

Relatório sobre direitos humanos



Sumário

Sobre este relatório	03
Resumo executivo	06
Gerenciamento de riscos relacionados a direitos humanos	09
1. Liberdade de opinião e expressão	10
2. Privacidade	13
3. Igualdade e não discriminação	14
4. Vida, liberdade e segurança pessoal	15
5. Melhor interesse da criança	16
6. Participação pública, direito de votar e de ser eleito	16
7. Liberdade de associação e reunião	17
8. Direito à saúde	17
Destaques dos problemas	18
Inovação com respeito aos direitos humanos	18
Promoção da acessibilidade por meio de tecnologia vestível	19
Continuação da inovação responsável em IA	19
Fraudes e golpes	22
Combate às organizações criminosas especializadas em golpes no Sudeste Asiático	22
Como nos preparamos e respondemos a crises	24
República Democrática do Congo	25



Envolvimento das partes interessadas	26
Parceiros confiáveis	27
Estudo de caso: Insights para a prevenção ao recrutamento de jovens na América Latina	28
Organizações internacionais	29
Transparência e reparação	30
Anexo	33
Como os direitos humanos são regidos e gerenciados na Meta	34
Treinamento dos funcionários da Meta sobre direitos humanos	34
Links dos relatórios de referência	34





Sobre este relatório

Este Relatório Anual sobre Direitos Humanos abrange insights e ações de 1º de janeiro a 31 de dezembro de 2025. Nós divulgamos informações sobre os serviços e os produtos da Meta, incluindo Facebook, Messenger, Instagram, WhatsApp, Threads e Reality Labs.

Ele se baseia no trabalho da Meta em prol dos direitos humanos e reflete o progresso alcançado em nossos compromissos com os [Princípios Orientadores sobre Empresas e Direitos Humanos das Nações Unidas](#) e nossa [Política Corporativa sobre Direitos Humanos](#). O relatório mostra como aplicamos esses princípios em toda a empresa em 2025 e indica onde encontrar informações mais detalhadas.



O conteúdo deste relatório está fundamentado em nossa [Avaliação Abrangente de Riscos Relevantes a Direitos Humanos](#), realizada em 2022. O objetivo da avaliação foi identificar e priorizar nossos potenciais impactos adversos mais significativos sobre os direitos humanos¹ das pessoas, incluindo usuários dos nossos produtos e terceiros que possam ser afetados por nossas ações. Este relatório apresenta esses potenciais riscos relevantes e exemplos das nossas ações e medidas de mitigação em 2025.

Os direitos humanos continuaram sendo um tópico de extrema importância para nossa empresa e para as partes interessadas em 2025. Reduzimos e simplificamos o relatório, mas ainda buscamos oferecer um panorama representativo do nosso trabalho desenvolvido por diferentes equipes e do envolvimento das partes interessadas no mundo todo.

Políticas e progresso

Além deste Relatório de Direitos Humanos, a Meta divulga anualmente informações sobre políticas e avanços usando os seguintes mecanismos:

[Relatório Anual](#)[Declaração de Procuração](#)[Responsible Business Practices Report](#)[Central de Transparência](#)[Relatório de Sustentabilidade](#)[Relatório sobre Mudanças Climáticas do CDP](#)[Pacto Global das Nações Unidas](#)

O presente relatório complementa o [Meta Responsible Business Practices Report](#). [Relatamos](#) separadamente nossos esforços para identificar e mitigar os riscos da escravidão moderna e do tráfico de pessoas em nossas operações comerciais e cadeias de suprimentos. Além disso, cumprimos as obrigações nacionais e da União Europeia de prestação de informações, disponíveis em nossa [Central de Transparência](#). Encontre os links para outras divulgações da Meta no [Anexo](#) deste relatório.

[Ir para o Anexo](#)

¹ O termo "impacto adverso sobre os direitos humanos" está em conformidade com os Princípios Orientadores sobre Empresas e Direitos Humanos das Nações Unidas e significa um impacto que ocorre quando uma ação remove ou reduz a capacidade de um indivíduo de usufruir de seus direitos humanos.



Nossa [Política Corporativa sobre Direitos Humanos](#) se aplica em toda a empresa. Cada serviço e entidade da Meta possui suas próprias políticas e procedimentos, que podem ter impactos diferentes sobre os direitos humanos. Este relatório menciona as ações realizadas pela Meta como empresa em relação a uma ou mais entidades da empresa. As declarações não têm a intenção de sugerir que a Meta tomou a mesma ação relativamente a todas as entidades e/ou em todas as circunstâncias.²



² A discussão deste relatório sobre moderação de conteúdo e ações relacionadas no Facebook e no Instagram não se aplica ao WhatsApp e, a menos que uma política ou uma ação seja especificada como aplicável ao WhatsApp, ela não se aplica ao WhatsApp. Além disso, embora muitas ações descritas neste relatório se apliquem ao Facebook e ao Instagram, há distinções intencionais nas políticas e nos procedimentos entre os serviços. Se uma política for identificada como pertencente ao “Facebook”, ela não poderá ser aplicada ao Instagram. Nenhuma declaração neste relatório tem a intenção de criar, ou deve ser interpretada como se criasse, novas obrigações (legais ou de outra natureza) em relação à aplicação de uma política ou procedimento a outros serviços ou entidades.



Sobre

Resumo executivo

Gerenciamento
de riscos

Problemas

Partes
interessadas

Transparência e reparação

Anexo

Resumo executivo





Este é o quinto Relatório Anual sobre Direitos Humanos da Meta. Ele apresenta insights sobre o trabalho realizado pela Meta em 2025 para gerenciar riscos relacionados a direitos humanos em escala e cumprir nossos compromissos com os [Princípios Orientadores sobre Empresas e Direitos Humanos das Nações Unidas](#).



Avaliação de Riscos Relevantes

Nossas prioridades em 2025 refletiram os riscos relevantes identificados na [Avaliação Abrangente de Riscos Relevantes a Direitos Humanos](#) de 2022: liberdade de opinião e expressão; privacidade; igualdade e não discriminação; vida, liberdade e segurança pessoal; melhor interesse da criança; participação pública, direito de votar e de ser eleito; liberdade de associação e reunião; e direito à saúde.

→ [Leia mais](#)

Avanço na inovação

A Meta buscou ampliar os limites da inovação em inteligência artificial (IA), incorporando as considerações sobre direitos humanos em nosso processo de desenvolvimento. Implementamos novos modelos, estruturas de segurança e produtos, como [os óculos Meta Ray-Ban Display junto com a Meta Neural Band](#). Essa pulseira utiliza sinais musculares para permitir o controle dos óculos com IA sem o uso das mãos. Desenvolvemos esses produtos por meio do nosso processo de produtos responsáveis com foco na [acessibilidade](#) e de acordo com o direito à igualdade e à não discriminação.

→ [Leia mais](#)

Combate a fraudes e golpes

Também ampliamos nossos esforços para combater fraudes e golpes. Baseamos [essas iniciativas](#) na aplicação de regras das plataformas e nas parcerias inter-regionais com organizações do setor, de autoridades policiais e da sociedade civil local. Muitas centrais de golpes recorrem ao tráfico de pessoas e ao trabalho forçado para praticar suas ações criminosas. Assim, nossos esforços de mitigação ajudaram a respeitar o direito à vida, à liberdade e à segurança, bem como a proibição da escravidão, da servidão e do trabalho forçado.

→ [Leia mais](#)

Segurança de crianças e jovens

Da mesma forma, mantivemos nosso compromisso com a [segurança de crianças e jovens](#). Entre outras iniciativas, anunciamos que o Instagram estava [reformulando](#) a experiência de conteúdo nas Contas de Adolescentes, inspirado pelos critérios de classificação indicativa de filmes para maiores de 13 anos e pelo feedback de pais, mães e responsáveis.* Além disso, desenvolvemos essa experiência atualizada de acordo com a orientação de especialistas e o princípio das capacidades em evolução da criança estabelecido na Convenção da ONU sobre os Direitos da Criança.

→ [Leia mais](#)

Resposta a crises

Nosso Relatório sobre Direitos Humanos de 2024 forneceu insights detalhados sobre como a Meta integra os princípios de direitos humanos à maneira [como nos preparamos e respondemos a crises](#), guiada pelo nosso [Protocolo de Política de Crise](#). Esse trabalho continuou em 2025. Aplicamos o Protocolo de Política de Crise para responder a diversas situações, desde aquelas na Índia (Caxemira) e na Indonésia até os “protestos da Geração Z” liderados por jovens em Madagascar, na Sérvia e em outros países. Neste relatório, apresentamos um estudo de caso sobre nossa resposta à crise na República Democrática do Congo.

→ [Leia mais](#)



Envolvimento das partes interessadas

Nossa [Política Corporativa sobre Direitos Humanos](#) serve de base para nossa interação contínua com as partes interessadas. Em 2025, mantivemos contato com diversos grupos de direitos humanos, comunidades vulneráveis, integrantes da sociedade civil, acadêmicos, think tanks e órgãos reguladores. Os principais tópicos incluíram a abordagem da Meta à IA e à moderação de conteúdo, o impacto das nossas [alterações em produtos e políticas anunciadas em janeiro](#) e nossos esforços para combater fraudes e golpes.

Continuamos trabalhando com nossos [parceiros confiáveis](#) globalmente para identificar tendências e entender melhor os impactos do conteúdo e do comportamento online nas comunidades locais. Eles forneceram insights e identificaram conteúdo potencialmente violador em mais de 20 países e regiões.

→ [Leia mais](#)

Comitê de Supervisão

Também fornecemos uma atualização sobre o trabalho do Comitê de Supervisão. Esse comitê é um órgão independente que emite decisões vinculantes e recomendações não vinculantes. Durante o período de geração do relatório, o Comitê de Supervisão considerou casos relacionados às nossas iniciativas para respeitar os direitos humanos, incluindo a liberdade de expressão, os direitos e os interesses dos jovens e o direito à igualdade e à não discriminação, entre outros tópicos.

Como um exemplo ilustrativo, em resposta a uma recomendação, a Meta [aumentou](#) a granularidade de notificações existentes, fornecendo um aviso mais detalhado aos usuários quando o conteúdo deles é restringido devido a uma denúncia formal do governo afirmando que ele viola leis locais.

→ [Leia mais](#)

Gerenciamento de solicitações governamentais

Durante o ano de 2025, mantivemos como referência nosso compromisso com a [Global Network Initiative](#) de respeitar a liberdade de expressão e a privacidade, inclusive ao responder a solicitações governamentais para restringir conteúdo. E essas solicitações governamentais estão aumentando. Além dos nossos relatórios de integridade habituais, publicamos [estudos de caso](#) relacionados a solicitações governamentais para restringir o discurso político na Áustria, na Colômbia, na Costa do Marfim, na Alemanha, na Índia, no Iraque, na Itália, no Cazaquistão, na Malásia, na Nova Zelândia, na Polônia, na Romênia, em Singapura, na África do Sul, na Tanzânia, na Turquia e no Vietnã. Em 2025, a Meta também passou por sua quarta avaliação da GNI (abrangendo o período de abril de 2022 a dezembro de 2024). O Conselho da GNI determinou, por unanimidade, que a Meta havia realizado “iniciativas de boa-fé para implementar os Princípios da GNI, com melhorias ao longo do tempo”.

[Ver estudos de caso](#)



Gerenciamento de riscos relacionados a direitos humanos

Os [Princípios Orientadores sobre Empresas e Direitos Humanos das Nações Unidas](#) deixam claro que as empresas devem identificar seus impactos adversos sobre os direitos humanos para preveni-los ou mitigá-los de forma eficaz.

Considerando a escala das operações da Meta e a variedade de direitos que podem ser afetados, prever e gerenciar nossos [riscos relevantes](#) é importante, mas complexo. Gerenciamos dois tipos de riscos inerentes: aqueles decorrentes das nossas próprias atividades e aqueles decorrentes das atividades de terceiros, inclusive de usuários das nossas plataformas.





A tabela abaixo apresenta nossos riscos relevantes inerentes relacionados a direitos humanos³, conforme definidos na Avaliação Abrangente de Riscos Relevantes a Direitos Humanos de 2022 (CSRA, na sigla em inglês), divulgada em nosso [Relatório sobre Direitos Humanos de 2022](#). Essa tabela inclui exemplos ilustrativos de como lidamos com os potenciais riscos em 2025. Mais adiante, detalhamos alguns desses exemplos e mostramos como gerenciamos riscos potenciais relacionados à inteligência artificial (IA), a eleições e a conflitos.

1. Liberdade de opinião e expressão

O [direito à liberdade de opinião e expressão](#) inclui o direito de buscar, receber e compartilhar informações e ideias de todos os tipos. Trata-se de um direito fundamental, essencial para a proteção da dignidade humana, da autonomia individual e da democracia. A liberdade de expressão é um dos pilares da nossa missão e reflete nosso compromisso de dar voz a todas as pessoas.

Exemplos de potenciais riscos inerentes relevantes em relação a direitos humanos identificados na CSRA

As políticas de moderação de conteúdo da Meta e suas respectivas aplicações podem limitar a liberdade de expressão.

Exemplos de como a Meta lidou com os potenciais riscos em 2025

Em janeiro de 2025, atualizamos nossa [Política sobre Conduta de Ódio](#) para permitir [mais discussão](#) sobre alguns tópicos que fazem parte do debate público, incluindo imigração e gênero.

Concentramos nossos sistemas automatizados no combate a violações ilegais de alta gravidade, como terrorismo, exploração sexual infantil, drogas, fraudes e golpes, e recorremos a denúncias de usuários para tomar medidas em relação a violações de política de menor gravidade. Além disso, ajustamos nossos sistemas para exigir, de forma geral, um grau muito maior de confiança antes que um conteúdo seja removido. Posteriormente em 2025, divulgamos informações sobre nosso progresso na precisão da aplicação de regras no Relatório de Aplicação de Padrões da Comunidade, como parte de nossos [Relatórios de Integridade](#).

Removemos alterações feitas em 2021 para reduzir a quantidade de [conteúdo cívico](#) que as pessoas veem e começamos a recomendar mais conteúdo político com base em sinais personalizados, como curtir um conteúdo.

³ Para saber a diferença entre risco "inerente" e "residual", veja o [Relatório sobre Direitos Humanos da Meta de 2024](#), página 11.



Exemplos de potenciais riscos inerentes relevantes em relação a direitos humanos identificados na CSRA

As políticas de moderação de conteúdo da Meta e suas respectivas aplicações podem limitar a liberdade de expressão. (cont.)

Limitações governamentais excessivamente amplas sobre conteúdo.

Exemplos de como a Meta lidou com os potenciais riscos em 2025

Encerramos nosso programa de verificação de fatos independente nos EUA e, em vez disso, começamos a desenvolver um programa de [Notas da Comunidade](#). No resto do mundo, continuamos a contar com [verificadores de fatos](#) independentes da Meta e certificados pela entidade apartidária [Rede Internacional de Verificação de Fatos](#) ou pela [Rede Europeia de Normas de Verificação de Fatos](#), na Europa, para combater a desinformação no Facebook, no Instagram e no Threads.

→ [Leia mais](#)

Buscamos cumprir nossos compromissos com a [Global Network Initiative](#). Em 2025, a Meta passou por sua quarta avaliação da GNI (abrangendo o período de abril de 2022 a dezembro de 2024). O Conselho da GNI determinou, por unanimidade, que a Meta havia realizado “iniciativas de boa-fé para implementar os Princípios da GNI, com melhorias ao longo do tempo”. Também divulgamos informações sobre nossas [respostas](#) a solicitações governamentais de dados ou de restrições de conteúdo ([aqui](#) e [aqui](#)). Em 2025, [estudos de caso](#) notórios sobre transparência envolvendo discurso político incluíram os casos na Áustria, na Colômbia, na Costa do Marfim, na Alemanha, na Índia, no Iraque, na Itália, no Cazaquistão, na Malásia, na Nova Zelândia, na Polônia, na Romênia, em Singapura, na África do Sul, na Tanzânia, na Turquia e no Vietnã.



Lançamento das Notas da Comunidade

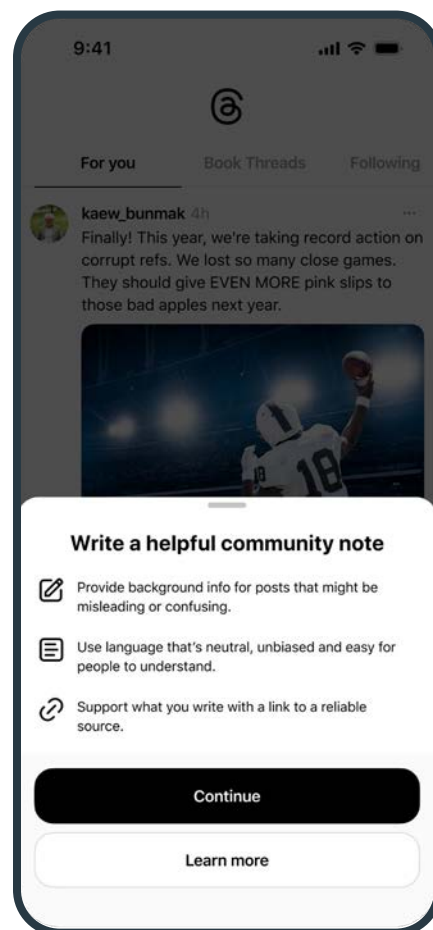
Em 2025, a Meta lançou as [Notas da Comunidade](#) nos EUA. Esse recurso é uma nova ferramenta para a comunidade adicionar contexto a um conteúdo que pode ser enganoso ou confuso. Nossa abordagem foi baseada em princípios de direitos humanos, incluindo a liberdade de expressão, e dedicada a melhorar o acesso a informações nas quais nossa comunidade de usuários confia amplamente, por meio de intervenções proporcionais e que protegem o discurso.

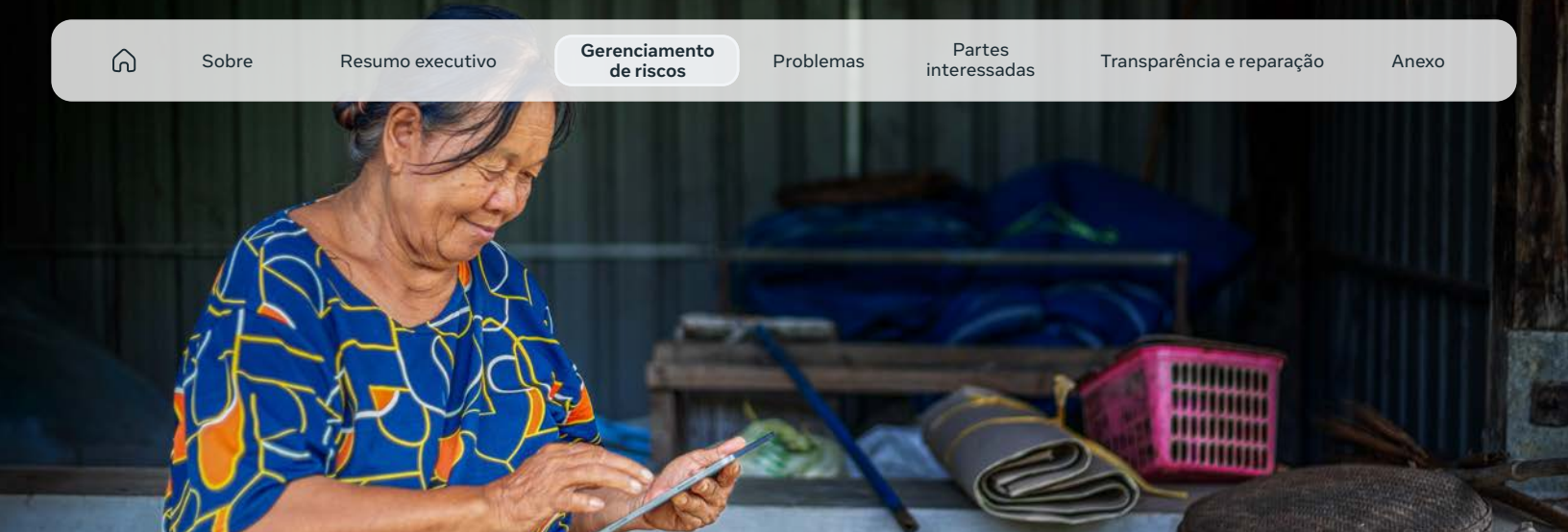
As notas são escritas e classificadas por usuários elegíveis — aqueles com 18 anos ou mais que têm uma conta verificada quanto à segurança e em situação regular. Uma nota é exibida apenas quando alcança um nível de “consenso útil”, ou seja, quando um número suficiente de colaboradores que historicamente discordavam passa a concordar que a nota é útil. Os nomes dos autores são ocultados para reduzir a pressão dos colegas e os riscos de retaliação e para incentivar a participação, inclusive em tópicos sensíveis.

As notas da comunidade podem aumentar a liberdade de expressão adicionando contexto, em vez de remover ou rebaixar a classificação de discurso legítimo, e ampliando a participação além de um pequeno grupo de profissionais de verificação de fatos. Esse recurso também foi criado com proteções para ajudar a prevenir abusos. Isso inclui requisitos de integridade da conta, proteções do URL e contra spam e aplicação dos Padrões da Comunidade nas notas.

A Meta solicitou a orientação do Comitê de Supervisão sobre os fatores nacionais que devemos considerar para decidir se um país deve ser omitido da expansão das Notas da Comunidade, incluindo baixos níveis de liberdade de expressão, falta de liberdade de imprensa e níveis de alfabetização digital.

Em novembro de 2025, o Comitê de Supervisão aceitou a [solicitação feita pela Meta de um parecer consultivo sobre política](#) para ajudar a definir nossa abordagem na expansão do programa de Notas da Comunidade fora dos EUA. O Comitê de Supervisão publicou seu [parecer](#) em março de 2026.





2. Privacidade

O [direito à privacidade](#) é uma condição necessária para a concretização de outros direitos humanos, como a liberdade de expressão, a liberdade de reunião e de associação e a liberdade de crença e de religião. Um dos princípios centrais descritos em nossa [Política Corporativa sobre Direitos Humanos](#) é manter a segurança das pessoas e proteger a privacidade.

Exemplos de potenciais riscos inerentes relevantes em relação a direitos humanos identificados na CSRA

Conteúdo ou comportamentos nos apps da Meta podem afetar negativamente os direitos à privacidade e à proteção dos dados.

Exemplos de como a Meta lidou com os potenciais riscos em 2025

Em 2025, desenvolvemos o [Processamento Privado para a Meta AI no WhatsApp](#), que foi criado com base em um Ambiente de Execução Confiável. O Processamento Privado permite que os usuários orientem a Meta AI a processar suas solicitações em um ambiente seguro e privado na nuvem, onde ninguém, nem mesmo a Meta ou o WhatsApp, pode acessar suas mensagens pessoais. Agora essa tecnologia é utilizada em alguns recursos opcionais da Meta AI no WhatsApp, como resumir mensagens e fornecer ajuda para escrever. Quando o Processamento Privado terminar de responder a uma solicitação, as mensagens não serão armazenadas. Leia mais sobre esse trabalho em nossa [Atualização sobre o progresso na privacidade](#).

Entramos com uma [ação judicial](#) contra a entidade por trás da CrushAI ou de apps de “nudificação”, que permitem que as pessoas criem imagens sexualmente explícitas ou de nudez de indivíduos geradas por IA sem o consentimento deles. Além disso, [anunciamos](#) medidas adicionais que tomamos para combater esses apps. Esse trabalho complementou nossos esforços antigos para evitar que imagens íntimas, sejam elas reais ou geradas por IA, sejam divulgadas na internet, por meio de ferramentas como o [StopNCII.org](#) e o [Take It Down do Centro Nacional para Crianças Desaparecidas e Exploradas](#).



3. Igualdade e não discriminação

O [direito à igualdade e não discriminação](#) garante proteção equitativa contra a discriminação. Com o propósito de respeitar esse direito, não permitimos condutas de ódio em nossas plataformas, conforme definido em nossa [política](#).

Exemplos de potenciais riscos inerentes relevantes em relação a direitos humanos identificados na CSRA

Alguns idiomas e dialetos podem ser mais difíceis de moderar do que outros.

Exemplos de como a Meta lidou com os potenciais riscos em 2025

[Implementamos](#), com código aberto, o [Meta Omnilingual Automatic Speech Recognition \(ASR\)](#), um conjunto de modelos que fornece funcionalidades de reconhecimento de fala automático para mais de 1.600 idiomas, incluindo 500 idiomas com poucos recursos. Fazendo com que um conteúdo falado possa ser transcrito em idiomas anteriormente com pouco suporte, essa tecnologia abre novas oportunidades de educação e aumenta ainda mais a liberdade de expressão, capacitando as pessoas a acessar o conhecimento e compartilhar conteúdo no idioma que elas conhecem melhor.

Conteúdo que afeta negativamente a igualdade e a não discriminação (por exemplo, conduta de ódio).

Antes das [alterações em produtos e políticas](#) anunciadas em janeiro de 2025, avaliamos os riscos das mudanças e levamos em consideração mitigações relevantes, como a disponibilidade de outras políticas e denúncias de usuários, para lidar com esse tipo de conteúdo. Em países que estejam passando por conflitos ou crises, podemos aplicar a detecção proativa por padrão em algumas áreas de políticas, incluindo violações por Conduta de Ódio. Por exemplo, usamos a detecção proativa na Líbia após a implementação do Protocolo de Política de Crise.



4. Vida, liberdade e segurança pessoal

O [direito à vida, liberdade e segurança pessoal](#) diz respeito à proteção contra danos físicos e privação de liberdade. Para a Meta, o respeito a esse direito humano inclui a mitigação do risco de que conteúdos possam provocar danos, incluindo exposição a violência e tráfico de pessoas, ameaças online patrocinadas por Estados e atuação de grupos não estatais envolvidos em violência ou ódio, ou que os promovam.

Exemplos de potenciais riscos inerentes relevantes em relação a direitos humanos identificados na CSRA

Agentes maliciosos que:

- Aproveitam os serviços e apps da Meta para provocar danos online e no meio físico
- Usam indevidamente serviços e apps para realizar ataques cibernéticos ou phishing
- Ameaçam e assediam defensores de direitos humanos, ativistas e outros grupos vulneráveis

Exemplos de como a Meta lidou com os potenciais riscos em 2025

Por meio de pesquisas aprofundadas e consultas com parceiros, a Meta identificou um potencial risco elevado de recrutamento de crianças-soldados no Sudão e desenvolveu mensagens sobre a segurança de jovens em árabe, francês e inglês adaptadas para a região. Em novembro de 2025, a Meta lançou duas campanhas complementares: uma rápida promoção veiculada a aproximadamente 1 milhão de jovens sudaneses (entre 13 e 18 anos de idade) no Facebook e no Instagram e uma mensagem personalizada em um pop-up interativo enviado a aproximadamente 40 mil jovens conectados a [Interrupções Estratégicas de Rede](#). As mensagens avisaram sobre os riscos de recrutamento e direcionaram os jovens para recursos externos de segurança.

[→ Leia mais](#)

Continuamos apoiando o Fundo de Defesa de Direitos Humanos e nosso [Programa de Parceiros Confiáveis](#) para aprimorar a resposta a emergências voltada a defensores de direitos humanos e a outras pessoas vulneráveis.

5. Melhor interesse da criança

A Convenção sobre os Direitos da Criança da ONU (UNCRC) estabelece que, em todas as ações relativas a crianças, “o melhor interesse da criança deve ser uma consideração primordial”. A [Estrutura do Melhor Interesse da Criança](#) da Meta se alinha aos valores fundamentais da UNCRC. Manter a proteção das crianças online é uma prioridade para a Meta. Oferecemos ferramentas para adolescentes, mães, pais e responsáveis, com proteções integradas, que conciliam a segurança com o exercício do direito à liberdade de expressão e ao acesso à informação.

Exemplos de potenciais riscos inerentes relevantes em relação a direitos humanos identificados na CSRA

Crianças podem ser expostas a conteúdos indesejados ou inadequados, bem como a comportamentos predatórios.

Exemplos de como a Meta lidou com os potenciais riscos em 2025

Com base nas [Contas de Adolescentes](#), adicionamos [novas proteções](#) e apresentamos [configurações de conteúdo atualizadas](#), inspirados nos critérios de classificação indicativa de filmes para maiores de 13 anos e no feedback de pais, mães e responsáveis. Isso significa que o objetivo do Instagram é que os adolescentes vejam conteúdo semelhante ao que veriam em filmes adequados à idade por padrão. Além disso, implementamos uma nova configuração mais rigorosa para pais, mães e responsáveis que preferem uma experiência mais restritiva para o adolescente. Também anunciamos novos controles para ajudar pais, mães e responsáveis a gerenciar como os [adolescentes interagem com a IA](#) em nossos apps, [tomamos medidas legais](#) contra empresas por trás de apps de “nudificação” e [fizemos parcerias com instituições de ensino para combater o bullying virtual](#).

6. Participação pública, direito de votar e de ser eleito

O [direito à participação pública, direito de votar e de ser eleito](#) em eleições livres e justas é um pilar da democracia. Proteger a integridade das eleições em nossos serviços e apps é uma das nossas maiores prioridades. Trabalhamos de forma contínua para proteger as eleições no ambiente online antes, durante e após os períodos eleitorais.

Exemplos de potenciais riscos inerentes relevantes em relação a direitos humanos identificados na CSRA

Conteúdos em violação podem afetar negativamente a participação pública, o exercício do voto ou a candidatura a cargos eletivos. Isso pode decorrer de atividades que incluem, entre outras:

- Ação coordenada de agentes maliciosos visando interferir em eleições
- Ameaças de violência e danos no meio físico contra candidatos
- Esforços individuais para impedir pessoas de votar, aumento de spam, interferência estrangeira ou relatos de conteúdos que violam nossas políticas

Exemplos de como a Meta lidou com os potenciais riscos em 2025

Mantivemos as iniciativas relacionadas a eleições, desenvolvendo nossa abordagem a eleições em escala e baseada em riscos, detalhada em nosso [Relatório sobre Direitos Humanos de 2024](#).

Em 2025, nossa preparação para eleições incluiu as ações na Austrália, nos Camarões, no Chile, na Costa do Marfim, no Equador, no Egito, na Alemanha, no Iraque, no Japão, em Mianmar, nas Filipinas, na Polônia, na Romênia e na Tanzânia, entre outros países.

7. Liberdade de associação e reunião

O [direito à liberdade de associação e reunião](#) é essencial para a democracia e está intimamente ligado a outros direitos garantidos pelas normas internacionais de direitos humanos, incluindo o direito à liberdade de expressão e à participação na condução dos assuntos públicos. Para a Meta, esse direito está relacionado aos nossos valores centrais de dar voz às pessoas e de promover conexões e comunidades.

Exemplos de potenciais riscos inerentes relevantes em relação a direitos humanos identificados na CSRA	Exemplos de como a Meta lidou com os potenciais riscos em 2025
Certos tipos de conteúdo nas plataformas da Meta podem desencorajar algumas pessoas de se conectar livremente, tanto nos apps da Meta quanto no mundo físico.	Implementamos nosso Protocolo de Política de Crise para ajudar a direcionar nossos esforços no enfrentamento de conteúdos violadores relacionados a manifestações em massa, incluindo os “protestos da Geração Z” no Irã, em Madagascar, no Nepal e na Sérvia, por exemplo.

8. Direito à saúde

O [direito à saúde](#) é o direito de todas as pessoas ao mais alto nível possível de saúde física e mental. A Meta respeita esse direito por meio da ampliação do acesso a informações sobre saúde, do incentivo à conexão entre indivíduos com problemas de saúde semelhantes e da promoção de decisões fundamentadas sobre saúde e bem-estar.

Exemplos de potenciais riscos inerentes relevantes em relação a direitos humanos identificados na CSRA	Exemplos de como a Meta lidou com os potenciais riscos em 2025
Conteúdo em violação às políticas que incita ou tem como objetivo causar danos no meio físico.	Anunciamos proteções para adolescentes em nossos recursos de IA, incluindo a configuração para que eles respondam de forma segura a comandos sobre autoagressão, suicídio e distúrbios alimentares. Além disso, anunciamos novos controles, como o lançamento de insights para pais, mães e responsáveis sobre os tópicos das conversas dos adolescentes com a IA. Nos EUA, também buscamos aumentar nossos esforços para combater o bullying virtual em escolas oferecendo a professores, educadores e administradores um sistema priorizado para denúncias de problemas de segurança para adolescentes.

[Sobre](#)[Resumo executivo](#)[Gerenciamento de riscos](#)[Problemas](#)[Partes interessadas](#)[Transparência e reparação](#)[Anexo](#)

Destques dos problemas

Inovação com respeito aos direitos humanos

As inovações em tecnologia vestível e em inteligência artificial (IA) ofereceram oportunidades importantes de promover os direitos humanos e apoiar a mitigação responsável de riscos.

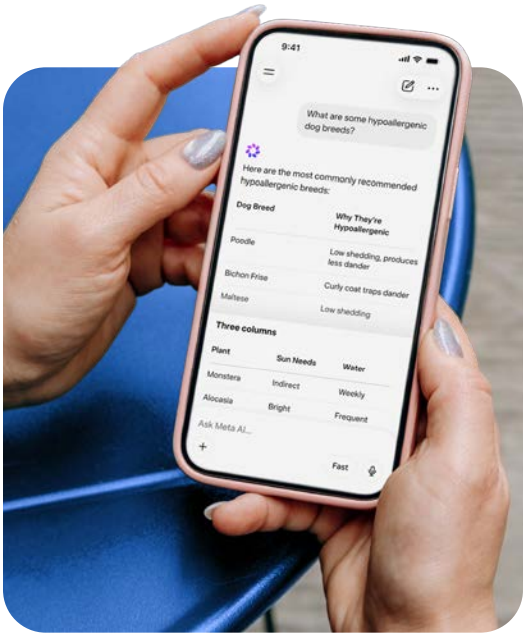
Promoção da acessibilidade por meio de tecnologia vestível

Em 2025, [lançamos o Meta Ray-Ban Display e a Meta Neural Band](#) — uma combinação de óculos com IA de última geração e uma pulseira que usa tecnologia de [eletromiografia de superfície \(sEMG\)](#) para converter sinais elétricos criados por movimentos sutis nos músculos em comandos, permitindo o controle sem usar as mãos.

Os protótipos da pulseira foram desenvolvidos com foco na [acessibilidade](#) e de acordo com o direito à igualdade e à não discriminação. Com base em pesquisas que envolveram quase 200 mil participantes, a pulseira permite [novas possibilidades](#) para pessoas com limitações de mobilidade devido a lesões na medula espinhal, derrames ou outros eventos limitadores, além de pessoas com tremores ou com menos de cinco dedos. Para proteger a privacidade, o machine learning que converte os sinais de EMG em comandos digitais é [processado localmente no dispositivo](#).



Junto com os óculos com IA, nossa tecnologia [capacita pessoas com deficiência](#) a ter mais independência no dia a dia. Recursos como a ligação sem usar as mãos, a tradução em tempo real e a legenda ao vivo eliminam barreiras de comunicação. Já nosso recurso de [“ligar para um voluntário”](#), criado em parceria com o [Be My Eyes](#), permite que pessoas cegas e com baixa visão se conectem com voluntários para receber descrições de seus arredores em tempo real. Criando essas ferramentas, dialogamos diretamente com partes interessadas em questões de pessoas com deficiência, incluindo a [Associação de Veteranos Cegos](#) e os [Centros de Reabilitação de Cegos](#) do Departamento de Assuntos de Veteranos dos EUA, para desenvolver guias de treinamento e apoiar casos de uso de acessibilidade.



Continuação da inovação responsável em IA

Lançamos o [assistente independente da Meta AI](#) junto com um grande modelo de linguagem de código aberto e atualizado, o [Llama 4](#). Também lançamos o [Processamento Privado](#), uma nova [tecnologia que utiliza um ambiente de execução confiável](#), para habilitar o assistente da Meta AI no WhatsApp e, ao mesmo tempo, preservar as proteções de privacidade, incluindo as comunicações criptografadas. Para saber detalhes, veja nossa devida diligência anterior que discute os [benefícios da criptografia de ponta a ponta para os direitos humanos](#).

Para mitigar os riscos de uso indevido e resultados potencialmente prejudiciais em nossos produtos e modelos de IA, o desenvolvimento do Llama 4 incluiu filtragem de dados pré-treinamento, mitigações pós-treinamento e mitigações no nível do sistema. Essas ações estão descritas em nosso [guia de uso sobre proteções da IA para desenvolvedores](#).

Também publicamos nossa [Estrutura de IA de Fronteira](#), focada nos riscos mais críticos nas áreas de ameaças à segurança digital e riscos gerados por armas químicas e biológicas. Nossa estrutura descreveu uma série de processos que seguimos para prever e mitigar riscos durante o desenvolvimento de sistemas de IA de fronteira, incluindo:

Identificar resultados catastróficos a serem evitados



Nossa estrutura identificou potenciais resultados catastróficos relacionados a riscos digitais, químicos e biológicos que nos empenhamos para evitar. Ela se concentrou em avaliar se esses resultados catastróficos eram possibilitados por avanços tecnológicos e, em caso positivo, identificar maneiras de mitigar esses riscos.

Realizar exercícios de modelagem de ameaças



Realizamos exercícios de modelagem de ameaças para prever como diferentes atores podem usar indevidamente a IA de fronteira para produzir esses resultados catastróficos, trabalhando com especialistas externos conforme necessário. Esses exercícios foram fundamentais para nossa abordagem orientada por resultados.

Estabelecer limites de riscos

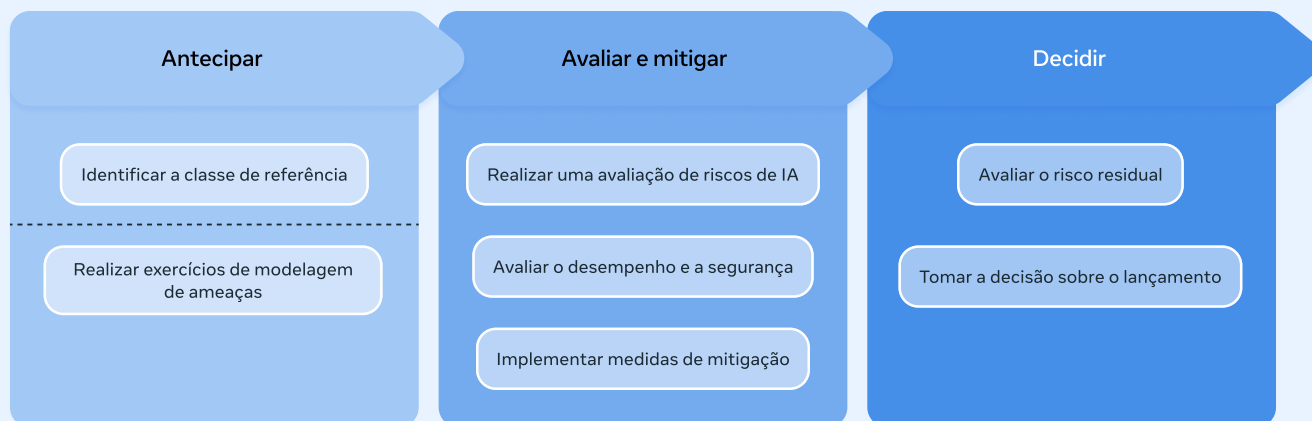


Definimos limites de riscos com base em até que ponto nossos modelos poderiam facilitar os cenários de ameaça. Temos processos em vigor para manter os riscos em níveis aceitáveis, incluindo a aplicação de mitigações.

Também continuamos a apoiar um ecossistema amplo para o desenvolvimento e a implementação responsáveis da IA. Para capacitar os desenvolvedores a criar apps seguros, ampliamos nosso investimento em [proteções no nível do sistema](#), incluindo o [Llama Guard 4](#) para proteções unificadas em modalidades de texto e imagem e o [Llama Prompt Guard 2](#) para detecção aprimorada de jailbreak e injeção de comando.

Nosso trabalho foi fundamentado em interações e consultas abrangentes com a sociedade civil, organizações internacionais, o meio acadêmico e o público sobre as dimensões da IA relacionadas aos direitos humanos. Isso incluiu [Fóruns da Comunidade](#), conduzidos em parceria com o Deliberative Democracy Lab da Universidade de Stanford, que reuniram perspectivas diversas sobre o futuro dos agentes de IA e sua governança.

Estrutura de IA de Fronteira de 2025





Ferramentas de bem-estar com código aberto para análise de conteúdo

Para que nossos analistas apliquem as políticas de forma precisa, eles também devem se concentrar em sua própria [resiliência e bem-estar](#).

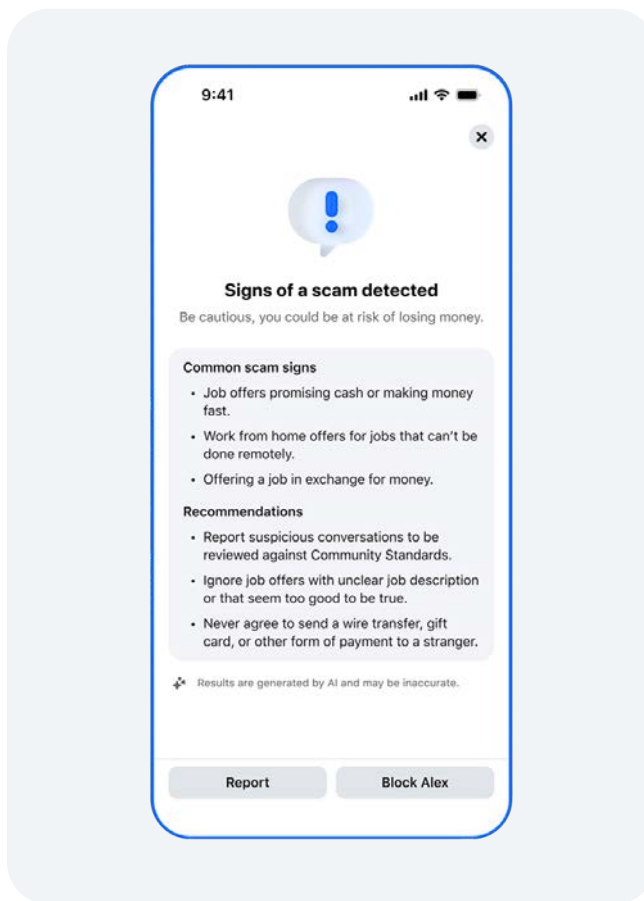
Em 2025, [abrimos o código-fonte de um conjunto de filtros de análise de conteúdo](#) criados para apoiar o bem-estar dos moderadores. Assim, outros desenvolvedores podem se beneficiar da nossa experiência nesse espaço. Esses filtros personalizáveis de imagem e vídeo, com recursos como desfoque, conversão para escala de cinza/sépia, redução de detalhes e telas de aviso preditivas habilitadas por IA, podem ajudar a reduzir os potenciais impactos adversos à saúde mental causados pela exposição a conteúdo explícito ou perturbador.

Fraudes e golpes

As redes de fraudes e golpes cresceram consideravelmente na última década e causaram um impacto negativo nos direitos humanos. Muitas centrais de golpes recorrem ao tráfico de pessoas e ao trabalho forçado para praticar suas ações criminosas, e as vítimas de golpes podem sofrer perdas financeiras significativas e outros danos. Em 2025, [removemos](#) mais de 159 milhões de anúncios fraudulentos por violarem nossas políticas e também [removemos](#) 10,9 milhões de contas no Facebook e no Instagram associadas a centrais de golpes.

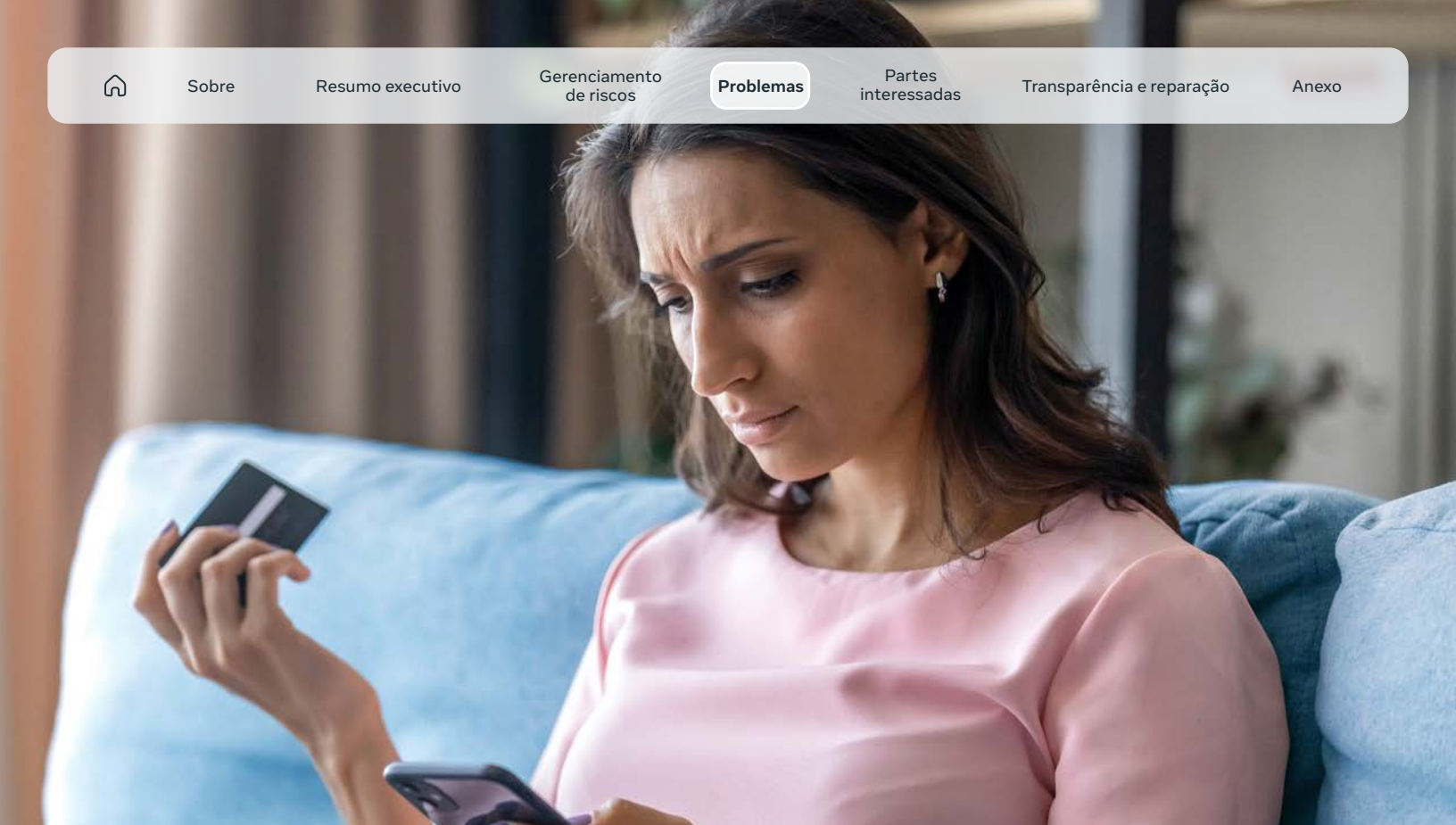
Priorizamos a segurança por meio de [defesas das plataformas](#), como [reconhecimento facial](#) para combater a falsificação de identidade, [avisos de detecção de golpe](#) no Messenger e [avisos de compartilhamento de tela](#) no WhatsApp. Nosso [trabalho](#) para impedir a ação de golpistas também envolveu discussão judicial e apoio a autoridades policiais para responsabilizá-los. Equilibrar a segurança e o bem-estar de nossas comunidades com mitigações que sejam proporcionais e respeitem os direitos pode ser desafiador. Realizamos essas ações com foco na privacidade e de acordo com as normas internacionais de direitos humanos, incluindo os [Princípios da Global Network Initiative](#).

Fraudes e golpes são desafios globais. O compartilhamento de sinais, a colaboração entre comunidades de proteção e uma abordagem que envolva toda a sociedade são necessários para um trabalho eficaz. Ao longo de 2025, a Meta realizou parcerias para o compartilhamento de sinais, como a [Global Signal Exchange](#), a [Global Anti-Scam Alliance](#) e a [Fraud Intelligence Reciprocal Exchange](#). Além disso, [defendemos](#) a publicação de listas de países que não tomam medidas adequadas e a aplicação de penalidades internacionais, levando em conta os direitos humanos e as consequências indesejadas para as populações vulneráveis, como pessoas traficadas forçadas a atuar em golpes.



Combate às organizações criminosas especializadas em golpes no Sudeste Asiático

Redes e organizações de golpes sofisticados e em escala industrial [despontaram](#) na última década. Na [região da Ásia Pacífico](#), essas organizações empregam centenas de milhares de pessoas e atuam praticamente impunes. As centrais de golpes atraem candidatos com falsos anúncios de vaga de emprego e os forçam a trabalhar praticando golpes online, muitas vezes sob a ameaça de agressão física. Essa prática criminosa viola a proibição da escravidão, da servidão e do trabalho forçado segundo as normas internacionais de direitos humanos.



[Removemos](#) mais de 6.400 contas e Páginas do Facebook entre janeiro e outubro de 2025 por [violarem](#) nossas políticas contra [fraudes, golpes e práticas enganosas](#) e [organizações e indivíduos perigosos](#). Uma rede interceptada proveniente de Camboja [se passava por](#) autoridades policiais globais para oferecer falsos serviços de recuperação de perdas. Esse trabalho exigiu uma colaboração internacional e um rigoroso monitoramento do impacto de mudanças geopolíticas, como o conflito em Mianmar e ao longo da fronteira no norte de Camboja. Discutimos essas iniciativas em mais detalhes em nosso [Adversarial Threat Report](#) do 1º semestre de 2026.



Nossas iniciativas de combate em 2025 resultaram em nossa participação em um programa piloto da [Joint Disruption Week](#) em Bangkok, no qual a Meta colaborou com o FBI, a Polícia Real Tailandesa e outras autoridades policiais. Esse piloto [resultou](#) na remoção de 59 mil contas, Páginas e grupos das plataformas da Meta e em seis mandados de prisão.



Além do combate aos golpes, atualizamos nossos [Padrões da Comunidade sobre Exploração Humana](#) para capturar mais conteúdo de recrutamento de possíveis organizações criminosas especializadas em golpes. A atualização restringe posts de recrutamento para vagas de emprego em locais com alto risco de exploração de mão de obra, segundo dados confirmados por autoridades policiais e organizações não governamentais locais.

A Meta também colaborou com o Escritório das Nações Unidas sobre Drogas e Crime na campanha [#TrappedInScamCrime](#) (“Presos por trás do golpe”) para ajudar a combater o recrutamento online e o tráfico de pessoas para atividades criminosas forçadas associadas às centrais de golpes no Sudeste Asiático. A campanha amplificou a mensagem de prevenção baseada no depoimento de sobreviventes para ajudar as pessoas a reconhecer falsas ofertas de emprego e a acessar recursos de suporte.

Também trabalhamos com a Associação de Nações do Sudeste Asiático a fim de desenvolver [princípios](#) para combater os golpes online. Esse trabalho se somou aos nossos [esforços mais amplos contra o tráfico](#).

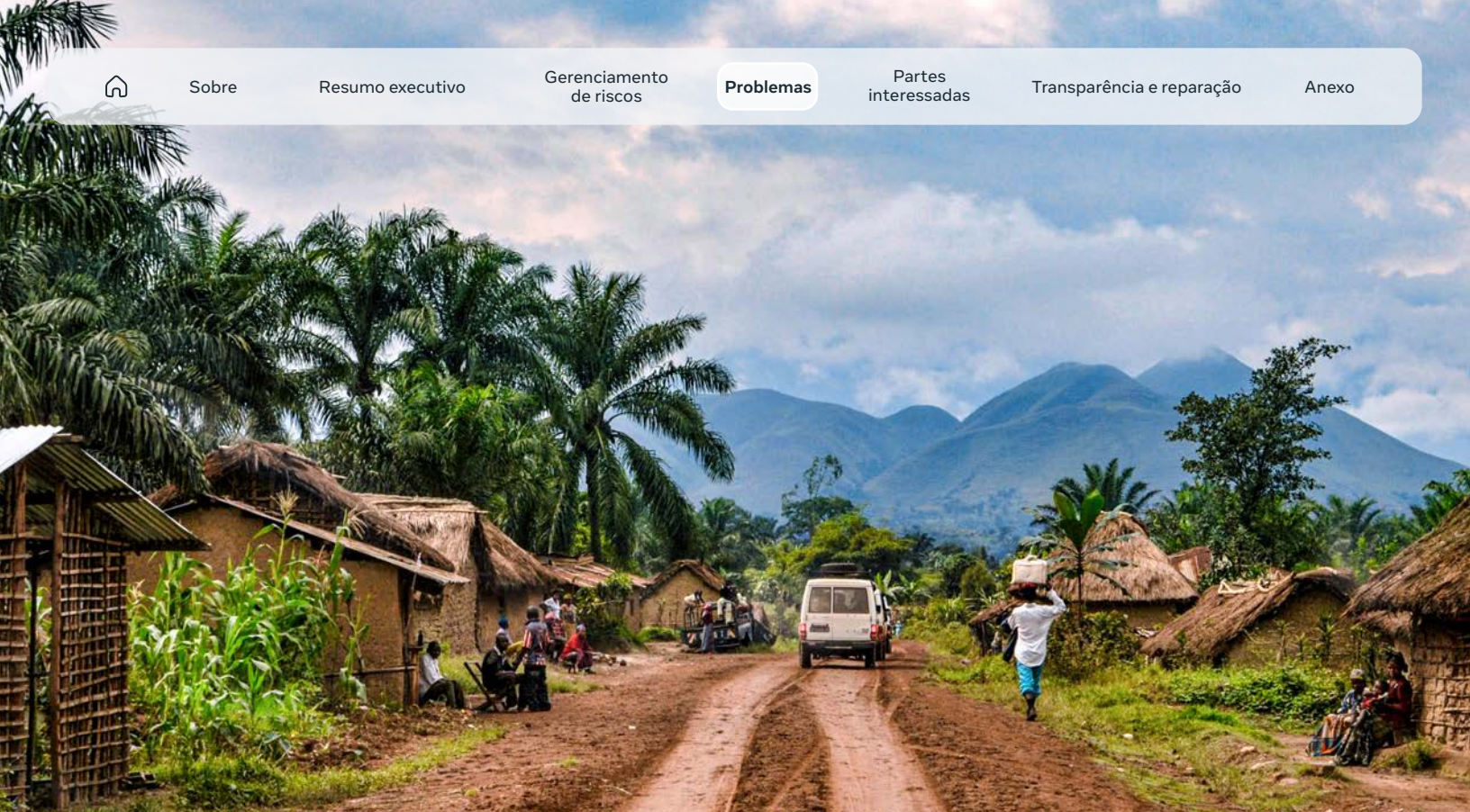
Como nos preparamos e respondemos a crises

Atuamos na preparação e na resposta a diversas crises em todo o mundo, como conflitos, violência intracomunitária, instabilidade civil, protestos em massa e desastres ambientais, além de ataques terroristas e tiroteios. Em 2025, iniciamos e coordenamos respostas a crises nos Camarões, na Índia (Caxemira), na Indonésia, na Líbia, em Madagascar, no Nepal, na Sérvia, na Tanzânia e entre a Tailândia e o Camboja, entre outros países e territórios. Demos continuidade aos nossos esforços em crises designadas no âmbito do [Protocolo de Política de Crise](#) para os conflitos na Ucrânia, no Sudão e no Oriente Médio. Abordamos essas crises designadas em nossos [Relatórios sobre Direitos Humanos anteriores](#).

Durante um período de crise, utilizamos o Protocolo de Política de Crise como uma importante ferramenta. O protocolo orienta o uso rápido de mecanismos para mitigar possíveis danos, conforme descrito em nosso [Relatório sobre Direitos Humanos de 2024](#). Os tipos específicos de resposta adotados são coerentes com os riscos observados. Eles se baseiam em intervenções anteriores em situações de crise, em princípios de direitos humanos e no direito dos conflitos armados.

Na próxima página, ilustramos como nos preparamos e respondemos a crises e áreas de alto risco e afetadas por conflitos, usando um exemplo de uma região não apresentada em nossos relatórios anteriores.





República Democrática do Congo

Em 2025, o conflito no leste da República Democrática do Congo (RDC) se intensificou quando o grupo armado M23, supostamente [apoiado](#) por Ruanda, tomou o território em Kivu do Norte, levando ao deslocamento em massa de populações civis e a [graves violações de direitos humanos](#). Parceiros confiáveis nos alertaram para a escalada da violência no leste do país e os riscos elevados de violência entre comunidades, levando a Meta a designar a crise nos termos do Protocolo de Política de Crise.

Uma equipe multifuncional de especialistas da Meta, incluindo as áreas de política, operações, segurança e política pública, vinha monitorando a situação no meio físico na RDC desde a volta do conflito no final de 2021, tomando medidas contra conteúdos violadores de acordo com nossas políticas sobre [desinformação e dano](#), [violência e incitação](#), [coordenação de atos danosos e incentivo à prática de atividades criminosas](#) e [organizações e indivíduos perigosos](#). Também havíamos implementado medidas de mitigação para lidar com potenciais riscos em conteúdo online, incluindo a designação de alegações potencialmente prejudiciais pré-analisadas nos termos da política sobre desinformação e dano e o uso da nossa [política sobre locais com alto risco temporário](#). Além disso, já havíamos designado grupos com risco de exposição relacionados a “ruandófonos, ruandeses, baniamulenges ou tutsis” para ajudar a prevenir o risco de dano no meio físico.

Com a escalada do conflito em 2025, avaliamos e mitigamos os riscos em conteúdos relacionados. Fornecemos esclarecimentos e orientações sobre políticas, analisamos palavras-chave para garantir uma detecção eficaz e uma aplicação precisa de nossas políticas e aplicamos nossa política sobre locais com alto risco temporário.

A Meta também consultou os parceiros confiáveis para ajudar a embasar a possível designação de novas alegações prejudiciais — declarações falsas ou enganosas que podem contribuir diretamente para um risco de danos físicos iminentes — pré-analisadas nos termos de nossa política sobre desinformação e dano. Os insights dos parceiros confiáveis nos permitiram detectar e remover com mais eficácia conteúdos violadores de políticas, como materiais com alta probabilidade de agravar divisões locais.

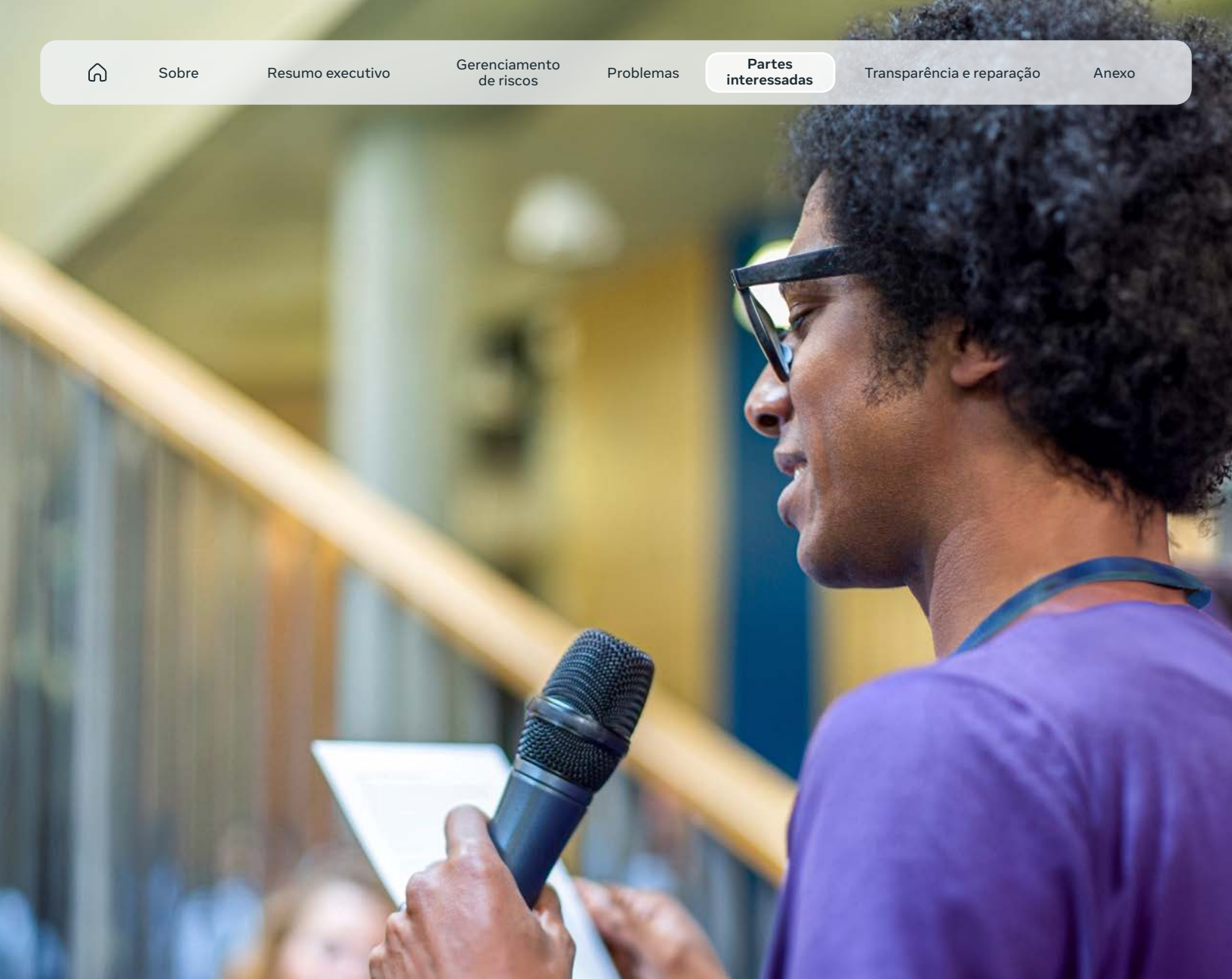
Envolvimento das partes interessadas

A [interação](#) proativa e estruturada com nossa comunidade global de usuários ajuda a moldar as políticas da Meta e é primordial para o gerenciamento de riscos em direitos humanos.

Em 2025, dialogamos com diferentes partes interessadas, incluindo membros da sociedade civil, acadêmicos, think tanks, especialistas em direitos humanos e órgãos reguladores. Questões importantes sobre políticas incluíram nossa abordagem à inteligência artificial (IA) e à moderação de conteúdo e o impacto das alterações em produtos e políticas apresentadas em janeiro de 2025, como as [Notas da Comunidade](#), além de fraudes e golpes.

Dialogamos com mais de 615 partes interessadas externas de 65 países para embasar o trabalho de desenvolvimento de políticas em diversas áreas, incluindo as políticas sobre organizações e indivíduos perigosos e sobre publicidade nas áreas de saúde e bem-estar. Consultamos outras centenas delas em fluxos de trabalho de previsão de riscos, análise de tendências, aplicação de regras e prevenção.

Além disso, dialogamos com a sociedade civil sobre diversas questões regulatórias e de liberdade de expressão, inclusive na Argentina e no Brasil, além da Malásia como parte da [Conferência da Assembleia de Direitos Digitais da Ásia Pacífico](#). Também realizamos treinamentos sobre nossas políticas para a sociedade civil. No Norte da África, por exemplo, treinamos grupos da sociedade civil sobre nossas políticas relacionadas a [desinformação e dano](#) e [conduta de ódio](#) para ajudá-los a identificar conteúdo potencialmente violador de acordo com essas políticas. O treinamento também ajudou a fornecer a defensores dos direitos humanos e a grupos vulneráveis ferramentas para proteger sua presença digital.



Parceiros confiáveis

Parceiros confiáveis são aliados importantes na identificação de violações de alta gravidade dos nossos Padrões da Comunidade. Continuamos trabalhando com nossos [parceiros confiáveis](#) para identificar tendências, entender melhor o impacto do conteúdo e do comportamento online nas comunidades locais e explorar formas de fortalecer nossos canais de escalonamento com a sociedade civil.

Em 2025, os parceiros confiáveis forneceram insights e identificaram conteúdo prejudicial em países e territórios que enfrentaram altos níveis de instabilidade ou realizaram eleições. Eles incluíram Afeganistão, Camboja, Costa do Marfim, República Democrática do Congo, Índia, Indonésia, Israel, Quênia, Líbano, Madagascar, México, Nigéria, Palestina, Somália, Sudão, Síria, Tailândia, Tanzânia, Ucrânia e Venezuela, entre outros países e regiões.

Removemos mais de 60 mil conteúdos violadores de políticas como resultado do nosso Programa de Parceiros Confiáveis. Um exemplo do impacto dos parceiros confiáveis na América Latina é apresentado na próxima página.

ESTUDO DE CASO

Insights para a prevenção ao recrutamento de jovens na América Latina



Na América Latina, os parceiros confiáveis desempenharam um papel fundamental na denúncia de conteúdo, incluindo glorificação, apoio e representação de [organizações e indivíduos perigosos](#). Eles também forneceram feedback sobre as iniciativas de prevenção ao recrutamento de jovens.

No Brasil, os parceiros confiáveis denunciaram conteúdo no Instagram que retratava membros de uma organização criminosa designada, incluindo representações de armas de fogo proibidas, atos de violência e drogas. Na Colômbia, os parceiros confiáveis denunciaram perfis que representavam grupos criminosos, no contexto de ameaças e ataques violentos desses grupos contra jornalistas no departamento de Guaviare. No Haiti, o conteúdo denunciado glorificava atos criminosos de indivíduos designados associados a gangues. Esses insights fortaleceram nossos esforços para remover conteúdo violador e mitigar o risco da atuação de organizações e indivíduos perigosos em nossas plataformas, bem como o de ataques físicos.

A Meta trabalhou com parceiros confiáveis e especialistas regionais para lançar campanhas de mensagens focadas nos jovens em toda a América Latina. Veiculadas para 80 mil usuários entre 13 e 25 anos de idade no México e mais 55 mil na Colômbia, no Haiti e no Brasil, essas campanhas foram criadas para informar os usuários sobre violações à plataforma, orientá-los sobre os riscos de se envolver com organizações e indivíduos perigosos e capacitá-los a fazer suas próprias escolhas mais conscientes na internet. A campanha levou a um aumento considerável na desativação de contas e na ação de desamigar usuários.



Organizações internacionais

A Meta continuou a colaborar com a ONU ao longo do ano. Participamos de discussões sobre políticas na 80ª Assembleia Geral da ONU (AGNU). Alguns tópicos foram a governança de IA e digital global, o uso da IA para a [segurança nacional](#) e maneiras de preencher as lacunas na capacidade da IA. Também participamos do lançamento do Diálogo Global da ONU sobre Governança da IA durante a Semana de Alto Nível da AGNU.

A Meta continuou a trabalhar em estreita colaboração com o [Alto Comissariado das Nações Unidas para os Direitos Humanos](#), inclusive no [Projeto B-Tech](#). Respondemos a solicitações de envios e consultamos a Comissão Africana dos Direitos Humanos e dos Povos e os Procedimentos Especiais de Direitos Humanos da ONU, formados por especialistas independentes em direitos humanos, incluindo os Relatores Especiais para a Liberdade de Opinião e de Expressão. Também dialogamos com outras agências da ONU e organizações humanitárias para ajudar a fundamentar nosso trabalho de resposta a crises.

Além disso, participamos do [Fórum de Governança da Internet](#) de 2025, do [Fórum da ONU sobre Empresas e Direitos Humanos](#) e do [RightsX Summit](#), em que falamos sobre casos de uso positivos da IA para defensores dos direitos humanos, entre outros tópicos. Continuamos a participar e a dialogar com as partes interessadas em vários outros fóruns multilaterais, incluindo o [Fórum Global da Internet para Combater o Terrorismo](#), o [RightsCon](#) e o fórum de [Tecnologia contra o Tráfico Humano](#).



Transparência e reparação



O [Comitê de Supervisão](#), como órgão independente, nos ajuda a responder a algumas das perguntas mais difíceis sobre o tema da liberdade de expressão online: o que remover, o que permitir e por quê. Ele analisa casos encaminhados pela Meta ou apelações apresentadas por indivíduos no Facebook, no Instagram ou no Threads que discordam das nossas decisões de moderação de conteúdo, proferindo decisões vinculantes sobre a remoção ou a manutenção de conteúdos ativos. O Comitê de Supervisão também apresenta recomendações para aprimorar nossas práticas de moderação de conteúdo e emite pareceres consultivos sobre política quando solicitado.



Em 2025, o Comitê de Supervisão analisou casos relacionados à aplicação das nossas políticas de conteúdo à luz da estrutura internacional de direitos humanos, incluindo a liberdade de expressão, os direitos das crianças e o direito à igualdade e à não discriminação, entre outros.

Exemplos de decisões do Comitê de Supervisão:

 O Comitê de Supervisão [revogou](#) as decisões da Meta de não rotular um clipe de áudio provavelmente adulterado de dois políticos curdos iraquianos discutindo a manipulação das eleições parlamentares. O Comitê considerou que, para ser consistente com as responsabilidades de direitos humanos da Meta, a plataforma deveria ter exibido um rótulo de “Alto risco” no conteúdo nesse caso, bem como em todos os conteúdos com esse clipe de áudio.

 A maioria do Comitê de Supervisão [manteve](#) a decisão da Meta de deixar ativos dois posts que incluíam vídeos com uma mulher transgênero sendo confrontada por usar um banheiro feminino e uma atleta transgênero que venceu uma corrida de atletismo.

Exemplos de ações que seguiram as recomendações do Comitê de Supervisão:

 Seguindo as [recomendações](#) no caso de uma [mulher iraniana confrontada na rua](#) e de acordo com nossos esforços para reduzir excessos na aplicação de regras, atualizamos nossa abordagem à aplicação mais reativa das políticas. Isso ajudou a mitigar os riscos de tomar medidas contra linguagem figurada ou brincadeiras que possivelmente não representam uma ameaça.

 Seguindo [recomendações](#) para fornecer mais detalhes sobre suas decisões relacionadas a conteúdo, a Meta [aumentou](#) a granularidade de notificações aos usuários, fornecendo um aviso mais detalhado quando o conteúdo do usuário é restringido devido a uma denúncia formal do governo afirmando que ele viola leis locais.

Saiba mais detalhes nos [Relatórios Semestrais da Meta sobre o Comitê de Supervisão](#) e em nossa [página de Impacto do Comitê de Supervisão](#).

Impacto: maior transparência por meio de notificações aos usuários

+9%

1º tri de 2024 – 1º tri de 2025

Aumento nas taxas de notificação sobre a aplicação elegível dos Padrões da Comunidade no Facebook

+16%

1º tri de 2024 – 1º tri de 2025

Aumento nas taxas de notificação sobre a aplicação elegível dos Padrões da Comunidade no Instagram

~100%

no 1º tri de 2025

Taxa de notificação sobre conteúdo restringido por violar leis locais no Instagram e no Facebook



Também apoiamos a [liderança de ideias](#) do Comitê de Supervisão relacionada à liberdade de expressão e à avaliação sistêmica de riscos, para que os órgãos reguladores pudessem aprender com a experiência do Comitê. Ao longo do ano, o Comitê participou de discussões sobre políticas na Assembleia Geral da ONU e dialogou com a Comissão Africana dos Direitos Humanos e dos Povos, entre outras partes interessadas.

***Aviso legal:** existem muitas diferenças entre redes sociais e filmes. Nós não trabalhamos com a MPA ao atualizar nossas configurações de conteúdo; ela não classifica nenhum conteúdo no Instagram e não endossa nem aprova nossas configurações de conteúdo de nenhuma forma. Em vez disso, usamos como referência as diretrizes públicas da MPA, que já são familiares para pais, mães e responsáveis, além do feedback deles. Nossos sistemas de moderação de conteúdo não são iguais a um órgão de classificação de filmes, portanto, a experiência pode não ser exatamente a mesma.

[Sobre](#)[Resumo executivo](#)[Gerenciamento
de riscos](#)[Problemas](#)[Partes
interessadas](#)[Transparência e reparação](#)[Anexo](#)

Anexo



Como os direitos humanos são regidos e gerenciados na Meta

Nossos especialistas em direitos humanos direcionam a implementação da nossa [Política Corporativa sobre Direitos Humanos](#), que é supervisionada pelo nosso diretor de assuntos globais e diretor jurídico.

As atribuições dos especialistas em direitos humanos incluem promover a integração da política nas políticas, nos programas e nos serviços existentes e em desenvolvimento, realizar auditorias e oferecer suporte ao treinamento dos funcionários sobre a política. A política orienta a criação de produtos que respeitem os direitos humanos, a resposta a crises emergentes e a atuação com rapidez e agilidade para incorporar os direitos humanos em larga escala.

Treinamento dos funcionários da Meta sobre direitos humanos

Na Meta, a forma como criamos é tão importante quanto o que criamos. Nosso treinamento sobre direitos humanos destaca os impactos potenciais e reais dos nossos serviços, políticas e decisões comerciais sobre os direitos humanos. O objetivo é fomentar uma mentalidade voltada para os direitos humanos em nosso trabalho diário, promovendo o respeito a esses direitos em benefício de todos que usam nossos serviços.

Lançamos nosso treinamento *Bigger than Meta: Human Rights* (Maior do que a Meta: direitos humanos) em 2022, e ele continuou ao longo de 2025. Os treinamentos em privacidade também reforçam nossos objetivos de capacitação em direitos humanos concentrando-se em desenvolver nossa capacidade coletiva de proteger indivíduos, especialmente aqueles pertencentes a categorias marginalizadas, contra danos resultantes do processamento dos dados das pessoas.

Links dos relatórios de referência

[Responsible Business Practices Reports](#)

[Relatórios de Sustentabilidade](#)

[Relatório sobre Direitos Humanos de 2024](#),
[Relatório sobre Direitos Humanos de 2023](#),
[Relatório sobre Direitos Humanos de 2022](#),
[Relatório sobre Direitos Humanos de 2021](#)

[Declaração contra Escravidão e Tráfico de Pessoas de 2026](#)

[Relatório sobre Minerais de Conflito de 2025](#)

[Relatórios de Transparência da Meta](#)

[Relatórios regulamentares e outros relatórios de transparência](#)

[Adversarial Threat Reports](#)

[Relatórios Semestrais da Meta sobre o Comitê de Supervisão](#)

