

Meta vai identificar, em suas redes sociais, imagens geradas por IA

Facebook, Instagram e Threads terão 'selo' para mostrar que fotos não são reais. Objetivo é combater desinformação

MONTECARLO

A Meta vai passar a identificar quaisquer imagens geradas por inteligência artificial (IA) que apareçam em suas redes sociais Facebook, Instagram e Threads. A iniciativa faz parte dos esforços para combater a desinformação em um ano de eleições importantes em vários países, incluindo Estados Unidos, Índia, África do Sul e Indonésia.

"Nos próximos meses, marcaremos as imagens que os usuários publicarem no Facebook, Instagram e Threads, desde que possamos detectar indicadores padrão do setor de que elas são geradas por IA", anunciou Nick Clegg, presidente de Assuntos Globais da Meta, no blog da empresa.

O aprimoramento das ferramentas de IA generativa

(capaz de criar textos, áudios e imagens) e o fato de elas agora estarem ao alcance de qualquer cidadão têm gerado temor de vídeos e áudios que usem a imagem de candidatos. Isso já ocorreu com o presidente americano Joe Biden, que busca a reeleição em janeiro, vários eleitores de New Hampshire receberam "telefonemas" de Biden pedindo que eles não votassem nas primárias que seriam realizadas no estado.

MINIMIZAR O PROBLEMA

A Meta já identifica imagens geradas por IA com a ajuda de sua própria ferramenta, a Meta AI, lançada em dezembro.

Segundo Clegg, a empresa quer fazer o mesmo com o conteúdo criado por meio de ferramentas de outras empresas, como Google, OpenAI (a criadora do ChatGPT), Mi-

crosoft, Adobe, Midjourney e Shutterstock.

"Estamos construindo essa ferramenta agora mesmo e, nos próximos meses, começaremos a aplicá-la em todos os idiomas disponíveis em cada aplicativo", explicou Clegg no blog.

Além do risco político, o desenvolvimento de programas de IA generativa pode levar a um fluxo incontrolável de conteúdo degradante, de acordo com muitos ativistas e órgãos reguladores, como deepfakes pornográficos de celebridades femininas, um fenômeno que também afeta muitos indivíduos anônimos.

Por exemplo, uma imagem falsa da cantora Taylor Swift foi vista 47 milhões de vezes na rede social X (antigo Twitter) no fim de janeiro, antes de ser removida. Segundo a imprensa ameri-



Terá alerta. A "foto" do papa Francisco de caça de caça enganou muita gente. Imagem assim sendo alvo da Meta

cana, a publicação permaneceu na plataforma por aproximadamente 17 horas.

Embora Clegg admita que essa marcação em larga escala, por meio de marcadores invisíveis, não eliminará totalmente o risco de produção de imagens falsas, ela "certamente minimizará" sua proliferação "dentro dos limites do que a tecnologia permite atualmente".

— Não é perfeita, a tecnologia ainda não está totalmente desenvolvida, mas, de todas as plataformas, é a tentativa mais avançada até agora de proporcionar transparência de forma significativa a bilhões de pes-

soas em todo o mundo — afirmou Clegg à agência de notícias AFP. — Espero sinceramente que, ao tomar a dianteira, incentivemos o restante do setor a trabalhar em conjunto e tentar desenvolver os padrões técnicos comuns de que precisamos. O executivo da Meta ressaltou que a empresa está disposta a compartilhar sua tecnologia aberta "da forma mais ampla possível".

PROVAR A AUTENTICIDADE

No Fórum Econômico Mundial, em Davos, no mês passado, Clegg alertou que a criação de padrões para identificar imagens geradas

por IA "é hoje a tarefa mais urgente" de todas.

Em entrevista à Bloomberg, o executivo da Meta afirmou que o fato de o sistema de identificação não ser perfeito "não é desculpa para não fazer nada".

Clegg disse ainda que, à medida que a internet for inundada com material gerado por IA, a indústria tecnológica será obrigada a também identificar o que é real.

— Teremos de promover um debate, em toda a sociedade ou, pelo menos, em toda a indústria, sobre como identificar aos usuários a veracidade ou autenticidade de conteúdo não gerado por IA.

Presidente do BC da Romênia é vítima de 'deepfake'

Işarescu, há mais de 30 anos no cargo, 'aparece' em vídeo promovendo falsos investimentos financeiros. Órgão alerta para fraude

BUENOS AIRES

O mais longo presidente de um banco central do mundo, Mugur Işarescu, da Romênia, foi alvo de um vídeo deepfake que o mostrava promovendo investimentos fraudulentos. O primeiro-ministro do país, Marcel Ciolacu, também "apareceu" recentemente divulgando falsos investimentos financeiros.

O episódio é mais um caso de desinformação criada para minar a credibilidade de instituições importantes. Na China, golpistas usaram a tecnologia deepfake para simular uma videocamada em grupo com o diretor financeiro de uma multinacional, que transferiu US\$ 26 milhões (cerca de R\$ 129 milhões) para cinco contas bancárias de Hong Kong.

O Banco Nacional da

Romênia, o BC do país, emitiu um alerta, lembrando aos cidadãos que nem Işarescu nem a autoridade fazem recomendações de investimentos. O vídeo usa a imagem e a voz de Işarescu para promover investimentos em ações e oferece às pessoas um link para uma plataforma financeira fraudulenta.

Estamos extremamente preocupados com o au-

mento significativo desses tipos de tentativas de fraude e pedimos que as pessoas sejam muito cuidadosas em

Mugur Işarescu. Não comande de BC romeno desde 1990

todas as transações que fizerem — disse o porta-voz do BC, Dan Suciu.

Os vídeos coincidem com um aumento no interesse pelo mercado de renda variável na Romênia, e pelos retornos acima da média oferecidos pela

Bolsa de Valores de Bucareste.

Especialistas em segurança

cibernética alertaram que os ataques se intensificaram este ano, com quatro rodadas de eleições — parlamentares, presidenciais, da União Europeia e local — programadas na Romênia.

Işarescu foi nomeado pela primeira vez para coman-

dar o BC romeno em 1990. Seu mandato atual expira este ano, e o premier já sinalizou apoio a outro candidato para o cargo.

Nas semanas passadas, o premier Ciolacu teve de trocar sua carteira de identidade depois que hackers

publicaram uma cópia e a publicaram na dark web.

(Da Bloomberg News)

Microsoft e Semafor criam serviço de notícias com IA

Plataforma, chamada de Signals, usa material de vários jornais. Gigante de tecnologia diz buscar formas de apoiar os jornalistas

DE EL PAÍS

Microsoft está fazendo uma série de acordos com organizações de notícias para implementar ferramentas de inteligência artificial (IA) generativa na produção de conteúdo. Entre as várias colaborações está uma aliança com a plataforma Semafor para criar uma fonte global de notícias de última hora, que a gigante de tecnologia batizou

de Signals, na qual jornalistas usariam ferramentas da Microsoft e da OpenAI, a criadora do ChatGPT.

"Nosso objetivo é encontrar maneiras de apoiar os jornalistas, não de substituí-los", afirmou a Microsoft em nota.

"Em um ano em que bilhões de pessoas votaram em eleições democráticas em todo o mundo, o jornalismo é fundamental para a criação de ecossistemas de informação saudáveis, e nossa missão, em

parceria com o setor, é garantir que as redações possam inovar para atender a este ano e aos anos seguintes."

À Semafor ressaltou que a Signals "responde às mudanças profundas e contínuas no cenário da mídia digital (...), bem como aos riscos e oportunidades apresentados pela inteligência artificial".

O Signals se baseia em diferentes fontes de mídia e resume seu conteúdo em pequenas publicações. Na segun-

da-feira, por exemplo, foram publicados resumos de notícias que citam fontes como The Times, The Washington Post, Der Spiegel e Financial Times, entre outras.

Muitos desses meios de comunicação cobram pelo acesso, que no Signals é gratuito.

A Microsoft diz que sua ideia é "treinar uma nova geração de repórteres nos melhores usos da IA e identificar maneiras pelas quais a IA pode ajudar a criar práticas comerciais ef-

icientes e ajudar a construir redações sustentáveis para as próximas gerações."

Os acordos da Microsoft com vários meios de comunicação e organizações ocorreram no momento em que a empresa, juntamente com a OpenAI, enfrenta uma ação judicial movida pelo New York Times pelo uso não autorizado de seu conteúdo para treinar tecnologias de IA.

A Microsoft também fechou um acordo com a Craig New-

mark Graduate School of Journalism, da Universidade Columbia em Nova York. Esta convidou jornalistas experientes para um programa gratuito, que vai explorar maneiras de incorporar a IA generativa ao trabalho nas redações.

Em outra frente, a Microsoft firmou parceria com a Online News Association (ONA), que lançou um programa para ajudar jornalistas e diretores de redação a navegar pelo ecossistema de IA; com o Ground Truth Project, que envia jornalistas para redações de todo o mundo; e com a Nota, uma startup dedicada a levar ferramentas de IA de alta qualidade às redações, para ajudar a melhorar suas operações.

INDICADORES

BOVESPA	+2,21%
FEVREIRO	-4,79%
em janeiro	

IMPOSTO DE RENDA	
Fevereiro de 2024	
Imposto sobre renda	19%
Imposto sobre renda	19%
Imposto sobre renda	19%
Imposto sobre renda	19%
Imposto sobre renda	19%

DÓLAR	
Comercial (Par)	5,382
Comercial (Par)	5,382
Comercial (Par)	5,382
Comercial (Par)	5,382
Comercial (Par)	5,382

OUTRAS MOEDAS	
Libra esterlina	5,382
Libra esterlina	5,382
Libra esterlina	5,382
Libra esterlina	5,382
Libra esterlina	5,382

ÍNDICES	
IPC-A	100,000
IPC-A	100,000
IPC-A	100,000
IPC-A	100,000
IPC-A	100,000

POUNÇA	
IPC-A	100,000
IPC-A	100,000
IPC-A	100,000
IPC-A	100,000
IPC-A	100,000

TR	
IPC-A	100,000
IPC-A	100,000
IPC-A	100,000
IPC-A	100,000
IPC-A	100,000

OUTROS ÍNDICES	
IPC-A	100,000
IPC-A	100,000
IPC-A	100,000
IPC-A	100,000
IPC-A	100,000

FUNDOS DE INVESTIMENTO	
IPC-A	100,000
IPC-A	100,000
IPC-A	100,000
IPC-A	100,000
IPC-A	100,000

ÍNDICES DE PREÇOS	
IPC-A	100,000
IPC-A	100,000
IPC-A	100,000
IPC-A	100,000
IPC-A	100,000

OUTROS ÍNDICES	
IPC-A	100,000
IPC-A	100,000
IPC-A	100,000
IPC-A	100,000
IPC-A	100,000