023 SCC OnLine Del 2361 (Delhi HC, April 2026, 2023).

celebrity was circulated.⁸ Following the incident, the Ministry of Electronics and Information Technology (MeitY) in India issued an advisory stating that such content must be taken down by platforms within 24 hours.⁹ In late 2025, draft rules to the IT Act, 2000 were released mandating takedown requirements for synthetically generated information (SGI), which were later notified earlier this year- they mandate labelling of SGI and shortened timelines for platform takedown to 2 hours.¹⁰

Despite the signalling of intent, and the regulatory interventions by the Indian government, the non-consensually created synthetic imagery (NCSI) continues to grow on the global and Indian internet. The current approach faces a 'whack-a-mole' problem where one piece of harmful content is taken down and another resurfaces.¹¹ In this report we present the more expansive set of interventions that regulators, tech companies as well as individuals need to take across the different layers of NCSI generation infrastructure to be effective in addressing NCSI. Such infrastructural mapping of NCSI moves beyond individual culpability, and also spotlights the accountability of other participants in the infrastructure.

This report builds on the journalistic investigation, desk research and contributions from an in-person workshop that Decode and Tattle hosted in June, 2026, that convened experts in tech, law and humanities to brainstorm interventions across the value chain.

A note on terminology:

The content category that we are concerned with in this report has been referred to by a number of terms such as gendered deepfakes, synthetic intimate imagery and AI generated non consensual intimate imagery (AIG-NCII). In this report we use the term non-consensual synthetic imagery (NCSI). This emerges from our research which shows that often the manipulated content need not contain nudity or sexual acts for it to be harmful- in the case of 'Babydoll Archi', many videos and images showed the persona doing everyday things. Without containing any scenes of intimacy, content can attack the dignity of the women depicted in the manipulated media through broader implied context. The more important qualifier is that to generate a manipulated video, someone's images and videos are used without their consent. Since the manipulations can be of many kinds- swapping faces on an existing video or generating entirely new images or videos- we use the term synthetic content as an umbrella term covering all forms of manipulations relying on deep learning techniques. Most public

8 Meryl Sebastian, "Rashmika Mandanna Calls for Action against 'Scary' Deepfake Video," BBC News, November 7, 2023, <https://www.bbc.com/news/world-asia-india-67305557>

9 "MeitY Reiterates Rules for Social Media Companies after Deepfake Controversy of Actress Rashmika," ANI News, November 7, 2023, <https://www.aninews.in/news/national/general-news/meity-reiterates-rules-for-social-media-companies-after-deep-fake-controversy-of-actress-rashmika20231107141943/>

10 Ministry of Electronics and Information Technology (India), Information Technology (Intermediary Guidelines and Digital Media Ethics Code) Amendment Rules, 2026, G.S.R. 120 (E), Gazette of India: Extraordinary, pt.II, sec. 3(ii), no.114, February 10, 2026.

11 Ding, Michelle L., Harini Suresh, and Suresh Venkatasubramanian. "How to Stop Playing Whack-a-Mole: Mapping the Ecosystem of Technologies Facilitating AI-Generated Non-Consensual Intimate Images." Preprint, arXiv, revised May 15, 2026. <https://doi.org/10.48550/arXiv.2602.04759>.

evidence shows NCSI directed towards women and girls, but a few cases of it being used to target men and boys have also emerged.¹² NCSI is another form of image-based abuse and tech-facilitated gender-based violence (TFGBV).

NCSI as an Emergent Harm of the Current Paradigm of AI:

This report focuses on the infrastructure that enables the creation and manipulation of a person's images and videos, without their consent, using AI. This infrastructure is not distinct from the "wide variety of pervasive and enabling network resources" spanning hardware devices, software applications and communication protocols that enable the training of general-purpose models and building of AI products.¹³ Every aspect of the AI infrastructure is also implicated in the creation of non-consensual synthetic imagery. One perspective looks at NCSI as a negative use of AI as a tool. From another perspective, NCSI is an inevitable outcome of the current paradigm of AI development that relies on data of unknown provenance, venture-capital-based funding subsidising development costs, and model performance defined by proximity to reality. In permitting training models on data of unknown provenance, we concede the possibility of models using non-consensually collected imagery such as that on adult content websites or people's social media profiles. VC based funding models lead to AI applications that prioritize user growth at all costs, including production of non-consensual imagery. Finally, there is uncritical acceptance that AI development must imply better mimicry of human reality. Mimicking human appearance, behaviour and experience is viewed as a benefit that is weighed against the cost of some of the applications producing non-consensual imagery. We recognise the need for reimagining AI from the ground up, into one that does not rely on ambiguous data access terms, unsustainable resource usage and a unidimensional chase towards approximating human reality. In such reimaginings we might not have to worry about NCSI as the shadow twin of visual AI development.

This report, while recognising the need for this long-term project, is written from the perspective of actions that regulators, AI companies and internet users can take immediately to address the deluge of NCSI that the current paradigm enables. These recommendations are meant to be actionable in the short term.

13 Geoffrey C. Bowker et al., "Toward Information Infrastructure Studies: Ways of Knowing in a Networked Environment," in *International Handbook of Internet Research*, ed. Jeremy Hunsinger, Lisbeth Klastrup, and Matthew Allen (Dordrecht: Springer, 2009), 97–117, https://doi.org/10.1007/978-1-4020-9789-8_5.

NCSI's Multi-layered Infrastructure

For analytical purposes we map the infrastructure into the following six layers:

Datasets

Datasets are the base layer of any AI model. All models, including text-to-image and text-to-video, are trained on large quantities of web data. Several research studies have documented that the popular datasets used in model training, such as LAION, contain non-consensual imagery and child sexual abuse material.^{14,15} No matter how complex or large the model, it produces outputs based on the data it was trained on.

Machine Learning Models

Models can refer to two things:

- The base foundation models, which are trained on large datasets require large amounts of compute.¹⁶ Only a handful of entities in the world have the resources to train the foundation models.
- Fine-tuned models are adapted from base foundation models using techniques like Low-Rank Adaptation (LoRA). Fine-tuning can be done by individuals on powerful personal computers.

Machine learning models lie on a spectrum of openness. Fully proprietary or closed-source models are those where no aspect of the training chain— datasets, training code or model weights— is public. Access to the model is controlled by the entity that developed it. Open-weight models are more transparent and allow for greater investigation and adaptability for different use cases, and help prevent market concentration. Because the weights are public, their default behaviour can be changed through fine-tuning.

Model Access Providers:

¹⁴ Thiel, David. "Identifying and Eliminating CSAM in Generative ML Training Data and Models." Stanford, CA: Stanford Internet Observatory, December 20, 2023. <https://doi.org/10.25740/kh752sm9123>.

¹⁵ Abeba Birhane, Vinay Uday Prabhu, and Emmanuel Kahembwe, "Multimodal Datasets: Misogyny, Pornography, and Malignant Stereotypes" (preprint, arXiv, October 5, 2021), <https://doi.org/10.48550/arXiv.2110.01963>.

¹⁶ Compute refers to graphical processing units (GPUs) that are needed to train models

Model access providers provide an interface to discover, test and modify the models. These could be for-profit platforms such as Hugging Face, GitHub and Civitai or individuals or anonymous outfits providing an API for other developers to build applications using a model they have hosted. Model access providers are primarily targeted towards technically sophisticated users who can tinker with models and/ or build software applications. For proprietary models, the model access providers and model developers are the same.

Graphical User Interface:

A graphical user interface is an accessible platform that enables users to upload an image and/or write a text prompt to manipulate the original to get a modified output. Graphical user interfaces abstract some of the complexity of model access providers making them easier to use for the average internet user. These could be mobile phone apps, Telegram bots, websites or platforms like X that make it easy to interact with the Grok model.

Discovery Channels:

Discovery channels are the entry point through which people interested in AI services for specific tasks, such as creating nudes, find platforms and services in the other layers of the stack. Discovery channels architect information- they amplify certain information and hide others. Nearly all the discovery channels- Google search, app stores, social media platforms- offer their service for free and derive revenue from advertisements. The exception is AI services, some of which are offered on subscription-based.

Core Infrastructure:

The core infrastructure refers to the foundational elements of the internet such as Domain Name Servers, Internet Service Providers and Cloud Services for hosting data and models. Every other layer of the infrastructure stack relies on this baseline layer.

The datasets, ML models and the core Infrastructure are necessary for production of a synthetic media item. The model access providers, discovery channels and GUI enable the monetisation and amplification of the content produced.

Infrastructure Tech Stack

How can I use AI to generate nudes on my comput... ✓

Discovery Channels

Platforms through which people can find the services that nudify or generate sexual scenes.

Graphical User Interface

The easy to use modality via which non technical users can upload images to generate synthetic content

Model Access Providers

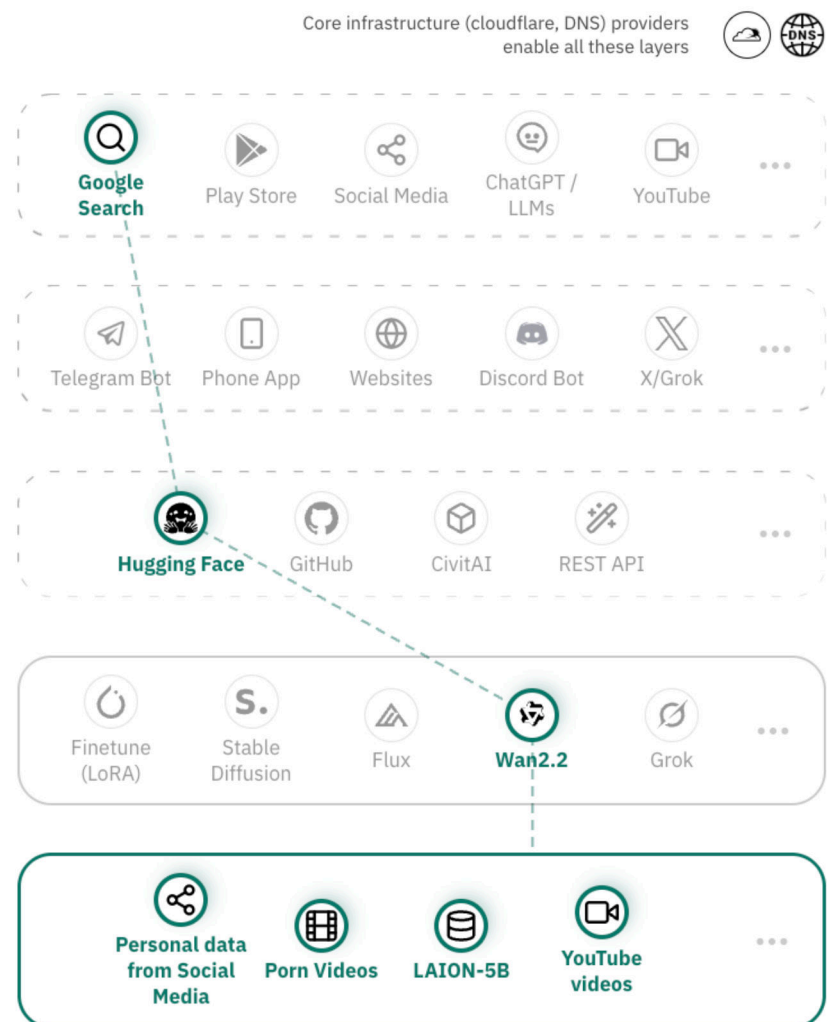
Interfaces through which more technical users can access models and build on top of AI models. Providers might also support model downloads, gamification and monetisation features.

Machine Learning Models

Models are trained on individual/corporate hardware and datasets. Large base media models are fine tuned on specific sexual scenes.

Datasets

Images and videos are collected and annotated. Datasets enable base model training and fine-tuning



* Discovery channels, Graphical User Interface and Model Access Providers are optional layers that enable access to ML Models that rely on Datasets.

One can discover NCSI-generating services on general LLM chatbots, platforms, and search engine listings. This can manifest through search engine queries (“free image generation tools”), tutorials (YouTube videos on “how to create realistic AI images”), social media (targeted ads on AI tools that help you get a million followers), and prompting (“help me create a realistic AI influencer persona who looks like [insert image]”). Bots on Telegram and Discord, apps that have interfaces to upload images to create NCSI, all allow one to request and receive NCSI outputs in the same channel. Developers can download and access the

repositories of model code on platforms like Hugging Face, GitHub, Civitai, or directly through RestAPI. The models themselves, open or closed-source, may be trained on sexual content through datasets scraped from different sources. Payment processors enable these activities to be monetised: they allow people to (i) buy NCSI services through platform credits and agents, and (ii) get paid for posting NCSI content through subscription-based offerings on platforms. The core infrastructure of the internet, such as DNS servers and CDNs like Cloudflare, also support these services.

Salient Findings from the Investigation

The original journalistic investigation covers findings from ground reporting and documents the real world human impact of NCSI. It also covers the responses from platforms. Below we detail some of the infrastructure specific findings:

- **Discovering new apps and services is trivial**

Apps and websites that can generate nudes or intimate content were found on the first page of all the search engines we checked- Google, DuckDuckGo, Brave. AI chatbots such as Claude and ChatGPT also suggested apps and Telegram bots.

- **The infrastructure spans continents**

We found websites hosted in Europe, apps hosted in China, API documentation maintained by developers in Bangladesh. Each of these was available to users in India. When monetised, the services relied on intermediaries who converted Unified Payments Interface (UPI) payments from Indian users to cryptocurrency or credit card transactions that the services accepted.

- **NCSI generating Telegram bots are widespread and fast growing**

Telegram is a common platform to access NCSI-generating services. Many are targeted explicitly at Indians. They offer similar features to websites in Indian languages and take UPI payments. Telegram also provides information for Indians to become agents who can accept UPI payments from other users to convert them into credits required by NCSI-generating bots.

- **The Backend Infrastructure is Centralised**

Many of the Telegram bots were enabled by the same backend or API- bots advertised the ability to clone themselves. New bots and websites names also appeared simultaneously across the Telegram channels, Discord servers and Reddit threads devoted to AI generated content that we were tracking. The coordination suggests centralisation behind the plurality of front-end interfaces.

- **Apps found creative ways to bypass app store restrictions**

While the Google Play Store doesn't allow apps that contain or

promote sexual content or profanity, including that created by AI, we found several apps that contained such content. One app with a registered address in Hong Kong and backed by venture capital in China, though appearing benign in its description and its home page, had design features that nudged users towards sexual content including that depicting minors. The in-app menu also contained a link to a Discord channel where images of women and their modified nudes were shared. Within a week of installation, the link disappeared from the menu on the app but the communication continued on the Discord channel. While the language of the Discord server was Chinese, the images were of South Asian women. Despite our reporting it for hosting CSAM, the app is still available on Google play store.

- **Models generating sex scenes trend on Hugging Face**

At the time of our research, 67% of the top 50 trending models on Hugging Face that enabled the generation of a video from an image using a text prompt were for creating sexual scenes. Nearly all of them were fine-tuned on the Wan series of models.

- **The wide prevalence of CloudFlare**

Two-thirds of the 31 websites we studied that enabled the creation of AI images and videos used CloudFlare as their Content Delivery Network (CDN). In all cases the administrator's name and location were "redacted for privacy". Accessing details about the administrators of these services will require cooperation from CloudFlare.

- **Google is the predominant SSL provider**

Twenty-three websites used Google as the SSL provider. The remaining 8 used a service called Let's Encrypt.

Undress API 1.0.0 OAS 3.1

/openapi.json

API for image undressing.

Rate limits:

- Unauthorized users: 10 requests per minute
- Authorized users: 60 requests per minute

Demo webhook handlers: [Github Gist](#)

Support: [Telegram](#)

Authorize

User

GET /api/v1/me Get User Info

Photo

POST /api/v1/photos/undress Undress

GET /api/v1/photos/poses Get Photo Poses

POST /api/v1/photos/poses/undress Undress Photo With Pose

Video

POST /api/v1/video/undress Undress Video

GET /api/v1/video/poses Get Photo Poses

 **Hugging Face**

Search models, datasets, users...

Models

Datasets

 **Aisha-AI-Official/wan2.2-standing-in-lust**

Like 7

Image-Text-to-Video lora nsfw Not-For-All-Audiences

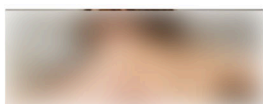
Model card Files and versions xet Community

Standing in lust (standing sex) [Real]

This was supposed to be just a hug 😊

Sorry baby, your pussy called me 😊

(Download links at the end of this page)



About this app

AI Image & Video Generator: Create AI anime, 3D art with 100,000+ filters in 5s!

Entertainment

Art & Design

Ai art

10M+
Downloads

Parental guidance



Data safety

OpenArt
openart.ai > features > ai-clothes-remover

The Best AI Clothes Remover (Free, Online)

Remove clothes from images easily with our free AI-powered tool. Fast, simple, and online — no downloads or editing skills needed.

Perfect Corp
ycc.perfectcorp.com > home > features/remove-and-r...

AI Clothes Remover – Replace Outfits & Remove Objects from Photos Online

December 19, 2025 - Instantly remove clothes in your photos and replace with new ones using **AI Clothes Remover**. Upload your photo, enter a text prompt, and let AI create a realistic transformation—no editing...

coqui-ai / TTS

Code Issues 15 Pull requests Discussions Actions Proj

Nude AI Generator: Top 5 Generators With F #4211

ghost started this conversation in General

ghost Apr 26, 2025

AI nude generators have gained attention for their ability to create or manipulate images using advanced artificial intelligence. These tools leverage machine learning and computer vision to produce realistic results, often sparking user curiosity. While some explore these apps for artistic or experimental purposes, their controversial nature cannot be overlooked. In these posts, I have written about the 5 best AI nude generators that I use on my fashion projects.

5 Best AI Nude Generators Of 2025

1. Undress.app

Legal Foundation For Action Against NCSI in India

Indian law has in the recent past sought to tackle the ramifications of image-based abuse. These laws have the scope to reach each actor in the chain, from the creator to the platform or tool enabling the content, although the basis for liability differs. This report examines the applicability of these laws to the services across the infrastructure used to manipulate and distribute images of real people using AI, without their consent.

A primary point of contention in AI regulation is how data should be understood as a contributor to the final value chain. In the case of AI models that can be used to create NCSI, the training data is the images and videos depicting women in intimate settings. The image that is finally morphed into a face swap or a video of a sexual act is also data that may be added back to the training dataset. There is nearly no visibility into the training datasets used to create these models, whether or not they are open-weight models. It is reasonable to assume that a significant proportion, if not the majority, of the nude and intimate images that the models are trained on were not collected with the explicit consent of those depicted in the media. The violation of consent is more explicit for those whose images are manipulated using the model. There are two traditions of legal philosophy—one emerging from the view of data as an extension of personhood and the other of data as property— which recognise the violation of consent. There is a third tradition, rooted in social amorality, that is concerned only with the appropriateness of the outcome. Below we describe the legal precedent in India around each of these traditions and how they have been applied to online images. We then describe specific interventions that can clarify the applicability of these precedents to NCSI.

The IT Act, 2000 defines data as information, including images, which are represented in a formal manner with the intent of being processed by a computer system.¹⁷ With the advent of e-commerce, the concepts of fraud and cheating were incorporated into cybercrime provisions penalising identity theft and impersonation, in addition to provisions addressing the violation of privacy, voyeurism, and the dissemination of explicit imagery of adults and children. In 2017, with the recognition of informational privacy, the link between privacy and consensual use of individuals' personhood was established as an expression of their dignity and autonomy under Article 21, where it was observed that privacy allows an individual to decide how to exercise their liberties.¹⁸ The autonomy centric approach flows into cases of NCII as well. The Delhi High Court, pursuant to a case filed in 2021 delivered a judgment addressing the structural liability of intermediaries surrounding NCII content, and defined it as:

*"sexual content that is distributed without the consent of those who are being depicted in the said content. The content may or may not be taken with the consent of the individual involved, however, its dissemination is largely meant to be non-consensual and comes under the larger umbrella of cyber harassment."*¹⁹

Reaffirming that privacy is inextricably linked to dignity, the Court stated that individuals have a reasonable expectation of privacy in relationships, including where an intimate image is shared with the understanding that it will not be shared elsewhere. Non-consensual synthetic generation of persons' likeness as an offence has been read into existing provisions on voyeurism²⁰, stalking²¹, cheating by impersonation²², violation of privacy²³, dissemination of obscene content²⁴, defamation²⁵ under the IT Act, 2000 and the Bharatiya Nyaya Sanhita, 2023, with penal consequences.²⁶ For responses to the presence of the content itself online, amendments have focused on takedowns and watermarking of the content.²⁷

With the rise of unauthorised use of persons' likeness, public figures have sought legal action, resulting in the recognition of 'personality rights', stemming from principles such as privacy and passing off. Personality rights protect public figures' images, voice, persona, name and other attributes of their personality from being commercially exploited without authorisation.²⁸

¹⁷ Information Technology Act, 2000, § 2(o)(India).

¹⁸ Justice K.S. Puttaswamy (Retd.) v. Union of India, (2017) 10 SCC

¹⁹ Mrs X v. Union of India, 2023 SCC OnLine Del 2361 (Delhi HC, April 2026, 2023).

²⁰ Bharatiya Nyaya Sanhita, 2023, § 77 (India).

²¹ Bharatiya Nyaya Sanhita, 2023, § 78 (India).

²² Information Technology Act, 2000, § 66D (India).

²³ Information Technology Act, 2000, § 66E (India).

²⁴ Information Technology Act, 2000, § 67 (India).

²⁵ Bharatiya Nyaya Sanhita, 2023, § 356 (India).

²⁶ <https://www.pib.gov.in/PressReleaseDetail.aspx?PRID=2291671®=1&lang=1>

²⁷ Information Technology Act, 2000

²⁸ Arijit Singh v. Codible Ventures, 2024 SCC OnLine Bom 2445 (Bom HC, July 26, 2024).

Images of people are also distinct assets and have proprietary rights attached to them, namely copyright, the ownership of which typically rests with the creator of the image. Where the image was taken by the person themselves, they can file a DMCA takedown notice for effective removal of the content since most major platforms hosting the content have servers in the US and opt for DMCA compliance; this has been used as a tool for faster takedown by support workers in tackling online harms.²⁹ However, the known limitation with copyright enforcement is that images may not be taken by the person themselves, in which case the copyright holder has the onus to report, and the victim is left to approach other remedies.

²⁹ Pillai, Prabhakar and Kalidindi, *Make It Real*, 39.

3. Image as Societal Amorality

Beyond the conception of the harm as being perpetrated between two parties, the law also considers whether such content may appeal to prurient interests, a legal test for obscenity. Obscenity provisions in India, derived from Victorian perceptions on public morality which were introduced in law during the colonial era, criminalise behaviour that is 'lascivious' and 'deprave', and has historically been used as a tool for moral policing. The community standards test is applied to evaluate whether content would be considered as 'prurient' to a reasonable person, and is not simply graphic.³⁰ Any individual can initiate legal action against such content, and obscenity provisions have been used as a mechanism for legal recourse to online harms as well.³¹

³⁰ *Bobby Art International v. Om Pal Singh Hoon*, (1994) 4 SCC 1

³¹ 'Experiencing Technology-Facilitated Gender-Based Violence in India'. Equality Now and Breakthrough Trust

The multi-layered infrastructure precludes the possibility of one simple intervention that addresses the challenge of NCSI. For every layer in the infrastructure there are multiple actions that different actors can take to limit the creation of NCSI. A systematic approach that is effective in addressing NCSI will involve interventions across all the six layers. The actions may be legally mandated: as described in the section above there are sufficient foundations existing in Indian law on image based abuse that can be applied to NCSI, but their applicability to actors across the infrastructure stack needs to be clarified. But individual actors across the value chain may also take actions of their own volition or through standards set by industry bodies. We describe four broad strategies, each encompassing a number of interventions that can be used to prevent generation, distribution and monetisation of NCSI.

Strategy 1: **Clarify Applicability of Laws to Model Training**

Actors: Courts, MeitY

Much of the content shared by individuals, such as that on their social media profiles or web presence, was shared before their utility for training visual models was known. Even now users' expectation of privacy in sharing an image on social media profiles is not always clear. This lack of clarity extends to enforceability of consent and attribution of liability from use of personal images for model training. Most charging provisions address the stages of 'capturing', 'publishing' and 'transmitting' real and explicit content. These provisions do not directly apply to manipulations using AI. The following clarifications can help introduce liability for developers whose models are used to create NCSI:

- **The applicability of Digital Personal Data Protection Act, 2023 (DPDPA) provisions for digital personal data that may be publicly available but has been used for illegal activities (such as creating NCSI):³²**

The NCSI output violates the law; however, the use of an image of an individual online (say a profile picture on a social media account) itself may be permissible since it was available publicly. Personal data made publicly available by the user themselves is excluded from the ambit of the DPDPA, and a clarification on liability for non-consensual processing of public personal data may be considered.³³

- **Copyright infringement action on models**

In many cases, profile pictures or social media images used for NCSI may have been taken by the victim themselves. Legal action may be initiated on the grounds of infringement for ingestion of copyrighted data, including images, for model training.³⁴ In the context of text-based models, a recent interim judgment of the Delhi High Court considered news data being used for model training as constituting 'fair dealing' and therefore not infringement.³⁵ However, the merits of the case are yet to be heard. The applicability of copyright to training data should also be interpreted in the context of large visual models that can be used to produce likeness of real-world individuals.

³² It has been reported that MeitY has requested formal explanations from tech platforms on the scraping of public data in light of concerns around CSAM, amongst others. ANI. 'Centre to Examine Instagram Data Scraping Concerns after Responses to Notices, Says MeitY Secretary S Krishnan'. ETTelecom.Com. Accessed 24 August 2026. <https://telecom.economictimes.indiatimes.com/news/policy/meity-to-investigate-instagram-data-scraping-issues-amid-growing-public-concerns/132285737>.

³³ Digital Personal Data Protection Act, 2023, § 3(iii) (c) (India).

³⁴ ANI Media Pvt Ltd v. OpenAI OpCo LLC, I.A. 45300/2024 in CS (COMM) 1028/2024 (Delhi HC, July 24, 2026).

³⁵ ANI Media, I.A. 45300/2024.

Strategy 2: Take Down NCSI and Services that Enable It

Actors: Platforms, Intermediaries for implementation, Courts, MeitY for enforcement

For NCSI, as with other categories of harmful content, the most common regulatory and platform response is taking down the content. This ex post facto intervention has taken precedence in rule making and enforcement in India. It has been used as a cross-cutting measure to tackle websites, content, repositories, platforms, and even Content Delivery Networks (CDNs).³⁶ The demand to take down content has come through regulations such as the IT Rules, 2021 as well as court rulings on cases of individual harm. Courts have frequently granted interim and permanent injunctions in light of the extent of damage that can be caused to a victim-survivor from delays in action against the dissemination of such content online.

In India, regulations and court rulings have effected takedowns at all layers of the tech infrastructure described in this report. While these have not always been motivated by action on NCSI, they provide precedents for takedown of content, applications, websites and models to address NCSI. Below we describe the precedents for takedowns at different layers of the tech infrastructure:

• Discovery Layer (Search engines and social media platforms)

Platforms are required to carry out proactive filtering once an actual order has been received.³⁷ De-listing is also mandated on search engine results.³⁸ The IT Rules, 2021 mandate 2 hour takedowns of NCII and SGI content by platforms enabling the creation, generation, modification, alteration, publication, transmission, sharing or dissemination of such content.³⁹ Platforms must similarly engage in proactive detection and filtering of NCSI content, including through the use of automated tools for takedown.⁴⁰ A general stance held by MeitY has been that violation of IT Rules, 2021 would result in the loss of safe harbour protection for platforms, however any such consequence must be consistent with procedural safeguards under the IT Act, and requirements of legality, proportionality, and necessity of such measures.

• Graphical User Interfaces (Apps, Telegram and Discord Bots, websites)

In 2026, the Government of India blocked access to Telegram upon a determination that the ongoing takedown enforcement process was inadequate.⁴¹ This was the first time that the messaging platform had

36 Ministry of Electronics and Information Technology, Standard Operating Procedure to Curtail Dissemination of Non-Consensual Intimate Imagery (NCII) Content, ver.1 (New Delhi: Government of India, October 2025), <https://www.meity.gov.in/static/uploads/2025/11/a2c9500ef5f8b62a43bfc68747de592d.pdf>. Content (other than that hosted upon social media platforms) may be actioned through intimation to DoT (Department of Telecommunications, Ministry of Communications) for appropriate action.

37 Mrs X, 2023 SCC OnLine Del 2361

38 Indian Courts have directed de-listing of results in sensitive matters; Mrs X, 2023 SCC OnLine Del 2361

39 Mrs X v. Union of India, 2023 SCC OnLine Del 2361 (Delhi HC, April 2026, 2023).

40 IT Rules, 2021, r3(3).

41 Telegram FZ LLC & Anr. v. Union of India & Ors., W.P.(C) 8259/2026 (Delhi HC, June 19, 2026)

been temporarily blocked nationwide. Stringent action may be taken against GUIs, including banning the apps themselves, however any such measure must be imposed only if it satisfies the tests of proportionality and follows procedural safeguards.

- **Model Access Providers (Repositories)**

Reporting repositories to GitHub/Hugging Face for violating their Acceptable Use policies and contravening intermediary liability provisions results in the removal of the repositories.^{42,43}

- **Model Layer (Personality rights infringement takedowns)**

Model tools engaging in unauthorised use of public figures' likeness are restricted for infringing personality rights.⁴⁴

- **Core Infrastructure**

Action can be taken on content not hosted on intermediaries by intimating the Department of Telecommunications through the national cybercrime portal.⁴⁵

As our research showed, many aspects of the infrastructure are hosted outside India. The Indian government can request for takedown of services hosted outside its borders, since extra-territorial jurisdiction does not require all actors in the chain to be located within India's jurisdiction, so long as one is.⁴⁶ Indian courts may exercise this power to summon a foreign actor.

Limits of Takedown:

Despite the long precedent for content and service takedowns, there are known issues with enforcement and efficacy of Takedowns in one part of the infrastructure chain to prevent overall proliferation. Here are some known gaps in the effectiveness of takedowns:

- Removing access to code repositories does not address already accessed or locally hosted and modified models. It may only reduce future access. The effect on 'forks' or modifications of the model may also be limited.
- Takedown measures cannot prevent abusive content from resurfacing en masse and replicating across platforms into less common networks; self-cloning bots may ultimately outpace takedowns.⁴⁷
- Enforcing takedowns across jurisdictions is especially challenging. Actors located in foreign jurisdictions may not comply with extra-territorial judicial requests, resulting in long-draw-out processes of

⁴² GitHub Docs. 'GitHub Acceptable Use Policies'. Accessed 24 August 2026. <https://docs-internal.github.com/en/site-policy/acceptable-use-policies/GitHub-acceptable-use-policies>.

⁴³ Jul 25, ET Bureau Last Updated: 'Govt Orders GitHub to Remove Access to Bitchat App in India'. The Economic Times, 25 July 2026. <https://economictimes.indiatimes.com/news/india/govt-orders-GitHub-to-remove-access-to-bitchat-app-in-india/articleshow/132613661.cms?from=mdr>.

⁴⁴ Arijit Singh v. Codible Ventures, 2024 SCC OnLine Bom 2445 (Bom HC. July 26, 2024).

⁴⁵ Ministry of Electronics and Information Technology, SOP on NCII Content.

⁴⁶ Information Technology Act, 2000, §75 (India).

⁴⁷ Bose, Adrija, Tarunima Prabhakar, and Tora Agarwala. Desi Faces, Foreign Servers: Inside the AI Economy Undressing Women.

initiating Mutual Legal Assistance Treaties (MLATs) and resultant enforcement. MLATs may not be as effective in cases of NCSI where rapid dissemination is the norm.

- Automated approaches for detection of NCSI or services that enable it will inevitably make mistakes. Mandates for faster timelines for takedowns are complied with through greater reliance on automation. These don't guarantee effective takedowns, and present concerns around over-moderation by platforms of genuine self-expression.
- A significant contributor to the inefficacy of the takedown mechanism is that violation of existing rules carry limited and inconsistent consequences. In 2026 it was reported that Grok was used to create sexual images of millions of women on the X platform at the rate of 6700 images per hour.⁴⁸ This was a direct violation of the IT Rules' previously set 24-hour takedown requirement for sensitive content. In response, the Indian government issued X a 72-hour ultimatum to take down the content. While a warning was given to the platform, the threat of revoking the platform's safe harbour immunity was not acted on.

⁴⁸ Cecilia D'Anastasio, "Musk's Grok AI Generated Thousands of Undressed Images Per Hour on X," Bloomberg, January 7, 2026, <https://www.bloomberg.com/news/articles/2026-01-07/musk-s-grok-ai-generated-thousands-of-undressed-images-per-hour-on-x>.

A number of interventions taken by model developers and individual users can prevent the creation of NCSI in the first place.

3.1 Datasets:

Intervention	Limits to Adoption	Actors
Removing Nude/ Sexually Explicit Content from the Training Data This makes the models trained on the data more tamper-resistant.⁴⁹ Even if the models have open weights, the ability to manipulate them to produce images imitating those that are not in the training data is limited.⁵⁰	• Filtering web-scale data is challenging. It is costly and prone to filtering errors.⁵¹ • Removing nudity from the dataset also reduces the overall performance of the models. That is, it not only reduces the ability to produce sexual content but also non-sexual videos and images. Model developers have found the degradation in overall performance significant enough to weaken the filtering of the datasets.⁵² • Training data of models is rarely made public. Even if this technical measure was legally mandated, it would be difficult to gauge compliance.	Model Developers
Poison Images Individual users, to protect their data being used in the training of models or being manipulated by them, could 'poison' or add noise to their images.⁵³	• Adding noise to an image is beyond the technical reach of the average internet user. • Models can get better at detecting techniques of poisoning data. While detecting poisoning attacks on large datasets is economically costly, once it is known that a media item is poisoned, model developers can remove the poisoning artefacts.	Individuals/ Users

⁴⁹ <https://deepignorance.ai/>

⁵⁰ <https://alignment.anthropic.com/2025/pretraining-data-filtering/>

⁵¹ UK AI Security Institute, "Managing Risks from Increasingly Capable Open-Weight AI Systems," AI Security Institute (blog), August 29, 2025, <https://www.aisi.gov.uk/blog/managing-risks-from-increasingly-capable-open-weight-ai-systems>.

⁵² <https://stability.ai/news-updates/stablediffusion2-1-release7-dec-2022>

⁵³ This idea emerges from technical projects such as Glaze and Nightshade "that transform images into 'poison' samples, so that models trained on them learn unpredictable behaviors that deviate from expected norms." <https://nightshade.cs.uchicago.edu/whatis.html>

3.2 Models:

Intervention	Limits to Adoption	Actors
Safety Evaluations Prior to release, models should undergo thorough safety testing. The process and outcome for safety evaluations pertaining to intimate and nude content should be publicly disclosed.	• Safety evaluations are currently voluntary. Systematic and public safety evaluations are lacking. There are currently limited mechanisms for mandating safety evaluations in companies across countries. • With open-weight models any guardrails provided by default can be removed at the fine-tuning stage. Building tamper-proof guardrails in open-weight models is an open area of research.	Model Developers, Regulators
Release under a Responsible AI License Model developers can explicitly exclude the use of their models to produce non-consensual and/or sexual content in the license of the model. Those infringing license terms can be held accountable under law.	License enforcement is limited. Although open-source licenses can be hard to enforce, they leave open the option of legal action or community pressure to hold violators accountable. 54	Open-weight Model Developers

54 Madhulika Sri Kumar, Kasia Chmielinski, and Jiyou Chang, Risk Mitigation Strategies for the Open Foundation Model Value Chain: Insights from PAI Workshop Co-hosted with GitHub (San Francisco: Partnership on AI, 2024), 6, https://partnershiponai.org/wp-content/uploads/dlm_uploads/2024/07/open-foundation-model-risk-mitigation_rev3-1.pdf.

Model Access Providers, as platforms that host content, and also enable the generation of content, pose a new challenge for safety and moderation.⁵⁵ Our own research and other investigations show that model access providers provide easy amplification of open-weight models that can be used for creating NCSI.⁵⁶ For proprietary models, the model access providers and model developers are the same. Below we list the interventions that model access providers, whether they be of open-weight models or proprietary service providers, can adopt to prevent the use of their service for the generation of NCSI.

⁵⁵ Gorwa, Robert, and Michael Veale. "Moderating Model Marketplaces: Platform Governance Puzzles for AI Intermediaries." *Law, Innovation and Technology* 16 (2): 341–91. 2024. doi:10.1080/17579961.2024.2388914.

⁵⁶ Bouchaud, Paul, and Silvia Semenzin. "Unmoderated by Design: How Hugging Face Enables NCII. AI Forensics, July 2026.<https://aiforensics.org/uploads/NCII%20HF%20Report%20by%20AIF.pdf>.

Intervention	Limits to Adoption	Actors
User/ Input Blocking Proprietary models can be made safer by blocking offending input prompts or blocking users altogether.	Prompt-level blocking can be circumvented through creative prompting or jailbreaking.	Proprietary Model Access Developers/ Providers
Banning and de-monetising models It is impossible to technically verify if a model is used to manipulate an image with consent. Model access providers could remove models that are used to generate sexual scenes because of the risk that they produce non-consensual intimate content. At the very least they should not monetarily benefit from hosting these models. This would be an extension of takedown measures taken for services that can produce NCSI.	Even with policies in place, companies might fail at detecting violations due to inadequate capacity for technical and/or human oversight. As described in Strategy 2, takedowns will also not delete already downloaded or modified models.	Model Access Providers
Preventing use of real-world images on platform Since consent to use real world images can not be technically verified, model hosting companies can prohibit the use of images depicting real people. This was the rule Civitai introduced after reports of the platform being used for creation of NCSI. ⁵⁷	Detection of the likeness of real-world people, whether in the image uploaded or in the final output, is a non-trivial technical problem. There will be inevitable lapses in the enforcement of this measure.	Model Access Providers

⁵⁷ <https://civitai.com/content/rules/real-people>

The monetisation of the creation and distribution of non consensual synthetic imagery enables the sustained scale of the phenomenon. Monetisation takes place throughout the infrastructure chain and different actors and levers can reduce the monetisation of NCSI.

4.1 Disable use of Payment Services

Actors: Payment Gateway and Network Providers

While many monetised websites, bots and apps require payments in cryptocurrency to use the service, they also rely on established gateways such as Stripe, Paypal, PhonePe; and payment rails like Visa and UPI. A decision by payment networks to withdraw their services for a website can single-handedly bring the service down since there are few, if any, substitutes that the services can rely on. Civitai's decision to disable the use of the likeness of real-world people on their platform was, in-part, driven by a card processing company disabling the use of its service on the platform.⁵⁸ Instead of dissociating from sexual content altogether, financial intermediaries should rely on violation of consent as the underlying principle for withdrawing their service.⁵⁹ While some services contain terms of use mandating consensual use of imagery,⁶⁰ many do not. At the very least, payment intermediaries can dissociate from such services.

58 Maiberg, Emanuel. "Civitai, Site Used to Generate AI Porn, Cut Off by Credit Card Processor." 404 Media, May 20, 2025. <https://www.404media.co/civitai-site-used-to-generate-ai-porn-cut-off-by-credit-card-processor/>.

59 Mahdawi, Arwa. "The real OnlyFans scandal is the unaccountable power of platforms and banks." Guardian, August 28, 2021. <https://www.theguardian.com/commentisfree/2021/aug/28/onlyfans-scandal-platforms-banks-ban-sexual-content-sex-workers>.

60 The presence of terms of service doesn't guarantee compliance with the terms. The Decode and Tattle investigation details that the perpetrator used prominent AI visual editors for manipulating images of his ex-girlfriend, in contravention of the platform terms.

4.2 Penalize Monetisation Through Ad Revenue

Actors: MeitY

Websites, apps and bots rely on social media platforms and search engines to be discovered by people. When these services run ads, the social media platforms and search engines gain monetarily. Tech platforms can be fined for monetarily benefiting from running ads for NCSI services on the platforms.

Actors: MeitY, Financial Sector Regulators

Payment processors are required under law to maintain records on suspicious transaction activity.⁶¹ The intent of these stipulations is primarily centred around money laundering and terror financing. Similar recording mandates can be extended to flag proceeds from NCII and NCSI as well. Coordination may also be mandated between intermediaries, payment processors and advertisers for faster redressal. Any proposals on extending/ introducing reporting processes or automated filtering for proceeds of transactions from NCSI must duly consider the proportionality of obligations as responses.

⁶¹ Prevention of Money-laundering (Maintenance of Records) Rules, 2005, rule 3(1) (D).

Intervention Strategies

- Clarify Applicability of Laws to Model Training
- Take Down NCSI and Services that Enable It
- Preventing Generation of NCSI
- Prevent Monetisation

Infrastructure Layer	Recommended Intervention		
Datasets	Poisoning Remove Imagery from Dataset	Clarify DPDP Applicability	Copyright Applicability
Models	Safety Evaluations Use License		
Model Access Providers	Block Users/ Prompts Ban Models Prevent Use Of Real- World Image	Takedowns	
GUI			
Discovery	Penalise AD monetisation		
Core Infrastructures & Payment Services	Disable use of payment services Introduce reporting mandates		

Over the last few years, the Indian government, like other governments, has taken note of the growing challenge of non-consensual synthetic imagery. Its response has focused almost exclusively on content takedown, a strategy beset with implementation problems. Furthermore, content takedown is a post-facto measure, applied only after harm has occurred. Piecemeal takedowns cannot match the scale and complexity of the cross-border economy around NCSI. At present, regulatory attention is disproportionately focused on the tip of the iceberg—the prominent discovery platforms—and misses the larger edifice beneath it: the multilayered infrastructure of ML models, model hosting platforms and bots that enables the production and proliferation of NCSI.

In this report, we document a broader set of strategies and interventions that also address the generation and monetization of content. We note that, beyond poisoning their own data, a step that requires significant technical sophistication, there is little that individual users can do to prevent their images from being used in training of models that can produce NCSI, or from being manipulated by the models. *Addressing NCSI requires regulators, model developers and technology platforms to take significant steps in tandem.* To begin with, the Indian courts and regulators can clarify how existing laws apply to the use of Indians' images for model training purposes. Companies developing and providing access to models can take a range of steps to prevent NCSI generation in the first place, from cleaning data to conducting safety evaluations and instituting safeguards, to delisting offending models and blocking users and prompts. As a "middle-power" reliant on open-weight models, India should take a strong interest in the safety provisions of these models. Since most foundation models are developed by companies in a few countries, ensuring that they are adequately tested and released with baseline safeguards will require global advocacy and coordination. Beyond weights, regulators and technologists should also push for greater transparency on the datasets used

for training the models as this will help build understanding of (i) why models produce intimate and sexual scenes so easily; and (ii) how models can be built to degrade generation capabilities over sexual content without affecting their performance over others. In its push for building AI capabilities, the government can also invest in research to understand model performance and to build more tamper-resistant safeguards. Since off-the-shelf services used for generating NCSI rely on subscriptions, the payment networks and services, especially UPI, offer a distinct point of intervention to track and block the production of NCSI in India. No one intervention will be sufficient. But a concerted attack on the “deepfake” infrastructure can reduce the scale of NCSI and the resultant harm.

Finally, we end on a bold vision for a longer-term overhaul of the AI development process into one that relies on clear consent-based data practices, more sustainable and accountable funding models, and a more distributed research and innovation ecosystem that isn't evaluated solely on the basis of a unidirectional chase towards efficiency. An AI economy built on consent from the ground up might be one in which non-consensual imagery has no place to take root.

Deconstruct the Deepfake Stack: An Action Plan to Tackle NCSI in India

Date:

12 October 2026

Authors

Tarunima Prabhakar, Kaustubha Kalidindi & Adrija Bose

License

The text and visuals in this document are licensed under CC BY SA 4.0. The report may be cited as "Prabhakar, Kailidindi, Bose. Deconstructing the Stack: An Action Plan for NCSI. Tattle Civic Technologies. 12 October 2026"