INDIAN GOVERNANCE AND POLICY PROJECT

Born Connected:

Keeping Childhood Safe in a Digital India

November 2025

BORN CONNECTED: KEEPING CHILDHOOD SAFE IN A DIGITAL INDIA

Published by
Indian Governance and Policy Project (IGAP)

Authored by
Soumya AK and Ananya Agrawal

Designer
Manoj Murali

About IGAP
The Indian Governance and Policy Project (IGAP) is a policy, business advisory, and research studio working at the intersection of governance, technology, markets, and national development.

Grounded in a clear understanding of how state capacity, market forces, and emerging technologies shape India's strategic trajectory, IGAP addresses key questions that define the country's future — from the governance of AI and digital infrastructure to financial innovation, sustainability, and national security.

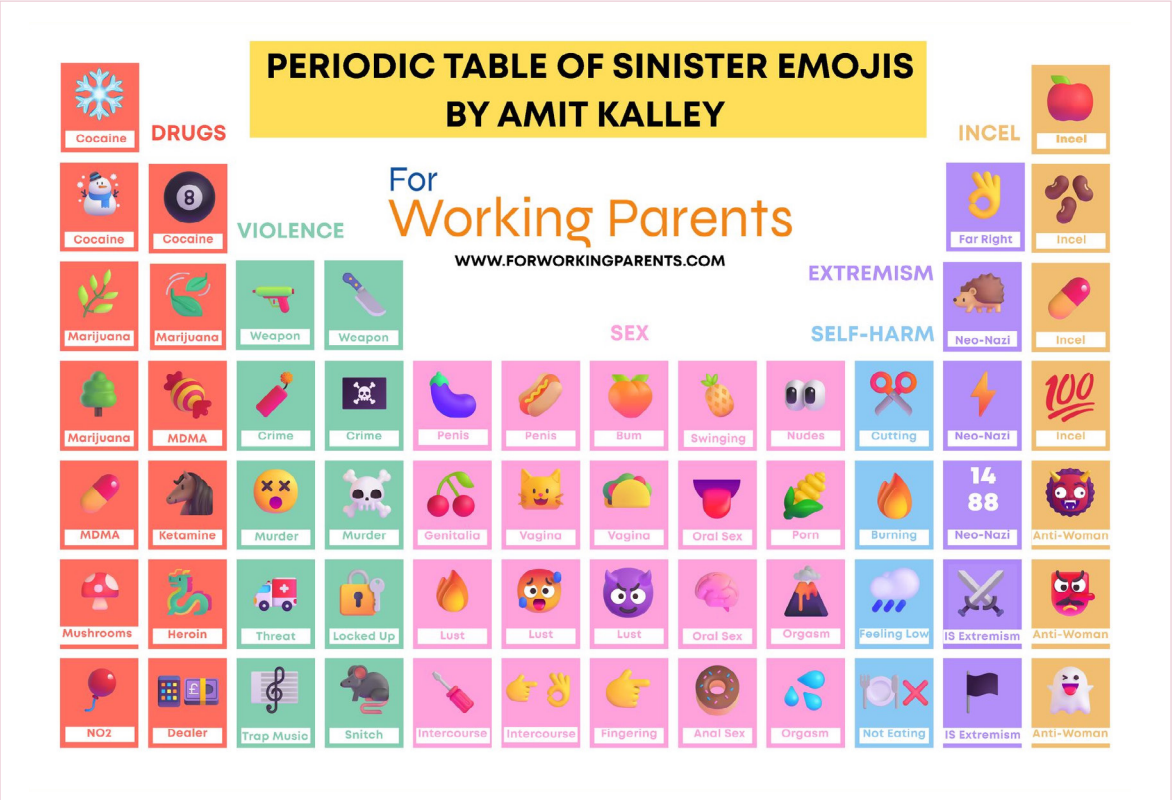
Bringing together lawyers, policy thinkers, and strategists with deep business and geopolitical insight, IGAP delivers solutions that balance India's developmental and security priorities with its democratic values and constitutional principles.

TABLE OF CONTENTS

01 Introduction

1.1 The New Digital Adolescence

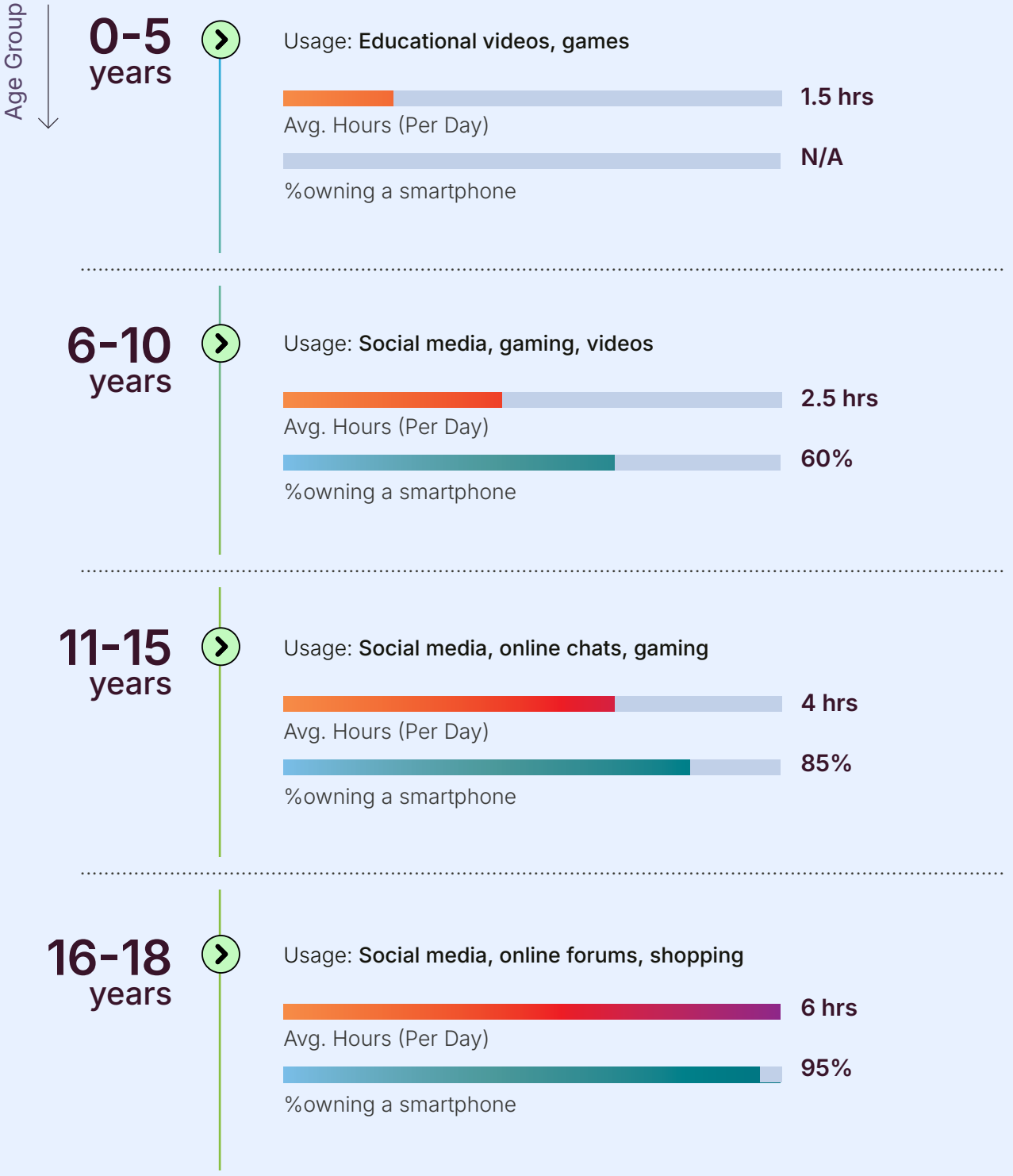
A scene from Netflix's acclaimed series *Adolescence* captures a chilling truth about what it means to grow up online.[1] A group of teenagers exchange what seems like harmless banter, yet the language is coded - insults are disguised as compliments, manipulation is embedded in memes, and cruelty hides behind irony. Within this altered linguistic landscape, harm travels unnoticed, until it surfaces offline as bullying, exclusion, or violence.



Credits: Periodic table of sinister emojis by Amit Kalley

This blurred boundary between the digital and the physical defines modern adolescence. As technology becomes inseparable from education, entertainment, and social life, young users navigate a landscape filled with both opportunity and risk.[2] According to UN estimates (2023), around 77 percent of people aged between 15 and 24 used the Internet, with many engaging with digital devices from their earliest years.[3]

In India, this digital transformation has been particularly rapid. The Telecom Regulatory Authority of India estimates **989 million active internet users**,[4] including a substantial and growing number of adolescents and school-aged children. The COVID-19 pandemic further accelerated this shift, as classrooms moved online and millions of children received their first devices for remote learning.

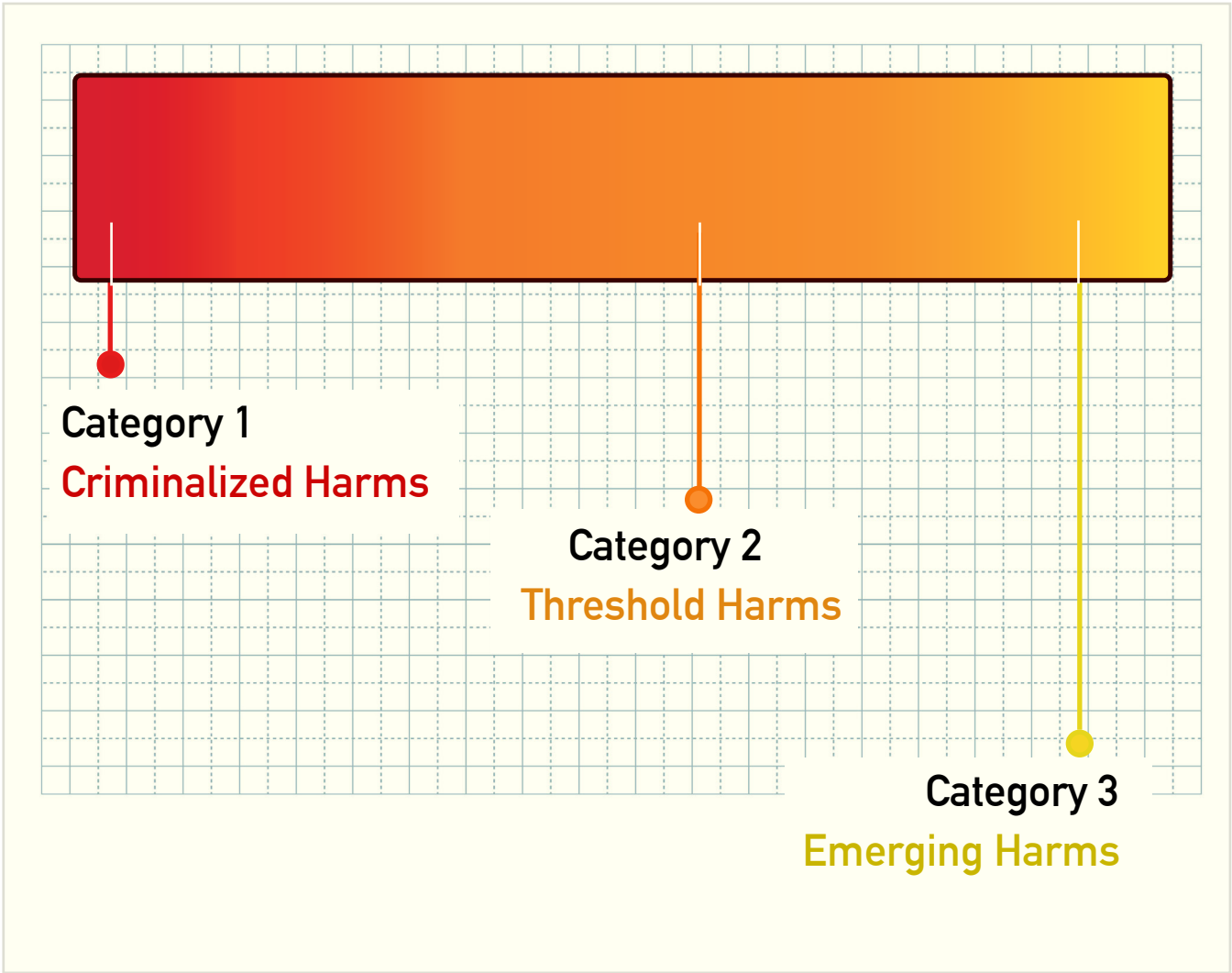


Source: NITI Aayog report on 'Online Safety for Children: Protecting the Next Generation from Harm[5]

1.2 The Double-Edged Sword of Digital Access

While digital access enables learning, creativity, and connection, it also exposes children to exploitation, harmful content, and psychological risk.[6] These risks extend beyond criminalized harms. Children increasingly encounter lawful yet harmful experiences including exposure to distressing content, manipulative design features, datafication of their online behaviour, and non-criminal forms of cyberbullying. Rising rates of anxiety, depression, and body-image disorders have been correlated with intensive platform use.[7] Children's digital experiences exist along a spectrum of online harms that requires differentiated responses. International frameworks, including the OECD's Typology of Online Harms (2021), the EU's Strategy for a Better Internet for Children (2022), and UN Special Rapporteur recommendations,[8] recognize this complexity and urge countries to combine legal enforcement with preventive and educational measures.

India's approach to child online safety must acknowledge three distinct but interconnected categories, each demanding different stakeholder responses:



Category 1: Criminalized Harms

Before turning to emerging harms, it is essential to anchor the discussion in the traditional domain of criminal offences. These offences remain the sharpest edge of online risk and continue to grow in scale and sophistication.

Clearly unlawful acts such as child sexual exploitative and abuse material ***(CSEAM)***, trafficking, and exploitation, prosecutable under laws like India's POCSO Act and IT Act. Data from the National Crime Records Bureau ***(NCRB)*** show a 32 percent rise in cybercrimes against children between 2021 and 2022. Overall crimes against children surged by more than 80 percent between 2014 and 2022, with POCSO cases alone accounting for 63,414 incidents in 2022. Experts caution that these figures likely represent only a fraction of actual victimization, given the cultural stigma and the under-reporting of digital abuse.

Category 2: Threshold Harms

Behaviours that may start as non-criminal interactions but can escalate into offences once they cross legal thresholds. Examples include grooming that progresses into sexual solicitation or exploitation, and cyberbullying which, though not defined as a distinct offence, may attract liability when its conduct satisfies the ingredients of , for instance, criminal intimidation or stalking under BNS. These occupy a grey zone requiring heightened vigilance from parents, educators, and platforms, alongside law enforcement readiness for timely intervention.

Category 3: Emerging Harms

Experiences and interactions that may not attract penal consequences yet can cause significant psychological, emotional, or developmental harm to children. These include algorithmic amplification of harmful content, manipulative design patterns, exposure to age-inappropriate violent or sexual content, body-image pressures, and exposure to extremist ideologies. Addressing them requires stronger platform accountability, informed parental and institutional engagement, and preventive regulatory measures focused on design integrity and digital well-being rather than prosecution.

While all three categories require coordinated multi-stakeholder action, the locus of primary responsibility shifts: law enforcement leads in Category 1, while prevention networks—schools, families, platforms, and civil society—lead in Categories 2 and 3.

This report provides a high-level overview of the spectrum of online harms affecting children and adolescents, outlining how each category manifests and where primary responsibility to act lies. Using case studies, it shows why no single actor—law enforcement, platforms, schools, or families, can address these harms in isolation. The report highlights coordination challenges that cut across categories and shape real outcomes for children. It closes by emphasizing the need for a coherent, multi-stakeholder framework that brings together prevention, response, and long-term safeguards in the context of India's new digital adolescence.

02 Legal and Institutional Frameworks

India's approach to child online safety has developed through a mix of legal instruments, judicial interventions, and institutional initiatives. This chapter provides an overview of the existing landscape across three broad categories of harm: Criminalized Harms (Section 2.1), Threshold harms (Section 2.2), and Emerging harms (Section 2.3). Each category calls for a distinct but shared response across stakeholders. For criminalized harms, law enforcement plays a central role alongside platforms, families, and support services. Threshold harms require early detection, coordinated intervention, and strong collaboration between schools, parents, platforms, and child protection actors. Emerging harms call for systemic prevention that involves platform accountability, digital literacy efforts, safer design practices, and community support.

2.1 Criminalized Harms

Clearly unlawful acts such as CSEAM, trafficking, and exploitation, prosecutable under laws like India's POCSO Act and IT Act fall under this category. Data from the NCRB show a 32 percent rise in cybercrimes against children between 2021 and 2022.[9] Overall crimes against children surged by more than 80 percent between 2014 and 2022, with POCSO cases alone accounting for 63,414 incidents in 2022.[10] Experts caution that these figures likely represent only a fraction of actual victimization, given the cultural stigma and the under-reporting of digital abuse.

India's response to such crimes is anchored in multiple, interlocking statutes, including the POCSO Act, the IT Act, and relevant provisions of the Bharatiya Nyay Sanhita ***(BNS)***.

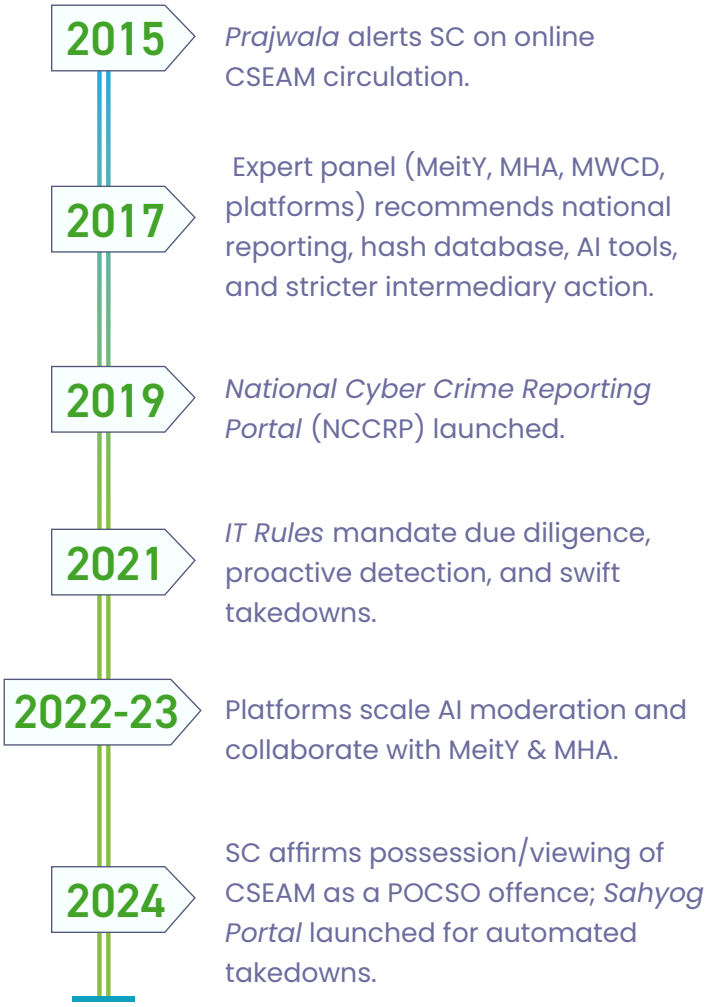
Law	Relevant Provisions	Scope of Protection
Information Technology Act, 2000	Sec. 67 – Publication or transmission of obscene material; Sec. 67A – Publication of electronic material containing sexually explicit acts; Sec. 67B – Publishing or transmitting material depicting children in sexually explicit acts; Sec. 66E – Violation of privacy; Sec. 69A – Blocking of websites/apps hosting illegal content.	Criminalizes online CSAM, empowers the government to block unlawful content.
Protection of Children from Sexual Offences Act, 2012 (POCSO)	Sec. 11–12 – Sexual harassment and online grooming; Sec. 13–15 – Use of child for pornographic purposes; Sec. 16 – Attempt to engage in sexual solicitation or grooming; Sec. 19–20 – Mandatory reporting obligations for intermediaries.	Defines and penalizes online sexual exploitation and grooming; creates duties for service providers to report offences.
Bharatiya Nyaya Sanhita, 2023 (BNS)	Sec. 98-99 - Child prostitution, 143-145 - Trafficking of minors, Sec. 294–295 - obscene content, Sec. 75 - sexual harassment, Sec. 78 - stalking; Sec. 96 - Procuring a child; Sec. 79 - Insulting modesty online.	Extends liability to digital/online offences against children including trafficking and harassment.
Juvenile Justice (Care and Protection of Children) Act, 2015	Sec. 75–77 – Cruelty to children, use of children for begging or drug trade; covers digital platforms facilitating such exploitation.	Addresses online facilitation of child exploitation.
Indecent Representation of Women (Prohibition) Act, 1986 (amended scope online)	Prohibits depiction of women in indecent or derogatory manner, which can include minors on digital platforms.	Supplements IT Act/ POCSO in tackling online indecent representation.

Stakeholder Coordination in Criminalised Harms

The legal framework establishes clear criminal liability for online exploitation, but law alone does not guarantee protection. Its effectiveness depends on coordination across institutions that each hold a piece of the enforcement chain. Police cyber cells require functional digital forensic capacity; platforms must detect and act on abusive content; prosecutors and judges need familiarity with technological evidence; and child protection agencies must ensure trauma-informed support for survivors. On the ground, however, these systems often operate in silos, producing uneven enforcement and significant gaps in response.

This fragmentation was at the heart of Prajwala v. Union of India, which began with a letter alerting the Supreme Court to the widespread circulation of rape and sexual assault videos involving women and children. By treating the letter as a suo motu petition, the Court acknowledged that piecemeal actions were insufficient. The case catalysed a coordinated push to curb CSEAM circulation and strengthen accountability across state agencies and online intermediaries, underscoring that criminal provisions require institutional alignment to work in practice.

Understanding the context and implementation of existing criminal law provisions on online offences is a necessary starting point before turning to the other categories of harm, for three reasons:



01 **It defines the baseline for child protection.**

Criminal provisions establish the *floor,* not the *ceiling,* of child safety online. They target the most severe harms but were never designed to address the broader spectrum of digital risks children face today.

02 **It exposes systemic enforcement gaps.**

Persistent weaknesses in digital forensic capacity, law-enforcement training, and the uneven functioning of Special POCSO Courts continue to limit the effective investigation and prosecution of offences.[11] These challenges are compounded by cross-border complexities and the chronic under-reporting of child sexual offences, reflecting deeper institutional constraints that also shape India's broader response to non-criminal online harms.

03 **It underscores the limits of reactive approaches.**

Even where strong laws exist, intervention occurs *after* harm has taken place. Effective protection requires a shift toward prevention and early risk mitigation, rather than reliance on post-facto criminal enforcement alone.

These insights also underscore the need to give equal attention to **threshold and emerging harms**— they have the potential to affect vastly more children, operate outside prosecutorial reach, and demand fundamentally different response mechanisms centered on prevention, coordination, and systemic accountability.

2.2 Threshold Harms

Between clearly criminal acts and lawful-but-harmful experiences lies a particularly challenging category: threshold harms. These are behaviours that approach or in some cases cross criminal thresholds, yet often evade prosecution due to their subtle nature, gaps in legal definitions, or evidentiary challenges and insufficient systems to detect and prevent.

Threshold harms are distinct because:

- Unlike criminalized harms, they often lack the explicit evidence or clear legal elements required for prosecution
- Unlike emerging harms, they involve intentional targeting of individual children and carry direct risk of escalation to serious criminal acts
- They operate in legal gray zones where harm is evident but remedies are unclear

This ambiguity makes threshold harms particularly dangerous. By the time behaviour crosses into clearly criminal territory, significant psychological damage has often occurred. Effective protection requires early detection and intervention before harms escalate, rather than waiting for prosecutable offences to materialize. Both begin as seemingly ordinary online interactions that can escalate into significant psychological harm or even criminal conduct, underscoring why these are considered threshold harms and why coordinated vigilance across stakeholders is essential.

Child Grooming: The Unaddressed Threat

Child grooming[12] is a deliberate and manipulative process through which an adult builds an emotional connection with a child to gain their trust and so they can make them do what they want and abuse them. It can lead to sexual abuse, exploitation and trafficking.

It typically unfolds in identifiable stages:[13] gradual psychological conditioning, desensitising the child to inappropriate behaviour, isolating them from protective relationships, and instilling dependence or secrecy. Unlike overt sexual abuse, grooming operates subtly, often masked as friendship, mentorship, or affection.

Studies[14] have shown that such offenders exploit children's emotional vulnerabilities, using compliments, shared interests, and emotional support to establish trust before introducing sexualised conversation or coercion. While existing Indian laws like the POCSO Act, 2012 and the IT Act, 2000 address explicit sexual offences against children, they do not expressly recognise "child grooming" as a standalone offence. The subtle process of gaining a child's trust through emotional manipulation, which often precedes any illegal act, remains largely unaddressed. Under POCSO, Sections 11(iv) and 11(vi) penalise acts showing sexual intent, while Section 67B(c) of the IT Act addresses online inducement of minors. However, both provisions apply after intent becomes evident or communication turns sexual.

The gap is conceptual as much as legal. Grooming inflicts deep psychological harm—guilt, fear, anxiety, difficulty trusting adults, disrupted relationships, even before any sexual act occurs. In severe cases, it correlates with self-harm and substance abuse. Yet because the harm is psychological rather than physical, and the process unfolds gradually rather than in discrete criminal acts, it often remains invisible until exploitation has occurred.

Cyberbullying: When Harassment Escalates

Most everyday online friction does not rise to the level of threshold harm. Occasional mean comments or short-lived disputes, while distressing, typically fall below the threshold. This becomes **threshold harm** when the behaviour is persistent, targeted, and produces measurable psychological distress (including self-harm) or functional impact. Yet prosecution is complex because perpetrators may also be minors, making the justice system reluctant to criminalise peer behaviour. Evidence is also dispersed across private chats, group messages, and offline interactions, which makes it difficult to establish a clear and continuous pattern of abuse. Compounding this, adults frequently minimize such conduct as harmless peer conflict or "kids being kids," further discouraging formal reporting and intervention.

The key challenge lies in distinguishing this from normative peer conflict. Indian law offers only partial remedies. **BNS** provisions on stalking (Section 78) and harassment (Section 75) primarily target adult offenders and hinge on proving criminal intent. The **POCSO Act** addresses sexual harassment of minors but excludes non-sexual bullying, however severe. The **IT Act** focuses on specific unlawful content, not on the cumulative harm caused by coordinated online abuse.

Institutional responses are often fragmented. When bullying crosses online–offline boundaries, schools, platforms, and law enforcement frequently pass responsibility between one another. Parents report incidents to schools; schools defer to platforms; platforms assess content in isolation; police decline to act absent a clear offence—while the affected child continues to experience escalating distress.

Stakeholder Coordination in Threshold Harms

The defining characteristic of threshold harms is that no single stakeholder can address them alone:

- Parents and educators are often the first to notice behavioural changes but may lack expertise to assess legal risk or know when to involve authorities.
- Platforms can detect patterns (age gaps in conversations, use of grooming language, sustained harassment campaigns) but cannot determine legal thresholds or provide victim support.

- Law enforcement has investigative authority but typically cannot act until behaviour becomes clearly criminal, and may lack resources for proactive monitoring.
- Civil society organizations can provide victim support and education but have no enforcement authority.

Effective response requires information-sharing protocols, clear escalation pathways, and coordinated prevention efforts that help all stakeholders recognize early warning signs and respond proportionately before harm escalates.

2.3 Emerging Harms

Beyond defined criminal offences lies a much broader ecosystem of online risks that affect children daily. Experiences that may be lawful yet psychologically damaging, or inadequately addressed by existing legal frameworks. These include algorithmic amplification of body-image disorders, exposure to age-inappropriate violent content, manipulative design features hijacking attention, and non-criminal forms of cyberbullying using coded language that evades moderation.

Emerging harms occupy a distinct space in the child online safety landscape. These harms are systemic rather than individual, often rooted in platform design choices, algorithmic systems, and commercial practices that shape how children interact with digital environments. Emerging harms also tend to be cumulative, developing gradually through repeated exposure rather than arising from a single incident. Moreover, they lack clear perpetrators, with responsibility dispersed across platform companies, advertisers, content creators, and the algorithms that determine what users see and engage with.

The ecosystem is complex, ever evolving, and includes lawful forms of harm that shape children's digital experiences. What begins as a seemingly innocuous interaction online — a comment on a post, a request, or a direct message, can rapidly escalate into psychological trauma, physical danger, or irreversible harm. From gaming chatrooms where predators cultivate trust before moving to encrypted apps, to algorithms that push pro-anorexia or self-harm content to vulnerable teenagers, children navigate digital spaces that blur the line between safety and exploitation. These online triggers often spill into real-world consequences: depression, eating disorders, violence, or in severe cases, trafficking and suicide.[15] Exposure to age-inappropriate content can also lead to premature sexualisation, desensitisation to violence, and distorted understandings of relationships and social norms.[16]

India has developed a framework of regulatory measures, self-regulatory codes, and educational initiatives to address these harms. The table below summarizes major categories of emerging harm affecting Indian children (detailed typology provided in **Appendix 1**):

Harm Category	Examples	Key Risks	Primary Response Mechanisms
Content-Related Harms	Age-inappropriate content, cyberbullying, body-image content, extremist ideological content, deceptive influencer content	Desensitization, anxiety, premature sexualization, eating disorders, polarization	IT Rules due-diligence, content moderation, age-gating, school protocols, media literacy
Commercial Exploitation & Design Harms	Manipulative advertising, dark patterns, excessive data collection, gaming monetization	Overspending, loss of autonomy, privacy loss, compulsive use	CCPA/ASCI enforcement, DPDP Act restrictions, platform design audits, parental controls
Behavioural & Wellbeing Harms	Excessive screen time, compulsive use, geosocial risks	Sleep loss, attention issues, academic decline, safety risks	Time-use tools, wellbeing prompts, school routines, privacy defaults

Stakeholder Coordination in Emerging Harms

While these frameworks exist, their effectiveness depends on coordination. The Delhi High Court's Y V v. Kendriya Vidyalaya17 ruling illustrates why consultative, coordinated approaches matter even at the school level. Rather than imposing a blanket smartphone ban, the Court recognized that effective digital safety requires policies developed with input from parents, teachers, and experts, including:

- Regulated and monitored use rather than blanket bans
- Safe storage facilities during school hours
- Digital literacy education on responsible use and risks
- Fair, proportionate disciplinary measures

The principle that protection requires coordination across stakeholders rather than unilateral action—applies across all emerging harms. Platforms must design safely; schools must educate; parents must engage; regulators must enforce; and researchers must provide evidence. When these systems work in isolation, children fall through the gaps. When they coordinate, protection becomes possible. The case studies in Chapter 3 illustrate how these emerging harms manifest in children's lives, why they matter despite their legal status, and what coordinated responses might achieve.

03 Understanding the Spectrum of Harms: A case-study approach

The three-category framework outlined in Chapter 2 becomes concrete when examined through real cases affecting Indian children. This chapter presents case studies organized by harm category, illustrating how different types of online risks manifest, escalate, and demand different responses.

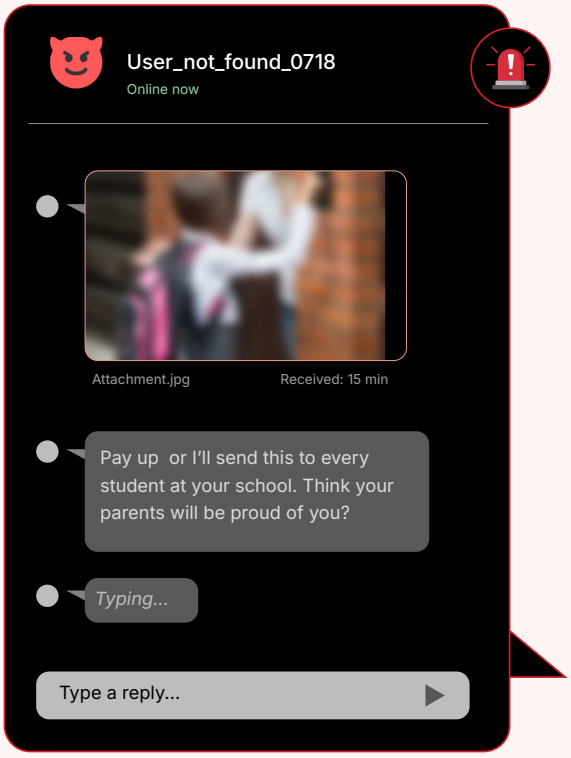
Each case identifies:

- Primary harm category (Criminal, Threshold, or Emerging)
- Key stakeholders whose coordination is needed
- Why it matters for policy and practice

Case: Sextortion and Financial Exploitation- AI

What Happened

A 16-year-old boy died by suicide after falling victim to an AI sextortion scam.[18] He received threatening messages demanding $3,000 to prevent circulation of an AI-generated nude image of him. His family described the perpetrators as "well-organized, well-financed, and relentless," warning that such scams increasingly target minors globally.



Why it matters

AI-powered sextortion represents a new threat vector: perpetrators no longer need actual compromising images; they can fabricate convincing deepfakes from innocuous social media photos, then extort victims with threats of distribution. This not only lowers the barrier for exploitation but also amplifies psychological harm. Victims struggle with both the fabricated material and the helplessness of defending against synthetic evidence.

Key stakeholders and coordination gaps

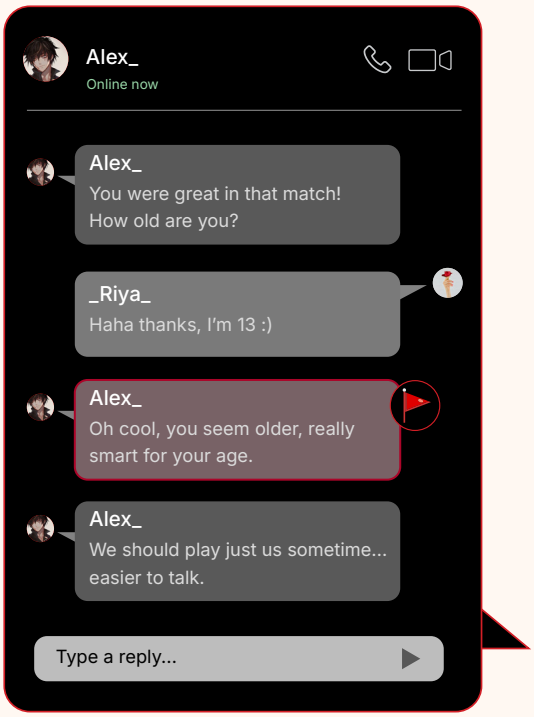
Effective response to sextortion requires coordination between law enforcement agencies (to investigate extortion networks that operate across borders and exploit anonymising technologies), social media and messaging platforms (to detect suspicious account activity, sudden shifts to private communication channels and patterns of synthetic image misuse), parents and educators (to help young users recognise early signs of coercion and safely disclose distress), and mental health professionals (to support victims experiencing panic, shame or suicidal ideation triggered by the threat of fabricated explicit content).

The gap: Platforms can identify unusual spikes in contact from unknown accounts or repeated attempts to move minors to encrypted channels, yet current systems rarely detect the use of AI-generated sexual images when they are shared privately as part of an extortion attempt. Law enforcement capability to trace perpetrators is limited by the speed with which scams operate and the global nature of these networks. Schools have little guidance on how to address the emotional fallout of sextortion attempts, particularly when no real image exists but the psychological impact is severe. Referral pathways for crisis counselling remain inconsistent, meaning children often confront threats alone until the situation escalates.

Child Grooming: The Unaddressed Threat

What Happened

Fourteen-year-old Riya began playing an online multiplayer game where she met "Alex," who claimed to be 16. Over several weeks, Alex complimented her gaming skills, shared personal stories about family problems, and offered emotional support when Riya mentioned stress about school exams. The conversations gradually shifted from game strategy to personal topics. Alex suggested they move to a private messaging app "to talk more freely without other players interrupting." Once on the encrypted platform, Alex revealed he was actually 28 and began sending increasingly inappropriate messages, eventually requesting photos and suggesting they meet in person. When Riya refused, Alex threatened to share their private conversations with her parents and school friends, manipulating screenshots to make it appear she had initiated the inappropriate exchanges.



Why it matters

Grooming operates in legal gray zones. The methods and dynamics of child grooming have evolved, with perpetrators exploiting social media, gaming platforms, and anonymous chat applications, to reach children with ease and often without detection. Grooming inflicts deep psychological and emotional harm even before any sexual act occurs. Victims often experience guilt, fear, anxiety, and difficulty trusting adults or forming healthy relationships. In severe cases, it can lead to self-harm or substance abuse.[19] As it operates in secrecy, disguised as harmless conversation or mentorship, even the most attentive parents may struggle to spot it. Many parents also lack the digital literacy to effectively monitor online interactions or recognise subtle psychological cues. Protecting children, therefore, demands a coordinated, multi-stakeholder effort involving families, platforms, educators, and law enforcement.

Key stakeholders and coordination gaps

Effective response requires gaming platforms (detecting age-disparate contact patterns and conversation escalation), parents and educators (recognizing warning signs like secretive device use and emotional withdrawal), law enforcement (investigating cross-platform grooming), and mental health services (supporting children experiencing grooming).

The gap: Platforms can detect grooming-like patterns, but detection is probabilistic and platforms' internal escalation practices vary; lack of clear protocols for when to escalate concerns to authorities. Schools have no standardized training to help teachers recognize grooming indicators. Mental health referral pathways between these stakeholders are largely absent, meaning intervention typically occurs only after criminal behaviour has materialized.

Online Harassment and Bullying

Ajey died by suicide after being subjected to sustained online bullying for wearing sarees and makeup in their social media videos.[20] The student, who identified as queer, regularly posted short clips expressing themselves through fashion and performance.

What began as personal and creative expression soon drew ridicule, with derogatory comments and homophobic slurs appearing across their social media handles. The harassment intensified as the videos were reshared in local social media circles and messaging groups. Much of the bullying took place in colloquial Hindi and regional slang, using coded expressions that automated moderation systems did not detect. The content spread rapidly, often stripped of context and sensationalised, turning the student into the subject of widespread ridicule.

Why it matters

The content, while not always explicitly illegal, mocked and targeted the student in ways that blurred the line between free expression and harm. As the videos spread, ridicule became sensationalised and viral, reinforcing social prejudice and silencing self-expression.

Key stakeholders and coordination gaps

Effective response to online harassment and bullying requires close coordination between social media platforms (to detect harassment trends, virality spikes, and targeted brigading), parents and educators (to notice behavioural changes such as withdrawal, fear of online spaces, or declining academic performance), law enforcement (to intervene when harassment crosses into criminal intimidation or abetment), and community-based mental health services (to support victims experiencing anxiety, shame, or trauma).

The gap: Most platforms' moderation systems struggle with colloquial abuse, regional slang, coded homophobic expressions, and the nuanced ways ridicule spreads through quote-posts, duets, and localised group chats. Platform escalation channels for repeated but "borderline" abusive content remain inconsistent and often depend on user reporting, which places the burden on already vulnerable children. Schools generally lack structured protocols for responding to cyberbullying that occurs outside the classroom but affects a student's wellbeing inside it. Coordination with law enforcement tends to occur only when harm becomes extreme, and mental health support rarely forms part of the initial response, leaving victims without early intervention despite visible signs of distress.

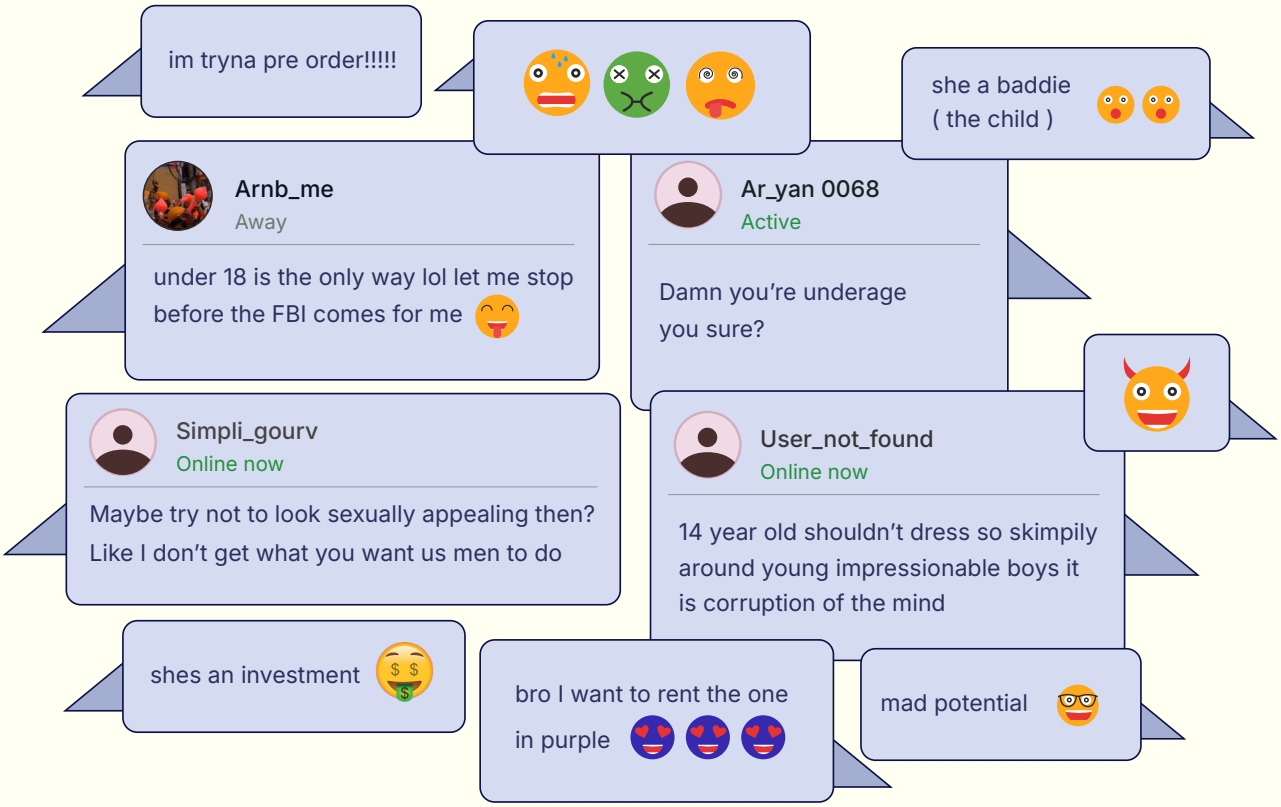
Objectification and Unsolicited Attention[21]

Minor girls frequently face inappropriate and unsolicited comments from adult men across social media platforms. These interactions often begin under harmless posts (selfies, dance clips, Get Ready With Me ***(GRWM)*** videos) but quickly turn invasive. While not always crossing the legal threshold for obscenity, the cumulative impact is corrosive.

Why it matters

A study collected and analysed a wide range of misogynistic comments and posts from social media platforms to identify recurring themes, language patterns, and narratives.[22] Content that mocks or targets minor girls can gain momentum, spread virally, and become sensationalized, drawing attention far beyond the original circle of communication. Cultural norms, gender biases, and the design of online platforms converge to create environments where minor girls are routinely objectified, and their self-expression is scrutinized, policed, and shamed.

These interactions signal emerging risks - patterns of abuse, coercion, and objectification, which have real-world psychological and social consequences. Addressing these harms requires a combination of platform responsibility, digital literacy education, parental and school engagement, and broader cultural interventions to challenge normalized harassment and protect the dignity of young users.



Key stakeholders and coordination gaps

Cooperation between social media platforms (to detect age-inappropriate interactions and identify repeated patterns of adult engagement with minors), parents and educators (to recognise early signs of discomfort, withdrawal, or changes in online behaviour), law enforcement (to investigate persistent predatory attention or escalating harassment), and mental health professionals (to support young users who experience fear, shame, or reduced self-esteem) is required.

The gap: Platforms can identify clusters of concerning behaviour, such as adults disproportionately commenting on or interacting with the accounts of minor girls, but current detection models often miss context-specific cues, subtle grooming-like attention, and comments that fall just short of obscenity. Reporting tools are fragmented and place significant responsibility on minors to flag behaviour that they may not fully understand or feel confident challenging. Schools typically lack structured guidance on how to address online objectification that occurs outside the classroom but affects a student's sense of safety and identity within it. Referral pathways to mental health support remain inconsistent, which means young users often internalise harm long before adults or authorities intervene.

Psychological Harm from Violent and Disturbing Content

Anisha, a middle-school student, began suffering from anxiety, nightmares, and difficulty concentrating after binge-watching crime and thriller videos online. She would stay up late consuming such content and developed a persistent fear of being attacked or followed.

What happened

Platform algorithms noticed Anisha's interest in crime videos and began recommending increasingly graphic material. These reels were designed to capture attention by using sensational headlines, dramatic text overlays, and emotionally charged thumbnails that prompted viewers to keep clicking. No alerts or safeguards flagged her excessive use or distress. Within weeks, constant exposure and sleep deprivation triggered anxiety severe enough to affect her schoolwork and daily life.

Why it matters

This harm lies outside the reach of criminal law.No content was illegal. No perpetrator exists to prosecute. No content violated any statute, and no perpetrator can be prosecuted. Yet a child suffered psychological harm — revealing how legal-but-harmful content and design features can cause real emotional injury.

Key stakeholders and coordination gaps

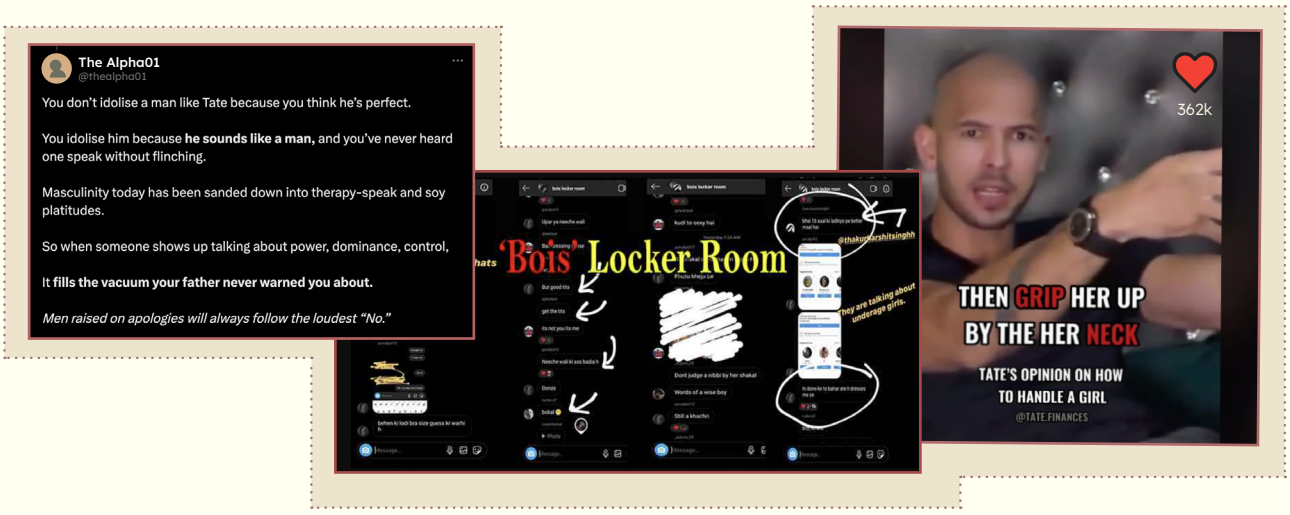
Psychological harm caused by violent and disturbing online content requires coordinated involvement from platforms (to identify binge-watching patterns, escalating recommendations and signals of distress), parents and educators (to observe changes in behaviour such as sleep disruption, irritability, fearfulness or avoidance), mental health professionals (to address anxiety, intrusive thoughts and the long-term effects of overstimulation), and child protection and digital literacy organisations (to build awareness about the risks of excessive exposure to sensational content).

The gap: Platforms can monitor viewing intensity, late-night usage and rapid consumption patterns, yet most systems do not flag when a child is repeatedly exposed to anxiety-inducing or graphic content. Schools lack structured guidance on how to recognise when online content consumption is affecting a student's emotional state or academic performance. Referral pathways to age-appropriate mental health support remain fragmented, resulting in children receiving help only after symptoms worsen and daily functioning is disrupted.

Exposure to Radical/Extremist Ideological Content

Across schools, teachers are noticing a growing range of harmful behaviours among students, including misogyny, sectarian attitudes, and radicalized political or ideological views. Teachers are witnessing a growing normalization of misogynistic behaviour among boys. Female teachers report being mocked, dismissed as "hysterical," and confronted with statements such as "the pay gap is a myth" or "women lie about assault."[23] Social media narratives often valorise dominance and independence, leading some young users to regard shared or domestic responsibilities as "beta" behaviours inconsistent with online ideals of masculinity.[24]

While boys in particular have been observed normalizing misogynistic behaviour,[25] similar patterns of extremist thought around identity, religion, or politics, are emerging across genders.



Why it matters

Understanding the rise of extremist behaviour among young students is essential because these incidents are rarely isolated—they reflect broader cultural reconditioning driven by digital ecosystems.[26] Social media and streaming platforms have become powerful sites of social learning, where figures such as Andrew Tate and self-styled "sigma males" promote a distorted ideal of masculinity rooted in dominance, emotional detachment, and contempt for women. Similar dynamics exist for other extremist content, where young users encounter narratives that reinforce ideological rigidity, intolerance, or glorified aggression. Algorithms reward outrage and extremity, allowing harmful ideologies to reach millions of impressionable users before counter-narratives or adult guidance can intervene[27]. Several clinical studies show that such digital conditioning manifests offline, through heightened aggression, normalization of sexist humour, and the trivialization of consent and gender equality in schools.[28]

Key stakeholders and coordination gaps

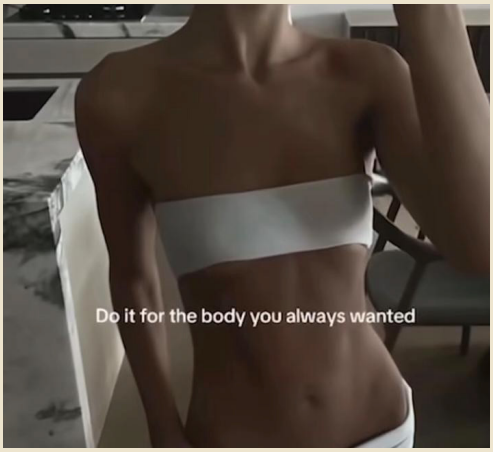
Effective response to the rise of extremist and misogynistic ideological content requires coordinated involvement from platforms (to detect harmful recommendation patterns and reduce the visibility of content that promotes hostility, intolerance or rigid gender norms), educators and school counsellors (to identify behavioural shifts such as increased aggression, dismissiveness, or sudden adoption of extremist slang), parents (to understand online influences shaping their child's worldview), and mental health professionals and community organisations (to support de-escalation, critical thinking, and emotional regulation in affected students).

The gap: Platforms can identify repeated engagement with highly polarizing or radical content, yet current detection systems often miss coded language, satire-laced misogyny and subtle ideological messaging that appeals to adolescents seeking identity and belonging. Schools typically lack structured guidance on how to respond when ideological content manifests as classroom disruption, sexist humour or identity-based hostility. Coordination with parents tends to be reactive, not preventative, and many families are unaware of the digital ecosystems shaping children's beliefs. Mental health referral pathways remain limited, which means interventions often occur only after extremist attitudes have hardened and begun influencing peer relationships, classroom dynamics and overall school climate.

Body Image Issues

What happened

Sixteen-year-old Meera began following fitness influencers on Instagram who promoted "clean eating" and intense workout routines. Within months, her feed filled with content glorifying extreme thinness—"thinspo" posts, before-and-after photos, and diet tips that were actually disordered eating behaviours disguised as wellness advice.[29] She started obsessively counting calories, skipping meals, and exercising excessively. Meera followed influencers with millions of followers promoting the same behaviours, arguing it was "healthy lifestyle choices." Her batchmate, sixteen-year-old Aryan returned from summer vacation visibly muscular, becoming an object of admiration among classmates. His high-protein, low-calorie diet—six eggs daily, protein supplements, no sweets—was quickly replicated by friends, causing their parents concern.




NEW YORK POST

LIFESTYLE

AI defines 'ideal body type' per social media – here's what it looks like

By Angelica Stabile, Fox News
Published May 16, 2023, 4:55 a.m. ET

37 Comments

Why it matters

Constant exposure to idealized and digitally altered imagery breeds unhealthy self-comparison, undermines confidence, and encourages obsessive control over diet, fitness, and appearance.[30] For adolescents still shaping their identities, this often manifests as body dysmorphia, depression, or disordered eating.

The rise of AI-generated influencers has magnified the problem, normalizing hyper-edited, unattainable beauty ideals that blur the boundary between reality and fabrication.[31] In striving to meet these illusions, many young users resort to extreme cosmetic interventions and internalize exclusionary notions of beauty. Recent studies indicate that 30-40% of people with eating disorders are male, driven by media-promoted ideals of lean muscularity and peer competition for physical attention.[32]

Key stakeholders and coordination gaps

Effective response to the rise of body image distress among adolescents requires coordination between social media platforms (to detect patterns of engagement), schools and educators (to recognise early warning signs), parents (to monitor emotional shifts and provide balanced guidance), and healthcare professionals

The gap: Platforms can identify clusters of posts involving restrictive diets, excessive workouts or idealised body transformations, yet content that is harmful while framed as wellness often escapes moderation. Recommendation systems frequently amplify unrealistic aesthetic standards, including AI-generated influencers who promote unattainable ideals without transparency. Schools rarely have structured protocols for addressing body image concerns or for integrating media literacy that helps students critically evaluate what they see online. Communication between families, educators and healthcare providers is inconsistent, which delays the recognition of early symptoms. As a result, intervention often happens only after physical health has deteriorated or psychological distress has intensified, reducing the chances of early recovery.

04 Cross-Cutting Gaps in India's Response to Online Harms

India has developed comprehensive legal and policy frameworks to address online harms affecting children—from criminal prosecution under POCSO and the IT Act, to platform regulations under IT Rules 2021, to educational initiatives through NCERT and NCPCR. Yet between policy and practice lies a persistent implementation gap. Gaps manifest differently across criminalized harms , threshold harms, and emerging harms, yet several systemic weaknesses cut across all categories, undermining protection regardless of harm type.

Coordination: The Central Challenge

The most fundamental gap is institutional fragmentation. MeitY, MWCD, MHA, Ministry of Education, MIB, state governments, platforms, schools, and civil society organizations operate parallel initiatives without systematic coordination mechanisms. This fragmentation manifests across harm categories:

For criminalized harms: The Prajwala case demonstrated what coordination can achieve—hash databases, automated detection, platform cooperation, improved prosecution rates for CSEAM. Yet when exploitation crosses platforms or jurisdictions, coordination breaks down. Mutual Legal Assistance Treaty (MLAT) requests can take months; platform-law enforcement information sharing remains ad hoc; victim support services operate separately from investigation processes.

For threshold harms: When a child is groomed across gaming platforms and messaging apps, or bullied across school WhatsApp groups and social media, institutional responses fragment. Parents report to schools; schools refer to platforms; platforms evaluate content in isolation; police decline to act absent clear criminal offenses. No entity owns the coordinated response before severe harm occurs.

For emerging harms: A teenager algorithmically guided toward pro-eating-disorder content crosses multiple domains—platform recommendation systems, school mental health services, parental oversight, healthcare intervention. Yet these systems rarely communicate. By the time harm surfaces, multiple opportunities for early intervention have passed.

Law Enforcement Capacity and Technical Gaps

Law enforcement agencies face significant capacity constraints that affect response to criminalized harms while also limiting detection of threshold harms before they escalate. Research examining India's fight against online child sexual abuse and exploitation documents how law enforcement operates under workload pressures, with uneven access to digital forensic tools, specialized training, and technical expertise across jurisdictions.[33] Many cyber cells lack infrastructure to trace perpetrators across encrypted platforms or recover digital evidence meeting evidentiary standards.[34]

These capacity gaps have cross-category effects:

- **Criminalized harms** remain under-investigated due to technical limitations, particularly those involving sophisticated encryption or cross-border elements.
- **Threshold harms** like grooming go undetected because law enforcement lacks resources for proactive monitoring and can only respond after behaviour becomes clearly criminal.
- **Emerging harms** receive minimal law enforcement attention since they fall outside criminal mandates, yet early identification could prevent escalation.

Platform Limitations in Content Moderation and Detection

While platforms have strengthened moderation systems, significant limitations persist across all harm categories.

Linguistic and cultural gaps: In India, platforms grapple with over 22 languages and numerous dialects, each with unique slang, code-switching patterns, and cultural references. Research on automated moderation pipelines for low-resource languages documents how automated tools frequently fail to detect harmful content in these vernaculars, leading to false negatives and inconsistent enforcement.[35] The rise of "algospeak" where users intentionally modify language to evade filters, further complicates detection.

This affects:

- **Criminal content** (CSEAM, exploitation) using coded language in regional dialects
- **Threshold harms** (grooming, severe bullying) that employ vernacular slang and coded expressions
- **Emerging harms** (objectification, extremist content) that normalize harmful attitudes through culturally-specific language

Infrastructural constraints:[36] Platforms often lack sufficient trained moderators to review vast content volumes. This shortage leads to delayed responses and potential oversight. Additionally, reliance on automated systems without adequate human oversight can amplify biases and fail to contextualize nuanced content.

Legal Definition and Framework Gaps

Current laws were designed primarily for criminalized harms , leaving threshold and emerging harms inadequately addressed. POCSO and IT Act provisions address explicit sexual solicitation but not the preparatory psychological manipulation characterizing grooming. This creates uncertainty about when intervention is legally justified. For cyberbullying, laws focus on specific illegal content rather than cumulative patterns of harm, leaving severe peer-to-peer harassment unaddressed when individual messages don't meet criminal thresholds.

There are no specific regulations on algorithmic transparency, design features causing compulsive use; age verification and parental control requirements exist but lack enforcement mechanisms or technical standards. Emerging technology risks like deepfake technologies enabling synthetic CSEAM, AI-driven manipulation through LLM-powered chatbots, generative AI lowering barriers to exploitation, affect all harm categories. Current frameworks lack provisions addressing these novel risks.

Awareness, Reporting, and Cultural Barriers

Cultural stigma and privacy concerns discourage reporting across all harm categories, limiting the effectiveness of even strong legal frameworks. Research on child digital safety policy in India notes how families avoid engaging formal complaint mechanisms.[37] NCRB data on cybercrimes against children likely represents only a fraction of actual victimization.[38] Without reporting, enforcement mechanisms cannot function.

NCERT/CIET cyber-safety initiatives and NCPCR school guidance provide frameworks, but implementation varies dramatically across states and between government and private schools. Rising digital literacy gaps hinder recognition of warning signs for grooming, ability to respond to cyberbullying, and understanding of emerging harm risks. Additionally, social issues such as objectification of girls, toxic masculinity narratives, attitudes normalizing harassment, require both platform moderation and cultural change around acceptable behaviour. Technical solutions alone are insufficient.

Institutional Fragmentation and Service Delivery Gaps

Even when harm is detected, fragmented service delivery undermines effective response.

For criminalized harms: Research on special courts for children documents persistent weaknesses in the uneven functioning of Special POCSO Courts—delays in case resolution, inadequate victim support infrastructure, inconsistent application of child-friendly procedures.[39]

For threshold harms: Children experiencing grooming or severe cyberbullying need mental health support, yet referral pathways between schools, platforms, law enforcement, and counseling services are largely absent. No coordination mechanism ensures holistic response.

For emerging harms: When algorithmic amplification contributes to eating disorders or self-harm ideation, response requires platform intervention, school recognition, parental engagement, and healthcare support, yet these systems rarely coordinate.

Civil society organizations providing victim support, education, and advocacy operate in fragments, competing for limited resources rather than coordinating responses. Schools lack teacher training and time for digital safety implementation. Healthcare services are not equipped to address digital-trigger mental health issues.

Evidence-to-Practice Translation Gaps

Research increasingly documents specific harms, yet translation to practice lags across categories.

Studies link social media intensity to mental health outcomes, document algorithmic amplification of eating disorders, show correlation between design features and compulsive use.[40] Despite this evidence, engagement-optimising algorithms remain widespread and digital-literacy programmes are often scaled without rigorous evaluation. India further lacks comprehensive, nationally comparable data on children's digital experiences, harm prevalence, or intervention outcomes — limiting the ability to design and target evidence-based policy. Harms usually evolve faster than institutional responses. By the time one manifestation is addressed, new forms emerge using different language, platforms, or techniques. This affects all categories but is particularly acute for emerging harms.

05 Building a Coordinated Framework - The Path Forward

The preceding chapters have documented India's evolving child online safety landscape—legal frameworks spanning criminalized harms, threshold harms, and emerging risks; case studies illustrating how these harms manifest; and systemic gaps revealing why existing measures fall short despite comprehensive policy development. The core finding is clear: **India's challenge is not absence of frameworks but fragmentation in their coordination.**

No single institution possesses the mandate, resources, or expertise to address the full spectrum of online harms. Law enforcement brings legal authority but struggles with technical capacity and cross-border coordination. Platforms possess technical sophistication, but the pace and complexity of emerging issues (evolving references, and circumvention of safety measures) mean that not all challenges can be addressed through platform governance alone. Schools interact with children daily but lack frameworks to address digital risks. Parents want to protect their children but feel overwhelmed by platforms they barely understand. Civil society organizations offer specialized services but operate in fragments, often competing for limited resources rather than coordinating responses.

Ensuring child safety online therefore requires a shift from isolated actions to an integrated protection architecture.

The WeProtect Model National Response Framework

The WeProtect Global Alliance, a global movement bringing together governments, private sector, and civil society to end child sexual exploitation and abuse online, has developed a Model National Response framework[41] that provides valuable architecture for such coordination. While not prescriptive, it offers a structured approach to assessing where gaps exist and how different stakeholders can work in concert.

The framework identifies six interconnected pillars, each requiring specific stakeholder engagement:

1. Policy and Legislation

Comprehensive legal frameworks to address emerging technologies, with clear coordination mechanisms across ministries to avoid fragmentation and duplication.

2. Criminal Justice Response

Law enforcement capacity (digital forensics, specialized training, victim-sensitive procedures), efficient prosecution, and judicial expertise in handling online exploitation cases.

3. Industry Engagement and Technology

Platform accountability through safety-by-design principles, robust content moderation, age-appropriate protections, and transparent cooperation with law enforcement.

4. Victim Support & Empowerment

Expand NGO capacity for victim assistance, hotlines, community education, and advocacy for systemic reforms, with coordination to avoid duplication. Provide specialized services for survivors, including trauma-informed counseling, legal aid, educational continuity, and long-term recovery support.

5. Society & Culture

Integrate digital citizenship into school curricula, deliver parent and community digital literacy programs, run awareness campaigns, and establish research infrastructure to generate evidence on effective interventions.

6. Research & Data

Develop systems to collect, analyze, and share data, informing policies, improving interventions, and guiding stakeholder actions.

Mapping Stakeholders to Coordinated Response

To operationalise this framework, clarity on who owns what responsibility is essential. The table below situates India's key actors within the six WeProtect pillars, highlighting their primary mandates and existing roles.

Stakeholders	Core Roles & Responsibilities	Illustrative Entities (non-exhaustive)
State / Government ***(Policy and Legislation)***	Set policy, notify rules/ standards, align centre–state action, monitor outcomes.	**MeitY** (IT Rules/intermediaries), **MWCD** (child protection policy), **MIB** (content/ad advisories), **DOE/School Ed. Depts.** (state), **NITI Aayog** (strategy), **MHA** (coordinates cybercrime cells/ portals), **I4C** (multi-agency coordination).
Law Enforcement ***(Criminal Justice Response)***	Investigate & prosecute offences , digital forensics, MLAT requests,	**State Police cyber cells, CBI/ IB/CID, Public Prosecutors, Judiciary, NCRB, CERT-In** (incident coordination)
Regulatory & Statutory Bodies ***(Policy and Legislation)***	Issue sector guidelines, enforce compliance, adjudicate complaints, order takedowns/ penalties.	**DPDPA Authority** (proposed), **CCPA** (ads & dark patterns), **NCPCR/SCPCRs, ASCI** (self-reg), **TRAI** (telecom).
Industry / Technology Platforms ***(Industry Engagement and Technology)***	Safety-by-design, moderation & takedown, age-gating/ parental tools, transparency reporting, lawful data practices, cooperation with LEAs.	**SSMIs** (social, search, video, messaging), **ISPs/TSPs, App stores, Cloud/CDNs, Gaming firms ,** Industry Associations such as **IAMAI, BSA, FICCI, NASSCOM**
Civil Society & Hotlines ***(Civil Society Mobilization and Awareness and Victim Support)***	Helplines/intake, victim support & referral, OSINT/flagging, community education, watchdog & advocacy, survivor networks.	**Childline 1098, ECPAT members, INHOPE partners, Bachpan Bachao Andolan, Prerana, HAQ, ARROW**

Schools & Educators ***(Society and Culture)***	Reporting pathways, digital citizenship curricula, parent engagement, peer-led programs.	**Govt & private schools, SCERTs, KVs/JNVs, Teacher training institutes.**
Parents & Caregivers ***(Society and Culture)***	Home safeguards (device settings, routines), early detection and response,participation in school programs.	**Parent–teacher associations, community groups, parental digital-safety initiatives (e.g., Google's Be Internet Awesome program with PTA involvement).**
Academia & Research ***(Research and Data)***	Evidence generation, measurement frameworks (risk/harms, efficacy), tool evaluations, ethics oversight.	**Universities, IIITs/IITs, Law Schools, think tanks**
International & Multilateral Partners ***(Research and Data)***	Best-practice transfer, capacity support, cross-border cooperation, research, funding.	**WeProtect, UNICEF, UNODC, INTERPOL, ITU/OECD, World Bank**
Media & Professional Bodies ***(Society and Culture)***	Public awareness, standards for reporting on child cases, professional ethics and training.	**Editors Guild, Press Council, newsrooms, broadcasters**

Conclusion

From Fragmentation to Coordination

Effective response requires addressing the distinct needs of criminalized harms, threshold harms, and emerging harms while building coordination infrastructure that enables all stakeholders to function together. The following priorities integrate category-specific actions with cross-cutting coordination mechanisms.

Criminalized Harms: Strengthening Detection, Prosecution, and Cross-Border Cooperation

It is essential to build technical capacity and improve coordination — across forensic labs, cyber cells, and courts. Priority should be given to systematic investment in cyber-investigation infrastructure, the establishment of specialized units for child-abuse material (CSEAM) cases, and developing expertise in emerging threats like encryption and synthetic content. India's Bureau of Police Research & Development (BPR&D) has developed training modules for police and forensic officers, but uptake and consistency remain uneven. Cross-border cooperation also needs more formalisation: India already receives CyberTipline reports (via the US's NCMEC) after its MoU, but deeper integration with global databases and faster MLAT processing would significantly accelerate investigations. Strengthening Special POCSO Courts is equally important — research shows persistent delays, shortage of judges, and inadequate victim-friendly infrastructure.[42] Improving coordination between courts, victim support services, and trauma-informed counseling, along with feedback mechanisms from judicial experiences, would help integrate investigation, prosecution, and support into a unified system rather than isolated silos.

Threshold Harms: Building Early Detection and Intervention Systems

Threshold harms require responses distinct from criminal prosecution. As Chapter 2 documented, these behaviours often lack the explicit evidence required for prosecution yet cause severe psychological harm and carry escalation risk. Addressing them demands early detection systems, clear referral pathways, and integrated mental health support. Platform behavioural analytics can identify grooming patterns—age-disparate contact, conversation escalation, attempts to move communication off-platform, before behaviour becomes criminal. But technical tools alone are not enough — platforms need protocols for escalation to child protection services or law enforcement, while respecting privacy.

In schools, teachers must be trained to spot digital exploitation warning signs (grooming, bullying) and know how to trigger formal referrals. While NCERT's cyber-safety guidelines provide a foundation, systematic teacher training, structured referral routes to counselors or authorities, and parent coordination are still weak. Legal clarity would strengthen responses: many countries have standalone grooming offences (for example, the UK's Sexual Offences Act 2003, Section 15),[43] but India lacks a similarly explicit preparatory offence. Finally, mental-health support must be trauma-informed and well-integrated: children exposed to grooming or harassment should have access to counselors, with clear communication channels between schools, platforms, law enforcement, and mental-health services.

Emerging Harms: Shifting Toward Prevention and Design Accountability

Emerging harms affect the largest number of children but receive the least systematic response. As Chapter 3 illustrated, algorithmic amplification of harmful content, manipulative platform design, and age-inappropriate content exposure cause cumulative psychological harm yet fall outside criminal frameworks. Addressing these issues requires shifting from reactive moderation to preventive approaches that include design accountability, strong digital-literacy systems, and improved linguistic and cultural capability across the online safety ecosystem.

The **DPDP Act (2023)** already lays a strong foundation: it requires verifiable parental consent to process children's data, bans behavioural tracking and profiling-based ads for minors, and obliges "significant data fiduciaries" to carry out impact assessments and audits.[44] But it does not explicitly require algorithmic risk assessments or limits on engagement-oriented recommendation systems, gaps that could be addressed through detailed design guidance or future regulation. The **EU Digital Services Act (DSA)** provides a useful benchmark. Under **Article 28**, very large platforms must assess systemic risks to minors, including risks arising from recommendation systems, and must implement mitigation measures.[45] Platforms should be encouraged to adopt similar principles: algorithmic impact assessments, transparency of recommender systems, and features designed to minimize harm for child users.

Age-assurance must also improve. While the DPDP Act empowers regulation for age-verification, there is a need for clear technical standards and regular compliance checks. Digital literacy also needs systematic scaling. Effective programs focus on practical skill-building, peer-led components, teacher training, and meaningful parental engagement. This is especially important in India's context because parental consent plays a central role in how children's data can be processed under the DPDP Act. Finally, linguistic and cultural capability remains a major operational gap. Automated systems struggle with the extensive linguistic diversity of India, and harmful content often circulates in dialect-specific forms. Strengthening multilingual moderation capacity including India-specific detection tools, local-language human moderators, and partnerships with researchers, will reduce blind spots and help contextualise content decisions.

An ecosystem approach is essential. Rather than relying solely on platforms, we should build an infrastructure where regulators, civil society, researchers, educators, and industry cooperate to evaluate, monitor, and improve safety measures. Continuous dialogue, data sharing (with strong privacy protections), and independent evaluation will help establish a resilient, adaptive system. ***Prevention will not eliminate all risk — but a shared, evidence-based system can substantially reduce harm.***

India's Path: Innovation Through Coordination

India faces unique challenges that create opportunities for innovation- linguistic diversity, scale, rapid digitization, and strong civil society networks enable approaches that could inform global practice.

- Vernacular moderation at scale is one opportunity. Developing AI and human moderation systems for 22+ languages positions India as a potential leader in multilingual content safety. Research collaborations between Indian technology institutes, platforms, and government could produce tools applicable across the Global South, where linguistic diversity is common but technical solutions scarce.

- Community-based support networks leverage India's extensive civil society infrastructure. Organizations like Childline 1098, with presence across districts, provide locally-grounded safety nets that complement centralized systems. Strengthening coordination between these community-level services, schools, platforms, and law enforcement creates responsive networks adapted to diverse cultural contexts—a model potentially more effective than top-down approaches designed for culturally homogeneous societies.

- Culturally-contextualized education recognizing India's family structures, social norms, and communication patterns could prove more effective than imported curricula. Research on culturally responsive digital citizenship education emphasizes adaptation to local contexts rather than universal frameworks. India's diversity demands this approach, creating opportunities to develop varied models applicable to different regions and communities.

Yet realizing these opportunities requires the coordination mechanisms discussed throughout this chapter. Innovation happens not through isolated technical solutions or policy pronouncements but through sustained collaboration among stakeholders with aligned incentives and clear accountability.

Appendices

Appendix 1

Emerging Harms	What it is	Typical manifestations (India context)	Key risks/ effects	Primary response levers (non-criminal)
Content Related Harm				
Age-inappropriate content	Material not illegal but unsuitable for a child's age (violence, sexual themes, profanity).	Short-video feeds, OTT profiles without age-gating, open search results.	Desensitization, anxiety, premature sexualization.	IT Rules due-diligence, age-gating/parental controls, school/ parent guidance, safer defaults.
Cyberbullying & online harassment	Repeated hostile behavior via digital channels.[46]	Class/peer WhatsApp groups, gaming voice chat, social DMs.	Depression, anxiety, school avoidance, self-harm risk.	Platform reporting & takedown, school anti-bullying protocols, counseling, digital citizenship curricula, School-level CBSE/NCERT advisories on cyber safety.
Body-image & self-harm content	Content normalizing extreme dieting, self-harm, or suicide.[47]	Hashtag challenges, "thinspiration," algorithmic recommendations.	Eating disorders, self-injury, suicidality.	Content demotion/ filters, well-being prompts, helpline signposting, media literacy modules.
Extremist/ ideological content	Propaganda/ recruitment that may not cross incitement thresholds.[48]	Meme pages, encrypted forwards, fringe forums.	Polarization, offline risk behaviors.	Risk-aware ranking, counterspeech resources, teacher training on critical media literacy.
Deceptive influencer/ edutainment content	Blurred ad/ editorial aimed at kids without disclosures.[49]	"Unboxings," study-hack channels with undisclosed promos.	Misleading beliefs, undue purchase pressure.	ASCI disclosures, platform label requirements, teacher/parent media-literacy tips.

Commercial Exploitation & Design Harms				
Manipulative advertising & dark patterns	Design nudges that steer children to spend/share/ subscribe.	"Limited-time" pop-ups, auto-opt-ins, influencer ads to kids.	Overspending, loss of autonomy, data over-sharing.	CCPA §8 (children), ASCI Ch. III enforcement, platform design audits, opt-in/ consent reforms.
Datafication & privacy loss	Excessive tracking/ profiling of minors without meaningful consent.	App SDK tracking, location sharing, contact syncing.	Commercial profiling, safety risks, long-term data trails	Data-minimization by default, parental dashboards, school BYOD policies, DPIA-style reviews for kid features.
Gaming-related harms[50]	Non-illegal risks from mechanics, chats, or monetization	Loot-box–like rewards, pay-to-win pressure, toxic chat.	Impulsive spend, exposure to abuse, displacement of activities.	"Responsible gaming" codes, spend/time limits, age ratings, chat filters/moderation.
Behavioural & Wellbeing Harms				
Excessive screen time & compulsive use	Overuse that disrupts sleep, study, or social life.[51]	Infinite scroll, streaks, variable rewards in games.	Sleep loss, attention issues, academic decline.	Time-use tools, friction/pauses, school/home routines, well-being nudges.
Geosocial risks	Location/ stranger-contact exposure that isn't overtly criminal.	Open profiles with location tags; public groups.	Stalking risk, doxxing, unwanted contact.	Privacy-by-default for minors, location-sharing off by default, safety prompts, reporting flows.

Appendix 2

Instrument / Body	What it covers for children	How it works in practice
IT Rules, 2021 (Intermediary Guidelines & Digital Media Ethics Code)	Intermediary **due-diligence:** publish rules; bar users from posting prohibited content (incl. *paedophilic* content); notice-and-takedown; grievance redressal; "**significant**" intermediaries to **endeavour** automated tools for sexual-abuse material; parental-control/age-verification features encouraged.	Sets baseline *platform processes* that reduce children's exposure to harmful content/design without invoking penal law in every instance.
CCPA Guidelines (2022) — Prevention of Misleading Ads & Endorsements	**Section 8:** prohibits ads that **exploit children's vulnerability**, encourage dangerous behavior, or make **unsubstantiated health/educational claims**; limits use of celebrities/influencers in kid-directed ads.	Statutory consumer-protection guardrails for kid-facing marketing across media (including online).
ASCI Code (Self-regulation; 2024 PDF)	**Chapter III (Children)** — ads addressed to children must **not** cause **physical, mental or moral harm,** or **exploit their vulnerability**; specific clauses on unsafe imitation, dealing with strangers, etc.	Industry **self-regulation** used by brands and platforms; widely cited in practice notes and enforcement narratives.
NCERT/CIET (2024) — Cyber-safety in schools	National study + programs to build **cyber-safety awareness** among students/ teachers; supports **school protocols** and classroom practice.	Education-system response: awareness, curricula, trainings; complements platform/advertising measures.
NCPCR guidance for schools	Broad **school-safety** guidance from the child-rights commission; used by states/ DoE to anchor cyber-safety advisories and response protocols.	Basis for school-level circulars (e.g., cyber safety advisories in Delhi schools).

The **Digital Personal Data Protection Act, 2023 (DPDP Act) also plays a role in this space, by imposing restrictions on processing children's personal data, including bans on targeted advertising, profiling, and data use without verifiable parental consent under Section 9.*

Endnotes

1 Marie-Claire Dorking, The secret emoji code highlighted by Adolescence that parents must know about, https://uk.style.yahoo.com/parents-secret-emoji-meaning-adolescence-netflix-children-152155057.html?guccounter=1&guce_referrer=aHR0cHM6Ly93d3cuZ29vZ2xlLmNvbS8&guce_referrer__sig=AQAAAIJAOCR4Kdo7CHw1DNuUrOnY6pzDVhKb706rikHrnz2v6FRqL_2yvgVrbOzNuxAwCYh4zwz4qT 9kRwJP6SJ6QpgOoificQ4ELiuXyGqfun9BpsCMArqyCvUKwM__hFh2F7swpX8EUYoXPGX22wEflj3eaCBC68ZIe63-CB8wGoi.

2 Hee Jeong Yoo, *Evolution of Digital Natives and the New Role of Research,* 32 J. Korean Acad. Child & Adolesc. Psychiatry 127 (2021), https://pmc.ncbi.nlm.nih.gov/articles/PMC8499039

3 United Nations, Child and Youth Safety Online, https://www.un.org/en/global-issues/child-and-youth-safety-online

4Telecom Regulatory Authority of India, Highlights of Telecom Subscription Data, (August 2025), https://www.trai.gov.in/sites/default/files/2025-10/PRNo.104of2025__0.pdf

5 Sadhna Singh, *Online Safety for Children: Protecting the Next Generation from Harm*, NITI Aayog, https://www.niti.gov.in/sites/default/files/2025-06/Online-safety-for-children-protecting-the-next-Generation-from-harm.pdf.

6 Office of the Surgeon General, U.S. Dep't of Health & Hum. Servs., *Social Media and Youth Mental Health: The U.S. Surgeon General's Advisory* (2023), https://www.ncbi.nlm.nih.gov/books/NBK594761/

7 Abderrahman M. Khalaf, Abdullah A. Alubied, Ahmed M. Khalaf, et al.,"The Impact of Social Media on *the Mental Health of Adolescents and Young Adults: A Systematic Review," Cureus,* vol. 15, no. 8, e42990, (August 5,2023), https://pmc.ncbi.nlm.nih.gov/articles/PMC10476631/

8 Mohua Das, *"Child cyber crime surges 32% reveals NCRB data underlining vulnerability to online risks,"* The Times of India, (January 26, 2024),https://timesofindia.indiatimes.com/india/child-cyber-crime-surges-32-reveals-ncrb-data-underlining-vulnerability-to-online-risks/articleshow/107168056.cms" The Times of India, (December 5,

9 Deepak Lavania, "*Maximum number of POCSO cases registered each day,*2023), https://timesofindia.indiatimes.com/articleshow/105739493.cms

10 Ranjana Ferrao, *Special Courts for Children, Lessons from India,* 2024,https://iacajournal.org/articles/10.36745/ijca.485

11 Khalaf, A. M., Alubied, A. A., Khalaf, A. M., et al. (2023). *The impact of social media on the mental health of adolescents andyoung adults: A systematic review.* Cureus. https://pmc.ncbi.nlm.nih.gov/articles/PMC10476631/.

12 Lakshmi Sravanti et al.,*"Childhood Digital Exposure to Sexual Content: Through the Lens of Developmental Psychopathology" Indian Journal of Criminology and Criminal Justice* (2025),https://journals.sagepub.com/doi/10.1177/26318318251322555.

13 Nicole Valdes, Kerry Breen,*"A teen died after being blackmailed with A.I.-generated nudes. His family is fighting for change," CBS News,* (May 31, 2025),https://www.cbsnews.com/news/sextortion-generative-ai-scam-elijah-heacock-take-it-down-act/.

14 Felipa Schmidt, Filippo Varese, Sandra Bucci, *"Understanding the prolonged impact of online sexual abuse occurring in childhood, " Frontiers in Psychology,* (2023), https://pmc.ncbi.nlm.nih.gov/articles/PMC10627921/

15 *Teen queer artist kills self after receiving homophobic comments on Instagram," Deccan Herald (2023),* https://www.deccanherald.com/india/madhya-pradesh/teen-queer-artist-kills-self-after-receiving-homophobic-comments-on-instagram-2785500

16 u/NotHowGirlsWork,"*Reels comments are sickening,* " Reddit, 2025,https://www.reddit.com/r/

NotHowGirlsWork/comments/17rxh3h/reelscomments are___sickening/;https://theweek.com/culture-life/personal-technology/child-influencers-Instagram.

17 Alexander Ornella,"*Tate, TikTok and Toxic Masculinity; Is social media to blame for this generation's Violence Against Women?" CrimSoc Hull (2024)*,https://crimsoc.hull.ac.uk/2024/10/04/tate-tiktok-and-toxic-masculinity-is-social-media-to blame-for-this-generations-violence-against-women/..

18 Dean Lewins, *"Make me a sandwich": our surveys disturbing picture of how some boys treat their teachers," The Conversation (2024)*,https://theconversation.com/make-me-a-sandwich-our-surveys-disturbing-picture-of-how-some-boys-treat-their-teachers-228891.

19 Inside the manosphere luring young Indian men and boys, February 2025,https:/ /www.thenewsminute.com/premium/inside-the-manosphere-luring young-indian-men-and-boys

20 UN Women, What is the Manosphere and Why Should We Care, (2025),https://www.unwomen.org/en/articles/explainer/what-is-the-manosphere-and-why-should-we-care

21 Steve Rose, *"The sad, stupid rise of the sigma male: how toxic masculinity took over social media," The Guardian (2024)*,https://www.theguardian.com/society/article/2024/jun/12/the-sad-stupid-rise-of-the-sigma-male-how-toxic-masculinity-took-over-social-media

22 Harriet Over, Carl Bunce, Jonathan Baggaley, David Zendle,*"Understanding the influence of online misogyny inschools from the perspective of teachers, "* PLoS One (2025), https://pmc.ncbi.nlm.nih.gov articles/PMC11864523/

23 Monica Romero Sanchez, J. Megias, Hugo-Carreto-Dios, Sexist Humor and Sexual Aggression Against Women:When *Sexist Men Act According to Their Own Values or Social Pressures, (2019)*, https://pubmed.ncbi.nlm.nih.gov/31738118/

24 Emily Hemendinger, *"Mounting research documents the harmful effects of social media use on mental health, including body image and development of eating disorders, " The Conversation (2023)*, https://theconversation.com/mounting-research-documents-the-harmful-effects-of-social-media-use-on-mental-health-including-body-image-and-development-of-eating-disorders-206170

25 Khushi Suhag, Shyambabu Rauniyar, *"Social Media Effects Regarding Eating Disorders and Body Image in Young Adolescents, " Cureus (2024*), https://pmc.ncbi.nlm.nih.gov/articles/PMC11103119/.

26 Angela Yang,*"Parents worry AI-generated influencers are promoting unrealistic beauty standards to kids," NBC News (2024)*,https://www.nbcnews.com/tech/internet/parents-worry-ai-influencers-promote-unrealistic-beauty-standards-rcna134814.

27 https://www.thehindu.com/life-and-style/fitness/rise-in-body-image-issues-among-men/article22489419.ece

28 Zhu Lili, He Along, Ma Lidingna, "Micro-dramas exhibit 'accelerated violence'," *CSSN (2025)*,http://english.cssn.cn/skw_research/journalism/202501/t202501085831856.shtml_

29 *"Chinese micro dramas slammed for' vilifying women' amid regulatory clampdown, " Channel NewsAsia (2025),* https://www.channelnewsasia.com/east-asia/china-micro-drama-criticism-harmful-stereotypes-women tighten-regulations-4946831

30 Natasha Randhawa, Media Opportunities: Vertical Micro Dramas, June 2025, https://mediacat.uk/media-opportunities-vertical-micro-dramas/?; How China's $7 billion micro drama industry is taking on the U.S. entertainment industry, July 2025, https://www.cnbc.com/2025/07/22/why-chinas-7b-micro-drama-industry-is-taking-over-social-feeds.html

31 https://economictimes.indiatimes.com/tech/technology/microdramas-major-appeal-draws-in-

foreign-ogs/articleshow/124296848.cms; https://economictimes.indiatimes.com/industry/media/entertainment/the-rise-of-micro-drama-content-in-india-a-new-era-in-entertainment/articleshow/124259293.cms

32 Prateek Sinha (ORF), *India's cyber-forensics push since 2020: building national capacity for digital investigations,* Observer Research Foundation, 24 June 2025, https://www.orfonline.org/expert-speak/india-s-cyber-forensics-push-since-2020-building-national-capacity-for-digital-investigations

33 Jharkhand Judicial Academy, *Cyber Crime Cases: Issues, Challenges & Solutions, 2025,* https://jajharkhand.in/wp-content/uploads/2025/02/Cyber-Crime-web.pdf; Data Security Council of India (DSCI), Encryption and the Digital Economy (policy brief on encryption & law-enforcement challenges), 2021, https://www.dsci.in/files/content/knowledge-centre/2023/Encryption-and-the-Digital-Economy.pdf?

34 Farhana Shahid, Mona Elswah & Aditya Vashistha, "Think Outside the Data: Colonial Biases and Systemic Issues in Automated Moderation Pipelines for Low-Resource Languages; Center for Democracy & Technology (CDT), *Moderating Tamil Content on Social Media* (research report), 2025.

35 Social & Media Matters, *Is Content Moderation Working in India?* (interview-based report on moderation capacity, language, and emotional burden), 2025, https://www.socialmediamatters.in/research/research-content-moderation-in-india

36 Cambridge Forum on AI, *AI Moderation and Free Speech: Ongoing Challenges in the Global South,* Cambridge Law and Governance, 2025, https://www.cambridge.org/core/services/aop-cambridge-core/content/view

37 Protsahan (report), POCSO — *10-year analysis of NCRB data on sexual crimes against children (2012–2022),* Protsahan/POCSO report, 2022

38 D. Manoj et al, *Behind the screens: Understanding the gaps in India's fight against online child sexual abuse and exploitation, 2025,* https://www.sciencedirect.com/science/article/pii/S2950193824000883

39 Centre for Child and the Law, National Law School of India University & Enfold Proactive Health Trust, "Study on the Working of Special Courts under the POCSO Act, 2012 (Andhra Pradesh)" , Enfold https://enfoldindia.org/wp-content/uploads/2024/09/CCL-NLSIU-Andhra-Pradesh-Report.pdf; *Project 39A and Enfold report* "The Verdict and Beyond: Judicial Trends and Survivor Narratives" , 2024.

40 L. Fassi et al.,"Social media use in adolescents with and without mental health conditions, *" Nature Human Behaviour, 2025.*

41 Model National Response to end child sexual exploitation & abuse online, https://www.weprotect.org/resources/frameworks/model-national-response/

42 Vidhi Centre, *A decade of POCSO, 2022,* https://mphc.gov.in/PDF/web_pdf/JJC/PDF/publication/Vidhi%20Report%20-%20A%20Decade%20of%20POCSO.pdf?

43 Explanatory Notes- Section 15 is intended to cover situations where an adult (A) establishes contact with a child through, for example, meetings, telephone conversations or communications on the Internet, and gains the child's trust and confidence so that he can arrange to meet the child for the purpose of committing a "relevant offence" against the child. The course of conduct prior to the meeting that triggers the offence may have an explicitly sexual content, such as A entering into conversations with the child about the sexual acts he wants to engage her in when they meet, or sending images of adult pornography. However, the prior meetings or communication need not have an explicitly sexual content and could for example simply be A giving the child swimming lessons or meeting her incidentally through a friend; https://www.legislation.gov.uk/ukpga/2003/42/notes/division/5/1/15

44 *Digital Personal Data Protection Act, 2023, s. 9*

45 Regulation (EU) 2022/2065, *Digital Services Act* ("DSA"), Art. 28

46 Medha Chawla & Rishabh Sharma, "Queer, teen and coming out on social media? What it takes to survive the trolls, " *India Today* (2023), https://www.indiatoday.in/sunday-special/story/pranshu-ujjain-queer-teen-suicide-lgbtq-social-media-online-trolling -bullying-2474080-2023-12-10; Saatvika Radhakrishna, *"Cyberbullying in India grows as schools, platforms, and the law fail to protect children" Frontline (2025),* https://frontline.thehindu.com/social-issues/cyberbullying-teenagers-india-law-mental-health-ugc-social-media/articl e69905344.ece.

47 Khushi Suhag, Shyambabu Rauniyar, "Social Media Effects Regarding Eating Disorders and Body Image in Young Adolescents, " Cureus (2024), https://pubmed.ncbi.nlm.nih.gov/38770510/; Emily Hemendinger, *"Mounting research documents the harmful effects of social media use on mental health, including body image and development of eating disorders, " The Conversation (2023),* https://theconversation.com/mounting-research-documents-the-harmful-effects-of-social-media-use-on-mental-healt h-including-body-image-and-development-of-eating-disorders-206170

48 Steve Rose, *"The sad, stupid rise of the sigma male: how toxic masculinity took over social media, " The Guardian (2024),* https://www.theguardian.com/society/article/2024/jun/12/the-sad-stupid-rise-of-the-sigma-male-how-toxic-masculin ity-took-over-social-media

49 Nicholas Gibbs, Timothy Piatkowski, "The Liver King Lie: Misrepresentation, justification, and public health implications (2023), https://www.sciencedirect.com/science/article/pii/S0955395923000282

50 "Akshay Kumar reveals shocking incident: 13-year-old daughter Nitara was asked for nude photos during online game, *" The Times of India (2025),* https://timesofindia.indiatimes.com/entertainment/hindi/bollywood/news/akshay-kumar-reveals-shocking-incident-13-year-old-daughter-nitara-was-asked-for-nude-photos-during-online-gamewatch-video/articleshow/124288082.cms

51 Sudheer Kumar Muppalla et al, *"Effects of Excessive Screen Time on Child Development: An Updated Review and Strategies for Management, " Cureus (2023),* https://pmc.ncbi.nlm.nih.gov/articles PMC10353947/

INDIAN GOVERNANCE AND POLICY PROJECT

The Indian Governance And Policy Project (IGAP) is an emerging think tank focused on driving growth, innovation, and development in India's digital landscape. Specializing in areas like AI, Data Protection, FinTech, and Sustainability, IGAP promotes evidence-based policymaking through interdisciplinary research. By working closely with industry bodies in the digital sector, IGAP provides valuable insights and supports informed decision-making. Core work streams include policy monitoring, knowledge dissemination, capacity development, dialogue and collaboration.

For more details visit: www.igap.in

Contact us: relations@igap.in