

Cartilha de  
Enfrentamento à  
Desinformação  
em Redes Sociais



# Cartilha de Enfrentamento à Desinformação em Redes Sociais

## Autoria

Ana Bárbara Gomes Pereira  
Fernanda dos Santos Rodrigues Silva  
Júlia Maria Caldeira Gertrudes  
Rafaela Ferreira Gonçalves da Silva

## Revisão

Abong  
Ação Educativa  
Ibase

## Apoio Institucional

Supremo Tribunal Federal

## Apoio<sup>1</sup>



Cofinanciado pela  
União Europeia

## Como referenciar em ABNT:

PEREIRA, Ana Bárbara Gomes. SILVA, Fernanda dos Santos Rodrigues; GERTRUDES, Júlia Maria Caldeira; SILVA, Rafaela Ferreira Gonçalves da. **Cartilha de Enfrentamento à Desinformação em Redes Sociais**. Belo Horizonte: Instituto de Referência em Internet e Sociedade, 2024. Disponível em: xxxxxx. Acesso em: dd mmm aaaa.

<sup>1</sup> Esta publicação foi produzida com o apoio da União Europeia. O conteúdo desta publicação é da exclusiva responsabilidade do Instituto de Referência em Internet e Sociedade (IRIS) e não pode, em caso algum, ser tomado como expressão das posições de nenhum de seus apoiadores ou parceiros.

## SUMÁRIO

PROJETO "ENTRE POSTS E POLÊMICAS" .....	4
O QUE É ESTA CARTILHA? .....	5
AFINAL, O QUE É MODERAÇÃO DE CONTEÚDO? .....	6
QUAIS DIREITOS A MODERAÇÃO DE CONTEÚDO AJUDA A PROTEGER? .....	7
E O QUE É DESINFORMAÇÃO? .....	8
COMO USAR A MODERAÇÃO DE CONTEÚDO PARA ENFRENTAR A DESINFORMAÇÃO? .....	11
E SE EU FOR VÍTIMA DE UMA MODERAÇÃO DE CONTEÚDO ERRADA? .....	13
ANOTOU TUDO? .....	15



## PROJETO “ENTRE POSTS E POLÊMICAS”

A **Cartilha de Enfrentamento à Desinformação em Redes Sociais** é um dos produtos do projeto “Entre posts e polêmicas”, desenvolvido pelo Instituto de Referência em Internet e Sociedade (IRIS), com apoio da Google. Este estudo propõe uma **análise aprofundada sobre o sistema atual de resolução de controvérsias sobre o conteúdo moderado em plataformas digitais no Brasil e em outras experiências internacionais significativas, bem como o mapeamento de alternativas nessa seara.** O objetivo central é não apenas compreender as estratégias vigentes, mas também investigar vias para aprimorar esse processo de gestão de conflitos.

Busca, assim, fornecer um entendimento mais nítido e embasado sobre como as polêmicas relacionadas à moderação de conteúdo podem ser geridas de maneira mais eficaz e transparente no contexto digital brasileiro, visando a uma maior segurança para usuários e segurança jurídica, legitimidade e confiabilidade às plataformas digitais. Essa iniciativa também busca aprofundar e dar mais um passo nas pesquisas que o IRIS vêm desenvolvendo sobre moderação de conteúdo, todas disponíveis em nosso site de forma gratuita.

Nesta edição, a cartilha também faz parte do Projeto Nanet,<sup>2</sup> realizado pela Ação Educativa, Abong e Ibase, com apoio da União Europeia, e que tem o objetivo de promover consciência pública sobre governança da Internet, direitos digitais e plataforma para públicos estratégicos como as organizações da sociedade civil (OSCs), redes e coletivos dos seguintes segmentos: profissionais de comunicação, blogueiras/os, mulheres negras ativistas, população LGBTQIAP+, indígenas, redes de comunidade e operadores do direito e de políticas públicas.

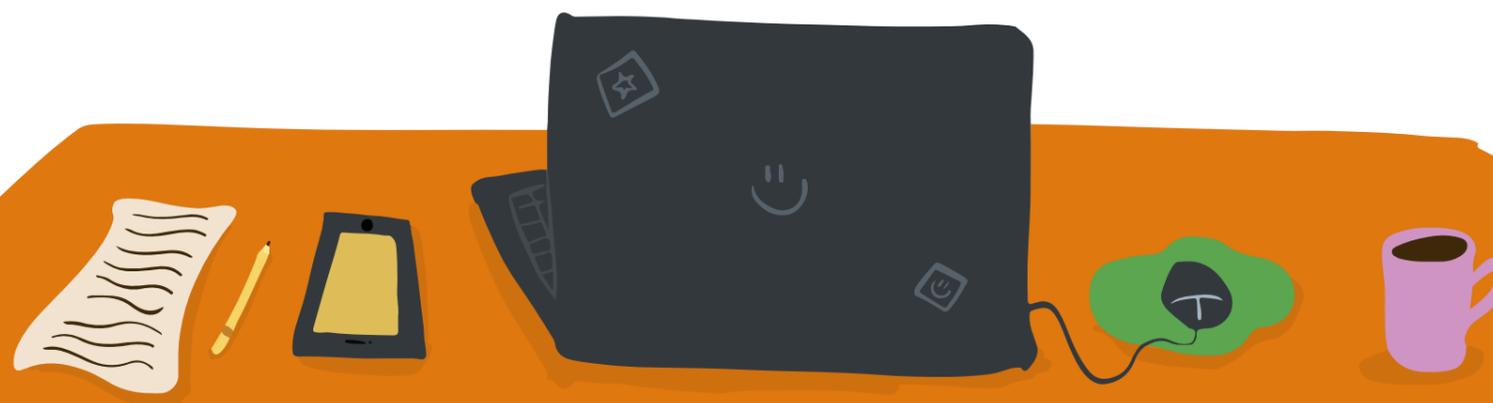
Este documento, por fim, também faz parte do Programa de Combate à Desinformação do Supremo Tribunal Federal,<sup>3</sup> que visa combater práticas que afetam a confiança das pessoas no sistema de justiça, distorcem ou alteram o significado das decisões e colocam em risco direitos fundamentais e a estabilidade democrática.

## O QUE É ESTA CARTILHA?

Esta Cartilha tem como objetivo apresentar formas de enfrentamento à desinformação em ambientes digitais, especialmente de redes sociais, para usuários de plataformas, organizações da sociedade civil e movimentos sociais. Também propõe uma conversa sobre moderação de conteúdo, para que se saiba do que se trata, como ela acontece e alguns desafios que surgem quando esse mecanismo falha.

Tudo mudou nos últimos anos, não é? A nossa forma de nos comunicar, de nos informar e, também, a forma como são produzidos e compartilhados os conteúdos que encontramos nas redes sociais. É também na internet onde parte do debate público acontece, onde nos informamos sobre a democracia, o papel do Estado, e os serviços que podemos acessar. Junto a toda essa transformação, também é necessário que estejamos preparados para interagir de forma responsável e bem informada nas redes, promovendo um ambiente digital seguro para todas as pessoas.

**Vem com a gente?!**



<sup>2</sup>Saiba mais em: <https://acaoeducativa.org.br/projeto/nanet-democratizando-a-tecnologia/>.

<sup>3</sup>Saiba mais em: <https://portal.stf.jus.br/desinformacao/>.

## AFINAL, O QUE É MODERAÇÃO DE CONTEÚDO?

Apesar do nome diferente, moderação de conteúdo nada mais é do que a forma com que plataformas digitais, como redes sociais, gerenciam o conteúdo postado em seus espaços.<sup>4</sup> Isso significa tanto as ações relacionadas à remoção ou suspensão de postagens e contas, como também a recomendação de conteúdo que aparece para você na sua linha do tempo.



Q Separamos esses conteúdos pra você!

Ou seja, tudo o que você vê (e o que você não vê) é determinado pela moderação de conteúdo realizada por plataformas. E elas fazem isso baseadas em suas próprias políticas e regras de comunidade, que são aqueles termos que elas perguntam se você concorda logo quando você cria a conta, lembra?

<sup>4</sup>Para saber mais sobre as ações e funções da moderação de conteúdo, veja SILVA, Fernanda dos Santos Rodrigues; GERTRUDES, Júlia Maria Caldeira; SILVA, Rafaela Ferreira Gonçalves da. Regulação de plataformas e devido processo na moderação de conteúdo: perspectivas em 5 continentes. Belo Horizonte: Instituto de Referência em Internet e Sociedade, 2024. Disponível em: <https://irisbh.com.br/publicacoes/devido-processo-na-moderacao-de-conteudo-em-5-continentes/>. Acesso em: 30 jul. 2024.

## QUAIS DIREITOS A MODERAÇÃO DE CONTEÚDO AJUDA A PROTEGER?

É importante lembrar que a moderação de conteúdo é uma atividade legítima das plataformas, sendo permitida, inclusive, pelo próprio Marco Civil da Internet. A moderação é fundamental, inclusive, para garantir que o ambiente virtual seja um espaço seguro e inclusivo para todas as pessoas.

Nesse sentido, a moderação de conteúdo protege o nosso direito de **liberdade de expressão** e de **acesso à informação**, buscando evitar também que as pessoas entrem em contato com conteúdos nocivos, como discurso de ódio e desinformação.



# E O QUE É DESINFORMAÇÃO?

Você já deve ter se deparado com uma informação sobre formas de conseguir dinheiro fácil na internet, ou ainda sobre a cura milagrosa para uma doença complexa. Esses são casos típicos de conteúdo desinformativo: feitos para enganar ou confundir os usuários, podem gerar danos financeiros, psicológicos e até colocar a vida de pessoas em risco.

O Tribunal Superior Eleitoral<sup>5</sup> entende que desinformação pode ser considerado um termo guarda-chuva para “diferentes conteúdos relacionados aos contextos de desordem informacional e manipulação informacional”, seja quando há a intenção ou não do agente que a divulga.

O material da UNESCO: “Jornalismo, Fake News & Desinformação: Manual para Educação e Treinamento em Jornalismo”<sup>6</sup> apresenta as diferentes formas que um conteúdo desinformativo pode tomar. Isso porque nem todo conteúdo desinformativo é uma mentira completa e mal-intencionada, podendo ser uma notícia antiga, fora de contexto ou uma interpretação equivocada que, mesmo que compartilhada com boa intenção, causa confusão ao ser passada para frente.



<sup>5</sup>TRIBUNAL SUPERIOR ELEITORAL (TSE). Programa permanente de enfrentamento à desinformação no âmbito da Justiça Eleitoral: Plano Estratégico Eleições 2022. Brasília: Tribunal Superior Eleitoral, 2022. Disponível em: <https://www.justicaeleitoral.jus.br/desinformacao/arquivos/programa-permanente-de-enfrentamento-a-desinformacao-novo.pdf>. Acesso em: 23 jul. 2024.

<sup>6</sup>Julie Posetti, Cheryl Iretton, Claire Wardle, Hossein Derakhshan, Alice Matthews, Magda Abu-Fadil, Tom Trewinnard, Fergus Bell e Alexios Mantzarlis. Jornalismo, Fake News & Desinformação. Manual para Educação e Treinamento em Jornalismo. UNESCO. 2019. Disponível em: <https://unesdoc.unesco.org/ark:/48223/pf0000368647/PDF/368647por.pdf.multi>. Acesso em: 23 jul. 2024.

Para ajudar a entender como identificar uma desinformação, os autores propõem dividi-las em sete formatos:

1. **Sátira:** É possível que, ao compartilhar um conteúdo no tom de humor ele possa ser mal compreendido e repassado como verdade, caracterizando um tipo de desinformação;
2. **Conexão falsa:** Sabe quando você vê uma notícia bombástica e quando vai ler se sente enganada porque na verdade não era aquilo que o título deu a entender? Esse é um tipo de desinformação que se apropria da informação com a intenção de atrair o maior número de leitores, mesmo que o conteúdo não seja exatamente aquilo que a chamada promete.
3. **Conteúdo enganoso:** Imagine que um influenciador digital promove um produto sem evidenciar que aquele é um conteúdo patrocinado. Esse é um exemplo de conteúdo enganoso. Apesar de parecer que assim o conteúdo será mais natural e “verdadeiro”, o CONAR (Conselho Nacional Autorregulamentação Publicitária) determina<sup>7</sup> que todo conteúdo publicitário seja identificado como tal. Outro exemplo de conteúdo enganoso é quando uma citação fora de contexto é utilizada para enquadrar uma notícia ou associar à uma personalidade. Isso seria apenas um recorte do todo e nos lembra da importância de verificar não apenas a veracidade das informações, mas também o seu contexto.
4. **Contexto falso:** Em 2024, assistimos a muitas imagens de enchentes e situações de emergência ambiental ao redor do mundo sendo compartilhadas como se estivessem acontecendo no estado do Rio Grande do Sul. Esse é um caso de contexto falso, em que a imagem compartilhada pode induzir ao erro e a uma compreensão falsa sobre a situação.

<sup>7</sup>Guia de Publicidade por Influenciadores Digitais. CONAR. Disponível em <[http://conar.org.br/pdf/CONAR\\_Guia-de-Publicidade-Influenciadores\\_2021-03-11.pdf](http://conar.org.br/pdf/CONAR_Guia-de-Publicidade-Influenciadores_2021-03-11.pdf)>, acesso em 23 de julho de 2024



5. **Conteúdo impostor:** Imagine um texto com conteúdo desinformativo, mas que supostamente tem a assinatura de uma grande jornalista, ou parece estar associado a um grande portal de informações. Esse é um conteúdo impostor, que se aproveita indevidamente da legitimidade de jornalistas sérios para compartilhar desinformação.
6. **Conteúdo manipulado:** É o caso de quando temos uma informação real, mas que determinados elementos são manipulados para confundir a interpretação.
7. **Conteúdo fabricado:** Esse tem sido um exemplo cada vez mais recorrente nos nossos feeds de redes sociais: imagens geradas por tecnologias de inteligência artificial fazem parecer que tal evento aconteceu de uma forma que não é real. No caso de uso de tecnologias para criar deep fake, um vídeo manipulado pode nos levar a acreditar que uma pessoa disse algo que ela jamais falou, muito menos registrou em vídeo.

### Combater a desinformação é uma postura cidadã e responsável na rede.

É dever de todo usuário zelar por uma internet segura para todos. Nas próximas páginas conversaremos sobre boas práticas que podemos adotar ao encontrar um conteúdo danoso ou desinformativo.

## COMO USAR A MODERAÇÃO DE CONTEÚDO PARA ENFRENTAR A DESINFORMAÇÃO?

A construção de um ambiente digital saudável e seguro para todos é uma responsabilidade compartilhada. Todas as pessoas devem ser responsáveis com os conteúdos que compartilham e terem uma postura cidadã ao encontrar um

conteúdo que pode oferecer dano a alguém. Se você, infelizmente, acabou consumindo uma foto, vídeo ou outro tipo de conteúdo nocivo ou desinformativo, **saiba o que fazer para que esse material não chegue a mais pessoas!**



1. **As plataformas costumam disponibilizar um menu com as tarefas que podemos realizar a partir daquele conteúdo.** Ele pode ter o formato de três pontinhos (...), vir junto ao botão de compartilhar, ou algum botão que indique a possibilidade de "saber mais"; depois de encontrar este menu, você deve buscar a opção "denunciar" e daí é só seguir o passo-a-passo.  
  
Você pode ser perguntado sobre o motivo da denúncia, ou seja, a razão pela qual o conteúdo parece nocivo, visto que isso ajuda na análise do conteúdo. A opção de denunciar também funciona para reportar aquele perfil que vive mandando desinformação nas mensagens diretas. Procure o botão que leva ao menu e denuncie.
2. **Outra possibilidade é que esta ferramenta esteja disponível quando você pressiona e segura um comentário que deseja denunciar.** Nesse caso, o menu de opções aparecerá e é só prosseguir com a opção de "Denunciar". Isso pode ser útil caso o seu conteúdo seja atacado com uma chuva de comentários indesejados, por exemplo.

## E SE EU FOR VÍTIMA DE UMA MODERAÇÃO DE CONTEÚDO ERRADA?

Obs.: Não se preocupe, **a conta denunciada não verá que você a denunciou**. No caso de algumas plataformas esse dado pode ser informado apenas se você estiver indicando que houve uma infração de propriedade intelectual, o que envolve um procedimento diferente. No combate à desinformação, é muito importante um compromisso coletivo em promover a integridade da informação. Ao se deparar com uma desinformação, denuncie o conteúdo! E não deixe de buscar informações seguras, consultar agências de checagens e se informar sobre os serviços públicos através de canais oficiais. **Seja uma boa influência, e só compartilhe informações verificadas!**



Como já conversamos, a moderação de conteúdo é muito importante para manter o ambiente digital seguro e inclusivo para todas as pessoas. Porém, essa atividade não está livre de falhas.

Há estudos que apontam, por exemplo, que pessoas negras e trans têm mais chances de terem conteúdo removido de forma errada quando compartilham experiências relacionadas à sua identidade.<sup>8</sup> Aqui mesmo no Brasil, já foram reportados casos em que ativistas tiveram conteúdos indisponibilizados após se manifestarem contra alguma forma de opressão.

Um exemplo bem conhecido foi quando o humorista, ator e apresentador Paulo Vieira fez uma publicação em rede social contra a chacina do Jacarezinho, em 2021, em que uma operação policial na favela de mesmo nome resultou na morte de 28 pessoas.<sup>9</sup> A postagem foi removida em um primeiro momento por supostamente violar as políticas de comunidade e o Conselho Nacional dos Direitos Humanos, por meio de sua Comissão Permanente de Direito à Comunicação e Liberdade de Expressão, encaminhou um pedido de informações à plataforma envolvida.<sup>10</sup>

Além disso, têm se tornado comum os casos em que pessoas se juntam para denunciar em massa alguma conta ou conteúdo específico, a fim de que a plataforma o remova, como em forma de protesto, por exemplo. No entanto, isso também pode ser utilizado de forma negativa, como para derrubar conteúdos legítimos e verídicos.

Já pensou se você ou a sua organização/movimento social tem um conteúdo de protesto removido de forma inadequada? Saiba o que fazer se isso acontecer!

<sup>8</sup>HAIMSON, Oliver et al. Disproportionate Removals and Differing Content Moderation Experiences for Conservative, Transgender, and Black Social Media Users: Marginalization and Moderation Gray Areas. *Proceedings of the ACM on Human-Computer Interaction*, v. 5, n. 466pp 1–35, out./ 2021. Disponível em: <https://dl.acm.org/doi/10.1145/3479610>. Acesso em: 16 jun. 2024.

<sup>9</sup>Post no Instagram de Paulo Vieira contendo print da mensagem de remoção da plataforma. Disponível em: <https://www.instagram.com/p/COIDkqdnjQv/>. Acesso em: 11 mar. 2024.

<sup>10</sup>Pedido de informações feito Conselho Nacional dos Direitos Humanos, por meio de sua Comissão Permanente de Direito à Comunicação e Liberdade de Expressão, à Meta. Disponível em: [https://intervozes.org.br/wp-content/uploads/2021/08/SEI\\_00135.215704\\_2021\\_92-3.pdf](https://intervozes.org.br/wp-content/uploads/2021/08/SEI_00135.215704_2021_92-3.pdf). Acesso em: 11 mar. 2024.

Nos casos de moderação de conteúdo errada, provavelmente a sua postagem ou conta já foi removida ou suspensa. Isso significa que, para tentar reverter essa situação, você precisará recorrer dessa decisão equivocada.

Em regra, as plataformas notificam o usuário quando uma moderação acontece, informando as razões para a suspensão ou remoção. A partir disso, normalmente existem três caminhos:

1. **A mensagem já traz junto algum ícone indicando a possibilidade de recorrer ou revisar a decisão.** Esse ícone pode ser um local escrito "recorrer da decisão", "apelar de decisão" ou outra mensagem similar, que vai te levar a opções sobre o porquê você quer apresentar um recurso. Selecione a opção que tem mais relação com o seu caso e peça para que a plataforma revise. Idealmente, as plataformas avisam em quanto tempo darão uma resposta. Caso a resposta continue negativa, você pode tentar buscar a ajuda de um advogado para tentar restaurar o conteúdo judicialmente.
2. **Você deve acessar as configurações da plataforma para solicitar suporte.** Isso pode acontecer caso você não tenha recebido a mensagem sobre a decisão de moderação ou deseja reclamar de alguma remoção de conteúdo que já aconteceu. Nesse caso, você deve buscar o ícone que leva às configurações da plataforma. Normalmente, ele está disponível na página do seu próprio perfil, que pode ser acessado através de ícones como  ou clicando na sua própria foto. O ícone para as configurações geralmente também está no formato de ,  ou . Após acessar, você deverá buscar pela opção de suporte ou ajuda, a fim de tentar identificar onde está o registro do conteúdo indisponibilizado e solicitar a revisão da decisão.
3. **Outras formas de solicitação de ajuda.** Se nenhuma das opções acima funcionar, você pode tentar buscar pela notificação em outros espaços. Em algumas plataformas, pode ser na aba ou caixa de entrada de notificações gerais. Também é possível que a rede social indique alguma opção mais direta como "conteúdo" ou, se o conteúdo removido ainda estiver disponível para você visualizar, você pode clicar nele mesmo e tentar recorrer da decisão.

**Pronto!** Assim que você tiver recorrido da decisão, é só esperar pelo período indicado para ter um retorno a respeito. Caso ela não responda ou a resposta não for suficiente, você pode procurar algum profissional especializado ou organizações que atuam pela defesa dos direitos digitais para verificar o que é possível fazer.

## ANOTOU TUDO?

Agora que você já sabe como proceder tanto para denunciar um conteúdo desinformativo, como para solicitar a revisão de algum conteúdo seu que tenha sido removido de forma errada, compartilhe as dicas dessa cartilha com alguém e amplie a rede contra desinformação na internet!

E não se esqueça da checagem **#IRIS** na hora de compartilhar alguma notícia online:

1. Investigue se os fatos são verdadeiros em veículos de comunicação confiáveis;
2. Recorra a instituições oficiais e agências de checagem de fatos para sanar dúvidas;
3. Informe um amigo ou amiga se ele ou ela estiver compartilhando algum conteúdo desinformativo; e
4. Suspeite de títulos sensacionalistas ou chamativos em excesso.

**Vamos juntos colaborar para um ambiente online mais saudável e democrático para todes!**



ISBN 978-65-993379-5-6

REALIZAÇÃO



PARCERIA



APOIO INSTITUCIONAL



APOIO



Cofinanciado pela União Europeia