

Inteligência Artificial

A Meta investe para proteger as eleições no ambiente online dentro e fora de períodos eleitorais, de forma contínua. Neste ano, mais de 150 milhões de brasileiros devem ir às urnas para eleger prefeitos e vereadores, e seguimos nos empenhando para garantir a integridade das Eleições nas plataformas da Meta. Desde 2016, investimos mais de US\$ 20 bilhões em segurança e integridade e temos cerca de 40 mil pessoas trabalhando nestas áreas globalmente. Estamos constantemente nos adaptando para estar à frente de novos desafios, como o uso de tecnologia de Inteligência Artificial (IA).

Este guia contém informações importantes sobre o uso de IA nos serviços Facebook e Instagram. Estas informações se referem às Eleições de 2024 e podem ser alteradas por novas atualizações dos serviços Facebook e Instagram.

Noções gerais sobre o uso de IA nos serviços Facebook e Instagram

Estamos atentos aos desafios que conteúdos gerados por IA podem representar quando usados por pessoas mal intencionadas. Temos trabalhado para assegurar o uso responsável da tecnologia de IA em nossos aplicativos, o que sabemos ser muito relevante em períodos eleitorais.

Nova regra publicada pelo TSE

Este ano, o TSE aprovou a **Resolução 23.732/24**, por meio da qual dispõe, entre outras questões, que os conteúdos editados com uso de recurso de IA para além de ajustes de qualidade, produção de elementos gráficos de identidade visual ou recursos de marketing costumeiro, precisam apresentar uma rotulagem ou uma indicação de que houve utilização dessa tecnolo-

gia (**art. 9-B, §1º, II e §2º, I ao III**).

Também de acordo com esta Resolução, os chamados “deep fakes” de vídeo e áudio estão expressamente vedados (**art. 9-C, §1º**).

Nova política já em vigor na Meta

Antes mesmo da publicação desta nova Resolução pelo TSE, anunciamos uma nova política para ajudar as pessoas a entenderem quando anúncios sobre temas sociais, eleições ou política no Facebook ou no Instagram foram criados ou alterados digitalmente, inclusive por meio do uso de IA. Essa política é aplicada globalmente. Os anunciantes terão que divulgar quando um anúncio sobre uma questão social, eleitoral ou política tiver uma imagem ou vídeo realista, ou áudio com som realista, que tenha sido criado digitalmente ou alterado para:

Descrever uma pessoa real dizendo ou fazendo algo que não disse ou fez; ou

Retratar uma pessoa, com aparência realista, que não existe ou um evento realista que não aconteceu, ou alterar a gravação de um evento real que aconteceu; ou

Descrever um evento real que supostamente tenha ocorrido, mas que não seja uma imagem, vídeo ou gravação de áudio verdadeira do evento.

Os anunciantes que veiculam esses anúncios não precisarão divulgar quando o conteúdo criado ou alterado digitalmente não forem relevantes ou substanciais para afirmação ou questão levantada no anúncio. Isso inclui ajustes no tamanho ou recorte de uma imagem, correção de cor ou nitidez da foto, a não ser que tais alterações sejam importantes ou significativas para uma afirmação ou questionamento levantado no anúncio. A Meta incluirá informações sobre esses anúncios quando um anunciante divulgar, durante o processo

de publicidade, que o conteúdo foi criado ou alterado digitalmente. Essas informações também aparecerão na Biblioteca de Anúncios. Se determinarmos que um anunciante não estiver divulgando conforme as orientações, não aceitaremos o anúncio e o descumprimento repetido dessa política poderá resultar em sanções contra o anunciante. Compartilharemos informações adicionais sobre o processo pelo qual os anunciantes passarão durante a criação de anúncios.

Como informar que uma mídia de anúncio foi criada ou alterada digitalmente

Para informar que um anúncio sobre temas sociais, eleições ou política contém uma imagem, vídeo ou áudio fotorrealista que foi criado ou alterado digitalmente:

1

No Gerenciador de Anúncios da Meta, verifique se Temas sociais, eleições ou política está selecionado na seção Categoria de anúncio especial.

2

Em seguida, na seção Criativo do anúncio, selecione a mídia neste anúncio que foi criada ou alterada digitalmente.

Quando um anunciante divulga o uso de mídia criada ou alterada digitalmente:

Um rótulo que diz “Informação de IA” aparecerá perto do rótulo “Pago por”

Na Biblioteca de Anúncios, a seção Detalhes do anúncio incluirá informações sobre mídia criada ou alterada digitalmente.

Importante frisar, ainda, que a IA tem um papel central em nosso trabalho de integridade para garantir a segurança e proteção dos usuários, sendo amplamente usada na detecção e remoção proativa em escala de conteúdos violadores de nossas plataformas, antes mesmo de denúncias da nossa comunidade.

Rótulos de IA em conteúdos orgânicos

Ajudamos as pessoas a compreender quando o conteúdo fotorrealista que estão vendo foi criado com nossas ferramentas de IA. Assim, para além dos rótulos em anúncios, conteúdo orgânico criado ou editado usando as ferramentas de IA e compartilhado no Instagram pode ser automaticamente rotulado como “informação de IA” acima ou no conteúdo.

A partir da evolução da indústria, aplicativos e ferramentas de edição ou criação de imagens e vídeos incluem em seus arquivos informações de uso de IA, seja para criação ou edição, quando aplicável. Essas informações, ao serem detectadas pela Meta, farão com que o conteúdo seja rotulado como “Informação de IA”. Isso inclui conteúdos criados ou editados em ferramentas de IA de terceiros. Quando os sinais de IA não são detectados pelo Instagram, talvez seja necessário rotular o conteúdo como “Informações de IA”. Em alguns casos, a rotulação é obrigatória.

A Meta exige que o rótulo seja utilizado caso um conteúdo gerado ou alterado através de criação digital ou ferramenta de IA - para além da melhoria da qualidade - possa ser considerado real, como vídeos, imagens ou áudios realistas.

Atuação e conteúdos violadores

Removemos conteúdos que violam nossas políticas, sejam eles criados por inteligência artificial ou por uma pessoa. Nossos parceiros de checagem de fatos independentes analisam e classificam desinformações virais e não permitimos que um anúncio seja veiculado caso seja classificado como “Falso”, “Alterado”, “Parcialmente Falso” ou “Sem Contexto”. Por exemplo, os parceiros de checagem de fatos podem classificar o conteúdo como “Alterado” se determinarem que ele foi criado ou editado de maneiras que possa enganar as pessoas, inclusive por meio do uso de inteligência artificial ou outras ferramentas digitais.