Rostos fabricados, danos reais

Um panorama dos *deepfakes* sexuais entre jovens brasileiros

Safer net

Safe Online

Tech Coalition

Relatório da **SaferNet Brasil** sobre violência sexual mediada por inteligência artificial contra crianças e adolescentes no Brasil

Qual o contexto desta pesquisa?

A inteligência artificial (IA) generativa mudou a violência sexual online contra crianças e adolescentes. Ferramentas que antes exigiam conhecimento técnico especializado agora permitem que qualquer pessoa com acesso à internet produza imagens hiper-realistas de abuso em segundos.

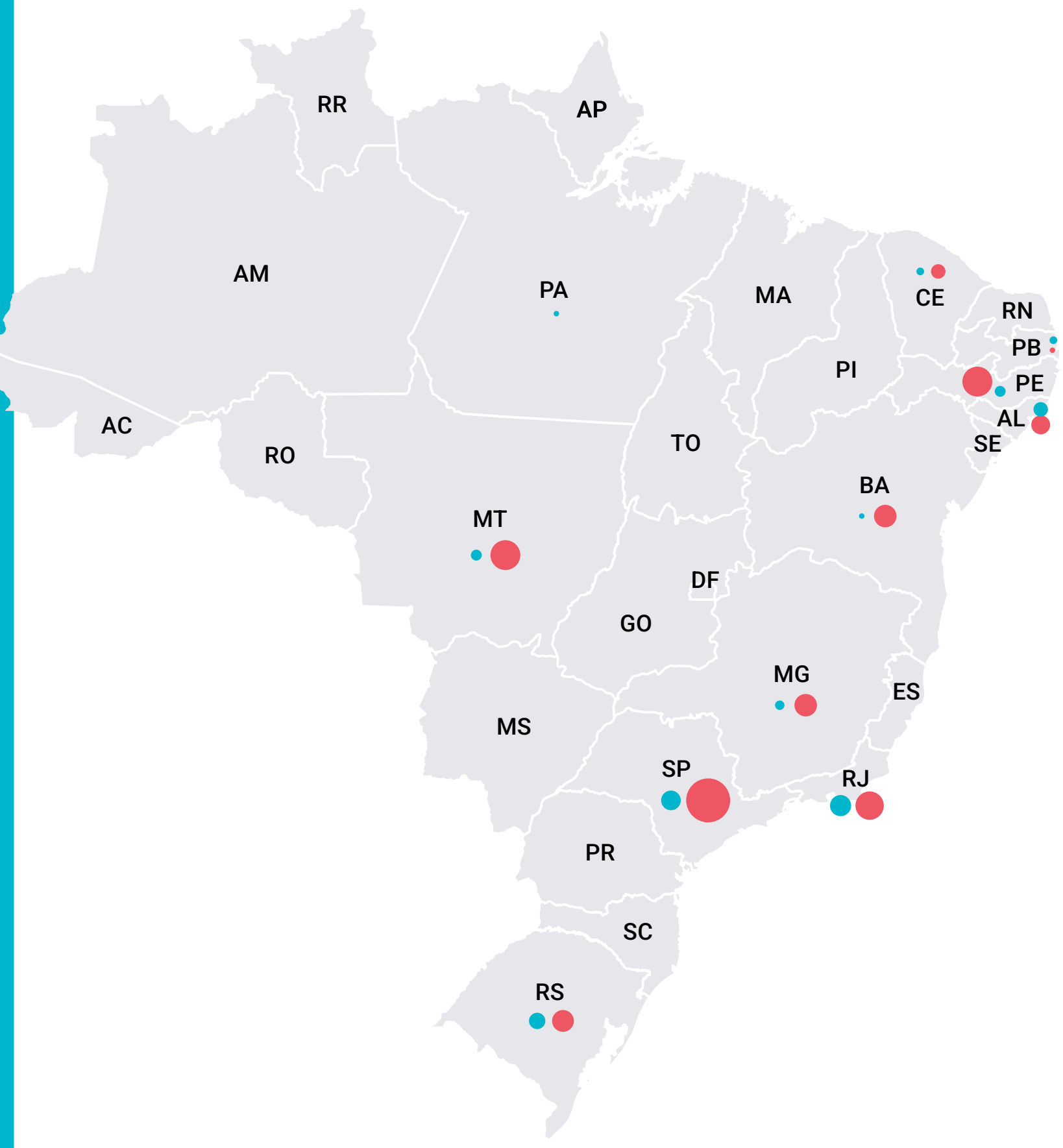
Este relatório é resultado de dez meses de pesquisa pela SaferNet Brasil, com apoio do fundo SafeOnline e da Tech Coalition. Ele documenta o fenômeno no Brasil em três frentes: os casos em escolas, o ecossistema digital que facilita a violência e as perspectivas dos próprios jovens brasileiros sobre o problema.

Quais os principais achados brasileiros desta pesquisa?

Deepfakes sexuais em escolas brasileiras

Entre agosto de 2025 e fevereiro de 2026, a SaferNet Brasil mapeou 17 episódios de *deepfake*s sexuais em escolas de 11 estados brasileiros (AL, BA, CE, MG, MT, PA, PB, PE, RJ, RS e SP), envolvendo 61 infratores e 222 vítimas. A totalidade das vítimas era do gênero feminino — estudantes de 12 a 17 anos e, em alguns casos, professoras. A totalidade dos infratores era do gênero masculino, da mesma faixa etária. A proporção de quase quatro vítimas para cada infrator indica práticas coletivas ou seriais de produção e circulação de imagens em canais fechados de mensagens. Casos foram registrados tanto em escolas públicas quanto privadas, em capitais e no interior, evidenciando que o fenômeno é nacional e sistêmico, e não circunstancial ou regional.

Mapeamento da SaferNet Brasil sobre casos de *deepfakes* sexuais em escolas brasileiras (dados até fevereiro de 2026).



- *Número estimado de infratores*
- *Número estimado de vítimas*

UF	Infratores	Vítimas	Ocorrências
AL	7	12	1
BA	1	17	1
CE	2	7	1
MG	3	17	1
MT	4	30	1
PA	1	Não informado	1
PB	2	1	1
PE	4	30	1
RJ	15	27	3
RS	9	16	1
SP	13	65	5
Total	**61**	**222**	**17**

A opinião dos jovens brasileiros

Uma enquete (*survey*) com 679 jovens (predominantemente entre 16 e 21 anos) revelou uma fratura de gênero profunda nas percepções sobre o problema.

Enquanto 86,8% das meninas cisgênero afirmaram com certeza que criar uma *deepfake* de alguém menor de 18 anos é crime, apenas 61,9% dos meninos deram a mesma resposta — e 9,4% deles afirmaram categoricamente que não é crime.

Dos 501 meninos cisgênero da amostra, 62 (12,4%) admitiram já ter criado uma *deepfake*; entre as meninas cisgênero, esse número foi zero. Ao todo, 42,6% dos jovens entrevistados declararam produzir, conhecer alguém que produz ou ter ouvido falar de colegas que produzem esse tipo de conteúdo. O Telegram foi apontado como a principal plataforma onde esses materiais são criados (42,6% dos respondentes), seguido por pesquisas nos buscadores (33,9%), outros sites (29,5%) e aplicativos de nudificação para celular (27,1%).

Quais as principais evidências desta pesquisa?

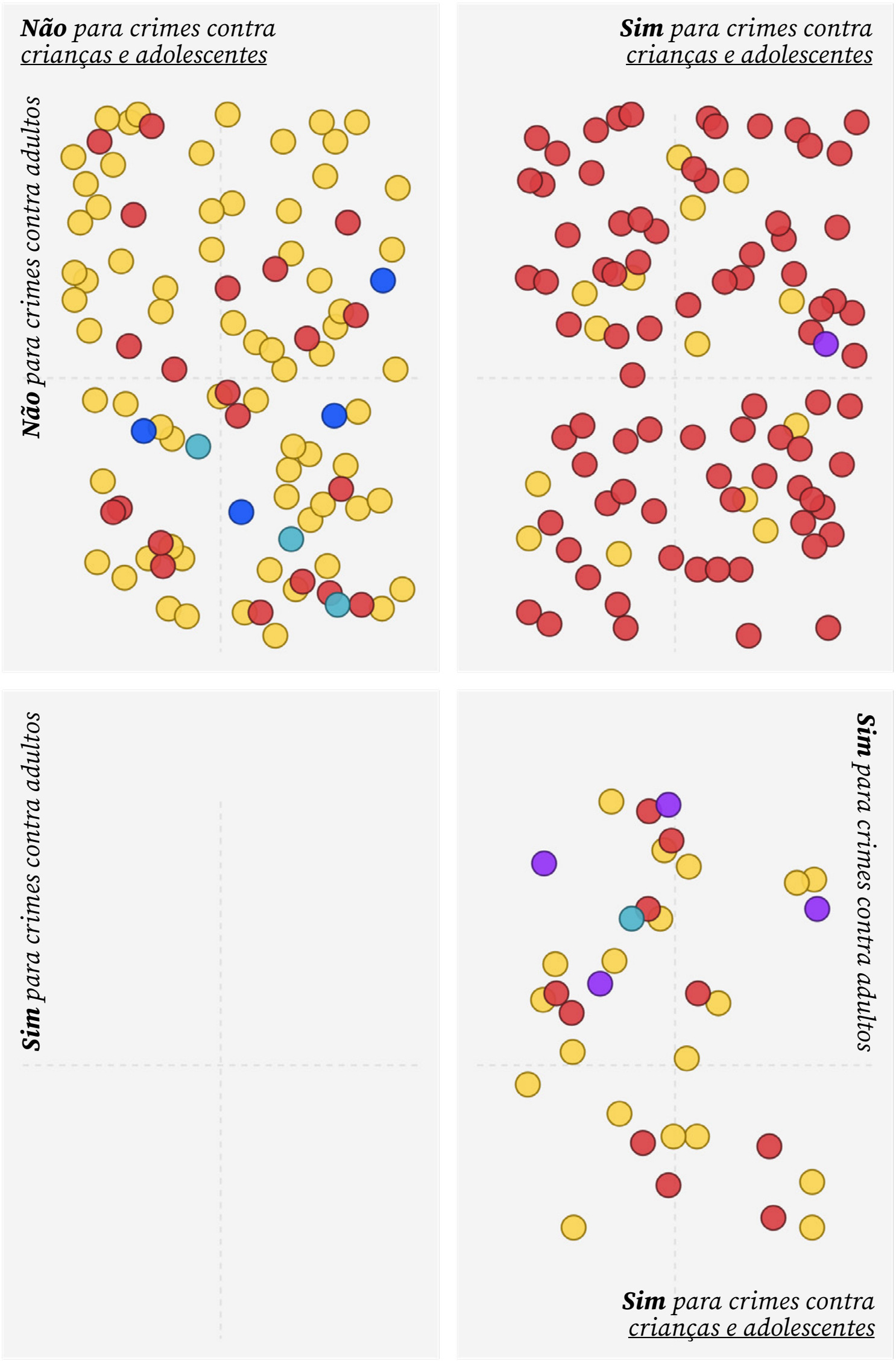
Um ecossistema digital de abuso

Para esta pesquisa, foram selecionadas 264 URLs submetidas ao canal de denúncias da SaferNet Brasil, relacionadas a suspeitas de abuso ou exploração sexual infantil. A categoria mais prevalente foi a de conteúdos com crimes contra crianças, identificada em 149 URLs (56,4% da amostra).

A análise aprofundada dos cinco estudos de caso identificou os principais vetores de risco. Por tipo de plataforma, a *dark web* apresentou a maior taxa de conteúdo ilícito (85,8% de suas 127 URLs).

Já os aplicativos de nudificação, apesar de terem poucos links na amostra, destacaram-se pela extrema gravidade: 100% deles continham *deepfakes* de crianças e crimes contra crianças.

Gráfico de dispersão com distribuição de análise de conteúdos vs. plataformas, crime ou não contra crianças e mulheres.



Nota: Os pontos foram levemente deslocados para facilitar a visualização de sobreposições.

Robôs de nudificação no Telegram

Entre 2023 e 2026, a SaferNet Brasil recebeu mais de 10 mil denúncias de potencial material ou conteúdo de violência sexual contra crianças e adolescentes com o domínio “t.me/”, do Telegram.

Bots de nudificação operam com modelos *freemium* — imagens de teste gratuitas seguidas de acesso pago via Telegram Stars, criptomoedas ou PIX. As Stars podem ser convertidas em criptomoedas pela TON Network (*blockchain* vinculada ao Telegram) após 21 dias, criando um circuito financeiro que entra por vias convencionais e sai em ativos de difícil rastreamento.

A cada transação, Apple e Google retêm US$ 3,00 e o Telegram, US$ 0,49 — três empresas lucrando sobre crimes contra crianças, com graus distintos de responsabilidade.

Fóruns e geradores na *dark web*

Dos 127 domínios da rede Tor analisados, 81,9% continham material ou conteúdo real de violência sexual contra crianças e adolescentes e 10,24% continham material artificial.

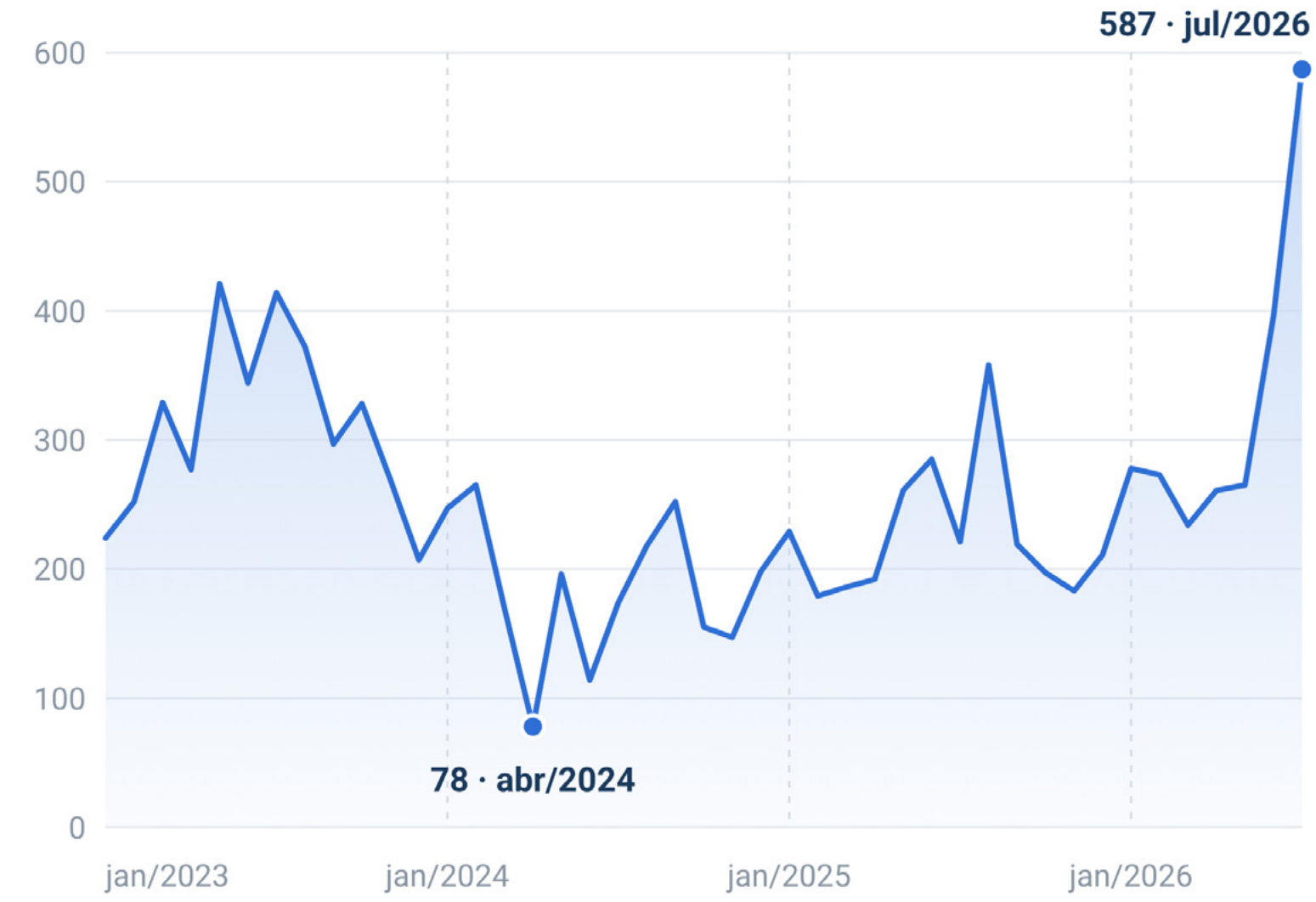
A SaferNet Brasil monitorou um fórum criado exclusivamente para produção e disseminação de material gerado por IA, cujas publicações já ultrapassavam 20 mil visualizações em março de 2026.

O fórum mantém centenas de guias e tutoriais — desde geração de textos descrevendo abuso até treinamento de modelos sob demanda —, com indicações que incluem ferramentas e vídeos acessíveis na internet aberta, evidenciando a porosidade entre a surface web e esses ecossistemas criminosos.

A retórica sustenta que imagens artificiais seriam uma "alternativa ética" ao abuso real, narrativa que funciona como mecanismo de recrutamento e normalização da exploração sexual de crianças e de adolescentes.

Evolução mensal das denúncias

por data da denúncia · base consolidada: URLs + entidades citadas nos comentários (14.782 registros datados, até jul/2026)



Registros por mês e ano

	2023	2024	2025	2026
jan	224	247	229	278
fev	252	265	179	273
mar	329	170	186	234
abr	277	78	192	261
mai	421	196	261	265
jun	344	114	285	396
jul	414	174	221	587
ago	372	218	358	—
set	297	252	219	—
out	328	155	197	—
nov	269	147	183	—
dez	207	198	211	—
Total	**3.734**	**2.214**	**2.721**	**2.294***

78 — 587 registros/mês — fora do recorte (ago–dez/2026)

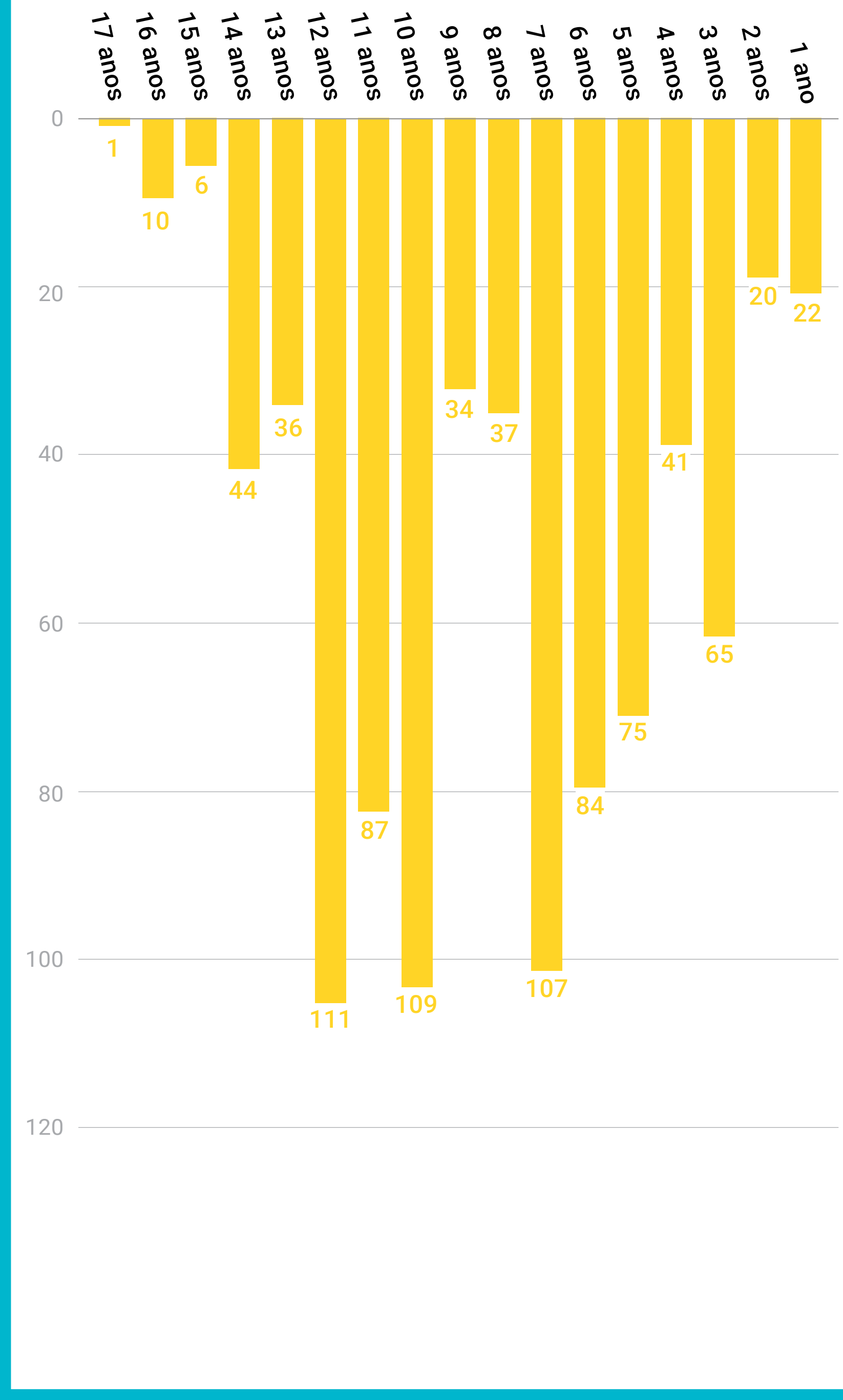
Base = 14.782 registros datados (URLs + entidades citadas), por data da denúncia.

* 2026 parcial (jan–jul); ago/2026 excluído por apuração incompleta.

Foi também analisado um gerador de imagens baseado em modelos de código aberto, dedicado exclusivamente à produção de material ou conteúdo de violência sexual contra crianças e adolescentes. Acesso anônimo (sem e-mail), com créditos acumulados ao avaliar mídias de outros usuários — mecanismo que configura aprendizado por reforço humano, aprimorando ativamente o modelo.

Em março de 2026, havia produzido mais de 510 mil materiais com a palavra *girl* (menina, em inglês). Conteúdos referentes ao Brasil cresceram 37,6% em 60 dias (4.401 para 6.056 registros), com as idades mais frequentes nos *prompts* sendo 12, 10, 7, 11 e 6 anos — e presença alarmante de idades de 1, 2 e 3 anos.

Distribuição de idade mencionada nos *prompts* coletados pela SaferNet Brasil.



MAIS DE

1.300

IMAGENS POR DIA

RETRATANDO MENINAS

OU APROXIMADAMENTE UMA
NOVA IMAGEM **A CADA MINUTO**

Plataformas de pornografia

Duas plataformas de conteúdo adulto foram analisadas como estudos de caso.

Na primeira, a SaferNet Brasil acumula 8,6 mil denúncias desde 2018 (de 3 em 2018 para 2.108 em 2025); 27% dos links disponíveis continham material ou conteúdo de violência sexual contra crianças e adolescentes, e buscas por termos como *AI teen* (adolescente de IA, em inglês) retornaram mais de 5 mil resultados.

Na segunda, foram identificados 263 *deepfakes* envolvendo 22 vítimas, das quais 19 tinham menos de 18 anos no momento da captura das imagens originais; a pesquisa documentou táticas de ofuscação (números substituindo letras nos nomes das vítimas) para burlar moderação e impedir que as próprias vítimas localizem os materiais.

Em ambos os estudos de caso, foram encontrados *deepfakes* de deputadas federais e estaduais brasileiras publicados

com nomes completos, sem qualquer moderação — instrumentos de intimidação política contra mulheres em posições de poder.

Nenhuma das plataformas exige verificação efetiva de idade para além da autoidentificação do próprio usuário.

Quais as recomendações prioritárias da SaferNet?

PODER EXECUTIVO E AUTORIDADES

- Criar mecanismo nacional (Banco Central, COAF, instituições financeiras) de monitoramento de transações suspeitas vinculadas a *deepfakes* sexuais de crianças e adolescentes.
- Integrar plataformas de criptoativos a sistemas de alerta e bloqueio de carteiras, em conformidade com padrões AML/KYC.
- Solicitar à Fragment informações sobre transações na TON Network originadas no Brasil; incluir campos de inteligência financeira no template do Art. 39 do Decreto nº 12.880/2026.
- Notificar Apple e Google para implementarem uma verificação de finalidade em pagamentos a bots no

Telegram e auditar apps de nudificação em suas lojas.

- Mapear a cadeia de desenvolvedores, distribuidores e intermediários; acionar cooperação internacional para bloqueio e confisco de ativos.
- Desenvolver e publicar protocolo nacional de resposta a *deepfakes* sexuais em escolas (MEC, MJ, MDH): identificação, acolhimento, capacitação de conselheiros tutelares e fluxo com a rede de proteção.
- Implementar programa nacional de capacitação de delegacias especializadas em crimes com IA generativa (análise forense, rastreamento de criptomoedas, infiltração em *dark web*).
- Expandir ferramentas de IA investigativas (Nêmesis, classificadores da Thorn) para todas as unidades da federação.

PODER LEGISLATIVO

- Criminalizar ferramentas de nudificação por IA e materiais instrutivos (guias/ tutoriais) para produção de imagens de abuso sexual infantil.

- Estabelecer responsabilização civil e criminal de proprietários, desenvolvedores e plataformas que hospedem, facilitem ou lucrem com imagens nudificadas não consensuais.

PODER JUDICIÁRIO (CNJ/CNMP)

- Criar marcadores, tags ou subclasses nas Tabelas Processuais Unificadas (TPU) para identificar processos envolvendo IA generativa em crimes da Lei nº 11.829/2008.
- Criar fluxo de validação dos assuntos processuais por servidores do Judiciário, corrigindo inconsistências do modelo atual.
- Viabilizar produção de dados estatísticos confiáveis sobre criminalidade digital com IA para orientar políticas públicas.

EMPRESAS DE IA E TECNOLOGIA

- Adotar o princípio de segurança desde a concepção (*safety by design*), com salvaguardas técnicas, jurídicas e éticas em todo o ciclo de vida dos produtos.
- Implementar avaliações obrigatórias de segurança infantil antes da publicação de modelos de geração de imagem, com testes adversariais focados em CSAM.
- Proibir explicitamente nas licenças de uso a geração de material ou conteúdo de violência sexual contra crianças e adolescentes e imagens íntimas não consensuais; responsabilizar distribuidores que redistribuam versões sem salvaguardas.
- Implementar avaliações de segurança infantil como requisito prévio à hospedagem de modelos e realizar auditorias retroativas; expandir filtros e integrar bases de *hashes* internacionais.

- Fortalecer controles AML/KYC no setor financeiro e meios de pagamento, adaptando modelos de risco para monetização de conteúdo sintético de abuso.
- Aplicar termos de uso de forma consistente (provedores de hospedagem, CDN, anúncios); suspender contas/domínios vinculados a conteúdo de abuso; criar canais especializados para notificações.
- Instituir governança dedicada à proteção infantil, designar pontos de contato, capacitar equipes e submeter-se a auditorias independentes com divulgação pública.

SETOR EDUCACIONAL

- Integrar nos currículos conteúdos sobre manipulação digital de imagens, consentimento, empatia e impactos da circulação não consensual de materiais sexualizados, de forma transdisciplinar (BNCC).
- Orientar alunos sobre proteção de dados pessoais, gestão da própria imagem e mitigação de riscos em plataformas digitais.
- Promover formação continuada de professores e gestores sobre IA generativa e apps de nudificação.
- Instituir e divulgar protocolos formais de denúncia, acolhimento de vítimas e encaminhamento, sem revitimização.
- Articular ações preventivas e gestão de incidentes com famílias, Secretarias de Educação, Conselhos Tutelares, MEC e sociedade civil.

Expediente

Publicação

SaferNet Brasil

Apoio

SafeOnline e Tech Coalition

Coordenação de pesquisa

Juliana Cunha

Pesquisa, análise de dados e gráficos

Sofia Schurig

Redação

Thiago Tavares, Juliana Cunha,
Sofia Schurig

Revisão

Renata Leal (Ibirá Comunicação)

Colaboração

Bianca Orrico, Isabella Ferro,
Júlia Fernandes, Guilherme Alves

Direção de arte e diagramação

Natália Durães Design (Ibirá Comunicação)

Data de publicação

Novembro/2026

ISBN

978-85-65151-12-2

SOCIEDADE CIVIL E PESQUISA

- Desenvolver estudos de prevalência de *deepfakes* sexuais em escolas brasileiras com recortes de classe, gênero e etnia, amostras representativas e metodologias padronizadas.
- Atuar como articuladoras entre escolas, famílias, autoridades e empresas; produzir materiais de referência e capacitar educadores e profissionais de segurança pública.
- Monitorar sistematicamente a evolução de ferramentas de IA generativa usadas para CSAM: novos modelos, técnicas de evasão e formas de monetização.
- Manter e ampliar redes de apoio a vítimas: orientação jurídica, acolhimento psicossocial e suporte para remoção de conteúdos.

Acesse a versão completa
da pesquisa em:

bit.ly/rostosfabricados

Safer net

Safe Online

Tech Coalition