

Germano Neto

A IMPLOÇÃO DA CONFIANÇA

¹A confiança no que vemos e ouvimos sempre foi um pilar da vida em sociedade e da forma como aprendemos, enquanto sujeitos, o mundo à nossa volta. Fotografias, vídeos e registros sonoros, mesmo sujeitos a interpretações, eram considerados evidências sólidas de eventos e falas. Mas o avanço das deepfakes – “conteúdos sintéticos relacionados a pessoas naturais, gerados ou manipulados por tecnologias de inteligência artificial (IA)”² – está corroendo essa base.

Quando o empresário Luciano Hang³ “aparece” oferecendo celulares por R\$179,90, quando o Padre Marcelo Rossi⁴ “apresenta” o remédio milagroso que o curou da depressão ou quando a jornalista Sandra Annenberg⁵ é “vista” anunciando indenizações em um site falso do Governo Federal, não estamos diante apenas de distorções técnicas: estamos testemunhando uma mudança profunda na relação entre percepção, prova e verdade, que afeta a todas as pessoas na sociedade. É um potencial danoso que reside na modelagem do comportamento visual e na espoliação da percepção, concretizando o que o filósofo e crítico de arte Jonathan Crary⁶ nomeia de “rebaixamento e rotinização” dos sentidos do corpo humano e da realidade; por consequência, na erosão do regime democrático e da vida comunitária.

¹ José Germano Neto (jose.germano@ufrn.br) é potiguar, bacharel, licenciado e mestre em Ciências Sociais pela UFRN, atualmente cursa o doutorado na mesma área e instituição. Bolsista CAPES DS. Pesquisador do LabMonDes - Laboratório de Monitoramento sobre Desinformação (ECT/UFRN) e do Marginália - Grupo de Estudos Transdisciplinares em Comunicação e Cultura (CNPq). Tem se dedicado nos últimos 8 anos aos estudos do digital e de suas relações com a sociedade.

² Conceito da Secretaria de Comunicação Social da Presidência da República (Secom/PR). BRASIL. Presidência da República. Secretaria de Comunicação Social. Secretaria de Políticas Digitais. **Portaria nº 1, de 22 de maio de 2025**. Disponível em: <https://www.in.gov.br/web/dou/-/portaria-n-1-de-22-de-maio-de-2025-631519909>. Acesso em: 26 ago. 2025.

³ **AOS FATOS**. Golpe do celular com ‘deep fake’ de Hang inunda redes e gera crise inédita para Havan. Disponível em: <https://www.aosfatos.org/noticias/golpe-celular-havan-deep-fake-luciano-hang/>. Acesso em: 14 ago. 2025.

⁴ **CORREIO BRAZILIENSE**. Padre Marcelo Rossi denuncia ter sido vítima de deepfake: “É triste isso”. Disponível em: <https://www.correiobraziliense.com.br/diversao-e-arte/2024/09/6951460-padre-marcelo-rossi-denuncia-ter-sido-vitima-de-deepfake-e-triste-isso.html>. Acesso em: 14 ago. 2025.

⁵ **AOS FATOS**. Posts usam ‘deepfake’ de Sandra Annenberg e falso site do governo para aplicar golpe. Disponível em: <https://www.aosfatos.org/noticias/golpe-resgata-brasil-deepfake-sandra-annenberg/>. Acesso em: 14 ago. 2025.

⁶ CRARY, Jonathan. **Terra arrasada**: além da era digital, rumo ao mundo pós-capitalista. Tradução de Humberto do Amaral. São Paulo: Ubu, 2023..



Figura 1. Capturas de tela de redes sociais.

A pergunta – o que acontece quando perdemos a confiança no que vemos e ouvimos? – guia a análise das deepfakes não como simples falsificações, mas como sintomas de uma cultura digital em que real e artificial se entrelaçam. A ameaça é tecnológica e também social, política e subjetiva.

A fusão entre real e artificial

O professor e teórico da comunicação Norval Baitello Junior⁷ descreve como a saturação de imagens, intensificada pela era das comunicações eletrônicas e digitais – marcada pela instantaneidade, pelo alcance global e pela centralidade da imagem na modelagem da percepção e da interação humanas – nos conduziu a uma “perda do presente”: passamos a viver em um circuito de representações que se referem mais entre si do que à experiência direta.

As deepfakes intensificam esse cenário. Ferramentas como o *Veo3*, modelo do *Google* lançado em maio de 2025 capaz de gerar cenas cinematográficas com áudio sincronizado, demonstram como o artificial se torna indistinguível do real e, por isso, potencialmente perigoso. A questão central, portanto, vai além de determinar se a imagem é verdadeira ou falsa; ela passa por compreender como esse conteúdo opera na esfera social.

⁷ BAITELLO JUNIOR, Norval. **A era da iconofagia**: reflexões sobre a imagem, comunicação, mídia e cultura. São Paulo: Paulus, 2014.

Quando um médico como Drauzio Varella⁸ “aparece” promovendo uma cura falsa para diabetes ou o presidente Lula⁹ “anuncia” um benefício que libera automaticamente até R\$8 mil reais via PIX, não se trata de um mundo artificial substituindo o real, é uma manipulação que se ancora na autoridade e na confiança de indivíduos reais.



Figura 2. Capturas de tela de redes sociais.

Essa capacidade de gerar conteúdo falso com alto grau de realismo compromete a credibilidade de fontes legítimas, criando um ambiente favorável à diversos danos, sendo um dos maiores deles a ameaça à integridade da informação. A integridade da informação, segundo a Organização das Nações Unidas (ONU)¹⁰, refere-se à exatidão, consistência e confiabilidade das informações. Ela é ameaçada por desinformação, informações equivocadas e discurso de ódio.

⁸ **O GLOBO**. Conar pede remoção de anúncio que prometia cura da diabetes com deepfake de Drauzio Varella. Disponível em: <https://oglobo.globo.com/blogs/lauro-jardim/post/2025/07/conar-pede-remocao-de-anuncio-que-prometia-cura-da-diabetes-com-deepfake-de-drauzio-varella.ghtml>. Acesso em: 14 ago. 2025.

⁹ **REDE NACIONAL DE COMBATE À DESINFORMAÇÃO**. Deepfake usa vídeo de Lula para promover golpe da falsa indenização. Disponível em: <https://rncd.org/deepfake-usa-video-de-lula-para-promover-golpe-da-falsa-indenizacao/>. Acesso em: 14 ago. 2025.

¹⁰ UNITED NATIONS. Our Common Agenda: **Policy Brief 8: Information Integrity on Digital Platforms**. New York: UN, jun. 2023. 28 p. Disponível em: <http://digitallibrary.un.org/record/4012857>. Acesso em: 26 ago. 2025.

Críticos acusaram recentemente o YouTube de manipular os vídeos *Shorts* com pós-processamento para que parecessem feitos por IA¹¹, preparando terreno para normalizar conteúdos sintéticos e reduzir a distinção com os vídeos do modo como realmente foram captados.

Rene Ritchie, o *creator liaison*¹² do YouTube, negou que o processamento em *Shorts* use IA generativa ou faça *upscaling* – técnica que faz com que as imagens pareçam melhores em uma resolução mais alta –, alegando tratar-se apenas de técnicas tradicionais de aprendizado de máquina, como as já utilizadas em smartphones.

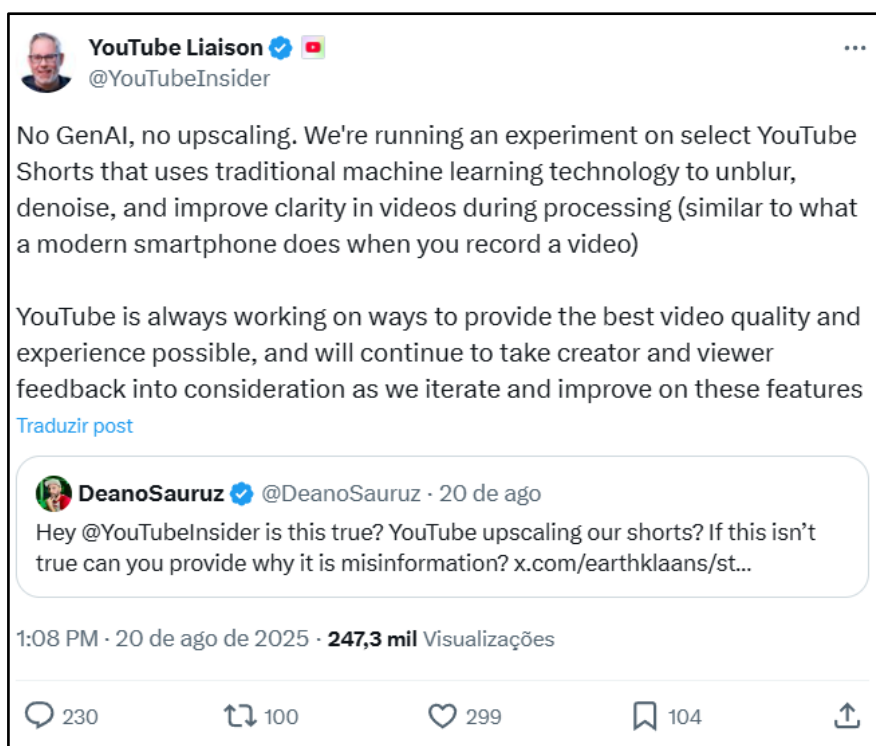


Figura 3. Captura de tela a partir do x.com¹³.

¹¹YouTube altera vídeos sem avisar e usuários questionam aparência de IA. **Tecnoblog**,. Disponível em: <https://tecnoblog.net/noticias/youtube-altera-videos-sem-avisar-e-usuarios-questionam-aparencia-de-ia/>. Acesso em: 6 set. 2025.

¹²“Sem GenAI, sem upscaling. Estamos realizando um experimento em vídeos selecionados do YouTube Shorts que usa tecnologia tradicional de aprendizado de máquina para desfocar, reduzir ruído e melhorar a clareza dos vídeos durante o processamento (semelhante ao que um smartphone moderno faz ao gravar um vídeo). O YouTube está sempre trabalhando em maneiras de fornecer a melhor qualidade e experiência de vídeo possível e continuará levando em consideração o feedback dos criadores e espectadores à medida que iteramos e aprimoramos esses recursos.” Tradução elaborada pela Google. YOUTUBE INSIDER. [@YouTubeInsider]. **Publicação de 20 ago. 2025**. Disponível em: <https://x.com/YouTubeInsider/status/1958199532363317467>. Acesso em: 6 set. 2025.

¹³ Em tradução livre: agente de ligação com criadores. Um profissional que atua como intermediário entre os criadores de conteúdo e uma empresa. Nesse caso, o YouTube.

A resposta tenta desarmar as críticas, mas ignora o ponto central: ao aplicar filtros automaticamente, a plataforma molda a percepção de qualidade e naturaliza efeitos de IA, o que pode reduzir ainda mais a capacidade do público de distinguir conteúdos artificiais.

Um terceiro elemento para essa discussão do aprimoramento técnico intervindo na relação entre artificialidade e realidade é o lançamento do *Gemini 2.5 Flash Image*¹⁴, anunciado pela Google em 26 de agosto de 2025. A ferramenta foi apresentada como um salto tecnológico em geração de imagens, permitindo que criadores produzam conteúdos complexos, personalizem campanhas e prototipem cenários rapidamente. Voltamos, assim, às discussões sensíveis que envolvem a reflexão sobre padronização de estéticas e deslocamento do valor da criatividade humana para interesses empresariais, levantando questões sobre concentração de poder e rastreabilidade ética das criações.

Uma pesquisa realizada pela Fundação Getúlio Vargas¹⁵ entre janeiro e maio de 2024 sobre os imaginários da inteligência artificial no debate digital brasileiro mostrou que, no Instagram, aproximadamente quatro em cada 25 posts foram classificados como casos de fraude e desinformação associados ao uso da IA. Esses dados evidenciam um cenário alarmante em que a tecnologia legitima e institucionaliza práticas fraudulentas, exigindo de nós uma reavaliação crítica sobre a maneira como interagimos com a informação e como construímos nossa compreensão da realidade.

Esse problema, porém, não se limita ao conteúdo orgânico: há também a circulação sistemática de anúncios fraudulentos veiculados e monetizados dentro da própria plataforma. Isso, de certo modo, revela uma cumplicidade ativa, na medida em que o Instagram extrai lucro direto dessas práticas. A retórica de compromisso com a integridade da informação, portanto, entra em choque com uma lógica econômica que prioriza o lucro mesmo diante da proliferação de danos sociais.

¹⁴GEMINI / GOOGLE AI STUDIO. Apresentamos o Gemini 2.5 Flash Image, nosso modelo de imagem de última geração. **Google Developers Blog**, 26 ago. 2025. Disponível em: <https://developers.googleblog.com/pt-br/introducing-gemini-2-5-flash-image/>. Acesso em: 7 set. 2025.

¹⁵No Instagram, a pesquisa analisou 40.940 posts, provenientes de aproximadamente 14 mil páginas. Fonte: RUEDIGER, Marco Aurelio; GRASSI, Amaro; SABBATINI, Leticia; PEREIRA, Laura; SILVA, Lucas Roberto da; BARBOZA, Polyana; CORDEIRO, Maria Sirleidy. **Imaginários sobre a inteligência artificial no debate digital brasileiro**. Rio de Janeiro: FGV Comunicação, 2024. Disponível em: <https://hdl.handle.net/10438/36793>. Acesso em: 16 ago. 2025.

Opacidade, escala e dano

Como alerta Cathy O’Neil¹⁶, matemática norte-americana, algoritmos deixaram de ser ferramentas neutras para se tornarem “armas de destruição matemática”¹⁷, usadas para controlar acesso a recursos (empregos, empréstimos, liberdade condicional), amplificar desigualdades (por vieses embutidos em seus códigos) e sustentar regimes de vigilância (como aplicações de *scoring* social). Para ela, três elementos se destacam: opacidade, escala e dano; e as deepfakes evidenciam de forma clara essas dimensões.

A opacidade diz respeito à dificuldade de compreender como esses conteúdos são gerados, já que dependem de algoritmos complexos e proprietários, sem transparência sobre seus processos internos. Em termos de escala, circulam em redes sociais e plataformas digitais com alcance global, impactando bilhões de pessoas em poucos instantes. Quanto ao dano, os efeitos podem ser diretos ou indiretos, variando da manipulação de dados pessoais à produção de impactos psicológicos e sociais.

Como observa Kate Crawford¹⁸, a IA é um “registro do poder”, moldado por interesses corporativos e militares, voltado ao controle das narrativas, dos corpos e do futuro. Nesse sentido, de modo contundente, as deepfakes exemplificam esse processo: conteúdos cuja origem não compreendemos plenamente, mas que se espalham em segundos e produzem prejuízos que vão da fraude financeira a ações contra a saúde pública e destruição de reputações. Exemplos, como expusemos anteriormente, mostram que a manipulação digital é parte de uma ecologia de informação em que imagens e sons são recursos estratégicos de influência, é algo muito maior que um fenômeno isolado.

Considerações para uma confiança crítica

Quando a confiança nas evidências visuais e sonoras é fragilizada, todo o sistema de validação social é afetado. A política torna-se mais vulnerável à desinformação e ao extremismo, o jornalismo enfrenta novos desafios de verificação e checagem, e indivíduos comuns podem sofrer danos irreversíveis às suas reputações. Nesse cenário, o que acontece quando já não acreditamos no que vemos e ouvimos?

¹⁶ O’NEIL, Cathy. **Algoritmos de destruição em massa**: como o big data aumenta a desigualdade e ameaça a democracia. Santo André: Rua do Sabão, 2020.

¹⁷ A pesquisadora faz uso de um jogo de palavras em inglês: *weapons of math/mass destruction*, onde matemática (*math*) e massa (*mass*) possuem grande similitude na escrita e na pronúncia.

¹⁸ CRAWFORD, Kate. **Atlas da IA**: poder, política e os custos planetários da inteligência artificial. São Paulo: Edições Sesc São Paulo, 2025.

A resposta é dupla: por um lado, abre-se espaço para a dúvida e a checagem crítica; por outro, emerge o risco de um ceticismo paralisante, em que nenhuma prova parece suficiente. Enfrentar esse problema exige medidas articuladas.

Em um nível macro, são necessárias leis que responsabilizem os criadores e disseminadores de deepfakes nocivas, além da implementação de ferramentas de detecção integradas às plataformas digitais. Em paralelo, devem ser desenvolvidos programas de educação midiática que estimulem habilidades de leitura crítica do mundo digital desde cedo, bem como pesquisas voltadas aos impactos psicológicos, sociais, políticos e econômicos dessas manipulações.

Todavia, nenhuma dessas medidas será eficaz se não reconhecermos que real e artificial, na cultura digital, são dimensões interdependentes. O desafio não é sobre como podemos restaurar uma pureza perdida, mas construir novos modos de validação e confiança nas informações que consumimos.

O que está em jogo é a preservação de um espaço público baseado em algum grau de certeza compartilhada, além da proteção contra os danos que esses conteúdos geram. Reclamar a autoria e a integridade dos conteúdos digitais que circulam na rede é, ao mesmo tempo, um desafio técnico, político e cultural. É também a única forma de evitar que a dúvida absoluta substitua a confiança mínima necessária para viver em sociedade.