

INÍCIO

1. DEFINIÇÕES
& TIPOLOGIAS
 2. VIOLÊNCIA SEXUAL
MEDIADA POR
IA EM ESCOLAS
BRASILEIRAS
 3. PONTOS DE VISTA
DE DIFERENTES
ATORES
ENVOLVIDOS
NO FENÔMENO
 4. ECOSISTEMA
DIGITAL QUE
FACILITA A
VIOLÊNCIA SEXUAL
MEDIADA POR IA
 5. RESPOSTAS
DE DIFERENTES
ATORES AO
PROBLEMA DA
VIOLÊNCIA SEXUAL
MEDIADA POR IA
 6. RECOMENDAÇÕES
DETALHADAS PARA
MÚLTIPLOS ATORES
 7. CONSIDERAÇÕES
FINAIS
- SOBRE ESTA
PESQUISA

CAPÍTULO QUATRO

Ecosistema digital que facilita a violência sexual mediada por IA

INÍCIO

1. DEFINIÇÕES
& TIPOLOGIAS
 2. VIOLÊNCIA SEXUAL
MEDIADA POR
IA EM ESCOLAS
BRASILEIRAS
 3. PONTOS DE VISTA
DE DIFERENTES
ATORES
ENVOLVIDOS
NO FENÔMENO
 4. ECOSISTEMA
DIGITAL QUE
FACILITA A
VIOLÊNCIA SEXUAL
MEDIADA POR IA
 5. RESPOSTAS
DE DIFERENTES
ATORES AO
PROBLEMA DA
VIOLÊNCIA SEXUAL
MEDIADA POR IA
 6. RECOMENDAÇÕES
DETALHADAS PARA
MÚLTIPLOS ATORES
 7. CONSIDERAÇÕES
FINAIS
- SOBRE ESTA
PESQUISA



criação, manutenção e exploração comercial de ferramentas de IA generativa capazes de reproduzir violência sexual contra crianças e adolescentes não decorrem de ações isoladas, mas de um ecossistema criminoso estruturado. Tal ecossistema opera sob a lógica de uma organização criminosa descentralizada, numa cadeia que abrange o desenvolvimento técnico de modelos generativos sem salvaguarda, bem como sua distribuição que extrai lucro financeiro dessas práticas. Por isso, a pesquisa da SaferNet Brasil buscou compreender o ecossistema cibercriminoso que sustenta e facilita essa violência — o que envolve analisar os fatores econômicos, culturais e sociais que estruturam as comunidades voltadas à produção desses crimes e ao desenvolvimento das ferramentas que os viabilizem.

Este capítulo trata dos principais vetores de risco associados a esse ecossistema, entendidos como canais e plataformas digitais cujas características facilitam a criação ou a propagação dos conteúdos abusivos. Os vetores primários

— de maior impacto e recorrência — são descritos nos subtópicos seguintes, acompanhados de cinco estudos de caso.* Esta seção da pesquisa contém descrições textuais de conversas entre criminosos que podem causar desconforto, mas não há imagens explícitas de violência ou exploração sexual infantojuvenil (real ou gerada por IA).

Os dados disponíveis evidenciam uma escalada sem precedentes. Segundo o *AI Incident Monitor* da Organização para a Cooperação e Desenvolvimento Econômico (OCDE), ao longo de 2020 os incidentes de segurança envolvendo IA e crianças mal ultrapassavam uma ocorrência mensal. Janeiro de 2026 rompe esse padrão: foram 169 incidentes registrados em um único mês — mais do que a soma de tudo o que havia sido contabilizado nos dois primeiros anos de monitoramento.

Embora há décadas exista um receio social sobre a possibilidade de criação de conteúdos artísticos, materiais escritos ou peças de pro-

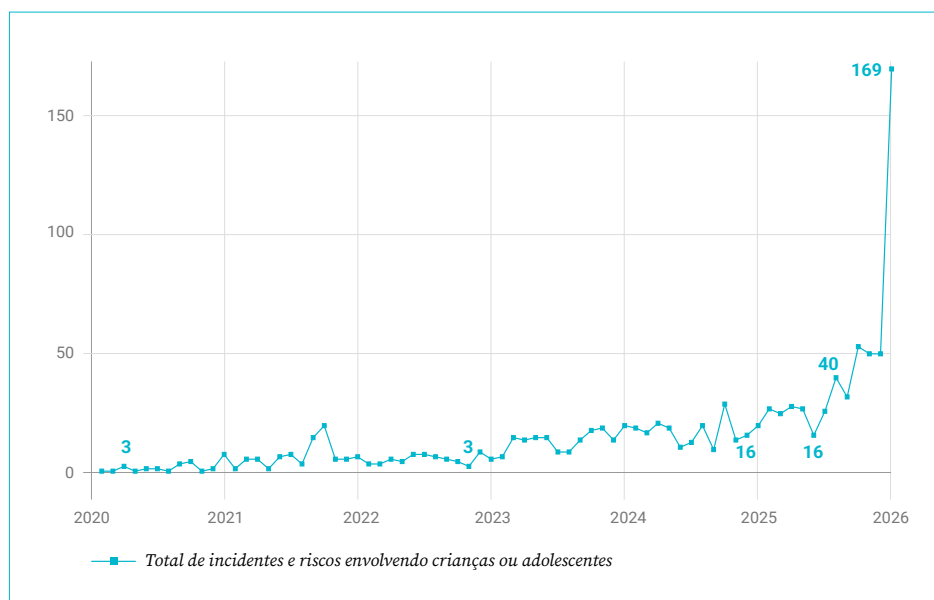
*Nota metodológica. Os dados apresentados neste capítulo não decorrem de amostragem estatística, mas de seleção qualitativa. A análise priorizou domínios recorrentes e ambientes de maior incidência identificados pela equipe técnica, sem a pretensão de cobrir de forma exaustiva todos os espaços onde tais crimes ocorrem.

INÍCIO

1. DEFINIÇÕES
& TIPOLOGIAS
 2. VIOLÊNCIA SEXUAL
MEDIADA POR
IA EM ESCOLAS
BRASILEIRAS
 3. PONTOS DE VISTA
DE DIFERENTES
ATORES
ENVOLVIDOS
NO FENÔMENO
 4. **ECOSSISTEMA
DIGITAL QUE
FACILITA A
VIOLÊNCIA SEXUAL
MEDIADA POR IA**
 5. RESPOSTAS
DE DIFERENTES
ATORES AO
PROBLEMA DA
VIOLÊNCIA SEXUAL
MEDIADA POR IA
 6. RECOMENDAÇÕES
DETALHADAS PARA
MÚLTIPLOS ATORES
 7. CONSIDERAÇÕES
FINAIS
- SOBRE ESTA
PESQUISA

FIGURA 16

Casos de incidentes e riscos classificados como “dano psicológico” e “violação de direitos humanos” envolvendo IA generativa e crianças, com base nos registros da Organização para a Cooperação e Desenvolvimento Econômico (OCDE).



paganda que enalteçam a exploração sexual infantil — tanto que esse risco já estava previsto no ECA em 2008 —, a tecnologia de IA generativa transforma esse receio em um dano real, não mais em uma mera possibilidade sem consequências. O monitoramento conduzido pela SaferNet Brasil para esta seção da pesquisa revela um cenário grave, que se estende para além da ausência de regulamentação da IA generativa entre adolescentes e do desenvolvimento descontrolado dessa mesma tecnologia em aplicações associadas à pornografia, à nudeificação não consensual e à violência sexual.

Em 2026, **as ferramentas necessárias para a criação desses conteúdos estão presentes tanto na surface web quanto na dark web**, em

A dark web é parte da internet acessível apenas com softwares como Tor Browser, que ocultam identidade e localização. É usada para privacidade e anonimato, mas também abriga atividades ilegais, mercados clandestinos e comunicação difícil de rastrear.

INÍCIO

1. DEFINIÇÕES & TIPOLOGIAS

2. VIOLÊNCIA SEXUAL MEDIADA POR IA EM ESCOLAS BRASILEIRAS

3. PONTOS DE VISTA DE DIFERENTES ATORES ENVOLVIDOS NO FENÔMENO

4. ECOSISTEMA DIGITAL QUE FACILITA A VIOLÊNCIA SEXUAL MEDIADA POR IA

5. RESPOSTAS DE DIFERENTES ATORES AO PROBLEMA DA VIOLÊNCIA SEXUAL MEDIADA POR IA

6. RECOMENDAÇÕES DETALHADAS PARA MÚLTIPLOS ATORES

7. CONSIDERAÇÕES FINAIS

SOBRE ESTA PESQUISA

modelos gratuitos e pagos, e podem ser localizadas com relativa facilidade por mecanismos de busca comuns — sobretudo quando os filtros de navegação segura estão desativados. Esse achado evidencia a fragilidade dos controles atualmente existentes e coloca em xeque a eficácia das salvaguardas voluntárias adotadas pelas plataformas. Soma-se a isso um padrão estrutural relevante: plataformas de pornografia adulta, historicamente caracterizadas pela ausência de moderação ativa, tornaram-se também espaços de hospedagem de materiais artificiais de exploração sexual infantojuvenil — uma continuidade entre os dois fenômenos que não pode ser tratada de forma dissociada.

Para esta análise, foram selecionadas 264 URLs submetidas ao canal de denúncias da SaferNet Brasil, relacionadas a suspeitas de abuso ou exploração sexual infantil. Esse número cresceu ao longo dos estudos de caso; a seleção inicial serviu para delimitar o escopo e constituir uma amostra quantitativa de partida. Os critérios de escolha foram qualitativos — como a presença de palavras-chave nos relatos dos denunciante, entre elas “deepfake”, “nudificação” e “IA” — além de evidências de pesquisas anteriores que já apontavam plataformas e aplicativos reconhecidos pela SaferNet Brasil como vetores

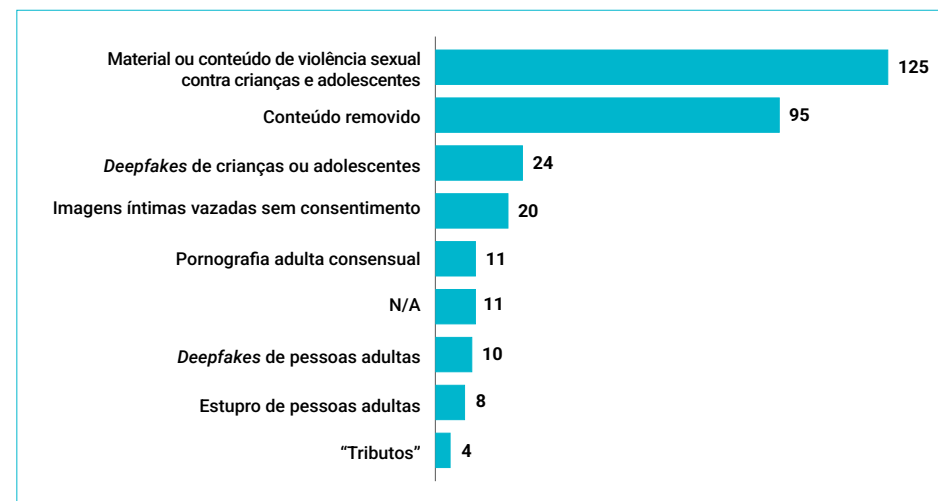
de risco. As denúncias foram extraídas em setembro de 2025, embora cubram registros acumulados sobretudo entre 2018 e 2025.

Das URLs analisadas, 149 (56,4%) apresentaram ao menos um tipo de conteúdo ilícito, ao

passo que 111 (42%) não continham nenhuma das categorias investigadas. A categoria mais prevalente foi a de conteúdos com crimes contra crianças — conteúdos artificiais de abuso sexual infantojuvenil e outros —, identificada em 139 URLs ou 52,7% de toda a amostra.

FIGURA 17

Total de análises das 264 URLs da amostra inicial. Há 308 menções porque uma mesma URL contém mais de uma análise (ex.: um site que hospeda tanto *deepfakes* sexuais de adultos como material artificial de abuso sexual infantil)



INÍCIO

1. DEFINIÇÕES
& TIPOLOGIAS
 2. VIOLÊNCIA SEXUAL
MEDIADA POR
IA EM ESCOLAS
BRASILEIRAS
 3. PONTOS DE VISTA
DE DIFERENTES
ATORES
ENVOLVIDOS
NO FENÔMENO
 4. ECOSISTEMA
DIGITAL QUE
FACILITA A
VIOLÊNCIA SEXUAL
MEDIADA POR IA
 5. RESPOSTAS
DE DIFERENTES
ATORES AO
PROBLEMA DA
VIOLÊNCIA SEXUAL
MEDIADA POR IA
 6. RECOMENDAÇÕES
DETALHADAS PARA
MÚLTIPLOS ATORES
 7. CONSIDERAÇÕES
FINAIS
- SOBRE ESTA
PESQUISA

Em outras 34 URLs (12,9%), havia conteúdos com crimes contra adultos. Em 20 URLs, havia crimes tanto contra pessoas adultas como crianças e adolescentes — por exemplo, um site que hospeda tanto material de exploração sexual infantojuvenil quanto filmagens de estupro de pessoas adultas.

No que diz respeito a *deepfakes* sexuais, 27 URLs (10,2%) continham ao menos um tipo de conteúdo sintético. *Deepfakes* envolvendo crianças foram detectados em 24 URLs (9,1%), ao passo que *deepfakes* envolvendo adultos apareceram em 10 (3,8%). Em sete URLs, foram identificadas todas as categorias analisadas simultaneamente — *deepfakes* de adultos e de crianças, além de crimes contra adultos e crianças —, configurando os casos de maior gravidade e sobreposição no conjunto de dados.

A distribuição por tipo de plataforma aprofunda essa leitura. As 264 URLs concentram-se em dois grandes grupos: *dark web*, com 127 URLs (48,1%), e plataformas de pornografia ou conteúdo erótico adulto, com 123 (46,6%). Os demais tipos — aplicativos de nudificação (cinco URLs), redes sociais (cinco) e sites diversos (quatro) — somam apenas 14 registros, mas sua análise qualitativa revela padrões que não podem ser ignorados.

A *dark web* apresentou a maior taxa de conteúdo ilícito: 109 de suas 127 URLs (85,8%) continham ao menos um tipo de crime ou material abusivo. A totalidade desses ilícitos dizia respeito a crimes contra crianças, enquanto crimes contra adultos apareceram em apenas 10 URLs (7,9%), sempre em sobreposição a material infantil. *Deepfakes* na *dark web* foram encontrados em 13 URLs e, dado relevante, todos os 13 casos envolviam exclusivamente crianças — nenhum registro continha *deepfake* de pessoa adulta. Esses números confirmam que a *dark web* opera, nesta amostra, como vetor predominantemente voltado à exploração infantil, com uma concentração temática que evidencia redes estruturadas de produção e distribuição.

Nas plataformas de pornografia, a taxa de conteúdo ilícito foi substancialmente menor — 37 de 123 URLs (30,1%) —, motivada pela maior taxa de remoção de conteúdo por serem sites hospedados na internet aberta (77 das 123 URLs, ou 62,6%, já haviam sido removidas no momento da análise), porém com perfil distinto: 23 URLs (18,7%) envolviam crimes contra crianças e 19 (15,4%), crimes contra adultos, estes últimos concentrados sobretudo em imagens íntimas vazadas sem consentimento. Como esperado, os aplicativos de nudificação,

ainda que apresentem poucas URLs da amostra original (inicialmente 5, com URLs adicionais analisadas durante o estudo de caso sobre o Telegram), destacam-se por conta da gravidade absoluta dos achados: a totalidade deles apresentou *deepfakes* de crianças e crimes contra crianças, e em 80% dos casos havia simultaneamente *deepfakes* de mulheres e crimes contra adultos. Trata-se do único tipo de plataforma em que todas as URLs analisadas concentraram todas ou quase todas as categorias de ilícito ao mesmo tempo — um dado que reforça o caráter intrinsecamente abusivo dessas ferramentas e torna difícil sustentar qualquer argumento de que possuam finalidade legítima predominante.

Este monitoramento identificou, ainda, a consolidação de comunidades organizadas de indivíduos que consomem, comercializam ou compartilham esse tipo de conteúdo. Muitas dessas comunidades operam em espaços de difícil acesso, onde são desenvolvidos e disseminados guias, tutoriais e sistemas de IA generativa treinados com material ou conteúdo de violência sexual contra crianças e adolescentes, frequentemente de forma gratuita, o que reduz as barreiras de entrada e amplia o alcance dessas redes.

INÍCIO

1. DEFINIÇÕES
& TIPOLOGIAS

2. VIOLÊNCIA SEXUAL
MEDIADA POR
IA EM ESCOLAS
BRASILEIRAS

3. PONTOS DE VISTA
DE DIFERENTES
ATORES
ENVOLVIDOS
NO FENÔMENO

4. ECOSISTEMA
DIGITAL QUE
FACILITA A
VIOLÊNCIA SEXUAL
MEDIADA POR IA

5. RESPOSTAS
DE DIFERENTES
ATORES AO
PROBLEMA DA
VIOLÊNCIA SEXUAL
MEDIADA POR IA

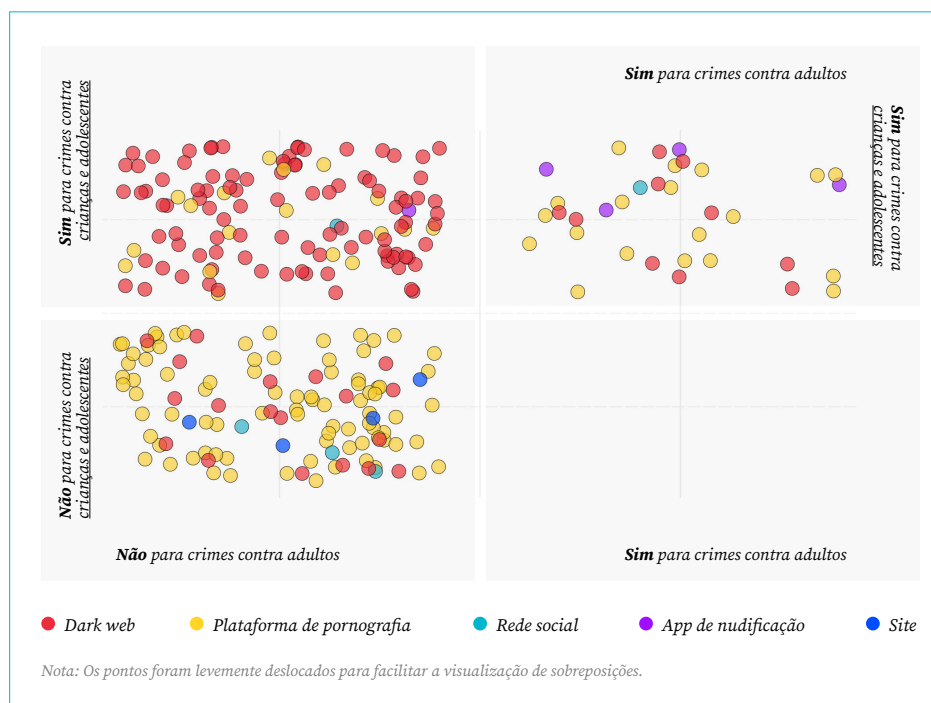
6. RECOMENDAÇÕES
DETALHADAS PARA
MÚLTIPLOS ATORES

7. CONSIDERAÇÕES
FINAIS

SOBRE ESTA
PESQUISA

FIGURA 18

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



INÍCIO

1. DEFINIÇÕES & TIPOLOGIAS
 2. VIOLÊNCIA SEXUAL MEDIADA POR IA EM ESCOLAS BRASILEIRAS
 3. PONTOS DE VISTA DE DIFERENTES ATORES ENVOLVIDOS NO FENÔMENO
 4. ECOSISTEMA DIGITAL QUE FACILITA A VIOLÊNCIA SEXUAL MEDIADA POR IA
 5. RESPOSTAS DE DIFERENTES ATORES AO PROBLEMA DA VIOLÊNCIA SEXUAL MEDIADA POR IA
 6. RECOMENDAÇÕES DETALHADAS PARA MÚLTIPLOS ATORES
 7. CONSIDERAÇÕES FINAIS
- SOBRE ESTA PESQUISA

4.1. VETOR DE RISCO N.º1 → APLICATIVOS, SITES E ROBÔS DE NUDIFICAÇÃO

O QUE ENCONTRAMOS NESTE VETOR DE RISCO



A análise de URLs de grupos e ferramentas de nudificação conduzida pela SaferNet Brasil indica que o maior risco não reside apenas na natureza do conteúdo, mas sobretudo em sua enorme acessibilidade — tanto para a produção quanto para a disseminação, que alcança todos os perfis de usuários e praticamente todas as plataformas.



Parte expressiva e estratégica dessas ferramentas opera em plataformas e redes sociais populares, como Telegram e Discord. A menção frequente a esses ambientes pelos jovens participantes da pesquisa, detalhada na seção 3, corrobora essa observação.



A monetização constitui um fator agravante central. Enquanto o objetivo primário da nudificação é, em grande parte dos casos, o lucro financeiro, outros espaços — como os da *dark web* — priorizam a exploração sexual como finalidade em si. Os crimes nesse ecossistema, portanto, articulam-se em torno de duas motivações principais: o interesse econômico e o próprio ato de explorar sexualmente crianças e adolescentes, mesmo quando mediados por IA.



Um agravante adicional da monetização diz respeito aos meios de pagamento: os aplicativos de nudificação aceitam modalidades amplamente difundidas, como PIX, carteiras digitais vinculadas a lojas de aplicativos e criptomoedas — estas últimas impulsionadas pelo crescimento das *exchanges* durante a pandemia —, operando diretamente em plataformas e aplicativos de uso cotidiano.



Entre os casos mais preocupantes estão aqueles que envolvem as Telegram Stars, moeda criada pela própria plataforma e de uso restrito ao seu ambiente interno. As Stars são adquiridas por meio dos métodos de pagamento já mencionados e facilitam o comércio internacional de conteúdos nudificados, *deepfakes* sexuais não consentidos e materiais artificiais de exploração sexual infantojuvenil.



O risco específico das Stars se intensifica porque, após um período de 21 dias, os chamados “criadores” — usuários que gerenciam serviços como *bots* dentro do Telegram — podem convertê-las em criptomoedas pela TON Network, blockchain desenvolvida pelo irmão do fundador do Telegram. Configura-se, assim, um circuito financeiro que permite a entrada de recursos por vias convencionais e sua posterior saída em ativos de difícil rastreamento.



O perfil das vítimas é heterogêneo, justamente em função da acessibilidade dessa tecnologia. A gratuidade ou o baixo custo dessas ferramentas tem permitido, por exemplo, que adolescentes produzam conteúdos nudificados de colegas de escola. Também é frequente o uso contra pessoas do convívio do agressor — namoradas, ex-namoradas, familiares —, mas igualmente comum o emprego contra desconhecidos, incluindo celebridades. Essa variedade de alvos reflete, em última instância, o fato de que a barreira técnica para a produção desse tipo de material foi praticamente eliminada.

INÍCIO

1. DEFINIÇÕES & TIPOLOGIAS
 2. VIOLÊNCIA SEXUAL MEDIADA POR IA EM ESCOLAS BRASILEIRAS
 3. PONTOS DE VISTA DE DIFERENTES ATORES ENVOLVIDOS NO FENÔMENO
 4. ECOSISTEMA DIGITAL QUE FACILITA A VIOLÊNCIA SEXUAL MEDIADA POR IA
 5. RESPOSTAS DE DIFERENTES ATORES AO PROBLEMA DA VIOLÊNCIA SEXUAL MEDIADA POR IA
 6. RECOMENDAÇÕES DETALHADAS PARA MÚLTIPLOS ATORES
 7. CONSIDERAÇÕES FINAIS
- SOBRE ESTA PESQUISA

Aplicativos, sites e robôs de nudificação são os primeiros vetores de risco mapeados na produção de material artificial de abuso ou exploração sexual infantojuvenil identificados nesta pesquisa.

HÁ DUAS TÉCNICAS PRINCIPAIS PARA CRIAR ESSAS IMAGENS ÍNTIMAS FALSAS SEM CONSENTIMENTO. NA NUDIFICAÇÃO, A IA REMOVE AS ROUPAS DE UMA PESSOA EM UMA FOTO COMUM, MANTENDO SEU ROSTO SOB UM CORPO NU GERADO DIGITALMENTE. NA TROCA DE ROSTO (FACE SWAP), OCORRE O INVERSO: SUBSTITUI-SE O ROSTO DE ALGUÉM EM UMA CENA PORNOGRÁFICA OU ERÓTICA PELO DA VÍTIMA.

O ato de “nudificar” pode ser feito através de produtos personalizados, como, por exemplo, robôs embutidos em plataformas, aplicativos e websites, mas sempre há de envolver uma pessoa real. Ou seja, a criação da nudificação sempre se baseia em uma ou mais mídias de uma pessoa real.

Entre os produtos personalizados capazes de criar esses conteúdos, os robôs de nudificação

são os mais comuns e serão o foco deste capítulo. Geralmente, esses sistemas funcionam com uma interface simples de comandos para que o usuário solicite o tipo de mídia, posições sexuais e outras funcionalidades como inserção de áudio, no caso de vídeos. Os robôs podem estar disponíveis em redes sociais, aplicativos ou geradores de imagem hospedados na *surface web*. Em muitos casos, o usuário pode criar uma ou duas imagens de teste para, posteriormente, seguir com um maior número de criações em troca de pagamentos.

A origem dessas tecnologias remonta aos primeiros sistemas de IA capazes de realizar *face swap* (em português, troca de rosto. O termo se refere à substituição do rosto de uma vítima pelo de uma pessoa da indústria pornográfica). Com o tempo, novos sistemas foram treinados especificamente para contextos sexuais ou de abuso, dispensando até mesmo o rosto da vítima: passaram a alterar posições corporais e replicar feições a partir de fotografias comuns. O resultado são ferramentas capazes de transformar uma foto de perfil, uma imagem escolar ou qualquer registro cotidiano em material sexualmente explícito — em segundos, sem qualquer conhecimento técnico.



A preocupação com essa disseminação massiva por meio das ferramentas de nudificação não é exclusiva da SaferNet Brasil. No *Dia da Internet Segura*, 10 de fevereiro de 2026, o INHOPE — rede internacional de canais de denúncias da qual a SaferNet Brasil faz parte — emitiu **um documento** conjunto com organizações de proteção infantil de todo o mundo pedindo o banimento global de aplicativos de nudificação.²⁶

“Ferramentas de nudificação usam IA para gerar imagens de nudez a partir de fotos vestidas. Embora comercializadas principalmente para adultos, estas ferramentas são frequentemente utilizadas indevidamente para criar imagens nuas não consensuais de pessoas e são cada vez mais utilizadas para produzir imagens ilegais de crianças”, diz a nota, da qual a SaferNet Brasil foi signatária.

O documento afirma que essas ferramentas não têm nenhum uso positivo, sendo utilizadas exclusivamente para humilhar, assediar e explorar — cada vez mais para produzir imagens ilegais de crianças, frequentemente por outras crianças, além de estarem associadas a casos de extorsão sexual. **“As empresas, desenvolvedores, e os indivíduos que os criam ou distribuem devem ser responsabilizados e enfrentar con-**

INÍCIO

1. DEFINIÇÕES & TIPOLOGIAS
 2. VIOLÊNCIA SEXUAL MEDIADA POR IA EM ESCOLAS BRASILEIRAS
 3. PONTOS DE VISTA DE DIFERENTES ATORES ENVOLVIDOS NO FENÔMENO
 4. ECOSISTEMA DIGITAL QUE FACILITA A VIOLÊNCIA SEXUAL MEDIADA POR IA
 5. RESPOSTAS DE DIFERENTES ATORES AO PROBLEMA DA VIOLÊNCIA SEXUAL MEDIADA POR IA
 6. RECOMENDAÇÕES DETALHADAS PARA MÚLTIPLOS ATORES
 7. CONSIDERAÇÕES FINAIS
- SOBRE ESTA PESQUISA

seqüências legais e criminais. A tecnologia que permite a criação de imagens de nudez e de abuso sexual têm um impacto insidioso na vida das crianças em todo o mundo. Imagens de crianças, incluindo sobreviventes de violência sexual na infância, estão sendo manipuladas de formas cada vez mais extremas e abusivas.”

“Esta tecnologia não é apenas uma ferramenta, é um novo vetor de exploração, colocando mulheres e crianças — especialmente as meninas — em um risco sem precedentes. Esse tipo de exploração não tem lugar em nenhuma sociedade que valorize a segurança e a proteção da infância.”

Embora plataformas de IA generalistas tenham investido progressivamente em salvaguardas, os aplicativos de nudificação — construídos sobre modelos de difusão de código aberto treinados em imagens pornográficas — removem esse obstáculo por completo. A barreira técnica, que antes funcionava como um limitador natural, deixou de existir: qualquer pessoa com acesso a uma fotografia e a um dispositivo conectado à internet é hoje capaz de produzir esse tipo de material. É essa combinação entre acessibilidade técnica, ausência de verificação de idade e ampla disponibilidade que torna os aplicativos

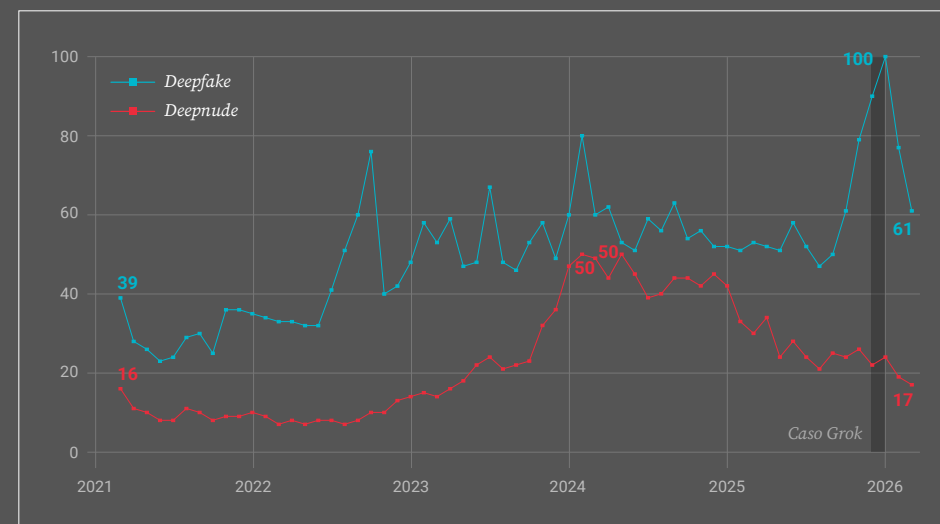
de nudificação o vetor de maior preocupação no ecossistema analisado por esta pesquisa.

O ecossistema é amplo e fragmentado: plataformas na internet aberta, aplicativos móveis, robôs em serviços de mensagens e marketpla-

ces alternativos coexistem e se retroalimentam. O modelo de negócio predominante é o freemium ou por créditos — a imagem de teste é gratuita, mas a remoção de marcas d'água, o aumento de resolução ou a produção em escala exigem assinaturas pagas ou moedas virtuais.

FIGURA 19

Gráfico com dados do Google Trends por pesquisas de brasileiros sobre termos como “deepfakes” e “deepnudes”. Os dados abrangem o período de 1 de março de 2021 a 16 de março de 2026.



INÍCIO

1. DEFINIÇÕES
& TIPOLOGIAS
 2. VIOLÊNCIA SEXUAL
MEDIADA POR
IA EM ESCOLAS
BRASILEIRAS
 3. PONTOS DE VISTA
DE DIFERENTES
ATORES
ENVOLVIDOS
NO FENÔMENO
 4. ECOSISTEMA
DIGITAL QUE
FACILITA A
VIOLÊNCIA SEXUAL
MEDIADA POR IA
 5. RESPOSTAS
DE DIFERENTES
ATORES AO
PROBLEMA DA
VIOLÊNCIA SEXUAL
MEDIADA POR IA
 6. RECOMENDAÇÕES
DETALHADAS PARA
MÚLTIPLOS ATORES
 7. CONSIDERAÇÕES
FINAIS
- SOBRE ESTA
PESQUISA

Dados do Google Trends confirmam a tendência: o volume atingiu picos em junho de 2024 e em janeiro de 2026 — este último coincidindo com o caso Grok. A média de buscas pelo termo *deepnude* no Brasil é de 21,87, e por *deepfake*, de 45,8, indicando interesse persistente ao longo de todo o período analisado.

O caso do Grok ilustra com precisão o que ocorre quando ferramentas de IA são integradas a plataformas de escala massiva sem salvaguardas adequadas.²⁷ O Grok é um robô de IA generativa criado pela xAI, empresa de desenvolvimento de IA e subsidiária da rede social do X (antigo Twitter). Ele possui uma versão de assinatura, mas antes poderia ser utilizado gratuitamente por usuários do X apenas marcando @grok em uma publicação; a funcionalidade foi limitada para somente usuários pagos após o caso. Em janeiro de 2026, usuários relataram que o Grok-2 produziu material ou conteúdo de violência sexual contra crianças e adolescentes a partir de *prompts* com imagens reais de crianças publicadas no X, que conta com mais de 500 milhões de usuários mensais e um algoritmo orientado à viralidade.²⁸ A funcionalidade foi suspensa temporariamente pela xAI após alguns dias e representantes da xAI admitiram publicamente a existência de “falhas nos mecanismos de proteção” no Grok. A empresa reconheceu que os filtros de segurança foram insuficientes para impedir pedidos que resultaram em imagens de crianças e adolescentes em contextos sexuais ou de exploração.²⁹

INÍCIO

1. DEFINIÇÕES
& TIPOLOGIAS
 2. VIOLÊNCIA SEXUAL
MEDIADA POR
IA EM ESCOLAS
BRASILEIRAS
 3. PONTOS DE VISTA
DE DIFERENTES
ATORES
ENVOLVIDOS
NO FENÔMENO
 4. ECOSISTEMA
DIGITAL QUE
FACILITA A
VIOLÊNCIA SEXUAL
MEDIADA POR IA
 5. RESPOSTAS
DE DIFERENTES
ATORES AO
PROBLEMA DA
VIOLÊNCIA SEXUAL
MEDIADA POR IA
 6. RECOMENDAÇÕES
DETALHADAS PARA
MÚLTIPLOS ATORES
 7. CONSIDERAÇÕES
FINAIS
- SOBRE ESTA
PESQUISA

4.1.1. ESTUDO DE CASO

ROBÔS DE NUDIFICAÇÃO NO TELEGRAM

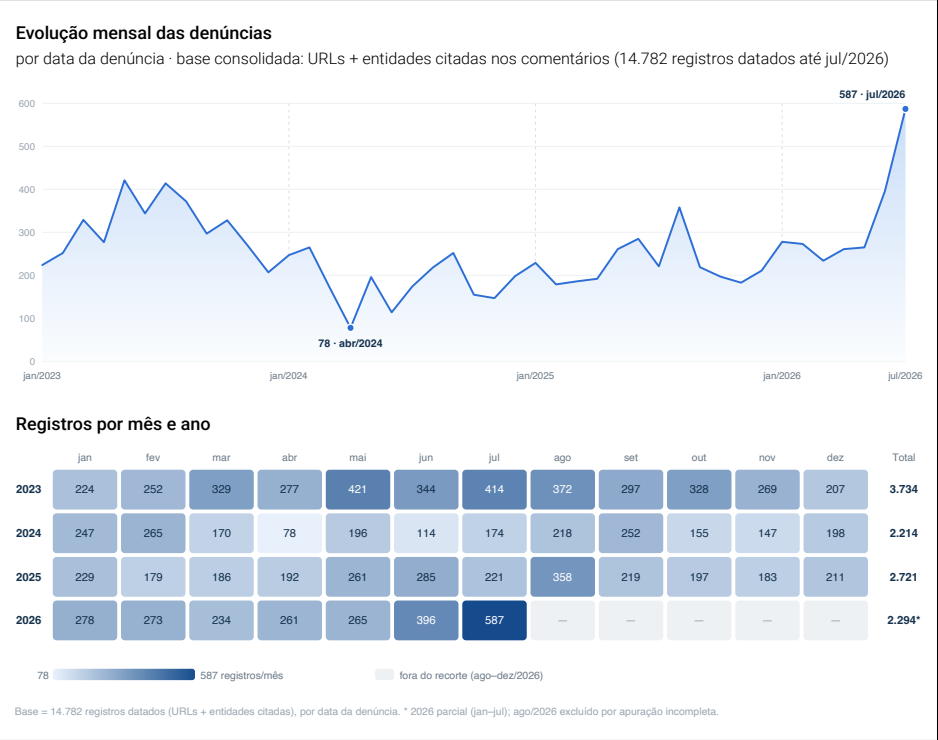
Os achados relativos aos aplicativos de nudificação — nos quais 100% das URLs analisadas apresentaram *deepfakes* de crianças e crimes contra crianças — evidenciam que essas ferramentas não são um vetor de risco entre outros, mas o elo mais diretamente associado à produção de conteúdo artificial de abuso sexual infantojuvenil. Parte significativa desse ecossistema opera por meio de robôs (*bots*) hospedados em plataformas de mensagens, que permitem a criação de imagens sintéticas de nudez a partir de fotografias reais, sem necessidade de conhecimento técnico por parte do usuário. É nesse contexto que o Telegram se impõe como estudo de caso incontornável: a arquitetura aberta de sua API (*Application Programming Interface*), a possibilidade de criação de robôs automatizados sem supervisão prévia e a postura historicamente refratária da plataforma à cooperação com autoridades criam um ambiente particularmente propício à proliferação dessas ferramentas.

Entre 2023 e 2026, a SaferNet Brasil recebeu mais de 8 mil denúncias de potenciais conteúdos de abuso ou exploração sexual infantil no Telegram — aplicativo de mensagens com mais de 950 milhões de usuários mensais. O dado, por si só, já seria suficiente para justificar atenção; mais revelador, contudo, é o sentido da curva: mesmo restrito ao primeiro trimestre de 2026, o mês de janeiro deste ano já havia superado em volume de denúncias todos os janeiros anteriores do período analisado — 249 registros, contra 203 em janeiro de 2025, 205 em 2024 e 212 em 2023.

INÍCIO

- 1. DEFINIÇÕES & TIPOLOGIAS
 - 2. VIOLÊNCIA SEXUAL MEDIADA POR IA EM ESCOLAS BRASILEIRAS
 - 3. PONTOS DE VISTA DE DIFERENTES ATORES ENVOLVIDOS NO FENÔMENO
 - 4. ECOSISTEMA DIGITAL QUE FACILITA A VIOLÊNCIA SEXUAL MEDIADA POR IA
 - 5. RESPOSTAS DE DIFERENTES ATORES AO PROBLEMA DA VIOLÊNCIA SEXUAL MEDIADA POR IA
 - 6. RECOMENDAÇÕES DETALHADAS PARA MÚLTIPLOS ATORES
 - 7. CONSIDERAÇÕES FINAIS
- SOBRE ESTA PESQUISA

FIGURA 20
Distribuição mensal de denúncias à SaferNet Brasil sobre potenciais materiais de abuso ou exploração sexual infantil com o domínio “t.me/”. Os dados abrangem o período de 1 de janeiro de 2023 a 30 de junho de 2026.



Uma pesquisa conduzida pela SaferNet Brasil em outubro de 2024, no âmbito do projeto *DISCOVER*, aprofundou esse diagnóstico ao identificar centenas de canais e grupos na plataforma com conteúdo e comportamento criminoso que extrapolavam a comercialização e a distribuição de material ou conteúdo de violência sexual contra crianças e adolescentes. Entre os conteúdos encontrados estavam imagens íntimas compartilhadas sem consentimento, cenas de zoofilia e material de exploração sexual infantojuvenil gerado por IA — evidências de que o Telegram opera não apenas como repositório de conteúdo ilícito preexistente, mas como ecossistema ativo de produção e circulação de novas formas de violência digital.

Em agosto de 2024, o fundador e CEO do Telegram, Pavel Durov, foi detido no aeroporto de *Le Bourget*, em Paris, e formalmente indiciado pelas autoridades francesas por cumplicidade na distribuição de conteúdo de exploração sexual infantil, além de facilitação de crime organizado, tráfico ilícito e fraude. Um ano após sua prisão, Durov disse “preferir morrer” do

INÍCIO

1. DEFINIÇÕES & TIPOLOGIAS

2. VIOLÊNCIA SEXUAL MEDIADA POR IA EM ESCOLAS BRASILEIRAS

3. PONTOS DE VISTA DE DIFERENTES ATORES ENVOLVIDOS NO FENÔMENO

4. ECOSISTEMA DIGITAL QUE FACILITA A VIOLÊNCIA SEXUAL MEDIADA POR IA

5. RESPOSTAS DE DIFERENTES ATORES AO PROBLEMA DA VIOLÊNCIA SEXUAL MEDIADA POR IA

6. RECOMENDAÇÕES DETALHADAS PARA MÚLTIPLOS ATORES

7. CONSIDERAÇÕES FINAIS

SOBRE ESTA PESQUISA

que autorizar terceiros a acessar mensagens privadas de usuários do seu aplicativo.³⁰ Ainda que, nos últimos anos, o Telegram tenha anunciado compromissos com a ampliação de seus esforços de moderação (estabelecendo parcerias com outros canais de denúncia, por exemplo, como a *Internet Watch Foundation*, o hotline do Reino Unido), o aplicativo permanece como um dos principais vetores de risco para atividade cibercriminal, incluindo exploração sexual infantojuvenil.

A persistência e o crescimento de denúncias à SaferNet Brasil nesse âmbito indica que as medidas adotadas até o momento estão aquém da dimensão do problema. A tendência de aumento de atividade criminosa no Telegram também é corroborada pela análise das denúncias e pela literatura especializada sobre o tema há ao menos cinco anos.³¹

Em novembro de 2025, uma influenciadora narrou anonimamente à SaferNet Brasil ter sido vítima de *deepfakes* sexuais sem o seu consentimento. A jovem havia acabado de fazer 18 anos.

As imagens fabricadas tiveram como base vídeos da influenciadora publicados na plataforma Twitch. Ela relatou: “Eu tinha 18 anos, mas ninguém garante que as pessoas que fizeram isso sabiam que eu tinha 18. E eu sei que se eu tivesse entrado em algum processo não teria dado em nada — tem várias situações que prejudicam a gente e essa é a situação mais grave e mesmo assim não dá em nada quando vai pra justiça.”

A crescente atratividade do Telegram para atividades criminosas decorre de um conjunto de fatores estruturais que se reforçam mutuamente. A plataforma é amplamente percebida como segura e avessa à cooperação com autoridades — reputação construída ao longo de anos de resistência a pedidos governamentais de acesso a dados —, o que transmite aos agentes criminosos uma sensação de impunidade. Somam-se a isso a facilidade para criar ou ingressar em comunidades, a permissão para canais com até 200.000 usuários e a possibilidade de criar contas com números de telefone descartáveis, o que torna a rastreabilidade dos responsáveis extraordinariamente difícil.

O funcionamento desses grupos segue, no geral, um padrão deliberadamente fragmentado e com camadas de acesso progressivo. Canais públicos ou semipúblicos funcionam como vitrines e repositórios de links, nos quais usuários normalmente não compartilham mídias diretamente, mas direcionam o tráfego para comunidades privadas — frequentemente pagas — onde a troca de conteúdo ilícito de fato ocorre. Os usuários que operam como intermediários ou promotores costumam apagar suas contas com rapidez após a divulgação, aproveitando-se do fato de que o Telegram permite que um mesmo número de telefone seja utilizado para criar e recriar contas repetidas vezes, sem qualquer registro permanente.

Diferentemente de outras comunidades que compartilham material de exploração sexual infantil, como fóruns da *dark web* — tradicionalmente voltados à curadoria e distribuição de conteúdo já existente —, alguns grupos que consomem conteúdo “real” vêm incorporando uma dimensão adicional: a promoção e circulação de ferramentas de IA capazes de sinteti-

INÍCIO

1. DEFINIÇÕES
& TIPOLOGIAS
 2. VIOLÊNCIA SEXUAL
MEDIADA POR
IA EM ESCOLAS
BRASILEIRAS
 3. PONTOS DE VISTA
DE DIFERENTES
ATORES
ENVOLVIDOS
NO FENÔMENO
 4. ECOSISTEMA
DIGITAL QUE
FACILITA A
VIOLÊNCIA SEXUAL
MEDIADA POR IA
 5. RESPOSTAS
DE DIFERENTES
ATORES AO
PROBLEMA DA
VIOLÊNCIA SEXUAL
MEDIADA POR IA
 6. RECOMENDAÇÕES
DETALHADAS PARA
MÚLTIPLOS ATORES
 7. CONSIDERAÇÕES
FINAIS
- SOBRE ESTA
PESQUISA

FIGURA 21

Grupo do Telegram denunciado à SaferNet Brasil por potencial conteúdo de abuso e exploração sexual infantil avaliado no âmbito desta pesquisa.



zar imagens. Nesse cenário, a dinâmica deixa de se limitar à troca e comercialização de acervos e passa a incluir a produção ativa de novos conteúdos. Um elemento distintivo, em comparação com ambientes mais restritos como a *dark web*, é a acessibilidade: plataformas como o Telegram são gratuitas e amplamente disponíveis a qualquer usuário com acesso a lojas de aplicativos ou a um computador, o que reduz barreiras de entrada e amplia o alcance dessas práticas. Essa mudança de paradigma eleva significativamente o risco, ao diminuir a dependência de acervos preexistentes e viabilizar a geração de material ilícito em escala.

No que diz respeito aos meios de pagamento, os conteúdos e as ferramentas ilícitas de nudeificação no Telegram são comercializados majoritariamente por criptomoedas — com destaque crescente para o Monero, cuja arquitetura dificulta o rastreamento em comparação com o Bitcoin —, além de PayPal e, mais recentemente, pelas próprias *Telegram Stars* (Estrelas do Telegram). **Essa diversificação de meios de pagamento não é acidental: ela reflete um**

esforço deliberado de adaptação dos operadores a diferentes perfis de usuário, combinando a opacidade das criptomoedas com a conveniência de sistemas integrados à própria plataforma.

As Estrelas são a moeda digital interna do Telegram, utilizadas para adquirir serviços de robôs ou jogos e para remunerar seus criadores. Somente Estrelas recebidas como pagamento por conteúdo ou serviços podem ser sacadas — Estrelas compradas por usuários para uso pessoal não são elegíveis para retirada. Para efetuar o saque, é necessário acumular um mínimo de 1.000 unidades e aguardar 21 dias para o recebimento. As Estrelas também podem ser empregadas para comprar anúncios no Telegram e itens colecionáveis ou convertidas em Toncoin — a criptomoeda da *The Open Network* (TON) — por meio da plataforma³² Fragment. A TON é uma blockchain originalmente desenvolvida por Nikolai Durov, cofundador do Telegram e irmão de Pavel Durov, e é caracterizada por oferecer alto grau de privacidade e anonimato, com uma infraestrutura descentralizada e supervisão mí-

INÍCIO

1. DEFINIÇÕES & TIPOLOGIAS

2. VIOLÊNCIA SEXUAL MEDIADA POR IA EM ESCOLAS BRASILEIRAS

3. PONTOS DE VISTA DE DIFERENTES ATORES ENVOLVIDOS NO FENÔMENO

4. ECOSISTEMA DIGITAL QUE FACILITA A VIOLÊNCIA SEXUAL MEDIADA POR IA

5. RESPOSTAS DE DIFERENTES ATORES AO PROBLEMA DA VIOLÊNCIA SEXUAL MEDIADA POR IA

6. RECOMENDAÇÕES DETALHADAS PARA MÚLTIPLOS ATORES

7. CONSIDERAÇÕES FINAIS

SOBRE ESTA PESQUISA

nima mesmo quando comparada aos padrões já permissivos do mercado de criptoativos.

Embora o Telegram tenha formalmente se afastado da TON em 2020, após um litígio da Comissão de Valores Mobiliários dos EUA (a SEC)³³, essa separação é cada vez mais tênue. Em maio de 2026, Pavel Durov anunciou que o Telegram substituirá a TON Foundation como principal força motriz da blockchain e se tornará o seu maior validador.³⁴ Na prática, isso significa que um operador de *bot* de nudificação pode receber pagamentos dentro do próprio ecossistema do Telegram, convertê-los em criptomoeda e sacá-los — tudo sem sair da infraestrutura vinculada ao aplicativo. O Telegram, aliás, retém uma comissão de 15% a 30% sobre cada saque de Estrelas, o que significa que lucra indiretamente com cada transação realizada por esses e outros operadores criminosos.

Uma avaliação dos termos de uso e de privacidade da Fragment em seu site — na versão mais recente, de abril de 2026, não deixa claro como a plataforma se diferencia do próprio Te-

legram além do nome empresarial Fragment Corp. como operadora, uma vez que todos os serviços oferecidos envolvem diretamente o aplicativo de mensagens. Os próprios termos de uso da Fragment confirmam essa arquitetura circular e revelam uma estrutura deliberadamente voltada ao anonimato. A empresa está registrada na República de Seychelles, jurisdição offshore conhecida por regulação financeira permissiva, sob o número de registro 221287. Seus termos de uso definem a plataforma como “não custodial”, o que significa que ela não detém fundos nem controla (ou modera) carteiras de usuários, o que difere completamente de outras plataformas de compra e venda de criptoativos.

O processamento da compra de Estrelas no Brasil pode ser feito via PIX ou por meio das lojas de aplicativos (Google Play Store ou Apple App Store). Essas lojas retêm uma comissão de 30% do valor total; o Telegram aplica uma taxa administrativa que varia entre 4% e 5%; e o restante — aproximadamente 65% — é repassado ao desenvolvedor ou criador de conteúdo.

Historicamente, Apple e Google impõem uma comissão de 30% sobre qualquer transação de bens e serviços digitais realizada dentro de aplicativos hospedados em suas plataformas. Essa política gerou disputas jurídicas globais de grande repercussão (como o caso *Epic Games vs Apple*).

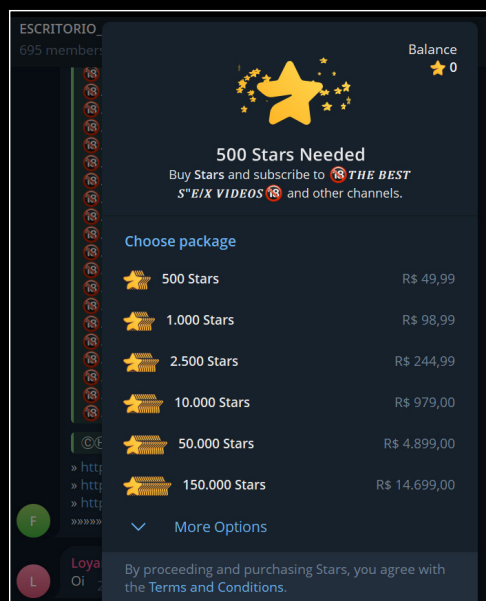
No caso do canal exemplificado a seguir, cuja assinatura mínima é de 500 Estrelas (US\$ 9,99), estima-se que o operador receba cerca de US\$ 6,50 por assinante — valor que, multiplicado por centenas ou milhares de membros pagantes, gera receita recorrente significativa. Em um contexto de comercialização de material ilícito, esse modelo fragmenta o fluxo financeiro entre ao menos três camadas — loja de aplicativo, plataforma e destinatário final —, dificultando tanto o rastreamento quanto a responsabilização.

INÍCIO

1. DEFINIÇÕES & TIPOLOGIAS
 2. VIOLÊNCIA SEXUAL MEDIADA POR IA EM ESCOLAS BRASILEIRAS
 3. PONTOS DE VISTA DE DIFERENTES ATORES ENVOLVIDOS NO FENÔMENO
 4. ECOSISTEMA DIGITAL QUE FACILITA A VIOLÊNCIA SEXUAL MEDIADA POR IA
 5. RESPOSTAS DE DIFERENTES ATORES AO PROBLEMA DA VIOLÊNCIA SEXUAL MEDIADA POR IA
 6. RECOMENDAÇÕES DETALHADAS PARA MÚLTIPLOS ATORES
 7. CONSIDERAÇÕES FINAIS
- SOBRE ESTA PESQUISA

FIGURAS 22 E 23

Canal de compartilhamento e venda de material real de abuso sexual infantil e conteúdos gerados por IA pede compra de *Telegram Stars* para acesso.



Star Amount	Price for Users	Apple / Google 30% Commission	Telegram Administrative Costs	Developer Proceeds
50 ★	\$0.99	\$0.30	\$0.04	\$0.65
75 ★	\$1.49	\$0.45	\$0.07	\$0.98
100 ★	\$1.99	\$0.60	\$0.09	\$1.30
150 ★	\$2.99	\$0.90	\$0.14	\$1.95
250 ★	\$4.99	\$1.50	\$0.24	\$3.25
350 ★	\$6.99	\$2.10	\$0.34	\$4.55
500 ★	\$9.99	\$3.00	\$0.49	\$6.50
750 ★	\$14.99	\$4.50	\$0.74	\$9.75
1000 ★	\$19.99	\$6.00	\$0.99	\$13.00
1500 ★	\$29.99	\$9.00	\$1.49	\$19.50
2500 ★	\$49.99	\$15.00	\$2.49	\$32.50

Tabela retirada do FAQ do Telegram sobre as Estrelas apontando as principais taxas de comissão quando as moedas são compradas por intermédio das lojas de aplicativo.

INÍCIO

1. DEFINIÇÕES
& TIPOLOGIAS

2. VIOLÊNCIA SEXUAL
MEDIADA POR
IA EM ESCOLAS
BRASILEIRAS

3. PONTOS DE VISTA
DE DIFERENTES
ATORES
ENVOLVIDOS
NO FENÔMENO

4. ECOSSISTEMA
DIGITAL QUE
FACILITA A
VIOLÊNCIA SEXUAL
MEDIADA POR IA

5. RESPOSTAS
DE DIFERENTES
ATORES AO
PROBLEMA DA
VIOLÊNCIA SEXUAL
MEDIADA POR IA

6. RECOMENDAÇÕES
DETALHADAS PARA
MÚLTIPLOS ATORES

7. CONSIDERAÇÕES
FINAIS

SOBRE ESTA
PESQUISA

Para lançar as *Telegram Stars* como a moeda interna oficial para o ecossistema de *Mini Apps*, robôs e conteúdos pagos, o Telegram foi obrigado a integrar os sistemas de pagamento *In-App Purchase* dessas corporações. No entanto, ciente de que a comissão de 30% limitaria sua margem de lucro e comprometeria parcela significativa da remuneração de criadores de conteúdo e desenvolvedores de robôs, o Telegram arquitetou e deliberadamente incentiva rotas alternativas de aquisição que barateiam o custo unitário da Star para o usuário final, conforme evidenciado ao lado.

FIGURA 24

Formas de Pagamento e Comissões para compra de Telegram Stars, segundo a documentação oficial da TON Wallet³⁵

Forma de pagamento e origem	Intermediário principal	Comissões e taxas	Contexto e regras da documentação
Compra integrada (iOS)	Apple App Store	~30% de comissão (Taxa padrão da Apple para produtos digitais, já embutida no valor final do pacote).	A carteira TON Wallet redireciona o usuário para a interface de compra da App Store. Disponível em pacotes fixos de 100, 250 ou 500 Stars.
Compra integrada (Android)	Google Play Store	~30% de comissão (Taxa padrão da Google para produtos digitais, já embutida no valor final do pacote).	A carteira TON Wallet redireciona o usuário para a interface de compra da Google Play. Disponível em pacotes fixos de 100, 250 ou 500 Stars.
Transferência de saldo geral	Telegram Messenger	0% de taxa para a transferência interna.	Permite mover Stars que já estejam no saldo geral da conta do Telegram para o saldo da TON Wallet.
Alternativa direta (Fallback)	Rede Blockchain TON	Variável (Taxa de rede calculada dinamicamente baseada na volatilidade do Toncoin).	Caso o usuário não possua Stars na TON Wallet, a documentação aponta que o usuário pode alterar a opção de pagamento da taxa para TON (Toncoin) diretamente na tela antes de confirmar o envio ou swap.

INÍCIO

1. DEFINIÇÕES
& TIPOLOGIAS
 2. VIOLÊNCIA SEXUAL
MEDIADA POR
IA EM ESCOLAS
BRASILEIRAS
 3. PONTOS DE VISTA
DE DIFERENTES
ATORES
ENVOLVIDOS
NO FENÔMENO
 4. ECOSISTEMA
DIGITAL QUE
FACILITA A
VIOLÊNCIA SEXUAL
MEDIADA POR IA
 5. RESPOSTAS
DE DIFERENTES
ATORES AO
PROBLEMA DA
VIOLÊNCIA SEXUAL
MEDIADA POR IA
 6. RECOMENDAÇÕES
DETALHADAS PARA
MÚLTIPLOS ATORES
 7. CONSIDERAÇÕES
FINAIS
- SOBRE ESTA
PESQUISA

Para compreender a engenharia financeira e regulatória que permite ao Telegram oferecer descontos significativos contornando o monopólio das lojas de aplicativos (Apple App Store e Google Play Store), é necessário analisar a arquitetura multicanal criada pela plataforma, que une soluções tradicionais de pagamento e a infraestrutura Web3 da blockchain TON (*The Open Network*).

A precificação das Telegram Stars varia de acordo com o nível de intermediação bancária e regulatória de cada canal de venda. Ao comprar Stars no cartão de crédito via app móvel, o KYC — *Know Your Customer*, prática regulatória desenhada para assegurar a identidade dos clientes, suas atividades financeiras e os riscos que eles podem representar — é herdado da conta do usuário na Apple Play Store ou Google Pay, que possuem cartões de crédito vinculados, nomes reais e CPFs/identidades rastreáveis). Para comprar Toncoin utilizando moedas como o Real (BRL) e depois gastar na Fragment, o usuário precisa transferir fundos de uma corretora (como Binance, OKX, Merca-

do Bitcoin). Essas corretoras aplicam KYC estrito (conferência de documentos e biometria facial) por força de regulação bancária.

Essas transações, geralmente feitas apenas no mobile, estão submetidas a níveis maiores de escrutínio e conformidade regulatória, ao contrário das transações efetuadas fora da interface mobile que se valem do ecossistema expandido do Telegram formado pela Fragment, pelo @PremiumBot e por diversas plataformas de gift cards, marketplaces de jogos (como Kinguin, LootBar, GGSe) e sites de recarga terceirizados (MooGold, Joytify) que operam por meio de um modelo paralelo conhecido como arbitragem de ecossistema.

O grande diferencial desse método é que ele não exige as credenciais de login (e-mail ou senha) do comprador, demandando exclusivamente o arroba (@username) da conta do Telegram. O funcionamento desse mercado paralelo, desde a compra até a entrega do saldo, é idêntico ao processo de compra de créditos para um jogo online (como Free Fire ou Roblox).

INÍCIO

1. DEFINIÇÕES
& TIPOLOGIAS

2. VIOLÊNCIA SEXUAL
MEDIADA POR
IA EM ESCOLAS
BRASILEIRAS

3. PONTOS DE VISTA
DE DIFERENTES
ATORES
ENVOLVIDOS
NO FENÔMENO

4. ECOSISTEMA
DIGITAL QUE
FACILITA A
VIOLÊNCIA SEXUAL
MEDIADA POR IA

5. RESPOSTAS
DE DIFERENTES
ATORES AO
PROBLEMA DA
VIOLÊNCIA SEXUAL
MEDIADA POR IA

6. RECOMENDAÇÕES
DETALHADAS PARA
MÚLTIPLOS ATORES

7. CONSIDERAÇÕES
FINAIS

SOBRE ESTA
PESQUISA

Aproximadamente, a estrutura de custos por unidade de Star se desenha da seguinte forma:

FIGURA 25

Comparativo de custos e mecanismos de desoneração por canal de compra.

Canal de compra	Meio de pagamento	Preço estimado por Star	Mecanismo de desoneração ou de desconto
In-App (iOS / Android)	Apple Pay / Google Pay	~\$0,025 (Preço Base)	Nenhum. Sofre a incidência integral da comissão de 30% das Big Techs.
@PremiumBot (Versões Direct/Web)	Cartão de Crédito / Processadores locais	~\$0,021 (~15% a 20% de desconto)	Parcial. Elimina os 30% das lojas mobile, mas ainda arca com taxas de adquirentes de cartão (Visa/Mastercard).
Fragment.com	Criptoativos Toncoin (\$TON)	~\$0,015 (Até 40% de desconto)	Total. Desintermediação bancária e de Big Techs. Transação direta peer-to-peer na blockchain, com possibilidade de ofuscamento da origem e pseudo animização das transações.

Na prática, o usuário acessa o site terceirizado (ex: Kinguin) e escolhe uma quantia de Stars (os pacotes variam de pequenas frações, como 50 ou 100 Stars, até volumes massivos de 50.000 ou 100.000 Stars). No momento do checkout, o site exibe um campo obrigatório para preenchimento: “Enter Telegram Username” ou “Player ID”. O usuário digita seu @nomedeusuário. O comprador realiza o pagamento utilizando os métodos integrados daquela plataforma específica (que podem incluir Pix, cartões de crédito nacionais, boleto, PayPal ou outras criptomoedas).

Como o Telegram não tem um sistema tradicional de cartões-presente com códigos de raspadinha de 16 dígitos, como acontece com os marketplaces, serviços de *streaming*, games e lojas de aplicativos, os vendedores dos sites que formam o ecossistema expandido do Telegram utilizam a plataforma Fragment e a automação via API para injetar o saldo na conta do comprador.

A Fragment possui uma ferramenta oficial que permite a qualquer pessoa comprar e enviar Telegram Stars para qualquer conta pública do

INÍCIO

1. DEFINIÇÕES & TIPOLOGIAS

2. VIOLÊNCIA SEXUAL MEDIADA POR IA EM ESCOLAS BRASILEIRAS

3. PONTOS DE VISTA DE DIFERENTES ATORES ENVOLVIDOS NO FENÔMENO

4. ECOSISTEMA DIGITAL QUE FACILITA A VIOLÊNCIA SEXUAL MEDIADA POR IA

5. RESPOSTAS DE DIFERENTES ATORES AO PROBLEMA DA VIOLÊNCIA SEXUAL MEDIADA POR IA

6. RECOMENDAÇÕES DETALHADAS PARA MÚLTIPLOS ATORES

7. CONSIDERAÇÕES FINAIS

SOBRE ESTA PESQUISA

Telegram, bastando digitar o @username do destinatário e pagar a fatura utilizando a criptomoeda Toncoin (\$TON) através de uma carteira Web3. Assim que o pagamento do usuário é aprovado no site de recargas (ex: Kinguin), o sistema do vendedor (ou um *bot* automatizado integrado à API da Fragment) é acionado.

Esses sites existem porque lucram com a diferença de preços (*spread*) gerada pelas taxas das Big Techs.

- Preço na Apple/Google (In-App): 100 Stars custam cerca de US\$ 2,99 (devido aos 30% de comissão).
- Preço na Fragment (Puro em Cripto): 100 Stars custam o equivalente a aproximadamente US\$ 1,50 a US\$ 1,60 em Toncoin.
- Preço nos Sites de Recarga (Kinguin/Joytify): as mesmas 100 Stars são vendidas por cerca de US\$ 1,90 a US\$ 2,10.

O site de recarga consegue oferecer um “desconto” atraente para o usuário em relação ao preço oficial do aplicativo mobile e, ao mesmo tempo, embolsa uma margem de lucro considerável em cima do preço de aquisição da cripto na Fragment.

Nesse modelo, o vendedor deve possuir reservas expressivas da criptomoeda Toncoin (\$TON) em carteiras institucionais, similar às contas bolsão de algumas fintechs utilizadas pelo crime organizado no Brasil e descobertas na operação Carbono Oculto. O sistema do vendedor converte o valor pago pelo cliente, acessa a Fragment de forma automatizada, insere o @username fornecido pelo comprador e liquida a transação na blockchain TON. O Telegram reconhece o pagamento feito em TON na Fragment e credita as Stars diretamente no saldo do usuário de forma quase instantânea.

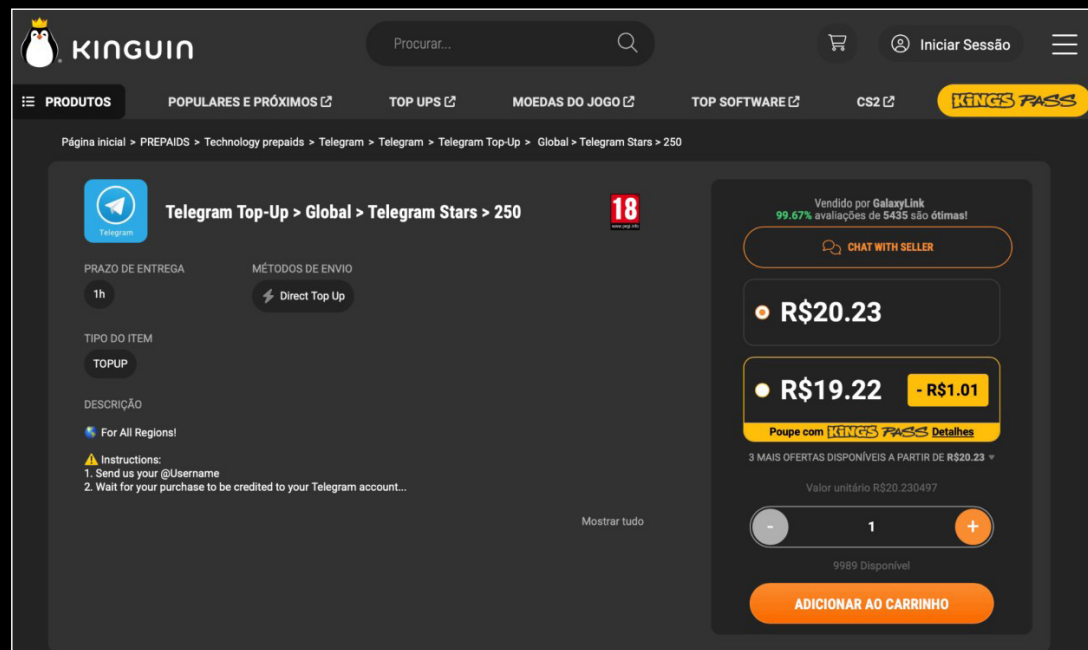
Desse modo, a arquitetura híbrida do Telegram (que mistura a Web2 do aplicativo e dos *bots* com a Web3 da blockchain TON e da Fragment), é estruturada com o objetivo de promover a desintermediação bancária, facilitando assim a prática de ilícitos através de ofuscamento da origem financeira, permitindo a um criminoso utilizar Stars para comprar material ilícito ou serviços de IA generativa voltados ao crime. Para fins didáticos, vejamos o seguinte exemplo: ao adquirir Stars no Kinguin pagando com Pix, o extrato bancário do usuário mostrará apenas uma transação legítima para uma empresa intermediadora de pagamen-

INÍCIO

1. DEFINIÇÕES & TIPOLOGIAS
 2. VIOLÊNCIA SEXUAL MEDIADA POR IA EM ESCOLAS BRASILEIRAS
 3. PONTOS DE VISTA DE DIFERENTES ATORES ENVOLVIDOS NO FENÔMENO
 4. ECOSISTEMA DIGITAL QUE FACILITA A VIOLÊNCIA SEXUAL MEDIADA POR IA
 5. RESPOSTAS DE DIFERENTES ATORES AO PROBLEMA DA VIOLÊNCIA SEXUAL MEDIADA POR IA
 6. RECOMENDAÇÕES DETALHADAS PARA MÚLTIPLOS ATORES
 7. CONSIDERAÇÕES FINAIS
- SOBRE ESTA PESQUISA

FIGURA 26

Captura de tela do site Kinguin com oferta de Estrelas do Telegram a partir de cartões-presente.



tos (gateways de e-commerce). Não haverá qualquer menção ao Telegram ou a criptoativos em seus registros bancários tradicionais.

Como esses sites não exigem senhas ou KYC para compras de baixo valor, o criminoso consegue abastecer sua conta do Telegram com fundos digitais sem precisar expor cartões de crédito pessoais em lojas oficiais como Google Play ou App Store. A investigação, nesse caso, passaria necessariamente pelo marketplace (como o Kinguin) pois é esta entidade que faz a custódia dos dados cadastrais, meios de pagamento e metadados associados (IP, e-mail, timestamp, etc) do usuário que comprou aquele pacote específico para o @username sob investigação. Essa camada adicional não apenas onera forças de segurança, como de fato pode retardar ou até inviabilizar investigações de diversos crimes se o marketplace não tiver representação legal no país, o que ensejaria pedido de cooperação internacional.

As compras financeiras de Stars através das lojas de aplicativos (Apple e Google) são finais

INÍCIO

1. DEFINIÇÕES & TIPOLOGIAS

2. VIOLÊNCIA SEXUAL MEDIADA POR IA EM ESCOLAS BRASILEIRAS

3. PONTOS DE VISTA DE DIFERENTES ATORES ENVOLVIDOS NO FENÔMENO

4. ECOSISTEMA DIGITAL QUE FACILITA A VIOLÊNCIA SEXUAL MEDIADA POR IA

5. RESPOSTAS DE DIFERENTES ATORES AO PROBLEMA DA VIOLÊNCIA SEXUAL MEDIADA POR IA

6. RECOMENDAÇÕES DETALHADAS PARA MÚLTIPLOS ATORES

7. CONSIDERAÇÕES FINAIS

SOBRE ESTA PESQUISA

e não reembolsáveis, conforme os Termos de Serviço do Telegram. No entanto, se o usuário transferiu um pacote de Stars do seu saldo do Telegram para a TON Wallet e não as utilizou, a plataforma permite realizar o estorno desse pacote completo (de 100, 250 ou 500 Stars) de volta para a sua conta geral do Telegram por meio do histórico de transações. Não são permitidos estornos parciais. Essa feature evidencia a existência de um sistema de *billing* e registro interno das transações. Dentro da TON Wallet, as Telegram Stars funcionam apenas para mitigar/pagar taxas de transação da rede e não podem ser convertidas ou trocadas (*swapped*) por outras criptomoedas dentro da interface da carteira, criando um sistema hermeticamente fechado.

Ao permitir que usuários comprem Stars pelas vias oficiais das lojas (mesmo que mais caro), o Telegram mantém seus aplicativos ativos e em conformidade com as diretrizes da Apple e do Google. Quando um usuário gasta suas Stars dentro de um jogo ou canal do Telegram, o desenvolvedor recebe essas Stars. O Telegram

permite que esses criadores saquem seus ganhos convertendo as Stars de volta para Toncoin (\$TON) pela própria Fragment.com.

Ao notar que comprar Stars na Fragment utilizando TON gera uma economia de até 40%, o usuário comum é financeiramente incentivado a aprender a usar uma carteira de criptomoedas, retroalimentando e valorizando o ecossistema da blockchain TON. Em suma: as regras das Big Techs são aplicadas apenas na “porta de entrada” convencional (*mobile*), enquanto o verdadeiro volume financeiro e de desconto é direcionado para a sua própria infraestrutura descentralizada e soberana na Web3.

Na avaliação da SaferNet Brasil, o modelo de negócio dos grupos de exploração sexual infantil vem se consolidando em torno de assinaturas recorrentes — em vez de transações isoladas —, o que gera receita previsível para os operadores. Os valores frequentemente não ultrapassam US\$ 100 mensais para usuários estrangeiros ou R\$ 50 para usuários brasileiros (valor máximo observado ao longo desta análise, ainda que variável), criando

uma barreira de entrada baixa que amplia o alcance dessas redes. A lógica é a mesma de qualquer serviço digital por assinatura: fidelização do usuário, previsibilidade de caixa e escala.

Embora a maioria dos operadores comercialize ferramentas pagas, há também aqueles que oferecem robôs de nudificação gratuitamente. Esses *bots* operam, em geral, no modelo freemium — permitindo a criação de algumas imagens sem custo ou limitando a geração a uma mídia por dia —, enquanto produções adicionais requerem a compra de moedas ou créditos internos. A gratuidade parcial não é um gesto de abertura: é uma estratégia de conversão, desenhada para criar familiaridade com a ferramenta e conduzir o usuário ao pagamento.

Muitos dos robôs também permitem o chamado check-in diário para acúmulo de moedas virtuais — mecanismo pelo qual o usuário precisa acessar o *bot* todos os dias para receber uma bonificação em créditos. Essa dinâmica, inspirada em práticas comuns de engajamento em jogos eletrônicos e aplicativos de fide-

INÍCIO

1. DEFINIÇÕES
& TIPOLOGIAS
 2. VIOLÊNCIA SEXUAL
MEDIADA POR
IA EM ESCOLAS
BRASILEIRAS
 3. PONTOS DE VISTA
DE DIFERENTES
ATORES
ENVOLVIDOS
NO FENÔMENO
 4. ECOSISTEMA
DIGITAL QUE
FACILITA A
VIOLÊNCIA SEXUAL
MEDIADA POR IA
 5. RESPOSTAS
DE DIFERENTES
ATORES AO
PROBLEMA DA
VIOLÊNCIA SEXUAL
MEDIADA POR IA
 6. RECOMENDAÇÕES
DETALHADAS PARA
MÚLTIPLOS ATORES
 7. CONSIDERAÇÕES
FINAIS
- SOBRE ESTA
PESQUISA

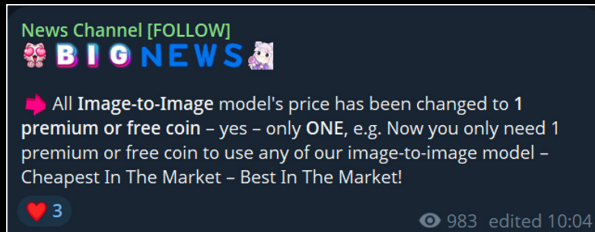


FIGURA 28

Apresentação de bot de nudificação no Telegram: exemplos de deepfakes sexuais, compra de créditos e minijogo para obter créditos gratuitos.

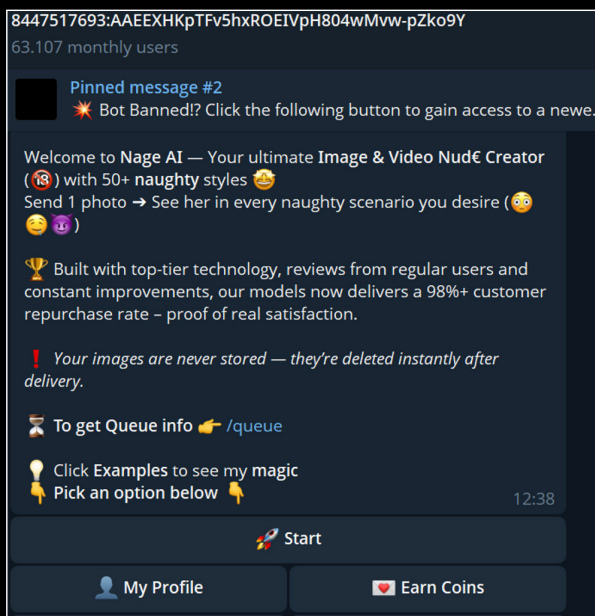


FIGURA 27

Mensagem compartilhada por robô de nudificação no Telegram promovendo novos modelos de criação de deepfakes sexuais.

dade, cumpre uma função estratégica no contexto criminoso: garante o retorno recorrente do usuário à plataforma, aumenta o tempo de permanência nas comunidades ilícitas, dificulta o abandono do ecossistema e, ao distribuir créditos gratuitos de forma periódica, reduz a fricção para que o usuário consuma — e eventualmente pague por — conteúdos de maior gravidade. O resultado é um ciclo de engajamento que reproduz, com precisão, as técnicas de retenção empregadas pela indústria de tecnologia, mas a serviço da exploração sexual.

No escopo principal desta pesquisa — a produção e distribuição de conteúdos criminosos em escolas brasileiras por adolescentes —, esse modelo adquire contornos particularmente preocupantes. Adolescentes, em geral, não dispõem de renda própria para adquirir ou assinar ferramentas digitais sem o conhecimento de seus responsáveis, o que os torna o público mais suscetível a ofertas gratuitas ou baseadas em mecânicas de acúmulo progressivo. A combinação entre gratuidade inicial, gamificação

e acessibilidade técnica configura, na prática, uma porta de entrada desenhada sob medida para quem tem menos de 18 anos — ainda que seus criadores eventualmente aleguem desconhecer o perfil etário de seus usuários.

Em abril de 2026, a organização europeia *AI Forensics* publicou relatório com base na análise de 2,8 milhões de mensagens em 16 grupos do Telegram ao longo de seis semanas, documentando um ecossistema de abuso em escala que alcançou aproximadamente 52 mil pessoas na Itália e na Espanha.³⁶ O relatório identificou que arquivos contendo imagens íntimas não consensuais — incluindo material ou conteúdo de violência sexual contra crianças e adolescentes — eram vendidos por valores entre 20 e 50 euros, e que alguns usuários atuavam como afiliados de bots de nudificação com comissões de até 40%. Segundo a organização, o Telegram viabiliza esse ecossistema por design: grupos fechados, uma vez removidos, eram recriados com o mesmo nome em questão de horas.

INÍCIO

1. DEFINIÇÕES
& TIPOLOGIAS
 2. VIOLÊNCIA SEXUAL
MEDIADA POR
IA EM ESCOLAS
BRASILEIRAS
 3. PONTOS DE VISTA
DE DIFERENTES
ATORES
ENVOLVIDOS
NO FENÔMENO
 4. ECOSISTEMA
DIGITAL QUE
FACILITA A
VIOLÊNCIA SEXUAL
MEDIADA POR IA
 5. RESPOSTAS
DE DIFERENTES
ATORES AO
PROBLEMA DA
VIOLÊNCIA SEXUAL
MEDIADA POR IA
 6. RECOMENDAÇÕES
DETALHADAS PARA
MÚLTIPLOS ATORES
 7. CONSIDERAÇÕES
FINAIS
- SOBRE ESTA
PESQUISA

Em escala global, o relatório da *Internet Watch Foundation* de 2026 registrou 8.029 imagens e vídeos gerados por IA avaliados como material realista de abuso sexual infantil ao longo de 2025, com um aumento de 26.385% nos vídeos em relação ao ano anterior — de 13 em 2024 para 3.443.³⁷ Do total de vídeos identificados, 65% foram classificados como Categoria A — a classificação mais grave segundo a legislação britânica, que abrange atos sexuais penetrativos, sadismo e práticas com animais.

O CONJUNTO DESSAS EVIDÊNCIAS — PRODUZIDAS POR ORGANIZAÇÕES DE DIFERENTES PAÍSES, COM METODOLOGIAS INDEPENDENTES — CONVERGE PARA UMA MESMA CONCLUSÃO: O TELEGRAM NÃO É APENAS UMA PLATAFORMA ONDE O ABUSO ACONTECE, MAS UMA INFRAESTRUTURA QUE, POR SUA ARQUITETURA TÉCNICA, SEU MODELO DE MONETIZAÇÃO E SUA POSTURA DE MODERAÇÃO, ESTRUTURA E VIABILIZA A PRODUÇÃO, A DISTRIBUIÇÃO E A COMERCIALIZAÇÃO DE MATERIAL OU CONTEÚDO DE VIOLÊNCIA SEXUAL CONTRA CRIANÇAS E ADOLESCENTES EM ESCALA INDUSTRIAL.

INÍCIO

1. DEFINIÇÕES & TIPOLOGIAS
 2. VIOLÊNCIA SEXUAL MEDIADA POR IA EM ESCOLAS BRASILEIRAS
 3. PONTOS DE VISTA DE DIFERENTES ATORES ENVOLVIDOS NO FENÔMENO
 4. ECOSISTEMA DIGITAL QUE FACILITA A VIOLÊNCIA SEXUAL MEDIADA POR IA
 5. RESPOSTAS DE DIFERENTES ATORES AO PROBLEMA DA VIOLÊNCIA SEXUAL MEDIADA POR IA
 6. RECOMENDAÇÕES DETALHADAS PARA MÚLTIPLOS ATORES
 7. CONSIDERAÇÕES FINAIS
- SOBRE ESTA PESQUISA

4.2. VETOR DE RISCO N.º2 → DARK WEB



O segundo vetor de risco identificado na produção de material artificial de abuso ou exploração sexual infantojuvenil é a *dark web*. Embora a internet aberta permaneça como o principal meio pelo qual usuários — com mais ou menos de 18 anos — acessam as ferramentas de IA generativa necessárias à produção de conteúdos ilegais, é indispensável compreender o papel dessa camada da internet na crise em curso.

Essa relevância decorre de ao menos três fatores. Primeiro, apesar de ser menos acessada por usuários comuns, é na *dark web* que surgem muitos dos sistemas de IA generativa mais avançados, capazes de gerar imagens com maior grau de realismo e violência. Segundo, é nesse ambiente que organizações criminosas estruturadas operam com capacidade de movimentar volumes financeiros significativos oriundos de cibercrimes, atravessando múltiplas jurisdições — dinâmica já amplamente documentada historicamente. Terceiro, a convergência entre o desenvolvimento de ferramentas ilícitas e o elevado realismo dos conteúdos ali produzidos amplia, na avaliação da SaferNet Brasil, significativamente o potencial de dano dessas operações.

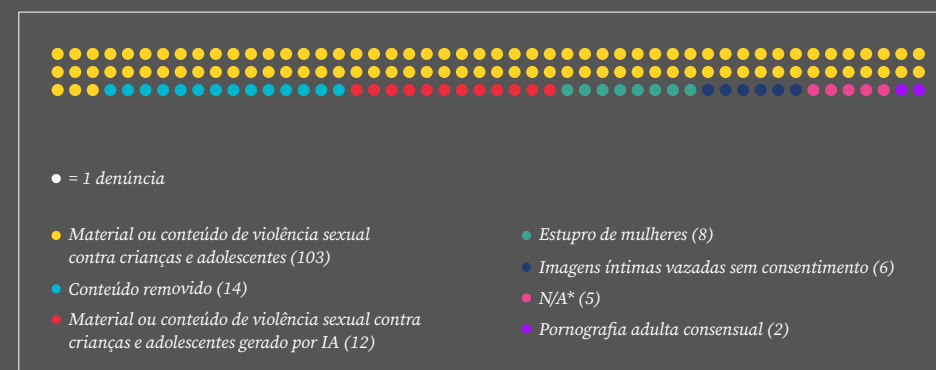
A dimensão do problema é evidenciada por dados recentes. Em março de 2025, uma operação coordenada pela Europol e liderada pela polícia da Baviera, na Alemanha, resultou no desmantelamento do *Kidflix*, uma das maiores plataformas de exploração sexual infantil já identificadas na *dark web*, com aproximadamente 1,8 milhão de usuários registrados em mais de 35 países. A plataforma hospedava cerca de 91 mil vídeos únicos, com uma média de 3,5 novos uploads por hora, e operava um sistema de

recompensas baseado em criptomoedas que incentivava a publicação de conteúdo abusivo.

Para dimensionar esse potencial, a SaferNet Brasil selecionou todos os domínios da rede Tor denunciados à central de denúncias de conteúdos de abuso ou exploração sexual infantil no mês de setembro de 2025. Ao todo, foram analisados 127 links — 108 originalmente denunciados à central e outros 19 identificados durante a própria análise. Entre os domínios

FIGURA 29

Distribuição da análise da SaferNet sobre links da rede Tor denunciados ao canal de denúncias no mês de setembro de 2025.



INÍCIO

1. DEFINIÇÕES & TIPOLOGIAS
 2. VIOLÊNCIA SEXUAL MEDIADA POR IA EM ESCOLAS BRASILEIRAS
 3. PONTOS DE VISTA DE DIFERENTES ATORES ENVOLVIDOS NO FENÔMENO
 4. ECOSISTEMA DIGITAL QUE FACILITA A VIOLÊNCIA SEXUAL MEDIADA POR IA
 5. RESPOSTAS DE DIFERENTES ATORES AO PROBLEMA DA VIOLÊNCIA SEXUAL MEDIADA POR IA
 6. RECOMENDAÇÕES DETALHADAS PARA MÚLTIPLOS ATORES
 7. CONSIDERAÇÕES FINAIS
- SOBRE ESTA PESQUISA

examinados, destacam-se fóruns de discussão frequentados por indivíduos autodeclarados pedófilos e ferramentas voltadas à criação de imagens de abuso sexual infantil. Do total de URLs “onion” analisadas, 81,89% (104) continham material real de abuso ou exploração sexual infantil e 10,24% (13) continham material artificial da mesma natureza. As categorias não são excludentes: oito URLs apresentaram os dois tipos de material simultaneamente.

Para aprofundamento, foram selecionados dois domínios específicos como estudos de caso neste trecho da pesquisa. O primeiro é um fórum dedicado à produção e disseminação de ferramentas de IA generativa e conteúdos artificiais de abuso sexual infantil; o segundo, um gerador de imagens por IA gratuito hospedado na *dark web*, voltado exclusivamente à criação desses mesmos conteúdos criminosos.

O QUE ENCONTRAMOS NESTE VETOR DE RISCO



Comunidades altamente estruturadas e organizadas em sites e fóruns oferecem diferentes formas de “apoio” aos seus membros.



Isso inclui suporte psicológico e emocional entre pares para persuadir membros, novos ou antigos, de que a produção de imagens de exploração sexual infantil utilizando IA não é tão grave por ser artificial, além de fornecerem sistemas e ferramentas para esse fim.



Diálogos entre cibercriminosos indicam que a crescente crise de materiais artificiais de abuso e exploração sexual infantil pode dificultar a atuação das autoridades de proteção à infância. Isso inclui o risco de sobrecarregar os serviços de investigação, ajuda e proteção, além de potencialmente atrapalhar a localização de vítimas reais em situações de emergência.



Comunidades criminosas envolvidas na produção, distribuição e comercialização de material ou conteúdo de violência sexual contra crianças e adolescentes estão atualmente segmentando os websites entre aqueles que veiculam conteúdo “real” e aqueles que veiculam e/ou produzem conteúdo artificial (*deepfakes*). Observa-se uma dissonância cognitiva entre os usuários que consomem ambos os tipos. Alguns defendem que o material artificial representa o “futuro” da exploração sexual e a “solução para seus transtornos”, pois acreditam que podem satisfazer seus desejos sem causar dano real a uma criança. Em contraste, os consumidores de conteúdo “real” manifestam oposição a essa visão.



O perfil das vítimas é diverso e varia de acordo com o propósito do site (como fórum de discussão ou gerador de imagens). No entanto, é recorrente a criação de “personas” de vítimas inexistentes de abuso sexual infantil por usuários nesses ambientes da *dark web*. Nesses casos, são construídas personalidades, histórias e traços específicos para essa criança ou adolescente fictício, e em seguida é produzida uma vasta quantidade de conteúdo diferente envolvendo essa pessoa irreal. É também comum que outros usuários acompanhem, sigam e interajam com essas criações.