DESMONTANDO A MÁQUINA:

* COMUNICAÇÃO DIGITAL
* DISPUTA DO DEBATE PÚBLICO

DESMONTANDO A MÁQUINA:
COMUNICAÇÃO, RADICALIZAÇÃO E DISPUTA DO DEBATE PÚBLICO
Estratégias digitais para mandatos, candidaturas, organizações e movimentos sociais de defesa de direitos e justiça social e ambiental

Uma cartilha desenvolvida pela
CODING RIGHTS
para o
LAMPEJO
Escola de Tecnopolítica, Cuidados Digitais e Comunicação Estratégica

Maio, 2026

Ficha técnica

Desmontando a Máquina:
Comunicação, Radicalização e Disputa do Debate Público

REALIZAÇÃO	Coding Rights
ANO	2026
AUTORIA	Joana Varon, Juliana Mastrascusa e Laila Almeida Braga
DESIGN E DIAGRAMAÇÃO	Carolina Simonian
ILUSTRAÇÃO	amandrafts
ORGANIZAÇÕES PARCEIRAS que colaboraram na revisão	Mulheres Negras Decidem, Instituto Democracia em Xeque, NetLab UFRJ e Sleeping Giants

Esta cartilha é material formativo e de pesquisa aplicada em comunicação digital, cuidados digitais e tecnopolítica, produzido em defesa de direitos humanos. O conteúdo destina-se a mandatos, candidaturas, organizações da sociedade civil, movimentos sociais, comunicadoras, comunicadores e jornalistas que atuam em defesa de direitos humanos, justiça social e justiça ambiental.

Esta publicação fez uso de ferramentas de IA generativa para facilitar sistematização e edição de texto. Todo conteúdo foi revisado e reeditado pelas autoras.

Como citar

CODING RIGHTS. Desmontando a máquina: comunicação digital e disputa do debate público. [Cartilha]. Rio de Janeiro: Coding Rights; Lampejo — Escola de Tecnopolítica, Cuidados Digitais e Comunicação Estratégica, maio 2026. Disponível em: https://lampejo.digital/biblioteca. Acesso em: _____.

ISBN: 978-65-992117-8-2

Índice

Módulo 1
Como a população brasileira se informa?

6. INTRODUÇÃO
10. SEÇÃO 1 | Desigualdades informativas e polarização política
12. SEÇÃO 2 | Os tipos de desinformação
12. SEÇÃO 3 | Desinformação e democracia

13. Desinformação no WhatsApp e Telegram
14. O QUE FAZER | Protocolo de verificação em 10 passos

Módulo 2
A Máquina: Como as plataformas foram desenhadas para favorecer conteúdos radicalizantes

15. SEÇÃO 1 | Mecanismos estruturais das redes sociais
16. SEÇÃO 2 | O ciclo de produção e amplificação
16. SEÇÃO 3 | A disseminação em cascata de acordo com os mecanismos de cada plataforma
13. ✦ As três camadas do ecossistema
    Nível 1 - Laboratório: Telegram como espaço de teste
    Nível 2 - Amplificação: WhatsApp como vetor de dispersão
    Nível 3 - Viralização mainstream: algoritmo como amplificador involuntário
18. Efeito bola de neve em redes articuladas
19. SEÇÃO 4 | Redes sociais como ferramenta: dados de desinformação plataforma a plataforma
    ✦ Dark Pooling - O mecanismo invisível
    ✦ Data brokers e microtargeting político
24. SEÇÃO 5 | Brasil na lanterna da transparência
26. O que isso significa para a desinformação eleitoral
26. O que o campo progressista pode exigir
27. O QUE FAZER | Monitorar, reagir e construir

Módulo 3
Psicologia da persuasão radical

28. SEÇÃO 1 | Os 8 gatilhos emocionais explorados
29. SEÇÃO 2 | Os 4 erros mais comuns na contra-narrativa
30. O QUE FAZER | Protocolo de contra-narrativa em 5 passos
31. O QUE FAZER | Monitoramento de narrativas
32. O QUE FAZER | Como aprender sem replicar

Módulo 4
Violência política de gênero como tática de silenciamento

34. SEÇÃO 1 | Os 7 tipos de violência política digital
35. SEÇÃO 2 | Censura algorítmica de corpos dissidentes
35. SEÇÃO 3 | Protocolo de documentação em 10 passos
36. O QUE FAZER | 10 dicas de higiene digital básica e onde pedir ajuda em caso de ataque
37. Maria d'ajuda - linha de ajuda em cuidados digitais

Módulo 5
IA e eleições: o novo campo de disputa

38. SEÇÃO 1 | IA generativa e agentes de IA
    ✦ O que é Inteligência Artificial generativa?
    ✦ O que são agentes de IA?
39. SEÇÃO 2 | As regras: o que o TSE decidiu para 2026
    ✦ Regra de ouro para candidaturas e mandatos
42. SEÇÃO 3 | Deepfakes: o que são, como funcionam e casos documentados em eleições pelo mundo
    ✦ **Os três tipos que mais circulam no contexto eleitoral**
    Deepfake de vídeo
    Deepfake de áudio: clonagem de voz
    Deepfake de imagem e deepnude
43. Como os deepfakes têm sido usados nas eleições - 4 padrões documentados
44. SEÇÃO 4 | Como identificar conteúdo sintético (IA)
    ✦ Como identificar conteúdo sintético (IA)
    ✦ Algumas ferramentas de verificação
46. SEÇÃO 5 | Quem monitora: observatórios e iniciativas de referência
47. O QUE FAZER | Protocolo de resposta a deepfakes
    ✦ Se você receber um deepfake sobre você ou sua candidatura?
    ✦ Se você ver um deepfake sobre outra pessoa
    ✦ Prevenção: o que fazer antes de ser alvo
49. SEÇÃO 6 | IA como ferramenta estratégica (desde que usada com consciência e cuidado)
    ✦ Como escolher uma ferramenta: tecnopolítica importa
    ✦ O que cada empresa de IA generativa representa?
56. O QUE FAZER | O que nunca inserir em ferramentas de IA?
    ✦ Configurações básicas de proteção
    ✦ Alguns usos estratégicos possíveis, depois de configurações de proteção
    ✦ Agentes de IA
    ✦ Cuidados antes de usar qualquer agente de IA
    ✦ Alguns riscos específicos em contexto eleitoral
    ✦ O que a IA nunca vai substituir?

Módulo 6
Comunicação progressista eficaz

60. SEÇÃO 1 | Os 6 princípios da comunicação progressista eficaz
61. SEÇÃO 2 | Framing: como o enquadramento define a batalha
62. SEÇÃO 3 | Formatos: escolhendo uma forma de comunicar
64. SEÇÃO 4 | Disseminação: passar pra frente também faz parte da mensagem
    ✦ Cuidados antes de usar qualquer agente de IA
    ✦ Alguns riscos específicos em contexto eleitoral
    ✦ O que a IA nunca vai substituir?

65. FONTES DA CARTILHA

INTRODUÇÃO

Em 2026, o Brasil vai às urnas pela primeira vez num cenário em que qualquer pessoa com um celular consegue usar IA para fabricar um vídeo com narrativas falsas, porém convincentes. Violência política de gênero, desinformação em massa, deepfakes e ataques coordenados tendem a aumentar, mirando em quem ousa disputar poder e fomentar narrativas que questionem velhas estruturas de opressão, principalmente mulheres, LGBTQIAPN+, pessoas negras, indígenas, periféricas ou representantes de outros grupos historicamente vulnerabilizados que atuam em mandados, candidaturas, movimentos sociais, no jornalismo e na comunicação em defensa de direitos fundamentais

Por outro lado, este momento eleitoral também abre uma oportunidade para o campo de defesa de direitos humanos: a de pensar a comunicação de forma estratégica e se apropriar das tecnologias como ferramentas de luta e espaço de disputa de poder. Não partimos do zero: há uma riqueza e criatividade imensuráveis na comunicação já produzida por nós, assim como imaginários políticos potentes que precisam transbordar para além das margens e influenciar a política institucional. Entender o funcionamento da máquina que processa nossas comunicações digitais é o primeiro passo para ampliar o alcance das nossas mensagens. E esse entendimento vai muito além de desvendar algoritmos, ele é tecnopolítico.

Tecnologia é política. A extrema direita global não vem alcançando maior espaço no debate público por ter melhores comunicadores. Ela aprendeu a explorar **a arquitetura das plataformas digitais das Big Tech, desenvolvidas para priorizar engajamento emocional, monetizar polêmicas e amplificar discurso de ódio e desinformação.**

Esta cartilha é uma resposta diante de uma constatação incômoda: os movimentos de mulheres, os movimentos negros, a comunidade **LGBTQIAPN+, todo o campo democrático, disputam o debate público em um tabuleiro que não foi construído por nós, nem para nós**. Foi concebido de acordo com a visão de um grupo muito pequeno de CEOs do Vale do Silício, homens brancos bilionários, liderando a lista das pessoas mais ricas do mundo, que cada vez mais se alinham com ideologias de extrema direita. São as escolhas de design dessas empresas que moldam o que circula, quem é amplificado e quem é silenciado.

Escolhas que vêm piorando ainda mais desde o início do segundo governo Trump, quando Elon Musk (X), Jeff Bezos (Amazon), Mark Zuckerberg (Meta), Sundar Pichai (Google), Tim Cook (Apple) e Sam Altman (OpenAI) ocuparam lugar de destaque no mandato e na posse do presidente de extrema direita, se alinhando mais explicitamente. Daí em diante se seguiram mudanças na política de moderação de conteúdo e valores dessas plataformas: Zuckerberg anunciou o fim das proteções contra discurso de ódio direcionado a pessoas LGBTQIAPN+ e mulheres e a Meta passou a substituir a moderação humana por moderação utilizando Inteligência Artificial, enquanto sabe que seus algoritmos favorecem conteúdo de ódio e desinformação. Amazon, Google, Microsoft e Meta também encerraram ou reduziram políticas de diversidade e inclusão, bem como suas equipes de segurança e revisão de risco. **O resultado é um tabuleiro sistematicamente desfavorável ao campo democrático, uma máquina de ódio e desinformação.** Não é uma conspiração. É uma arquitetura. E entender essa arquitetura é condição para disputá-la.

O tabuleiro está se tornando mais hostil. E é exatamente por isso que esta cartilha existe: para desmontar essa máquina. A resposta à tecno-oligarquia e à extrema direita não pode ser apenas defensiva. Precisa ser também ofensiva, disputando narrativas, construindo redes, ocupando plataformas com estratégia própria.

Por isso, esta cartilha faz uma engenharia reversa das táticas de comunicação que radicalizam para servir de material informativo para mandatos, candidaturas, movimentos sociais, jornalistas e pessoas comunicadoras cuja atuação em defesa de direitos humanos cada vez mais depende também de se apropriar das tecnologias digitais de maneira estratégica. Compilando análises e dados do ecossistema informacional produzidos por diferentes organizações da sociedade civil e jornalistas, bem como a experiência das autoras em análises tecnopolíticas, cuidados digitais e comunicação estratégica, esta cartilha oferece ferramentas para **entender os mecanismos, reconhecer as táticas e traçar estratégias de comunicação capazes de disputar o debate público nos meios digitais.** O jogo segue em aberto. E esta cartilha é uma ferramenta para jogá-lo melhor.

Para tal, dividimos a cartilha em seis módulos temáticos, finalizados com algumas dicas sobre o que fazer. O primeiro módulo traz um apanhado sobre estudos recentes que mapeiam **como a população brasileira se informa**, traçando diferenças significativas por gênero, classe, raça e escolaridade na maneira como se navega pelo ecossistema informacional atual. Já o segundo módulo **desmonta a máquina**, ao examinar como diferentes estratégias de comunicação política emergiram de plataforma a plataforma, explorando a arquitetura, os algoritmos e os incentivos de engajamento específicos de cada ambiente digital. Mapeia os tipos de desinformação, o ciclo de produção e o financiamento opaco que sustenta esse sistema ao expor a engrenagem de amplificação: a disseminação de narrativas é coordenada em cascata, do Telegram ao WhatsApp às plataformas mainstream, criando um efeito bola de neve que transforma conteúdo fabricado em pauta da mídia tradicional. O módulo também documenta a opacidade das Big Techs no Brasil, seja nas práticas de moderação de conteúdo e de lobby no Congresso Nacional, e expõe a censura algorítmica sistemática que silencia corpos dissidentes. Termina com orientações para monitorar, reagir e construir presença estratégica nas redes.

O terceiro módulo, **psicologia da persuasão radical**, parte de uma pergunta central: por que a desinformação funciona? A resposta está na psicologia, não nos fatos. O módulo desmonta oito gatilhos emocionais geralmente explorados de forma maliciosa, aponta os erros mais comuns no exercício da contra-narrativa e oferece um caminho prático: verificar antes de responder, monitorar antes de reagir, comunicar com emoção real em vez de medo fabricado. O quarto módulo, **violência política de gênero como tática de silenciamento**, trata de um dado estrutural: a violência política digital não é efeito colateral da desinformação, é uma tática deliberada de silenciamento. O módulo classifica sete tipos de violência política habilitada por tecnologias (de ameaças e *doxxing* ao *dogpiling* e invasão de contas), apresenta um protocolo de documentação em nove passos e oferece orientações de higiene digital básica. Termina com a indicação da Maria d'Ajuda, uma linha de atenção e acompanhamento feminista de resposta a incidentes e emergências em segurança digital, bem como a indicação de outras redes de apoio para quem enfrenta ataques.

O quinto módulo, **IA e eleições**, apresenta os desafios que os usos da IA generativa representam para o contexto eleitoral: como alguns de seus usos tem ampliando desinformação e ameaçado processos democráticos? Como lidar com deepfakes? Quais as regras atuais da Justiça Eleitoral sobre usos de IA? Onde denunciar? Como fomentar visão crítica para saber escolher e configurar algumas destas ferramentas para usá-las com cautela na produção de conteúdos, acessibilidade e monitoramento de narrativas?

Por fim, o módulo **comunicação progressista eficaz** mostra como o campo pode disputar o espaço público com estratégia própria, sem ser pautado por outros e sem se deixar levar pela ideia de uma credibilidade geralmente apoiada em figuras de autoridade, algo que tende a ser repelido como um didatismo autoritário e arrogante. O módulo parte de uma premissa central, nosso campo tem o que a desinformação não tem: solidariedade real, causa verificável e abrangente (pois cabe vida com dignidade para todo mundo no nosso projeto) e experiências de construção de confiança e poder coletivo. Do outro lado, sobram o ódio a amplos grupos sociais, um projeto excludente feito de bilionários para bilionários e uma credibilidade forjada por eles próprios com dinheiro e poder para comprar ou manipular os meios de difusão de informações. Compreendendo o cenário e nossas potências, este módulo oferece seis princípios de comunicação progressista eficaz: pautar e não apenas reagir, enraizamento comunitário, emoção legítima, simplicidade sem simplismo, consistência no tempo e representação como mensagem. A teoria do framing, como o enquadramento define qual batalha está sendo travada antes mesmo do debate começar, é traduzida em exemplos concretos de reenquadramento para os temas mais atacados pela extrema direita. O módulo ainda orienta sobre escolha de formatos adequados a cada pessoa e realidade, com um guia de cinco passos para testar, medir e construir uma rotina sustentável de produção de conteúdo e estratégias de disseminação que não dependem só do algoritmo.

Comunicação progressista eficaz não é a negação da comunicação da extrema direita. É algo diferente: mais enraizada em comunidades, mais consistente no tempo, mais honesta sobre emoções, inclusive não insistindo apenas na racionalidade, mas também dando abertura para colocar sobre a mesa outros afetos para canalizar insatisfações, raiva e frustração, que são sentimentos legítimos e devem achar seu caminho de expressão. Esperamos **que esta cartilha seja um lampejo de esperança e tática para a comunicação de quem** defende direitos humanos, justiça socioambiental e sonha, nas suas ações e ativismos do dia a dia, com um mundo mais justo, onde se possa viver com dignidade e em paz.

DESMONTANDO A MÁQUINA:

Módulo 1

COMO A POPULAÇÃO BRASILEIRA SE INFORMA?

Seção 1 Desigualdades informativas e polarização política

A pesquisa Desigualdades Informativas (Aláfia Lab, 2025), terceira edição de uma série histórica desde 2023, entrevistou uma amostra de 1512 pessoas para traçar um retrato atualizado do ecossistema informativo brasileiro, revelando onde as batalhas pela narrativa são travadas:

· **Redes sociais** lideram como principal fonte de informação geral (53,5%), seguidas pela televisão (52,5%) e portais de notícias (39,7%)

· **Aplicativos de mensagens** cresceram de 21,5% (2024) para 28% (2025), aumento de 30% em um ano

· **Ferramentas de IA** (9,7%) e newsletters (12,2%) já superam revistas (5,8%) e jornais impressos (9,5%)

· **As redes estão "menos sociais":** cai o acompanhamento de amigos e família, crescem algoritmos, influenciadores e perfis de humor

Outro dado estratégico sobre envio de mensagens: aplicativos de mensagens crescem conforme a escolaridade aumenta na direita e diminuem conforme aumenta na esquerda. Padrão inverso que revela ecossistemas opostos: quanto mais escolarizada a pessoa de direita, mais usa canais fechados e menos auditáveis.

O estudo do Aláfia Lab também investigou como gênero, idade, classe econômica e escolaridade influenciam hábitos informacionais de pessoas à direita e à esquerda no ambiente digital. Os achados revelam padrões que ajudam a entender por que a desinformação encontra terreno fértil entre determinados grupos. Segundo o relatório:

Por posição política: direita e extrema direita concentram seu consumo em Instagram, WhatsApp e YouTube – plataformas fechadas e menos auditáveis. Na direita, o consumo prioritário de informação política tem as redes sociais como local dominante em todas as classes, idades e níveis de escolaridade, acima de 52% em cada estrato. Esquerda e centro apresentam um ecossistema mais diversificado: TV, portais e redes sociais disputam espaço. Entre os 45+, a TV lidera com 53%. Enquanto G1 e O Globo lideram entre esquerda e centro, Record, R7, SBT, Band e Jovem Pan são majoritários entre direita e extrema direita.

Por raça: ainda segundo a pesquisa, entre pessoas negras, o consumo de portais de notícias cai drasticamente quando o assunto é política (de 37,3% para 25,1%), enquanto a TV recua de 53,2% para 44,3%. A lacuna de acesso ao jornalismo verificado é um gargalo de proteção contra desinformação.

Por renda: o uso de IA para se informar é 2x maior entre quem ganha acima de 10 salários mínimos (17,2%) do que entre quem recebe até 1 salário (7,9%).

Por idade: o X (Twitter) desabou entre jovens de 18 a 24 anos (de 15,7% para 5%), enquanto o TikTok se tornou a segunda principal fonte de informação nessa faixa (14,9%), uma mudança estrutural no ecossistema.

Por gênero: o Instagram é a principal fonte de informação para mulheres (49,3%) e homens (33%), mas com perfis diferentes: mulheres seguem mais perfis que misturam informação e entretenimento; homens, mais veículos jornalísticos tradicionais. No consumo de informação política, mulheres de esquerda dependem mais das redes (49%) que homens de esquerda (37%), que preferem TV (47%) e sites de notícias (41%). Homens de direita usam mais aplicativos de mensagens para informação política; mulheres de direita são as que mais consomem redes sociais e as que menos buscam informação política na televisão

Por escolaridade: quem tem ensino superior acessa portais de notícias em 53,5% dos casos; com ensino fundamental, apenas 28,1%. Na esquerda, o uso de sites jornalísticos vai de 19% (baixa escolaridade) a 52% (alta); na direita, de 20% a 42%, mas as redes seguem dominantes em todos os níveis.

A pesquisa confirma que o tabuleiro informativo é desigual antes mesmo da desinformação entrar em campo. Estratégias de comunicação progressista precisam considerar esses pontos de partida diferentes. (Aláfia Lab. Desigualdades Informativas, 2025)

Seção 2 Os tipos de desinformação

Desinformação

Conteúdo falso criado deliberadamente para enganar.

Ex.: *gráfico com estatística inventada sobre criminalidade.*

Mal-informação

Informação verdadeira usada fora de contexto para prejudicar.

Ex.: *foto real de protesto de 2013 apresentada como "de ontem".*

Misinformação

Conteúdo falso compartilhado sem intenção de enganar.

Ex.: *avô que reencaminha corrente de WhatsApp sem verificar.*

Conteúdo sintético (IA)

Imagem, áudio ou vídeo gerado por IA para parecer real.

Ex.: *deepfake de político ao lado de criminosos. Deepnude de candidatas.*

Observação: como muitos dos estudos que produzem dados sobre essas práticas não seguem esta tipologia, nesta cartilha vamos utilizar o termo desinformação para se referir tanto às táticas de desinformação, misinformação ou mal-informação.

Seção 3 Desinformação e democracia

2 em 3 **conteúdos políticos com IA** circularam sem identificação nas redes entre dez/2025 e fev/2026, violando regras do TSE.
Amostra: 137 casos monitorados.
Fonte: Observatório IA nas Eleições / Aláfia Lab, abril 2026

23 milhões é a soma de inscritos em 123 **canais do YouTube classificados como misóginos**, muitos deles foram denunciados, mas seguem ativos e com audiência crescente, sem remoção efetiva pelas plataformas.
Fonte: NetLab UFRJ, Misoginia no YouTube, 2025

4.321 anúncios fraudulentos nas plataformas da Meta divulgando um mesmo golpe que usa IA para retratar políticos divulgando uma suposta indenização a ser paga pelo Serasa para enganar consumidores. **Fonte:** Anúncios com IA usam imagens de políticos brasileiros para aplicar golpes. Netlab/UFRJ, março a maio de 2024.

Desinformação no WhatsApp e Telegram

O crescimento dos apps de mensagens como fonte de informação

1 em cada 3 brasileiros já usa WhatsApp ou Telegram como fonte habitual de informação. Nesses ambientes, o conteúdo circula sem curadoria editorial, sem verificação, e muitas vezes sem qualquer identificação de origem: ambiente ideal para a desinformação coordenada funcionar. O crescimento no uso de mensageria é o crescimento do vetor mais difícil de monitorar e mais eficaz para circular desinformação.

O relatório Abaixo do Radar (Aláfia Lab, coLab e Instituto Democracia em Xeque, 2025) é o estudo mais abrangente já realizado sobre circulação de desinformação em aplicativos de mensagens nas eleições brasileiras. No estudo, foram monitorados 35 grupos de WhatsApp e 22 grupos de Telegram com perfil de extrema direita durante todo o período oficial de campanha das eleições municipais de 2024, de 2 de setembro a 27 de outubro.

35 GRUPOS + 22 GRUPOS } Período oficial de campanha das eleições municipais de 2024

Abaixo do radar: o que o estudo revelou

As taxas de desinformação por tema no WhatsApp

40% dos links sobre o **STF** continham desinformação, a maior taxa registrada

28% dos links sobre **economia** continham desinformação

23% dos links sobre **política internacional e nacional** continham desinformação

Padrões de desinformação identificados

- Textos opinativos apresentados como matérias factuais, sem sinalização do gênero jornalístico
- Desinformação concentrada nos títulos: o texto da matéria era verdadeiro, a manchete mentia
- Ataques ao STF (23,6% dos links) foram o segundo tema mais presente no Telegram, contra apenas 5,2% no WhatsApp
- No Telegram circularam temas como ataques às universidades e "ideologia de gênero", ausentes nos grupos de WhatsApp monitorados.

A plataforma tem sido um dos principais palcos e ferramenta de mobilização nas chamadas guerras culturais e disputas de narrativas ideológicas.

Mesmo durante eleições municipais, que deveriam focar temas locais, a política nacional dominou o debate nos dois aplicativos, revelando como a polarização nacional penetra e contamina os pleitos territoriais. (Aláfia Lab, coLab e Instituto Democracia em Xeque. Abaixo do Radar, 2025)

O QUE FAZER | Protocolo de verificação em 10 passos

1. A **fonte** é um veículo jornalístico reconhecido?
2. O link abre para **site real**, não para uma cópia ou domínio suspeito?
3. Quem assina a matéria são pessoas reais?
4. O fato aparece em pelo menos **outro veículo** independente?
5. Li o artigo inteiro, a manchete **corresponde** ao texto?
6. O **dado** tem período, metodologia e fonte original?
7. Fiz a **checagem da informação** em agências que checam fatos, como Aos Fatos, Agência Lupa ou E-Farsas?
8. Há sinais de **IA**? (Proporções distorcidas, erros em textos, indícios de manipulação de conteúdo audiovisual)
9. Em caso de foto ou prints, **consigo procurar** a imagem na internet?
10. **Esperei** pelo menos 20 minutos antes de compartilhar?

DESMONTANDO A MÁQUINA:

Módulo 2

A MÁQUINA: COMO AS PLATAFORMAS FORAM PENSADAS PARA FAVORECER CONTEÚDOS RADICALIZADOS

Seção 1 Mecanismos estruturais das redes sociais

As **plataformas digitais** não são neutras. Foram **desenhadas para maximizar tempo de tela e o engajamento** e o **conteúdo que mais gera engajamento é aquele que aciona emoções fortes:** raiva, medo, indignação, escândalo.

Algoritmo de engajamento - Conteúdo que provoca reação emocional intensa e retém a atenção dos usuários é sistematicamente distribuído para mais pessoas, independente de ser verdadeiro ou não.

Monetização da polêmica - Sites falsos lucram com publicidade programática. Quanto mais viral o conteúdo falso, maior a receita. Sleeping Giants Brasil documenta o mecanismo de "dark pooling", pelo qual portais bloqueados continuam sendo monetizados. Além disso, plataformas também lucram com publicidade atrelada a outros conteúdos nocivos. Por exemplo, no caso do YouTube e conteúdos misóginos. Em todos esses casos, a renda de publicidade é sempre dividida entre plataformas e canais.

Microssegmentação - Ferramentas de impulsionamento de publicidade permitem direcionar desinformação para grupos específicos sem que os outros vejam.

Moderação deficiente - segundo estudo do NetLab UFRJ: aproximadamente 90% dos canais misóginos mapeados no YouTube em 2024 continuam ativos e com audiência crescente.

IMPLICAÇÃO ESTRATÉGICA | Saber que o tabuleiro foi construído assim não é motivo para paralisarmos, é o ponto de partida para uma estratégia mais inteligente. A extrema direita não é mais criativa: ela apenas aprendeu antes a jogar esse jogo específico.

Seção 2 O ciclo de produção e amplificação

Conhecendo os mecanismos estruturais das redes, podemos identificar que a produção de desinformação segue um roteiro previsível em cinco etapas:

1. **Criação:** sites falsos, perfis anônimos, canais pagos
2. **Seeding:** publicação inicial coordenada para parecer orgânico
3. **Amplificação:** bots e grupos de disparo simultâneo
4. **Mainstream:** mídia legítima ou influenciadores "reportam", mesmo que para desmentir
5. **Negação produtiva:** o desmentido vira nova notícia e amplifica a mentira original

Seção 3 A disseminação em cascata de acordo com os mecanismos de cada plataforma

A disseminação de desinformação (seeding e amplificação) não é espontânea nem descoordenada, é **arquitetada em camadas**. Pesquisadores descrevem esse sistema como um **ecossistema multiplataforma** estruturado em três níveis que se retroalimentam.

As três camadas do ecossistema

Nível 1 | Laboratório: Telegram como espaço de teste

O Telegram funciona como ambiente de produção e teste de narrativas. A plataforma serve como laboratório para testar quais são as palavras-chave e as narrativas que geram maior engajamento. Muitos dos memes e das palavras que se popularizam são testados ali para que se veja quais conteúdos podem pular para outras plataformas.

O Telegram atrai quem dissemina desinformação por características técnicas específicas:

Grupos com até 200 mil pessoas - no WhatsApp o limite é 1.024 pessoas
Canais com audiência ilimitada - onde só o administrador posta
Anonimato reforçado - é possível enviar mensagens sem expor número de telefone
Ausência de moderação efetiva - conteúdos banidos em outras plataformas encontram refúgio ali
Arquivamento permanente - uploads de arquivos grandes para acesso e compartilhamento contínuos

Nível 2 | Amplificação: WhatsApp como vetor de dispersão

O conteúdo testado no Telegram migra para o WhatsApp via grupos de disparo coordenado, criando a aparência de que múltiplas fontes independentes estão reportando a mesma informação simultaneamente, quando na verdade é coordenado.

O WhatsApp opera como uma "dobradiça" digital entre dois tipos de público:

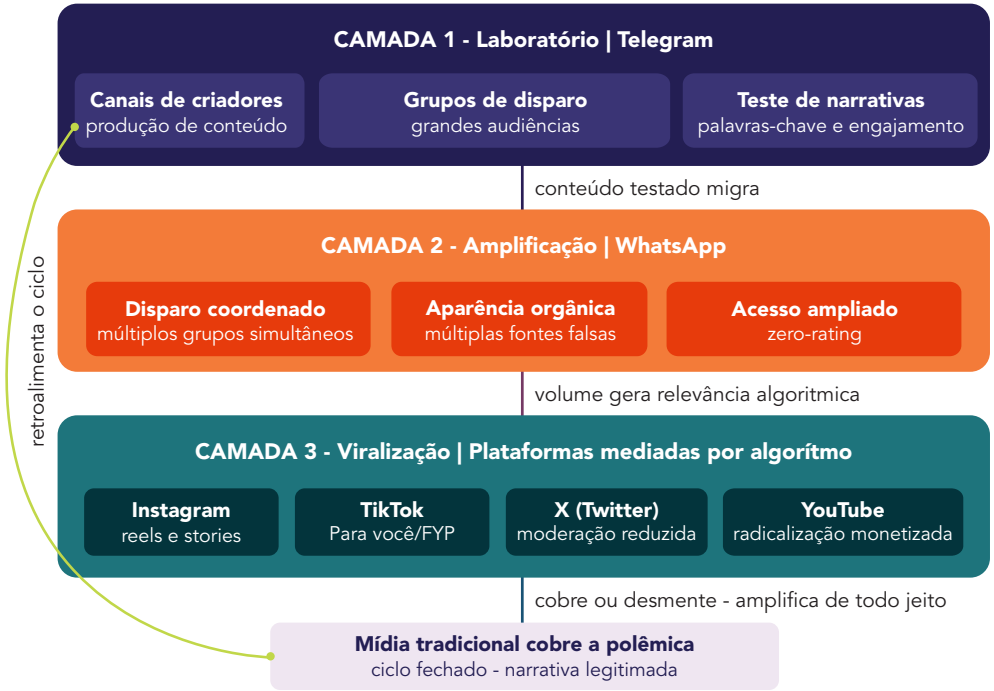
Públicos refratados: circulam em ambientes fechados e opacos, com efemeridade deliberada de postagens
Públicos dominantes: presentes nas plataformas mainstream (Instagram, TikTok, X)

No Telegram, os canais de maior relevância para a interconexão entre comunidades estão associados a influenciadores ativos em outras plataformas digitais, o que explica a velocidade com que narrativas migram de ambientes fechados para o debate público. O Telegram articula os dois mundos, permitindo que narrativas produzidas em ambientes restritos alcancem audiências massivas no WhatsApp e nas redes sociais sem que a origem seja rastreável.

Nível 3 | Viralização mainstream: algoritmo como amplificador involuntárioteste

O conteúdo, já com volume de interações suficiente, é captado pelo algoritmo de plataformas como Instagram, TikTok e X, que o distribuem organicamente para audiências que nunca tiveram contato com a origem da narrativa.
A mídia tradicional frequentemente fecha o ciclo ao reportar a polêmica, mesmo que para desmentir, ampliando ainda mais o alcance e legitimando a narrativa ao introduzi-la em novos círculos de audiência.

EFEITO BOLA DE NEVE EM REDES ARTICULADAS



Seção 4 Redes sociais como ferramenta: dados de desinformação plataforma a plataforma

As ferramentas do mestre jamais vão desmantelar a casa grande, como escreveu Audre Lorde. Sabemos que as redes sociais não são espaços criados por nós e para nós, e não são apenas o ambiente onde a desinformação circula: são a infraestrutura ativa que a amplifica. Mas é exatamente por isso que precisamos compreender a lógica de cada plataforma. A extrema direita aprendeu a usar cada uma de forma específica.

Instagram

DADOS DE DESINFORMAÇÃO

O Instagram concentrou 55% dos casos de conteúdo político com IA identificados pelo Observatório IA nas Eleições entre dezembro de 2025 e fevereiro de 2026, em uma base de 137 casos monitorados.
Fonte: Observatório IA nas Eleições, abril de 2026, disponível em: https://observatorioianaseleicoes.com.br/

PERFIL DE USO E PÚBLICO

Conta com aproximadamente 147 milhões de usuários no Brasil em 2025. É a rede com maior presença feminina proporcional entre as grandes plataformas: cerca de 57% do público são mulheres. A faixa etária predominante é de 18 a 34 anos. Nos últimos anos, a rede tem feito um giro para priorizar conteúdos em vídeo (reels). Também é um dos espaços mais afetados pela produção a partir de conteúdo sintético.
Fontes: Digital 2026 Brazil e Digital 2025 Global Overview Report, DataReportal/We Are Social & Meltwater. Disponíveis em https://datareportal.com/reports/digital-2026-brazil e https://wearesocial.com/wp-content/uploads/2025/02/GDR-2025-v2.pdf

PRINCIPAIS DIFERENCIAIS

Oferece ferramentas de impulsionamento que permitem uma micro segmentação demográfica precisa, a partir de dados da Meta (que controla Instagram, Facebook e WhatsApp). É uma das mais completas em possibilidades de postagens e combinação de espaços abertos (feed) e fechados (inbox).

TikTok

DADOS DE DESINFORMAÇÃO

O mesmo estudo do Observatório IA nas Eleições apontou que o TikTok concentrou 20% dos conteúdos sintéticos políticos da base de casos analisados. O algoritmo FYP (For You Page) distribui conteúdo para não-seguidores com base em comportamento de visualização, tornando a radicalização mais rápida e imprevisível.

PERFIL DE USO E PÚBLICO

Possui cerca de 131 milhões de usuários no Brasil. O público feminino é de 41,4%, enquanto o masculino de 58,6% extremamente jovem, embora a faixa acima dos 30 anos seja a que mais cresce, tendência observada em relatórios dos últimos dois anos.
Fonte: Relatório Digital 2026, 2025 e 2024 da We Are Social & Meltwater https://datareportal.com/reports/

PRINCIPAIS DIFERENCIAIS

É, provavelmente, o algoritmo mais refinado de todas as redes. A entrega de conteúdo é baseada em interesses e não em conexões sociais, o que facilita a viralização orgânica e a descoberta por novos públicos.

WhatsApp

DADOS DE DESINFORMAÇÃO

Vetor mais perigoso no Brasil: zero-rating (uso do app não desconta dados do plano de celular) amplia o alcance entre populações de menor renda; grupos de disparo coordenado criam aparência de múltiplas fontes.

PERFIL DE USO E PÚBLICO

É a rede mais utilizada no país, com 206 milhões de usuários ativos (quase a totalidade da população brasileira, cerca de 213 milhões). É a principal porta de entrada para a internet em classes mais baixas pelo zero-rating.
Fontes: Universidade de Viena & SBA Research, 2025 (https://arxiv.org/html/2511.20252v1); IBGE, Estimativas da População 2025 (Portaria nº 1.098/2025).

PRINCIPAIS DIFERENCIAIS

Comunicação direta e privada, o que gera alta percepção de confiança. A dinâmica de grupos permite que uma mesma informação pareça vir de várias fontes de confiança ao mesmo tempo.

Telegram

DADOS DE DESINFORMAÇÃO

No Telegram, ataques ao STF se destacaram como o segundo tema mais presente entre os links monitorados (23,6% do total), proporção significativamente maior do que no WhatsApp, onde o mesmo tema representou 5,2% dos links. Esse dado é proveniente do monitoramento de 22 grupos públicos do Telegram com perfil de extrema direita, realizado pelo Aláfia Lab durante o período oficial de campanha das eleições municipais de 2024 (2 de setembro a 27 de outubro). A amostra foi de 181 links de mídia hiperpartidária, com intervalo de confiança de 95% e margem de erro de 5%. Grupos públicos podem ser monitorados por pesquisadores; canais privados permanecem fora do alcance de qualquer auditoria externa. *Fonte: Aláfia Lab, coLab e Instituto Democracia em Xeque. 2025.*

PERFIL DE USO E PÚBLICO

Estima-se que esteja presente em 60% dos smartphones no Brasil (aproximadamente 80 milhões de usuários). O público é predominantemente masculino (cerca de 60%) e mais engajado em debates de nicho, política e tecnologia.
Fonte: Panorama Mobile Time/Opinion Box de 2022

PRINCIPAIS DIFERENCIAIS

Grupos de até 200 mil pessoas e canais de transmissão ilimitados. Por isso funciona tão bem para testar narrativas que depois migram para outras redes.

X (Twitter)

DADOS DE DESINFORMAÇÃO

No X (antigo Twitter), pesquisadores de plataformas digitais documentam redução significativa da moderação de conteúdo desde a aquisição por Elon Musk em 2022. Em 2025, durante a polêmica da liquidação do Banco Master, o NetLab UFRJ identificou mais de 80 mil publicações nas redes sociais com ataques ao Banco Central em um período de 30 dias, em coordenação com a contratação de influenciadores que questionaram a decisão.
Fonte: NetLab UFRJ, 2025; cobertura em Jornal Hoje, TV Globo, 2025.

PERFIL DE USO E PÚBLICO

Possui cerca de 17.1 milhões de usuários ativos no Brasil. É a plataforma com maior disparidade de gênero: aproximadamente 60% do público são homens. Concentra formadores de opinião, jornalistas e políticos.
Fonte: Relatório Digital 2026 da We Are Social & Meltwater https://datareportal.com/reports/digital-2026-brazil

PRINCIPAIS DIFERENCIAIS

Imediatismo e circulação de notícias em tempo real. A redução na moderação de conteúdo desde a compra da plataforma por Elon Musk tem facilitado ataques coordenados e a propagação de narrativas de urgência, além do cometimento de crimes e a proliferação de bots.

Youtube

DADOS DE DESINFORMAÇÃO

O YouTube tem sido apontado como um dos principais vetores de radicalização algorítmica: seu sistema de recomendação tende a direcionar usuários a conteúdos progressivamente mais extremos à medida que o tempo de visualização aumenta. O NetLab/UFRJ mapeou 137 canais do YouTube com conteúdo misógino em 2024. Em atualização publicada em março de 2026, identificou que 123 deles, aproximadamente 90%, seguem disponíveis na plataforma e, juntos, acumularam mais de 3,6 milhões de novos inscritos no período. O estudo apontou que 80% dos canais com conteúdo misógino analisados tinham ao menos um recurso de monetização ativo, sendo os recursos nativos do YouTube as principais formas utilizadas. A pesquisa dedica uma seção às demais estratégias de monetização desses canais. Recomendamos a leitura. *Fonte: NetLab UFRJ e Ministério das Mulheres, 2024 e 2026.*

Produtores de desinformação também são financiados nessa plataforma por meio de monetização via publicidade programática. Pior: a Sleeping Giants Brasil denunciou que mesmo portais bloqueados por anunciantes continuam recebendo receita por meio do mecanismo de dark pooling documentado pela organização.
Fonte: Sleeping Giants Brasil, 2026.

PERFIL DE USO E PÚBLICO

Alcança 150 milhões de brasileiros. O público é levemente mais feminino (52% mulheres). É a rede preferida para conteúdos de longa duração e tutoriais. *Fonte: Relatório Digital 2026 da We Are Social & Meltwater*
https://datareportal.com/reports/digital-2026-brazil

PRINCIPAIS DIFERENCIAIS

Motor de busca potente (o segundo maior do mundo) e possui um sistema de recomendação que direciona usuários a conteúdos progressivamente mais extremos para aumentar o tempo de visualização. Além disso, a monetização via publicidade programática permite que produtores de desinformação continuem lucrando mesmo após bloqueios.

O que é dark pooling: é uma prática fraudulenta que ocorre no universo da publicidade digital quando sites compartilham uma identificação que deveria ser individual. Esse comportamento tem como objetivo inflar números para que o processo automatizado que direciona publicidade recomende aquele site, que, ao ter a identificação compartilhada, também divide receita com os demais sites do grupo. Isso engana anunciantes sobre o espaço real em que seus anúncios serão exibidos e sua qualidade e reputação. Assim, um site que publica conteúdo falso ou extremista consegue receber dinheiro de publicidade de grandes marcas sem que essas marcas saibam.

Como funciona na prática: sites desinformativos são "empacotados" por casas de venda de inventário de publicidade junto com veículos jornalísticos reconhecidos e se disfarçam de portais legítimos dentro do sistema automatizado de publicidade. A partir então desse pacote ou lote de veículos comercializados por uma mesma casa de revenda ("pool"), todos passam a usar um mesmo número de identificação publicitária (Seller ID). Desta forma, mesmo quando alguns sites são denunciados por anunciantes e tem seu Seller ID bloqueado, continuam recebendo repasses financeiros usando essa brecha, pois ainda possuem um Seller ID compartilhado com esse pool. Como o Sleeping Giants Brasil descreve: *"a porta da frente está trancada, mas os anúncios continuam entrando pela porta dos fundos."*

Por que isso importa: a desinformação é um negócio lucrativo. Quanto mais viral o conteúdo falso, maior a receita publicitária, ou seja, o próprio sistema financeiro da internet incentiva a produção e circulação de mentiras. Sem cortar o financiamento, não se corta a produção. *Fonte: Sleeping Giants Brasil, 2026.*

Data brokers e microtargeting político: em 2018, a Coding Rights investigou como empresas especializadas em dados, chamadas data brokers, combinam dados públicos do IBGE com dados de operadoras de telefonia e empresas de cadastro de crédito para construir perfis detalhados de pessoas eleitoras. Números de WhatsApp e perfis socioeconômicos são usados em microtargeting político sem autorização das pessoas. Uma das empresas investigadas na ocasião, a Ponte Estratégia, era parceira da Cambridge Analytica, empresa foco de um escândalo sem precedentes. A empresa utilizou dados de 87 milhões de usuários do Facebook, coletados sem consentimento, para construir perfis psicológicos e direcionar propaganda eleitoral personalizada na campanha de Donald Trump e no referendo do Brexit. O caso inaugurou uma nova era de manipulação eleitoral em escala industrial. Essa indústria tem seu modelo de negócios baseado no desrespeito à proteção de dados pessoais.
Fonte: Coding Rights, 2018.

Seção 5 Brasil na lanterna da transparência

O relatório Data Not Found (NetLab UFRJ e Minderoo Centre for Technology and Democracy, Universidade de Cambridge, 2026) apresenta o primeiro índice sistemático de transparência de dados de redes sociais, avaliando 15 plataformas em três contextos regulatórios: União Europeia, Reino Unido e Brasil. A pesquisa mostra que o Brasil tem acesso muito mais limitado a dados de plataformas do que a UE e o Reino Unido. O Google impede buscas por palavras-chave em sua biblioteca de anúncios no Brasil. O X/Twitter bloqueia pesquisadores brasileiros de ferramentas disponíveis em outras regiões. Isso não é falta de capacidade técnica, é assimetria deliberada.

> **A transparência ainda é a exceção, não a regra. O Brasil registra consistentemente os níveis mais baixos de acesso a dados entre as três regiões avaliadas e as práticas das plataformas aqui dependem, em grande medida, de sua própria boa vontade.** *Fonte: NetLab UFRJ et al., 2026.*

O que o índice avaliou e o que encontrou

As 15 plataformas avaliadas foram: YouTube, Bluesky, X/Twitter, Telegram, Reddit, TikTok, LinkedIn, Facebook, Instagram, Discord, Kwai, Pinterest, Snapchat, Threads e WhatsApp. Cada uma foi pontuada de 0 a 100 em duas dimensões: acesso a conteúdo gerado por usuários e acesso a dados de publicidade, conforme abaixo:

CONTEÚDO DE USUÁRIOS

- Apenas YouTube e Bluesky oferecem APIs públicas classificadas como "Significativas" e de forma consistente nas três regiões analisadas, sem diferenciar o acesso por localização.
- Facebook e Instagram permitem acessar conteúdos públicos apenas via Meta Content Library, em ambientes controlados, sem permitir extração de dados desagregados para infraestruturas de pesquisa.
- X (antigo Twitter) oferece acesso a dados públicos, mas com estruturas de preços proibitivas. Seu repositório de anúncios, disponível apenas na União Europeia, não retorna resultados.
- TikTok e LinkedIn oferecem APIs para pesquisa no âmbito da regulação europeia (DSA, Lei de Serviços Digitais), mas Pinterest e Snapchat só aceitam solicitações manuais, modelo insuficiente para pesquisa sistemática.
- WhatsApp e Threads pontuam próximo de zero nesse aspecto em todas as regiões.

PUBLICIDADE

- Metade das maiores plataformas digitais não atende aos padrões mínimos de transparência em dados de publicidade.
- As ferramentas de acesso a dados de anúncios são excessivamente limitadas: plataformas como Meta, LinkedIn e TikTok permitem buscas por palavras-chave, mas X, YouTube e Pinterest só permitem busca por nome de anunciante, impedindo descoberta ampla.
- Dados de alcance e segmentação são divulgados apenas em faixas amplas de valores, impossibilitando auditoria real das campanhas.
- O Google impede buscas por palavras-chave em sua biblioteca de anúncios no Brasil. A Meta e o TikTok oferecem ferramentas que permitem esse tipo de pesquisa em outras regiões, mas as restringem no Brasil.

Por que o Brasil está na lanterna

O relatório explica a assimetria com precisão: a União Europeia tem o Digital Services Act (DSA), legislação que obriga plataformas a abrirem dados para pesquisa. O Reino Unido se beneficia indiretamente do chamado "Efeito Bruxelas": as plataformas adaptam suas práticas na região à regulação europeia. O Brasil não tem marco regulatório equivalente.
A transparência aqui depende da boa vontade das empresas. E as empresas escolhem não ser transparentes.

O QUE ISSO SIGNIFICA PARA A DESINFORMAÇÃO ELEITORAL

Sem acesso a dados de conteúdo e publicidade, pesquisadores e jornalistas brasileiros não conseguem:

- **Mapear** como narrativas de desinformação circulam e se amplificam nas plataformas
- **Identificar** quem financia anúncios eleitorais e para quem são direcionados
- **Auditar** de forma independente se as plataformas estão cumprindo as regras do TSE
- **Monitorar** campanhas de assédio coordenado contra candidaturas específicas

Isso não é falta de capacidade técnica: é uma assimetria deliberada que protege as plataformas e expõe a democracia.

O QUE O CAMPO DEMOCRÁTICO PODE EXIGIR

O relatório Data Not Found faz recomendações concretas que também servem como agenda de pressão:

- Acesso programático (via API), gratuito e aberto a dados de publicidade e conteúdo público em todas as regiões.
- Fim das práticas seletivas de transparência por jurisdição: o padrão mais aberto deve ser aplicado globalmente.
- Acesso a dados de conteúdo moderado e removido, essencial para entender como a desinformação é (ou não é) combatida.
- Para o Brasil especificamente: consolidar os princípios legais existentes em um marco regulatório dedicado à transparência de dados de plataformas, no mínimo equivalente ao Digital Services Act (DSA) europeu.

O QUE FAZER | MONITORAR, REAGIR E CONSTRUIR

Monitorar antes de reagir
Identificar narrativas em circulação no Telegram antes que cheguem ao mainstream dá tempo para preparar contra-narrativa proativa.

Não amplificar para desmentir
Cada resposta pública alimenta o ciclo. Usar o protocolo de contra-narrativa do Módulo 3 (quando e como responder).

Construir redes orgânicas próprias
Nosso campo precisa mais do que de uma boa estratégia de comunicação. Precisa também de canais orgânicos e uma rede de criadores que distribua os conteúdos.

Documentar padrões de coordenação
Quando o mesmo texto aparece em múltiplos grupos simultaneamente, é sinal de disparo coordenado. Em caso de conteúdo com IA, é interessante registrar e encaminhar ao Observatório IA nas Eleições.

Romper a bolha do próprio campo
É preciso estar onde outros campos políticos também estão, com estratégia e linguagem adequadas a cada plataforma. O diálogo é o único caminho possível para diminuir polarização.

Ocupar canais subutilizados
Para alcançar quem já está em outro tabuleiro, é preciso presença consistente em canais que nosso campo ainda subutiliza: YouTube (forte na direita de baixa escolaridade) e WhatsApp (canal dominante em todas as classes à direita).

DESMONTANDO A MÁQUINA:

Módulo 3

PSICOLOGIA DA PERSUASÃO RADICAL

Seção 1 Os 8 gatilhos emocionais explorados

Por que pessoas informadas e bem-intencionadas acreditam e compartilham desinformação? A resposta não está na falta de educação, mas na psicologia humana. Quem usa desinformação como estratégia compreendeu isso e otimizou sua comunicação para explorar mecanismos cognitivos que todos nós temos.

O Desinformante documenta que a **disseminação de conteúdos nas redes está diretamente ligada à mobilização de afetos, não de fatos.** A raiva, o medo, a indignação e o sentimento de pertencimento são os verdadeiros motores do compartilhamento. A seguir listamos algumas categorias comuns de gatilhos emocionais:

Medo

Ativa medo existencial da família, identidade e segurança, amplificado automaticamente pelo algoritmo.

Ex.: *narrativas de "invasão cultural", "ameaça a crianças", "ameaça à propriedade".*

Raiva

A emoção que mais leva ao compartilhamento, conteúdo indignante gera mais interações do que conteúdo positivo.

Ex.: *escândalo político real ou fabricado, suposta "injustiça".*

Pertencimento

Compartilhar se torna ato de lealdade ao grupo. Não compartilhar é traição.

Ex.: *"Compartilhe se você é patriota". "Quem não vê isso está com eles".*

Urgência

Urgência artificial desativa o pensamento crítico antes que ele funcione.

Ex.: URGENTE, COMPARTILHE AGORA, "vão censurar isso".

Identidade ameaçada

Narrativas que parecem atacar a identidade da pessoa ativam reação de defesa.

Ex.: *"Eles querem destruir nossa família", "atacam nossa fé".*

Autoridade falsa

Uso de linguagem técnica, títulos ou cenário de telejornal para criar aparência de credibilidade.

Ex.: *canais do YouTube com cenário de noticiário falso.*

Conspiração

Cria a sensação de acesso a uma verdade oculta, posicionando o público como parte de um grupo seleto que "enxerga além". Gera engajamento ao transformar o compartilhamento em revelação.

Ex.: *"Eles não querem que você saiba disso", "Estão escondendo isso de você".*

Falsa imparcialidade/O isentão

Adota uma postura de neutralidade aparente para ganhar credibilidade, apresentando-se como quem "só faz perguntas" ou "mostra os dois lados", enquanto direciona sutilmente para uma conclusão.

Ex.: *"Não sou de esquerda nem de direita, só quero a verdade", "Tô só perguntando...".*

Seção 2 Os 4 erros mais comuns na contra-narrativa

ERRO 1
Começar pela **negação**

NÃO FAÇA
"Não é verdade que eu disse X."

FAÇA
"O que eu disse foi Y, e fiz por estas razões."

ERRO 2
Repetir a **acusação**

NÃO FAÇA
"A mentira de que eu roubei Z é absurda."

FAÇA
"Meu histórico de gestão é este: [dados concretos]."

ERRO 3
Vocabulário **do adversário**

NÃO FAÇA
Usar termos como "kit gay", mesmo para criticar.

FAÇA
"desinformação", "ataque coordenado", seus próprios termos.

ERRO 4
Responder **só com texto**

NÃO FAÇA
Thread longa explicando o contexto completo.

FAÇA
Vídeo curto, infográfico simples, áudio direto.

O que fazer | Protocolo de contra-narrativa em 5 passos

REGRA FUNDAMENTAL: **verificar antes de responder.** ✓✓ Mandatos e organizações são frequentemente alvo de desinformação projetada para provocar resposta emocional rápida, que depois será usada contra eles. O protocolo de verificação não é lentidão: é proteção.

PASSO 1
PARAR (20 minutos)

O impulso de responder é frequentemente o que a desinformação quer provocar. Se é urgente demais para esperar, é sinal de que a urgência artificial está sendo usada para te manipular.

PASSO 2
IDENTIFICAR

Isolar a afirmação central: qual é o fato específico sendo afirmado? A desinformação mistura fatos verdadeiros com falsidades.

PASSO 3
VERIFICAR

Buscar fonte primária. Confirmar em agências de checagem como Aos Fatos, Agência Lupa e E-farsas. Se achar que o conteúdo envolve IA, observe as dicas do módulo IA e eleições.

PASSO 4
DECIDIR

Nem toda desinformação merece resposta pública. O conteúdo está ganhando alcance significativo? A resposta vai chegar a quem ainda não decidiu? Não responder pode ser a decisão certa.

PASSO 5
RESPONDER a partir da sua perspectiva

Se a decisão for responder: não comece repetindo a mentira. Comece pela verdade: o que é real, o que você fez, o que você defende.

O que fazer | Monitoramento de narrativas

Saber o que os adversários estão dizendo antes de reagir é uma vantagem estratégica. Não é possível construir uma contra-narrativa eficaz sem primeiro monitorar quais narrativas estão em circulação, qual o seu volume, quem as amplifica e como se distribuem por espectro político. Para isso, existem ferramentas e relatórios de referência:

SEMANAL DX | INSTITUTO DEMOCRACIA EM XEQUE

O Semanal DX é um relatório publicado semanalmente pelo Instituto Democracia em Xeque que analisa as narrativas políticas em circulação nas redes sociais no Brasil e seus impactos para a integridade democrática. Cada edição cobre um período de sete dias e entrega:

- Highlights dos principais movimentos e picos de engajamento da semana
- Mapeamento de ameaças à integridade democrática identificadas nas redes
- Métricas e vocabulários mais utilizados por eixo ideológico (esquerda, direita, imprensa)
- Temas relevantes por espectro político e desempenho de perfis presidenciáveis

Os dados são extraídos do Data Lake DX, ferramenta própria do Instituto que analisa publicações de perfis selecionados por critérios de relevância nas redes sociais, complementada pela plataforma Talkwalker. A edição de 05/05/2026, por exemplo, analisou 172.909 publicações que somaram 260.871.620 interações.

POR QUE É ESTRATÉGICO PARA O CAMPO DEMOCRÁTICO

- Permite identificar antecipadamente quais narrativas a extrema direita está construindo, antes que cheguem ao mainstream
- Revela o enquadramento (frame) que cada espectro usa para o mesmo fato: dado essencial para o re-enquadramento do campo democrático (ver Módulo 5)
- Mapeia hubs de influência e redes de disseminação, identificando quem amplifica o quê e com qual alcance
- Documenta ameaças à integridade democrática em tempo real: útil para denúncia e documentação

Acesso: institutodx.org/semanaldx. Publicado toda segunda-feira, com cobertura da semana anterior.

COMO USAR O SEMANAL DX NA PRÁTICA

Incorporar o monitoramento semanal à rotina de comunicação de mandatos e organizações:

- Leia a seção "Ameaças à integridade democrática" para identificar narrativas que precisam de resposta estruturada.
- Compare os vocabulários e identifique onde estamos perdendo terreno.
- Use os dados de volume e interação para calibrar prioridades: o que está ganhando escala merece resposta, o que ainda está pequeno pode ser ignorado.
- Combine com o passo 4 (Decidir) do protocolo de contra-narrativa deste Módulo 3: o Semanal DX ajuda a responder "o conteúdo está ganhando alcance significativo?"
- Arquive as edições relevantes: a série histórica documenta como narrativas evoluem ao longo da campanha, dado valioso para planejamento de longo prazo

O que fazer | Como aprender sem replicar

APRENDER (SEM REPLICAR)

Emoção legítima é poderosa: comunicar com emoção real (orgulho, cuidado, solidariedade, justa raiva) é mais sustentável e mais ético do que criar medo artificial.
Simplicidade é necessária: argumentos complexos não chegam a quem não está prestando atenção. Isso não é simplismo, é respeito pelo tempo de quem recebe a mensagem.
Constância importa: a repetição funciona para a verdade também. Nossas narrativas precisam de constância, não só de picos reativos.
Combater violência de gênero, em suas interseccionalidades, é combater as táticas de desinformação da extrema direita.

NÃO REPLICAR

Desmentir sem amplificar ou publicizar a fonte da desinformação. Ex.: Não fazer reacts usando o conteúdo falso ou desinformativo original, não mencionar os autores dos posts, nem incluir links para os perfis e plataformas falsas. Tudo isso faz com que os algoritmos atribuam mais visibilidade para o conteúdo que se quer desmentir.

DESMONTANDO A MÁQUINA:

Módulo 4

VIOLÊNCIA POLÍTICA DE GÊNERO COMO TÁTICA DE SILENCIAMENTO

A violência política digital não é um efeito colateral da desinformação, é uma tática deliberada. Seu objetivo é silenciar vozes específicas: mulheres negras, pessoas LGBTQIAPN+, periféricas e ativistas. O resultado é uma democracia mais estreita.

77

casos de violência política de gênero e raça no ambiente digital mapeados entre 2021 e 2025, focados em mulheres negras, LBTQIA+, periféricas e defensoras de direitos humanos.

Fonte: Instituto Marielle Franco, Justiça Global e Terra de Direitos. Regime de Ameaça, 2025.

71%

das ameaças mapeadas envolveram morte ou estupro. 63% fizeram referência direta ao assassinato de Marielle Franco como advertência a mulheres negras que disputam poder.

Fonte: Instituto Marielle Franco, Justiça Global e Terra de Direitos. Regime de Ameaça, 2025.

0

condenações definitivas em 4 anos da Lei de Violência Política contra a Mulher (14.192/2021), 62 ações penais mapeadas, nenhuma gerou condenação.

Fonte: Instituto Alziras. Monitor de Violência Política de Gênero e Raça, 2025.

O Instituto Marielle Franco, em parceria com a Justiça Global e a Terra de Direitos, concluiu na pesquisa Regime de Ameaça que "a violência política digital contra mulheres negras é um regime, não uma exceção", um padrão sistemático e reiterado de exclusão, não episódios isolados. A pesquisa classificou sete tipos de violência política digital, descritos a seguir:

Seção 1 Os 7 tipos de violência política digital

1. **AMEAÇAS E INTIMIDAÇÕES**
   Mensagens que evocam violência física ou sexual. 71% das ameaças mapeadas envolveram morte ou estupro.
2. **DESINFORMAÇÃO E FAKE NEWS**
   Conteúdo falso direcionado a descredibilizar trajetória, histórico ou proposta de pessoa específica.
3. **DISCURSO DE ÓDIO**
   Conteúdo que ataca identidade de gênero, raça, sexualidade ou origem combinado com ataque político.
4. **VIOLÊNCIA SIMBÓLICA E DISCURSIVA**
   Linguagem e imagens que humilham, infantilizam ou desumanizam, sem ameaça explícita.
5. **EXPOSIÇÃO DE DADOS (DOXXING)**
   Divulgação de endereço e dados familiares com intenção de facilitar ataque físico.
6. **ASSÉDIO DIGITAL (DOGPILING)**
   Bombardeio coordenado de mensagens para saturar e expulsar a pessoa da plataforma.
7. **INVASÃO DE REDES E CONTAS**
   Acesso não autorizado, vazamento de dados e sequestro de perfis. Tipo de ataque especialmente crítico em períodos eleitorais.

Fonte: Instituto Marielle Franco, Justiça Global e Terra de Direitos (2025, p. 119-120)

Seção 2 Censura algorítmica de corpos dissidentes

Além da violência de gênero propagada por terceiros, as próprias plataformas de redes sociais também violam nossos direitos nos seus processos de moderação de conteúdo. Por um lado não removem conteúdos de ódio e outros ataques, e por outro acabam censurando, ou não promovendo visibilidade de conteúdos feministas, antirracistas, indígenas e do movimento LGBTQIAPN+, enquanto outros conteúdos de ódio e desinformação ganham tração no algoritmo. Por isso, é importante não depender apenas de redes sociais para entregar conteúdo para seu público-alvo.

VISIBILIDADE SAPATÃO NAS REDES

A pesquisa da Coding Rights documentou como mulheres lésbicas enfrentam censura algorítmica ao visibilizar suas existências. A palavra "sapatão" era bloqueada ou filtrada no Facebook e Instagram, enquanto conteúdo homofóbico circulava livremente. O Google levou até 2019 para ajustar seu algoritmo para que "lésbica" deixasse de ser conectada automaticamente a conteúdo pornográfico. "Estamos usando um espaço do outro. Quando o outro quiser, nós podemos ser censuradas ou contidas." - Entrevistada, Coding Rights, 2020.

Seção 3 Protocolo de documentação em 10 passos

PONTO CRÍTICO: documentar ANTES de denunciar na plataforma. Ao denunciar para a plataforma, o conteúdo é frequentemente removido e a evidência desaparece.

1. ANTES de denunciar: fazer print da tela completa com URL visível.
2. Registrar data e hora do conteúdo e do print.
3. Salvar URL completa do post, perfil ou grupo.
4. Se necessário, pedir ajuda ou acompanhamento de alguém que você confia para documentar. Em casos de violência, é comum ficar nervosa e não se atentar a alguns detalhes. Além disso, pedir ajuda previne alguns episódios de revitimização.
5. Se vídeo: fazer screen recording antes de denunciar.
6. Identificar: é conta real, bot ou perfil recentemente criado?
7. Verificar se há padrão: mesmo texto em contas diferentes (coordenação)?
8. Guardar em pasta segura, preferir nuvem com 2FA (autenticação em dois fatores) ativado.
9. Se *doxxing*: documentar quais dados foram expostos especificamente.
10. Se ameaça grave: notificar pessoa de confiança imediatamente e denunciar para autoridades competentes (ver legislação aplicável a seguir).

LEGISLAÇÃO APLICÁVEL A ALGUNS TIPOS DE ATAQUES DIGITAIS EM CAMPANHAS

A cartilha Eleições e Internet (2020) mapeou e detalhou os instrumentos legais disponíveis para candidaturas atacadas nas eleições municipais do mesmo ano. Para 2026, o marco foi atualizado com novas normas:

- Marco Civil da Internet (Lei nº 12.965/2014): privacidade e responsabilidade de plataformas.
- Lei Geral de Proteção de Dados (Lei nº 13.709/2018): uso indevido de dados pessoais.
- Lei Carolina Dieckmann (Lei nº 12.737/2012): invasão de dispositivos e contas
- Lei Lola (Lei nº 13.642/2018): atribui à Polícia Federal a investigação de crimes digitais misóginos.
- Lei Maria da Penha (Lei nº 11.340/2006): violência doméstica e digital.
- Lei de Violência Política de Gênero (Lei nº 14.192/2021): criminaliza assédio, constrangimento, humilhação e ameaça a candidatas e mulheres com mandato.
- Código Eleitoral e Resolução TSE nº 23.610/2019: propaganda irregular e desinformação em campanha.
- Resolução TSE nº 23.755/2026: regras específicas sobre inteligência artificial, deepfakes e conteúdo sintético em campanhas. Atualizou a Resolução nº 23.732/2024 e Resolução nº 23.610/2019. Todas aqui detalhadas no Módulo 5.
- Decreto nº 12.976/2026 — estabelece deveres das plataformas digitais no enfrentamento à violência contra mulheres na internet, com tratamento prioritário para casos de violência política de gênero e contra mulheres com exposição pública (como jornalistas e parlamentares).
- Decreto nº 12.975/2026 - impõe deveres de prevenção contra a circulação massiva de conteúdos criminosos, incluindo violência contra mulheres, com responsabilização das plataformas em casos de falha sistêmica.

PARA FICAR DE OLHO:

- Planos de conformidade das plataformas (art. 125-B da Resolução TSE nº 23.610/2019, incluído pela Resolução nº 23.755/2026): plataformas digitais são obrigadas a elaborar e apresentar à Justiça Eleitoral planos de conformidade detalhando suas estratégias de controle sobre o uso de IA, moderação de conteúdo e combate à desinformação eleitoral.

O QUE FAZER | 10 DICAS DE HIGIENE DIGITAL BÁSICA E ONDE PEDIR AJUDA EM CASO DE ATAQUE

1. Ativar autenticação em dois fatores (2FA) em TODAS as contas.
2. Usar senha diferente para cada conta. Exemplos de gerenciadores: Bitwarden (gratuito) e KeePassXC (gratuito e software livre).
3. Revisar permissões de aplicativos a cada 3 meses.
4. Não usar Wi-Fi público sem VPN para comunicações sensíveis.
5. Criar e-mail dedicado para contas públicas.
6. Ter número de telefone dedicado para redes sociais públicas.
7. Fazer backup criptografado de dados sensíveis mensalmente.
8. Configurar alertas de login em todas as contas importantes.
9. Usar Signal para comunicação interna sensível.
10. Ter plano de resposta definido caso conta seja sequestrada.

Maria d'Ajuda: Linha de atenção e acompanhamento feminista de resposta a incidentes e emergências em segurança digital.

A Maria d'Ajuda é um serviço gratuito e seguro de assistência emergencial em casos de ameaças digitais. Atende mulheres, pessoas dissidentes de gênero, organizações da sociedade civil, ativistas e defensores de direitos humanos na América Latina e Caribe, através de metodologia feminista de acolhimento e suporte técnico com escuta ativa e incentivo a autonomia digital. Operada pela MariaLab.

O que a Maria d'Ajuda atende

- Segurança digital em redes sociais: orientações em casos de perda de contas, ataques de ódio, assédio e outras ameaças digitais que causem insegurança.
- Segurança digital organizacional: estratégias de mitigação de danos para ataques à infraestrutura digital de organizações e outras ameaças no contexto institucional.
- Repressão, perseguição e censura: formas de proteção para quem enfrenta restrições de acesso à internet, perseguição ou censura em função da causa que defende.
- Outros casos: avaliados individualmente, conforme a demanda.

Como pedir ajuda: mariadajuda.org

Outras linhas de apoio

- Central de Atendimento à Mulher: ligue 180
- Disque Direitos Humanos: ligue 100
- Centro de Valorização da Vida: ligue 188 ou cvv.org.br/chat
- Safernet: safernet.org.br/helpline
- Plantão de Apoio Colmeia: específico para casos de violência política de gênero e raça, serviço do Instituto E Se Fosse Você? https://www.acolmeia.org/
- Arquivo de danos digitais | ADD+: plataforma da CTRL+Z para colher relatos e conectar vítimas de suspensões, bloqueios e violações de direitos por plataformas digitais, e viabilizar indicações de suporte jurídico.

DESMONTANDO A MÁQUINA:

Módulo 5

IA E ELEIÇÕES: O NOVO CAMPO DE DISPUTA

Seção 1 IA generativa e agentes de IA

A inteligência artificial mudou a velocidade, a escala e o custo da desinformação eleitoral. Deepfakes, áudios sintéticos e imagens geradas por IA são produzidos em minutos e distribuídos para milhões, sem que eleitoras e eleitores tenham ferramentas simples para identificá-los. Mas a IA também chegou ao cotidiano da comunicação política como ferramenta de produção, alcance e monitoramento, e ignorá-la não é neutralidade, é desvantagem. A democracia brasileira enfrenta esse campo com regras novas, mas ainda insuficientes, e com um campo de defesa dos direitos humanos que precisa aprender a navegar pelos dois lados dessa disputa.

O que é inteligência artificial generativa?

Inteligência artificial generativa é um tipo de sistema capaz de criar conteúdo novo (texto, imagem, áudio e vídeo) a partir de instruções em linguagem comum. Quando você digita uma pergunta para o chat de IA e recebe uma resposta, está usando IA generativa. Quando um vídeo com o rosto ou a voz de uma pessoa é fabricado digitalmente (um deepfake) a tecnologia por trás é a mesma.

O que mudou nos últimos anos não foi só a qualidade do resultado, mas o custo e a velocidade de produção. Antes, criar um deepfake convincente exigia equipe técnica especializada. Hoje, qualquer pessoa com internet faz isso em minutos com ferramentas gratuitas.

O que são agentes de IA?

Agentes de IA são sistemas que executam tarefas, ou seja, ao invés de pedir uma resposta, você pede uma ação. Eles têm muito mais autonomia: agem em sequência e em nome da pessoa que os usa. Portanto, requerem muito mais cautela com seu uso, pois ainda não se sabe o quanto eles podem sair do controle. Os cuidados com esse tipo de ferramenta são tratados na Seção 6 deste módulo.

Seção 2 As regras: o que o TSE decidiu para 2026

O TSE aprovou a Resolução nº 23.755/2026, o principal marco regulatório sobre IA nas eleições de 2026, que altera a Resolução nº 23.610/2019, incorporando e ampliando as inovações introduzidas pela Resolução nº 23.732/2024. Mais de 158 milhões de eleitores estão aptos a votar, e as novas regras buscam evitar que a tecnologia se torne uma ameaça à democracia. A resolução complementa as anteriores, traçando as seguintes regras do jogo no que diz respeito a IA e eleições:

IA e Eleições

OBRIGAÇÕES PARA QUEM PRODUZ PROPAGANDA ELEITORAL
(candidatas, partidos, pessoas apoiadoras):

- **Obrigatoriedade de identificação de conteúdo gerado por IA:** todo conteúdo de propaganda eleitoral criado ou alterado por IA deve exibir aviso explícito, visível e de fácil compreensão. O objetivo é evitar que eleitoras e eleitores sejam enganados por montagens que simulam situações reais. (ref. Art. 9º-B)
- **Identificação de chatbots:** o uso de chatbots, avatares e conteúdos sintéticos como artifício para intermediar a comunicação de campanha com pessoas também submete-se às regras de rotulagem, ou seja, deve-se informar explicitamente que não se trata de interação humana. (ref. Art. 9º-B, § 3º e Art. 28, § 1º-C)
- **Proibição de pagamento de influenciadores:** é proibido contratar pessoas físicas ou jurídicas para publicar conteúdo político-eleitoral em seus perfis em troca de dinheiro ou qualquer vantagem econômica. (ref. Art. 29, § 8º)

- **Proibição de desinformação:** é proibido conteúdo fabricado ou manipulado para difundir fatos notoriamente inverídicos ou descontextualizados com potencial para causar danos ao equilíbrio do pleito ou à integridade do processo eleitoral. (ref. Art. 9º-C § 1º)

- **Proibição de deepfakes que prejudicam ou favorecem candidaturas:** é proibido o uso, para prejudicar ou para favorecer candidatura, de conteúdo sintético em formato de áudio, vídeo, ou combinação de ambos, que tenha sido gerado ou manipulado digitalmente, ainda que mediante autorização, para criar, substituir ou alterar imagem ou voz de pessoa viva, falecida ou fictícia. (ref. Art. 9º-C § 1º)
- **Janela de silêncio:** proibição total de divulgar conteúdo sintético nas 72h antes e 24h após o pleito, mesmo que rotulado. (ref. Art. 9º-B, § 3º-A)

IA e Eleições OBRIGAÇÕES PARA PLATAFORMAS DIGITAIS

- **Obrigatoriedade de identificação:** disponibilizar campo específico para declaração de uso de IA nos fluxos de impulsionamento pago. (ref. Art. 9º-B, § 5º)
- **Remoção de conteúdo:** remover imediatamente conteúdo que descumpra rotulagem ou incida nas vedações, por iniciativa própria ou por ordem judicial. (ref. Art. 9º-B, § 4º)
- **Remoção de perfis falsos:** perfis comprovadamente falsos, apócrifos ou automatizados que pratiquem de forma reiterada crimes eleitorais ou espalhem desinformação reconhecida pela Justiça Eleitoral poderão ser removidos. No caso de perfis automatizados, diretamente pelos próprios provedores. Nos demais casos, dependem de ordem judicial. (ref. Art. 38-A)
- **Proteção da integridade do processo eleitoral:** remover autonomamente, sem ordem judicial, conteúdo que descredibilize a urna eletrônica, incite crimes contra o Estado Democrático, fomente ruptura institucional ou represente violência política contra a mulher. (ref. Art. 28, § 4º-A e 4º-B)

- **Proibição de recomendação de voto por IA:** provedores de sistemas de IA estão proibidos de ranquear, recomendar, priorizar ou emitir opiniões sobre candidaturas e partidos. (ref. Art. 28, § 1º-C, incisos I e II)

- **Sistemas de IA não podem criar conteúdo de violência política contra a mulher:** é vedado aos sistemas de IA criar conteúdo sexual envolvendo candidatas ou candidatos, também fica proibido que esses sistemas formulem publicidade que represente violência política contra a mulher. (ref. Art. 28, § 1º-C, incisos III e IV)
- **Canais de denúncia:** plataformas ficam obrigadas a implementar canais específicos para que candidatas, candidatos, partidos, federações e coligações denunciem infrações. (ref. Art. 9º-E, § 2º)
- **Janela de silêncio digital:** bloquear impulsionamento de conteúdo sintético nas 72 horas antes e 24 horas depois do pleito. (ref. Art. 9º-B, § 3º-A).
- **Obrigação de elaborar planos de conformidade:** são planos destinados à prevenção e mitigação de riscos à integridade do processo eleitoral, detalhando como a plataforma vai cumprir as obrigações acima. O plano é requisito para credenciamento no TSE, necessário inclusive para poder monetizar propaganda eleitoral. (ref. Art. 125-B)

Embora tenhamos avançado na regulação de IA nas eleições, ainda temos algumas lacunas críticas. A principal delas é que a resolução não especifica mecanismos de verificação independente de que as plataformas estão cumprindo as regras. É a própria plataforma que declara conformidade, sem auditoria externa obrigatória. O estudo Data Not Found do NetLab UFRJ documenta exatamente esse problema: as plataformas não fornecem dados suficientes para que pesquisadores ou a Justiça Eleitoral possam verificar o cumprimento de forma independente. Para que planos de conformidade não virem documentos meramente declaratórios é necessário assegurar:

- a publicidade obrigatória dos planos, sem a qual não há controle social possível;
- a definição de indicadores auditáveis externamente, e não apenas relatórios internos das próprias plataformas;
- a clareza no modelo de responsabilidade solidária.

PERÍCIA ESPECIALIZADA

Pela Resolução TSE 23.755/2026, os Tribunais Eleitorais podem firmar parcerias com universidades e órgãos especializados para identificar se um conteúdo foi manipulado. Em alguns casos, a Justiça pode inverter a responsabilidade: em vez de quem denuncia precisar provar que o vídeo é falso, quem postou terá o dever de provar que é verdadeiro. Organizações como a WITNESS têm mapeado ferramentas, boas práticas e desafios para esse tipo de perícia (mais informações na seção 3, sobre Deepfakes).

REGRA DE OURO PARA CANDIDATURAS E MANDATOS

- Antes de compartilhar qualquer conteúdo sobre adversários ou sobre si: verifique a origem.
- Antes de responder a qualquer ataque com conteúdo sintético: verifique a autenticidade.
- Antes de usar IA para criar conteúdo de campanha: inclua a identificação obrigatória exigida pelo TSE. O descumprimento pode gerar remoção judicial imediata do conteúdo e resposta da Justiça Eleitoral.

Seção 3 Deepfakes: o que são, como funcionam e casos documentados em eleições pelo mundo

Deepfake é um conteúdo (vídeo, imagem ou áudio) em que a aparência ou a voz de uma pessoa real foi substituída ou gerada artificialmente por IA de forma convincente. A palavra vem da combinação de deep learning (aprendizado profundo) com fake (falso). O que era restrito a laboratórios especializados em 2018 é hoje acessível a qualquer pessoa com um smartphone e uma assinatura de aplicativo. Como alertou a WITNESS, organização de referência global em verificação de mídia, o maior perigo trazido pela IA generativa audiovisual é que ela abriu a possibilidade da negação plausível. Qualquer coisa pode ser alegada como deepfake, mesmo quando é real. Isso é tão perigoso quanto o deepfake em si.

Os três tipos que mais circulam no contexto eleitoral

DEEPFAKE DE VÍDEO
Substituição ou geração do rosto e movimentos de uma pessoa em vídeo. É o mais impactante visualmente. Inclui técnicas como troca de rosto (face-swap) e geração de âncoras de noticiário sintéticas.

Casos documentados: figuras públicas que já faleceram fazendo declarações políticas em redes sociais; candidaturas "anunciando" desistência em deepfake coordenado na véspera do pleito; deepfakes de âncoras de telejornais conhecidos espalhando escândalos eleitorais.

DEEPFAKE DE ÁUDIO: CLONAGEM DE VOZ
Replicação da voz de uma pessoa a partir de amostras de áudio, às vezes com apenas alguns segundos de gravação original. É o tipo mais barato, mais rápido de produzir e mais difícil de detectar. Pesquisa da University College London (2023) concluiu que humanos não conseguem detectar áudios deepfake 27% dos casos, mesmo quando estão alertas.

Caso documentado: áudio de líder pró-Europa na Eslováquia "discutindo compra de votos", vazado dias antes da eleição, o partido perdeu para uma legenda pró-Rússia (2023).

DEEPFAKE DE IMAGEM E DEEPNUDE
Geração de imagens falsas realistas, incluindo deepnudes: imagens sexuais falsas de pessoas reais criadas por IA sem consentimento. Usadas sistematicamente contra candidatas mulheres para humilhação e silenciamento político.

Como os deepfakes têm sido usados nas eleições – 4 padrões documentados

1. **Golpes financeiros**
Candidatos com seus rostos usados para promover esquemas de investimento falsos. Na Romênia (maio 2025), deepfakes de candidatos presidenciais no Facebook promoviam oportunidade de investimento inexistente. No Canadá (abril 2025), deepfake de candidato a primeiro-ministro promovia esquema de criptomoedas.

2. **Supressão de votos**
Deepfakes lançados horas antes da eleição para confundir eleitores. Na Argentina (maio 2025), dois deepfakes afirmando a desistência de candidatos foram distribuídos horas antes da abertura das urnas. Na Irlanda (outubro 2025), vídeo deepfake anunciando falsamente a retirada de candidata foi lançado dias antes da eleição.

3. **Desinformação com aparência jornalística**
Deepfakes usando logos e formatos de canais reais para parecer notícia legítima.Na Alemanha (fevereiro de 2025), uma rede coordenada usou vídeos com IA imitando emissoras de TV e agências de inteligência (incluindo material falsamente atribuído ao MI6 britânico) para difundir falsas ameaças terroristas sobre as eleições. No Equador (fevereiro 2025), âncoras sintéticos com logos da CNN e France 24 espalharam acusações de fraude.

4. **Envenenamento de chatbots**
Na Austrália (maio 2025), uma rede coordenada publicou milhares de artigos falsos projetados não para enganar humanos, mas para contaminar os dados que alimentam os chatbots de IA. Testes mostraram que 16,6% das respostas dos chatbots amplificaram narrativas falsas plantadas pela rede.

DONA MARIA:
PERSONAGEM DE IA DISSEMINADORA DE DESINFORMAÇÃO

Dona Maria é uma personagem fictícia criada com IA, representada como uma mulher negra e idosa de linguagem direta. Presente no Instagram, TikTok, Facebook, YouTube e X, o perfil acumulou mais de 100 milhões de visualizações. O conteúdo dissemina desinformação contra o presidente Lula, ministros do STF e a esquerda, incluindo a invenção de um "imposto da reciclagem" para catadores de latinha e ataques à segurança do Pix. O perfil gerou renda direta para seu criador de extrema direita.

O Instituto Democracia em Xeque, em parceria com o Departamento de Comunicação da PUC-Rio, criou o projeto Rotulando IA nas Eleições para identificar como plataformas digitais informam o uso de IA generativa. Os três primeiros boletins de 2026 registraram mais de 9.300 publicações com marcações de IA no Instagram, YouTube, TikTok e X, gerando cerca de 5,5 milhões de interações apenas no terceiro período. O Instagram concentrou a maior parte do material, e o uso predominante foi menos para conteúdos realistas e mais para reforço estético e retórico de peças políticas, especialmente paródias e caricaturas. No espectro político, a direita respondeu por 53% dos posts e 75% das interações no primeiro boletim.

Seção 4 Como identificar conteúdo sintético (IA)

50% de aumento mensal na média diária de casos de conteúdo sintético político sem identificação, acompanhando a aproximação das eleições de 2026. 2 em cada 3 conteúdos com IA circularam sem qualquer aviso ao público.

Fonte: Observatório IA nas Eleições, abril 2026

COMO IDENTIFICAR CONTEÚDO SINTÉTICO (IA)

- **Imagens:** aspecto plastificado, proporções distorcidas, erros em textos e cartazes, mãos com dedos extras.
- **Vídeos deepfake:** sincronização imperfeita de lábios/áudio, piscadas ausentes, bordas do rosto borradas e mais "cartunescas" e texturas sintéticas.
- **Áudio sintético:** pronúncia excessivamente perfeita, ausência de "tiques" naturais de fala e variação na emoção, qualidade de áudio muito uniforme.
- **Contexto:** conta criada recentemente, sem identificação de IA (obrigatória pela Resolução TSE 23.755/2026), padrão de compartilhamento muito rápido e em volume incomum.

ALGUMAS FERRAMENTAS DE VERIFICAÇÃO ✓✓

- **Enviar para agências de checagem brasileiras:** como Aos Fatos (aosfatos.org), Agência Lupa (piaui.folha.uol.com.br/lupa), E-farsas (e-farsas.com) ✓✓
- **Busca reversa de imagem:** a busca reversa permite rastrear onde uma imagem apareceu pela primeira vez na internet. Ferramentas como a TinEye (tineye.com) funcionam fazendo o upload da imagem ou colando o link, e a ferramenta mostra todos os lugares onde aquela foto já apareceu, com datas. Se uma imagem que supostamente mostra um evento de hoje aparece em sites de outros países há meses ou anos, é sinal claro de desinformação. A busca reversa também ajuda a identificar imagens geradas por IA. Para vídeos, plataformas como o WeVerify (plugin de navegador desenvolvido em parceria entre portais de notícias e centros de pesquisa europeus) fazem função similar, extraindo quadros para verificação quadro a quadro. ✓✓

Ferramentas de detecção têm limites sérios

O Reuters Institute (Universidade de Oxford) e a WITNESS conduziram testes sistemáticos com as principais ferramentas de detecção de uso de IA, para além de deepfakes. As ferramentas testadas incluem Hive Moderation, Optic, V7, Deepware Scanner, Illuminarty e outras. Os testes documentaram limitações importantes que qualquer pessoa usando essas ferramentas precisa compreender:

O que as ferramentas fazem bem

- Detectar deepfakes de figuras públicas com ampla presença online
- Reconhecer padrões deixados por ferramentas de IA já conhecidas
- Funcionar como ponto de partida em uma verificação mais ampla
- Identificar fotos de perfil geradas por IA em contas falsas

Onde as ferramentas falham

- O modelo não reconhece pessoas com pouca presença digital
- Áudio regravados perdem metadados (gravação de tela)
- Imagens comprimidas (como as que circulam no WhatsApp)
- Deepfakes feitos com ferramentas novas, ainda fora dos conjuntos de treinamento
- Imagens com blur (desfoque) ou motion effects (efeitos de movimento)

CASOS DOCUMENTADOS DE FALHA DAS FERRAMENTAS

✦ Um deepfake do ex-presidente Barack Obama teve sua detecção anulada apenas com redução de resolução e corte de alguns segundos do vídeo, o resultado passou de "deepfake detectado" para "nenhum deepfake detectado".

✦ Um robocall com voz falsa do então presidente Joe Biden foi detectado corretamente no arquivo original, mas a mesma gravação, feita com um celular apontado para o alto-falante, foi classificada como "altamente provável que seja real".

✦ Screenshots de avatares de IA usados para apoiar um golpe militar na África Ocidental foram classificados como "não gerados por IA" em mais da metade dos casos.

Conclusão da WITNESS: *"Até agora, não encontramos nenhuma ferramenta que não tenha falhado em nossos testes e que oferecesse resultados transparentes e acessíveis."*A detecção tecnológica é um apoio, não um substituto, para verificação humana contextual. *(WITNESS e Reuters Institute, 2024)*

Seção 5 Quem monitora: observatórios e iniciativas de referência

Os casos de usos de IA em contexto eleitoral se proliferam rapidamente, para além do que esta cartilha é capaz de documentar. Para atualizações, aconselha-se acompanhar os observatórios e sistemas oficiais do TSE e AGU.

SIADE (TSE)

Sistema de Alertas de Desinformação Eleitoral do TSE. Qualquer pessoa pode denunciar conteúdos falsos ou fora de contexto que possam causar danos ao equilíbrio do pleito ou à integridade eleitoral. tse.jus.br/eleicoes/sistema-de-alertas

Observatório da Democracia (AGU)

Grupo criado pela Advocacia-Geral da União para produzir estudo que auxilie eleitores e instituições a se precaver contra desinformação com IA nas eleições de 2026. gov.br/agu

Observatório IA nas Eleições

Repositório em tempo real de casos de uso de IA generativa nas eleições brasileiras. Iniciativa do Data Privacy Brasil, Aláfia Lab e Desinformante. Permite envio de denúncias e consulta pública. observatorioianaseleicoes.com.br

Rotulando IA

O Instituto Democracia em Xeque, em parceria com o Departamento de Comunicação da PUC-Rio, publica relatórios quinzenais sobre uso de IA nas redes sociais, através do projeto Rotulando IA nas Eleições para identificar como as plataformas digitais informam o uso de IA generativa. institutodx.org/rotulandoia/

O que fazer | Protocolo de resposta a deepfakes

Se você receber um deepfake sobre você ou sua candidatura:

PASSO 1 — DOCUMENTAR ANTES DE TUDO

Faça print com URL visível, registre data e hora, salve o link completo, grave o vídeo ou áudio com gravação de tela (screen recording), ou use ferramentas de arquivamento online, como o Web Archive ou o archive.ph. Lembre: ao denunciar à plataforma, o conteúdo pode ser removido e a evidência desaparece. Documentar é sempre o primeiro passo, antes de qualquer outra ação.

PASSO 2 — NÃO AMPLIFICAR

Resista ao impulso de compartilhar para desmentir. Cada compartilhamento aumenta o alcance. O desmentido frequentemente amplifica o conteúdo original. A WITNESS alerta: responder publicamente pode causar efeito Streisand: o conteúdo que poucos viram se torna viral após a resposta. Use o protocolo de contra-narrativas do Módulo 3 para decidir se vale responder e como.

PASSO 3 — DENUNCIAR FORMALMENTE

Denuncie à plataforma (depois de documentar). Em período eleitoral, acione o SIADE do TSE (tse.jus.br/eleicoes/sistema-de-alertas). Para casos de deepnudes ou ameaças graves, registre boletim de ocorrência e considere acionar suporte jurídico especializado.

PASSO 4 — RESPONDER (SE DECIDIR RESPONDER)

Se o conteúdo tem grande alcance e a resposta é necessária: comece pela verdade, não pelo deepfake. Mostre o que é real: um vídeo seu, uma declaração autêntica, dados concretos. Nunca repita o conteúdo falso no início da resposta. Conforme recomenda o GIJN: "o conteúdo verdadeiro deve dominar o topo da resposta". Mantenha-se firme: o ataque é também uma evidência de que sua candidatura incomoda.

Se você ver um deepfake sobre outra pessoa

- Não compartilhe, nem para perguntar se é verdade
- Avise a pessoa diretamente antes de qualquer outra ação pública
- Documente e encaminhe ao SIADE se for conteúdo eleitoral (tse.jus.br/eleicoes/sistema-de-alertas)
- Sinalize às agências de checagem (Aos Fatos, Lupa, etc.) se o conteúdo está ganhando alcance. Para casos de alto impacto, a WITNESS opera o Deepfake Rapid Response Force, que conecta jornalistas e organizações a especialistas em forense de mídia e IA para análise aprofundada e confiável. Se for comunicar os resultados, sempre explicite qual ferramenta foi usada, quais são seus limites conhecidos e como os resultados foram interpretados. Nunca apresente a detecção automática como prova definitiva.
- Se for deepnude ou ameaça: apoie a pessoa em documentar e acionar suporte jurídico especializado

Prevenção: algumas precauções possíveis

- Avise sua equipe e base de apoio sobre o risco de deepfakes antes da campanha se intensificar
- Defina internamente quem é responsável por monitorar conteúdo e quem autoriza respostas públicas
- Conheça as ferramentas de denúncia antes de precisar usá-las sob pressão
- Considere o uso de Content Credentials e marcas d'água digitais para autenticar conteúdo oficial da campanha, como recomendado pelo CETaS, do Alan Turing Institute
- Ao usar IA para criar conteúdo de campanha: sempre inclua a identificação obrigatória exigida pelo TSE e produza conteúdo autêntico e consistente que crie referência clara para eleitoras, eleitores e checadores

Recursos de referência

- **WITNESS Deepfake Rapid Response Force:** suporte especializado em forense para organizações e jornalistas: gen-ai.witness.org/deepfakes-rapid-response-force
- **Reuters Institute e WITNESS:** análise de como ferramentas de detecção funcionam e onde falham: reutersinstitute.politics.ox.ac.uk
- **GIJN:** guia prático com dicas para identificar, investigar e verificar deepfakes de áudio em eleições: gijn.org/resource/tipsheet-investigating-ai-audio-deepfakes
- **CETaS / Alan Turing Institute:** análise de deepfakes eleitorais, golpes financeiros e envenenamento de chatbots em 2025: cetas.turing.ac.uk
- **SIADE / TSE:** Sistema de Alertas de Desinformação Eleitoral do TSE. Qualquer pessoa pode enviar conteúdos falsos ou fora de contexto que possam causar danos ao equilíbrio do pleito ou à integridade eleitoral: tse.jus.br/eleicoes/sistema-de-alertas
- **Observatório IA nas Eleições:** canal de denúncias e repositório de casos: observatorioianaseleicoes.com.br/reportes

Seção 6 IA como ferramenta estratégica (desde que usada com consciência e cuidado)

A inteligência artificial generativa chegou ao cotidiano da comunicação política antes que tivéssemos tempo de discutir se queremos, como queremos e em que termos queremos usá-la. Enquanto isso, essas ferramentas já foram apropriadas com certa escala para fabricar deepfakes, produzir desinformação e automatizar ataques. Ignorar a ferramenta não é neutralidade: é desvantagem.

Mas usar IA sem critério também tem custo. As plataformas de IA não são neutras, foram construídas por empresas com visões de mundo, vínculos políticos e modelos de negócio que precisam ser conhecidos antes de qualquer decisão de uso.

Como escolher uma ferramenta?

Tecnopolítica importa

Toda escolha de plataforma de IA é também uma escolha política. Antes de adotar qualquer ferramenta, vale perguntar:

- Quem é a empresa por trás? Qual é sua visão de mundo e seus vínculos políticos?
- A empresa tem contratos militares ou com governos autoritários?
- Os dados que você insere são usados para treinar os modelos? Com que finalidade?
- A empresa tem políticas de proteção a comunidades vulneráveis, ou as removeu?
- Existe uma alternativa de código aberto ou com governança mais transparente?

Ferramenta	Útil para	Quem fez	Vínculos políticos Visão de mundo	Escândalos e controvérsias	Dados treinam o modelo?	Dá para desativar?	Legislação de proteção de dados aplicável	Gratuito? (muitas vezes qndo o acesso é gratuito, o pagmnto são seus dados)
Chatbots | Texto								
ChatGPT EUA	Texto, roteiros, pesquisa, síntese	OpenAI	CEO Sam Altman doou US$1 milhão pessoalmente à posse de Trump (dez/2024). OpenAI assinou contrato militar com o Pentágono em fev/2026 - horas após Trump banir a Anthropic - concordando com uso para 'todos os cenários lícitos', revertendo política de 2023 que proibia acesso militar. Parte do grupo de Big Techs que esteve na posse de Trump.	Contrato militar com o Pentágono (fev/2026) sem restrições para armas autônomas; doação pessoal do CEO à posse de Trump; casos documentados de uso em operações de influência (Rússia, China, Israel); vazamento de títulos de conversas (2023).	Sim, por padrão	SIM ✦Config ✦Controles de dados ✦Desativar ✦'Melhorar o modelo'	Dados nos EUA sujeitos à legislação de vigilância americana (FISA, Patriot Act), que permite acesso governamental a dados de usuários sem notificação. Revelações Snowden (2013) documentaram colaboração de Big Techs com programas de vigilância da NSA.	Versão básica gratuita
Claude EUA	Texto, análise, síntese, roteiros	Anthropic	Foco declarado em segurança de IA. Tinha contrato de US$200 milhões com o Pentágono (jul/2025) - o Claude foi usado em operações militares americanas, inclusive durante a guerra com o Irã. O conflito surgiu quando o Pentágono quis ampliar o uso para armas autônomas e vigilância doméstica em massa, o que a Anthropic recusou. É uma empresa americana com investimento da Amazon que quer continuar tendo contratos militares - com limites que ela própria define.	Tinha e quer continuar tendo contratos militares com o Pentágono. Trump declarou a Anthropic 'risco à cadeia de suprimentos' após a empresa recusar uso para armas autônomas e vigilância em massa (fev/2026); Anthropic processou o governo e obteve bloqueio parcial judicial. Investimento da Amazon.	Não, por padrão	Sim ✦Config ✦Privacidade	Dados nos EUA sujeitos à legislação de vigilância americana (FISA, Patriot Act), que permite acesso governamental a dados de usuários sem notificação. Revelações Snowden (2013) documentaram colaboração de Big Techs com programas de vigilância da NSA.	Versão básica gratuita
Gemini EUA	Texto, integração com ferramentas Google	Google	Google participou do Project Maven (2018) — contrato de IA militar que gerou protestos internos e demissões em massa. Assinou novo contrato militar em 2026. Parte do grupo de Big Techs presente na posse de Trump. Lobby documentado contra regulação de plataformas no Brasil (2023).	Project Maven (contrato militar com IA para drones, 2018); lobby contra PL das Fake News no Brasil (2023); imagens com imprecisões históricas documentadas; novo contrato militar em 2026.	Sim, por padrão	Sim ✦Atividade do app Gemini ✦Pausar	Dados nos EUA sujeitos à legislação de vigilância americana (FISA, Patriot Act), que permite acesso governamental a dados de usuários sem notificação. Revelações Snowden (2013) documentaram colaboração de Big Techs com programas de vigilância da NSA.	Versão básica gratuita

GUIA DE FERRAMENTAS DE IA – O que você precisa saber antes de usar

Ferramenta	Útil para	Quem fez	Vínculos políticos Visão de mundo	Escândalos e controvérsias	Dados treinam o modelo?	Dá para desativar?	Legislação de proteção de dados aplicável	Gratuito? (muitas vezes qndo o acesso é gratuito, o pagmnto são seus dados)
Chatbots | Texto								
Meta IA | Llama EUA	Não recomendada para uso organizacional	Meta	Parte do grupo de Big Techs presente na posse de Trump (2025). Removeu proteções contra discurso de ódio a LGBTQIAP+ e mulheres (2025). Modelo Llama open-source foi usado por instituições ligadas ao exército chinês sem autorização da Meta.	Remoção de proteções contra discurso de ódio (2025); substituição de moderação humana por IA; escândalo Cambridge Analytica (2016/2018); Llama open-source usado por instituições ligadas ao exército chinês; participação na posse de Trump.	Sim	Limitado	Dados nos EUA sujeitos à legislação de vigilância americana (FISA, Patriot Act), que permite acesso governamental a dados de usuários sem notificação. Revelações Snowden (2013) documentaram colaboração de Big Techs com programas de vigilância da NSA.	Gratuito
DeepSeek China	Não recomendada para uso organizacional	High-Flyer	Empresa chinesa sujeita à Lei de Inteligência Nacional da China (2017). O risco de acesso governamental é estruturalmente similar ao das Big Techs americanas - com contextos políticos distintos. Banido em vários países por preocupações de segurança.	Banido em vários países (Itália, Austrália, EUA, entre outros) por preocupações de segurança nacional; dados sujeitos à lei chinesa de inteligência.	Sim	Limitado	Dados na China sujeitos à Lei de Inteligência Nacional chinesa (2017), que obriga empresas a cooperar com o governo quando solicitado. Risco estruturalmente similar ao das empresas americanas - contextos políticos distintos.	Gratuito
Le Chat (Mistral) França	Texto, análise, tradução	Mistral AI	Empresa europeia com apoio do governo francês (Macron). Foco em IA aberta e soberania europeia. Sem vínculos com governos autoritários conhecidos.	Críticas por uso de dados sem licença no treinamento (conjunto LAION). Sem escândalos graves documentados.	Depende open-source	Sim ✦Config da conta	Sujeito ao GDPR europeu, considerado um dos marcos mais rigorosos de proteção de dados do mundo. Dados de usuários têm proteções legais robustas contra acesso governamental.	Versão básica gratuita
Mistral | Ollama França / Internacional	Texto com máxima privacidade, uso local	Mistral AI / código aberto	Empresa europeia sem vínculos políticos conhecidos. Código aberto auditável pela comunidade.	Sem escândalos conhecidos. Código aberto permite auditoria independente.	Não, roda localmente	N/A ✦Dados não saem do computador	Roda localmente - nenhum dado enviado para servidores externos. Máxima soberania sobre os dados.	Gratuito

GUIA DE FERRAMENTAS DE IA – O que você precisa saber antes de usar

Ferramenta	Útil para	Quem fez	Vínculos políticos Visão de mundo	Escândalos e controvérsias	Dados treinam o modelo?	Dá para desativar?	Legislação de proteção de dados aplicável	Gratuito? (muitas vezes qndo o acesso é gratuito, o pagmnto são seus dados)
Greação de imagem								
Midjourney EUA	Geração de imagens para comunicação visual	Midjourney Inc.	Sem vínculos políticos declarados. Empresa privada americana.	Processos judiciais por uso de imagens de artistas sem autorização para treinamento do modelo.	SIM imagens usadas para treinamento	Apenas planos pagos empresariais	Dados nos EUA sujeitos à legislação de vigilância americana (FISA, Patriot Act), que permite acesso governamental a dados de usuários sem notificação. Revelações Snowden (2013) documentaram colaboração de Big Techs com programas de vigilância da NSA.	Pago (a partir de US$10/mês)
DALL-E EUA	Geração de imagens	OpenAI	Mesmos vínculos da OpenAI - contrato militar com Pentágono, doação à posse de Trump.	Mesmos escândalos da OpenAI. Processos por uso de imagens de artistas sem autorização.	SIM	Limitado	Dados nos EUA sujeitos à legislação de vigilância americana (FISA, Patriot Act), que permite acesso governamental a dados de usuários sem notificação. Revelações Snowden (2013) documentaram colaboração de Big Techs com programas de vigilância da NSA.	Pago (via ChatGPT Plus)
Stable Diffusion Reino Unido / código aberto	Geração de imagens com máxima privacidade	Stability AI / código aberto	Empresa passou por crise financeira (2024). O modelo de código aberto é independente da empresa e pode ser usado localmente.	Processos por uso de imagens de artistas sem autorização; Stability AI em crise financeira (2024). Modelo open-source é independente dessas questões quando usado localmente.	NÃO, roda localmente	N/A ✦ Uso local	Roda localmente - nenhum dado enviado para servidores externos. Máxima soberania sobre os dados.	Gratuito

NOTA: O risco de acesso governamental a dados é estruturalmente similar entre países - tanto a legislação americana (FISA, Patriot Act) quanto a chinesa (Lei de Inteligência Nacional, 2017) permitem acesso governamental aos dados armazenados nessas jurisdições. A diferença está no contexto político, não na estrutura do risco. Ferramentas de código aberto rodando localmente eliminam esse risco independentemente do país de origem.

Fonte: Informações baseadas em termos de uso públicos, relatórios de privacidade e casos documentados até maio de 2026. Verifique atualizações antes de adotar qualquer ferramenta.

O QUE NÃO FAZER | O que nunca inserir em ferramentas de IA?

Mesmo usando plataformas com políticas mais cuidadosas, há informações que nunca devem ser compartilhadas com ferramentas de IA:

- Nomes, endereços ou dados de identificação de ativistas, pessoas apoiadoras ou pessoas em situação de vulnerabilidade
- Documentos estratégicos internos: planos de campanha, listas de contato, estratégias de comunicação confidenciais
- Relatos de violência ou casos de ataque que possam identificar vítimas
- Dados financeiros da organização ou da campanha
- Senhas, tokens de acesso ou credenciais de qualquer tipo
- Demais dados ou informações que você considerar sensíveis

REGRA GERAL

Trate qualquer ferramenta de IA como um ambiente público. Se você não publicaria a informação nas redes sociais, não insira em uma ferramenta de IA.

Os motivos:

- O que você digita pode ser lido por humanos que revisam conversas para melhorar o modelo
- Dados inseridos hoje podem aparecer em respostas geradas para outras pessoas no futuro
- Plataformas podem ser hackeadas ou sofrer falhas técnicas que expõem conversas
- Em alguns países, dados armazenados em servidores americanos estão sujeitos à legislação de vigilância dos EUA

Nota: todos esses itens são baseados em casos que já aconteceram com diferentes plataformas.

CONFIGURAÇÕES BÁSICAS DE PROTEÇÃO

Antes de usar qualquer plataforma de IA generativa, configure:

- Desative o uso das suas conversas para treinamento
- Use contas separadas: não use sua conta pessoal ou institucional principal. Crie uma conta específica para uso de IA, sem vínculo com dados sensíveis
- Prefira modo sem histórico quando estiver trabalhando com conteúdo estratégico ou sensível
- Nunca use Wi-Fi público sem VPN ao acessar plataformas de IA com conteúdo organizacional

Alguns usos estratégicos possíveis

(depois de configurações de proteção)

Com esses cuidados, a IA pode ser uma aliada real para:

PRODUÇÃO DE CONTEÚDO

Redigir roteiros, adaptar e revisar textos para diferentes plataformas, criar variações de uma mesma mensagem para públicos distintos. Nunca pedir informações factuais, estatísticas para serem incorporadas em materiais oficiais de campanha sem fontes oficiais e sem processo de checagem humana e criteriosa, pois há um enorme risco de receber informações inventadas. Por isso lembre-se: a voz final deve sempre ser sua. A IA é ponto de partida, não produto acabado.

ACESSIBILIDADE E ALCANCE

Transcrição automática, legendas, tradução, audiodescrição. Ferramentas como o CapCut permitem alcançar mais pessoas com menos recurso e tornam a comunicação mais inclusiva.

RESUMO E SÍNTESE

Processar grandes volumes de texto, relatórios, atas ou monitoramento de narrativas para extrair o essencial rapidamente, lembrando que é sempre importante checar os conteúdos diretamente na fonte para validar o resumo.

PREPARAÇÃO PARA ENTREVISTAS E DEBATES

Simular perguntas difíceis, treinar respostas, antecipar enquadramentos e ataques previsíveis com base em narrativas em circulação.

PROGRAMAÇÃO DE FERRAMENTAS E PÁGINAS WEB

Criar formulários, automatizar planilhas, montar landing pages ou pequenos sites de campanha mesmo sem equipe técnica. A IA pode gerar e explicar o código, mas é essencial testar tudo antes de publicar e, em projetos sensíveis, contar com revisão de alguém da área.

Lembre-se: *nenhum conteúdo gerado por IA generativa ou agentes de IA deve ir a público sem revisão humana criteriosa.*

AGENTES DE IA

O uso de agentes de IA é muito mais arriscado, pois operam com mais autonomia do que chatbots convencionais. Muitos deles estão em fases iniciais de teste, sem documentação completa de suas vulnerabilidades e riscos. Se optar por adotar qualquer agente, é importante entender o que você está autorizando a ser realizado em seu nome.

Ninguém dá procuração a ninguém para agir em seu nome sem limites e transparência. Imagine fazer isso para uma IA, onde até a responsabilização pelo que ela fizer ainda está em uma área cinza de regulação.

Se ainda assim, optar por testar essas ferramentas, seu uso requer também um perfil mais técnico na equipe, com habilidades suficientes para avaliar detalhes dos termos de uso, das configurações técnicas e da ética da empresa fornecedora, além de saber monitorar o agente, para garantir que ele não esteja fugindo do controle e saber desativá-lo rapidamente em caso de urgência.

Cuidados antes de usar qualquer agente de IA

Deve-se poder garantir:

✦ **Confiabilidade da empresa desenvolvedora:** pesquise quem está por trás da ferramenta, onde os dados são armazenados e quais são as políticas de privacidade antes de qualquer configuração. Seus dados podem estar treinando a ferramenta e serem acessados, inclusive por moderadores de conteúdo, o que em contexto de campanha não é ideal.

✦ **Teste antes do uso real:** rode o agente em ambiente controlado antes de ativá-lo em situação real. Agentes podem cometer erros factuais, de tom ou de julgamento político, e em período eleitoral esses erros têm custo alto.

✦ **Escopo limitado de configuração e acesso:** conceda apenas as permissões estritamente necessárias para a tarefa. Um agente que monitora e alerta é muito diferente de um que monitora, decide e publica automaticamente.

✦ **Supervisão humana em tudo que vai a público:** agentes podem redigir, categorizar e alertar, mas uma pessoa deve revisar e aprovar antes de qualquer publicação, resposta ou envio.

✦ **Proteção do que é sensível:** listas de militantes, documentos estratégicos, sistemas de arrecadação e contas críticas não devem ser acessíveis ao agente. Ele age em seu nome e as consequências políticas e legais também são suas.

✦ **Monitoramento frequente:** revise periodicamente os registros de atividade. Erros de julgamento político podem só aparecer quando já causaram dano.

✦ **Agentes configurados por terceiros: empresas de consultoria, por exemplo, podem ter acesso a dados sensíveis da campanha. Exija transparência total sobre o que é coletado e armazenado.**

✦ **Agentes que publicam automaticamente em redes sociais podem violar as regras do TSE sobre propaganda eleitoral se não houver supervisão humana no processo.**

✦ **Um agente comprometido ou mal configurado pode vazar informações estratégicas, responder de forma inadequada a ataques ou amplificar narrativas equivocadas no pior momento possível**

O QUE A IA NUNCA VAI SUBSTITUIR

O VÍNCULO COMUNITÁRIO

Nenhuma IA é capaz de replicar a confiança construída em território.

O JULGAMENTO POLÍTICO

A IA não sabe o que é estratégico para o seu contexto.

A RESPONSABILIDADE EDITORIAL

O conteúdo gerado por IA é sua responsabilidade ao publicar. Por isso, revise tudo e cheque fontes: as ferramentas de IA alucinam e tê m vieses, principalmente de gênero, raça, mas também culturais, entre outros.

A TRANSPARÊNCIA

Use IA, mas sinalize quando o fizer. É uma prática de transparência e integridade que todos precisamos adotar.

DESMONTANDO A MÁQUINA:

Módulo 6

COMUNICAÇÃO PROGRESSISTA EFICAZ

Este módulo faz a virada: como o campo progressista pode disputar o espaço público com estratégia própria, sem jogar apenas no tabuleiro da direita. Comunicação progressista eficaz não é a negação da comunicação da extrema direita. É algo diferente: mais enraizada em comunidade, mais consistente no tempo, mais honesta sobre emoções reais.

O CAMPO PROGRESSISTA TEM O QUE A DESINFORMAÇÃO NÃO TEM

- **Comunidade real:** redes de confiança que a fabricação não consegue replicar.
- **Causa verificável:** pautas que afetam vidas concretas como habitação, saúde, educação, segurança.
- **Credibilidade construída:** histórico de luta e entrega que a desinformação não pode fabricar.

Seção 1 Os 6 princípios da comunicação progressista eficaz

1. **Pautar e não apenas reagir:** cada vez que reagimos a uma provocação, validamos a pauta da direita. Comece com a pergunta: o que eu quero que as pessoas pensem, sintam e façam, independentemente do que a direita disser hoje? Que pauta é realmente relevante para o meu público olhar hoje?

2. **Enraizamento comunitário:** a comunicação mais eficaz vem de quem está na comunidade. Saiba dialogar com os anseios do seu público. Mulheres Negras Decidem é um exemplo de comunicação que parte de quem vive a política no corpo, no território, no cotidiano.

3. **Emoção legítima, não manipulação:** comunicar com emoção real (orgulho, cuidado, solidariedade, justa raiva) é mais sustentável e ético do que criar medo artificial. O que move as pessoas a longo prazo é a identificação, não o medo.

4. **Simplicidade sem simplismo:** encontrar a entrada mais simples para temas complexos. "O que isso muda na vida de quem acorda cedo?" organiza a comunicação sem perder profundidade. As pessoas querem entender as coisas, mas não se esqueça que elas também estão imersas em feeds intermináveis de disputa de atenção.

5. **Consistência no tempo:** a direita usa repetição como arma. O campo progressista tende a se comunicar em picos reativos e silenciar. Narrativa consistente tem mais impacto do que comunicação explosiva e descontínua.

6. **Representação como mensagem:** a comunicação é composta por emissor, mensagem e receptor. Nenhum símbolo se faz sozinho. Estimule a imaginação e a crítica no seu público. Quem fala é tão importante quanto o que se fala. O Instituto Alziras afirma que mulheres negras em candidaturas representam "uma imaginação política radical que propõe novos caminhos para a democracia".

Seção 2 Framing: como o enquadramento define a batalha

Framing é a escolha de qual ângulo usar para apresentar um tema. Muitas vezes, a extrema direita tem sido sistematicamente melhor em framing do que o campo progressista, e quando negamos um frame, frequentemente o reforçamos.

TEMA	FRAME DA DIREITA	REENQUADRAMENTO | PROGRESSISTA
Educação sexual	"Doutrinação ideológica nas escolas"	Proteger crianças com informação, para que reconheçam abuso
Cotas raciais	"Privilégio para uns - discriminação para outros"	Correção de 400 anos de exclusão sistemática rumo à igualdade real
Direitos LGBTQIAPN+	"Agenda gay", "ataque a família"	Proteger pessoas reais de violência real: é garantir o direito de existir
Políticas sociais	"Gasto público desnecessário"	Investimento que retorna: menos doença, menos crime, mais economia
Mulheres na política	"Cota política", "politicamente correto"	Democracia real precisa representar a maioria da população

Seção 3 Formatos: escolhendo uma forma de comunicar

Ninguém acha que a Glória Maria ou o William Bonner poderiam apresentar o mesmo programa que a Hebe ou o Faustão. Então, por que nas redes sociais tentamos plastificar o conteúdo sem considerar características próprias de quem está produzindo? Na tentativa de viralizar, caímos na lógica de replicar trends que podem ser desconfortáveis ou desajustadas. É preciso considerar a linguagem da internet, especialmente as orientações de boas práticas das plataformas (como não postar vídeos com mais de 3 minutos no Instagram), mas também é preciso manter sua autenticidade.

ALGUNS FORMATOS COMUNS

REACT - FALA POVO - VLOG - 1X1 COM A CÂMERA - TRENDS - CARROSSÉIS INFORMACIONAIS - TUÍTES

PASSO 1
CONSTRUA REPERTÓRIO

Boas produções acontecem a partir de boas referências. Que tal criar uma pasta ou um perfil só para reunir conteúdos com os quais você se identifica e te servem de inspiração para reproduzir? Também é importante não manter sua cabeça só em produtores de conteúdo do seu nicho. Existe muita gente produzindo bons conteúdos em outras áreas (moda, saúde, audiovisual, humor e por aí vai) e até em outros países. Não se limite.

PASSO 2
ELENQUE REFERÊNCIAS A PARTIR DA SUA REALIDADE

É claro que um vídeo com uma edição super elaborada vai chamar a atenção. Mas se você não tem condições de fazer algo parecido, isso deve ficar em uma caixinha das referências para o futuro, não para o agora. Escolha alguns formatos que você se vê fazendo, seja por disponibilidade material (espaço adequado para gravar, cenário, apps de edição) ou por características suas (se você não se sente bem falando diretamente com a câmera, por que segue tentando fazer isso?).

PASSO 3
FAÇA TESTES E FALE COM A SUA AUDIÊNCIA

A partir das referências elencadas, escolha duas ou três para fazer testes. Tente reproduzir pensando no seu nicho, com o seu conteúdo e as suas características. Anote também quanto tempo você levou para executar cada etapa da produção e pense se esse tempo será sustentável dentro da sua rotina. Se algo levou um mês para ficar pronto, por mais legal que seja, provavelmente não irá se encaixar em uma produção constante para redes sociais. Depois de postado, tente conversar com a sua audiência sobre o conteúdo. Pergunte para algumas pessoas mais próximas em quem você confia se elas gostaram, o que poderia melhorar, se a mensagem está sendo passada.

PASSO 4
MONITORE E PRODUZA DADOS

Observe os resultados da postagem, sejam as métricas das redes sociais ou métricas subjetivas, como a aceitação do seu público. Guarde isso, de preferência em algum documento, e tire uns minutinhos para pensar nos resultados obtidos. Se o conteúdo foi mal, o que pode ter interferido? Horário ruim? Em descompasso com o tema da rede no momento? Falta de ritmo? Muito longo? Uso de palavras com ban? Se for possível, republique o conteúdo mudando algumas coisas, para ter base de comparação. Depois de alguns testes, se você concluir que o formato não performa bem, é hora de partir pra outra.

PASSO 5
ACHEI ALGO QUE FUNCIONA PRA MIM, O QUE FAZER?

Com certeza existe um formato que funciona para você e para o seu público. Lembre-se de que, além de likes, ele precisa gerar conexão com as pessoas e também ser executável no seu dia a dia. Produção de conteúdo é trabalho. Organize sua rotina para ter uma produção constante e diária, incluindo tirar alguns minutos para medir o impacto.

Seção 4 Disseminação: passar pra frente também faz parte da mensagem

Produzir um conteúdo e contar só com a boa vontade dos algoritmos para que ele chegue nas pessoas é arriscado, por melhor que seja o seu vídeo. Por isso, é importante que sua rotina também inclua algumas outras tarefas:

- Se for vídeo, poste em reels de teste com algumas alterações
- Mapeie perfis que podem querer publicar em collab com você
- Prepare uma correntinha no WhatsApp ou Telegram para enviar o link para as pessoas
- Fale em atividades ao vivo sobre o conteúdo que você está produzindo (“essa semana mesmo fiz um vídeo no meu Instagram sobre essa situação, vejam lá, @fulana de tal”)

Esta cartilha foi feita com a convicção de que entender o tabuleiro é o primeiro passo para disputá-lo e que desmontar a máquina começa por conhecer suas engrenagens. Nossa intenção foi oferecer ferramentas para pensar e agir, não fórmulas prontas, para que candidaturas, mandatos e organizações progressistas construam comunicação mais estratégica, mais enraizada e mais resistente aos ataques.

Mas o trabalho não termina aqui. Termina quando a informação vira ação. Se usar este material em cursos, formações ou processos internos da sua organização, conta pra gente?

Queremos saber como ela foi usada,
o que funcionou e o que falta

lampejo.digital

FONTES

AGÊNCIA LUPA — OBSERVATÓRIO DA DESINFORMAÇÃO. I Panorama da Desinformação no Brasil: padrões, estratégias e impacto. Rio de Janeiro: Agência Lupa, fev. 2026. Disponível em: https://uploads.agencialupa.org/2026/02/Panorama_da_Desinformacao_no_Brasil_LUPA-1.pdf. Acesso em: maio 2026.

ALAFIA LAB. Abaixo do radar: desinformação em grupos de extrema direita no WhatsApp e no Telegram nas eleições de 2024. Autoras: Andressa Costa, Ellen Guerra, Inara Almeida, Lizete Nóbrega, Natália Silva, Nina Santos, Patrícia Santos, Tatiana Dourado, Viktor Chagas. Coordenação: Rodrigo Carreiro e Maria Paula Almada. Salvador: Aláfia Lab; coLab; Instituto Democracia em Xeque, 2025. Disponível em: https://alafialab.org/wp-content/uploads/2025/01/Abaixo-do-radar-Desinformacao-em-grupos-de-extrema-direita-no-WhatsApp-e-no-Telegram-nas-eleicoes-de-2024.pdf. Acesso em: maio 2026.

ALAFIA LAB. Desigualdades informativas: entendendo os caminhos informativos dos brasileiros na internet 2024. Salvador: Aláfia Lab, 2025. Disponível em: https://alafialab.org/wp-content/uploads/2025/05/Desigualdades-informativas-2024.pdf. Acesso em: maio 2026.

ALAFIA LAB. Desigualdades informativas: entendendo os hábitos de consumo de informação dos brasileiros 2025. Autoras: Ellen Guerra, Larisse Pontes, Lizete Nóbrega, Vivian Peron. Direção executiva: Maria Paula Almada e Rodrigo Carreiro. Salvador: Aláfia Lab, 2025. Disponível em: https://alafialab.org/wp-content/uploads/2025/12/Desigualdades-Informativas-2025-.pdf. Acesso em: maio 2026.

ALAFIA LAB. Polarização política e consumo de informação no Brasil: uma análise sobre gênero, idade, classe e escolaridade. Autora: Vivian Peron. Coordenação de pesquisa: Rodrigo Carreiro e Maria Paula Almada. Salvador: Aláfia Lab, 2025. Disponível em: https://alafialab.org/wp-content/uploads/2026/01/Polarizacao-politica-e-consumo-de-informacao-no-Brasil-Uma-analise-sobre-genero-idade-classe-e-escolaridade-1-1.pdf. Acesso em: maio 2026.

ALAFIA LAB; ARTIGO 19; IRIS. Nota técnica sobre a Resolução TSE 23.755/2026. mar. 2026. Disponível em: https://alafialab.org. Acesso em: maio 2026.

ALVES, Marcelo; Ferreira, Douglas da Silva. Rotulando IA nas eleições. Boletim #01 – 01 a 14 de março de 2026. Instituto Democracia em Xeque, 2026. Disponível em: https://institutodx.org/rotulandoia/14032026/ Acesso em: maio 2026.

ANLEN, Shirin; VÁZQUEZ LLORENTE, Raquel. Spotting deepfakes in an election year: how AI detection tools work and where they fail. Reuters Institute for the Study of Journalism / WITNESS. Publicado também em: GIJN, abr. 2024. Disponível em: https://reutersinstitute.politics.ox.ac.uk/news/spotting-deepfakes-year-elections-how-ai-detection-tools-work-and-where-they-fail. Acesso em: maio 2026.

AZMINA. Tecnologia e redes de acolhimento enfrentam a violência política digital de gênero. Autora: Natali Carvalho. 22 abr. 2026. Disponível em: https://azmina.com.br/reportagens/tecnologia-redes-acolhimento-enfrentam-violencia-politica-digital-genero. Acesso em: maio 2026.

BRASIL. Lei nº 14.192, de 4 de agosto de 2021. Estabelece normas para prevenir, reprimir e combater a violência política contra a mulher. Brasília: Presidência da República, 2021. Disponível em: https://www.planalto.gov.br/ccivil_03/_ato2021-2024/2021/lei/l14192.htm. Acesso em: maio 2026.

BRASIL. Advocacia-Geral da União. Observatório da Democracia vai elaborar estudo sobre IA para as eleições. Brasília: AGU, abr. 2026. Disponível em: https://www.gov.br/agu/pt-br/comunicacao/noticias/observatorio-da-democracia-vai-elaborar-estudo-sobre-ia-para-as-eleicoes. Acesso em: maio 2026.

BRASIL. Tribunal Superior Eleitoral. Resolução nº 23.610, de 18 de dezembro de 2019. Dispõe sobre propaganda eleitoral, utilização e geração do horário eleitoral gratuito e condutas ilícitas em campanha eleitoral. Brasília: TSE, 2019. Disponível em: https://www.tse.jus.br/legislacao/compilada/res/2019/resolucao-no-23-610-de-18-de-dezembro-de-2019. Acesso em: maio 2026.

BRASIL. Tribunal Superior Eleitoral. Resolução nº 23.732, de 27 de fevereiro de 2024. Altera a Resolução nº 23.610/TSE, de 18 de dezembro de 2019, que dispõe sobre propaganda eleitoral, introduzindo a regulação do uso de inteligência artificial nas eleições municipais de 2024. Brasília: TSE, 2024. Disponível em: https://www.tse.jus.br/legislacao/compilada/res/2024/resolucao-no-23-732-de-27-de-fevereiro-de-2024. Acesso em: maio 2026.

BRASIL. Tribunal Superior Eleitoral. Resolução nº 23.755, de 2 de março de 2026. Altera a Resolução nº 23.610/TSE, de 18 de dezembro de 2019, que dispõe sobre propaganda eleitoral, ampliando a regulação sobre inteligência artificial para as eleições gerais de 2026. Brasília: TSE, 2026. Disponível em: https://www.tse.jus.br/legislacao/compilada/res/2026/resolucao-no-23-755-de-2-de-marco-de-2026. Acesso em: maio 2026.

BRASIL. Tribunal Superior Eleitoral. Por dentro das eleições: conheça as regras sobre uso de IA na campanha eleitoral de 2026. Brasília: TSE, abr. 2026. Disponível em: https://www.tse.jus.br/comunicacao/noticias/2026/Abril/por-dentro-das-eleicoes-conheca-as-regras-sobre-uso-de-ia-na-campanha-eleitoral-de-2026. Acesso em: maio 2026.

CANALTECH. Instagram vai reorganizar menu do app e facilitar acesso aos Reels, 2025. Disponível em: https://canaltech.com.br/apps/instagram-vai-reorganizar-menu-do-app-e-facilitar-acesso-aos-reels/. Acesso em: maio 2026.

CETIC.BR. TIC Domicílios 2024: análises. São Paulo: NIC.br, 2025. Disponível em: https://cetic.br/pt/pesquisa/domicilios/analises/. Acesso em: maio 2026.

CODING RIGHTS. Dados como ferramenta de influência política nas eleições brasileiras. Autora: Joana Varon. [S.l.]: Coding Rights; Tactical Tech Collective, 2018. Disponível em: https://codingrights.org/en/library-item/data-as-a-tool-for-political-influence-in-the-brazilian-elections. Acesso em: maio 2026.

CODING RIGHTS. Internet e eleições: guia para proteção de direitos nas campanhas eleitorais. Autoras: Ladyane Souza e Joana Varon. Coding Rights; Coalizão Direitos na Rede; Rede Transfeminista de Cuidados Digitais, set. 2020. Disponível em: https://codingrights.org/docs/eleicoes&internet.pdf. Acesso em: maio 2026.

CODING RIGHTS. Visibilidade sapatão nas redes: entre violência e solidariedade. Autoras: Ivanilda Figueiredo e Joana Varon.Coding Rights, 2020. Disponível em: https://codingrights.org/docs/visibilidade_sapatao.pdf. Acesso em: maio 2026.

DANTAS, Glenda; RUDNITZKI, Ethel. Dois em cada três conteúdos políticos com IA circulam sem aviso nas redes. Observatório IA nas Eleições / Desinformante, abr. 2026. Disponível em: https://desinformante.com.br. Acesso em: maio 2026.

DATAREPORTAL. Digital 2025: Brazil, 2025. Disponível em: https://datareportal.com/reports/digital-2025-brazil. Acesso em: maio 2026.

DESINFORMANTE. IAí? Inteligência artificial e desinformação.Desinformante, 2026. Disponível em: https://desinformante.com.br/observatorio-ia. Acesso em: maio 2026.

DESINFORMANTE; ALAFIA LAB. Qual é o cenário da desinformação? Desinformante, [2026]. Disponível em: https://desinformante.com.br/cenario. Acesso em: maio 2026.

DIP, Andrea; VARON, Joana; CALDEIRA NETO, Odilon; CAPONE, Letícia; PECORARO, Caroline; BERNARDI, Ana Julia; FRANCO, Agnes; DOURADO, Tatiana; GARRIDO, Fabiano. Democracia sob pressão: reflexões sobre a extrema direita com as chaves do passado, presente e futuro. Organização: Marilene de Paula e Manoela Vianna. Rio de Janeiro: Fundação Heinrich Böll Brasil, nov. 2025. 64 p. (Série 25 Anos). ISBN 978-65-87665-26-9. Disponível em: https://br.boell.org. Acesso em: maio 2026.

ESTADÃO. Instagram se consolida como habitat da IA e políticos disputam engajamento com perfis sombrios, 2026. Disponível em: https://www.estadao.com.br/politica/instagram-se-consolida-como-habitat-da-ia-e-politicos-disputam-engajamento-com-perfis-sombrios/. Acesso em: maio 2026.

FOLHA DE S.PAULO. Campanhas ampliam ataques e alimentam guerra judicial após IA inundar redes. São Paulo, maio 2026. Disponível em: https://www1.folha.uol.com.br/poder/2026/05/campanhas-ampliam-ataques-e-alimentam-guerra-judicial-apos-ia-inundar-redes.shtml. Acesso em: maio 2026.

GEGENHUBER, Gabriel K.; FRENZEL, Philipp É.; GÜNTHER, Maximilian; ULLRICH, Johanna; JUDMAYER, Aljosha. Hey there! You are using WhatsApp: enumerating three billion accounts for security and privacy. Viena: University of Vienna; SBA Research, 2025. Preprint. Disponível em: https://arxiv.org/pdf/2511.20252. Acesso em: maio 2026.

GIJN — GLOBAL INVESTIGATIVE JOURNALISM NETWORK. How to identify and investigate AI audio deepfakes, a major election threat. Autor: Rowan Philp. 26 fev. 2024. Disponível em: https://gijn.org/resource/tipsheet-investigating-ai-audio-deepfakes. Acesso em: maio 2026.

INSTITUTO ALZIRAS. Convergências Democráticas América Latina. Instituto Alziras, [2026]. Disponível em: https://convergenciasdemocraticas.org. Acesso em: maio 2026.

INSTITUTO ALZIRAS; OBSERVATÓRIO NACIONAL DE MULHERES NA POLÍTICA. Monitor de Violência Política de Gênero e Raça 2025. 2. ed. Instituto Alziras, ago. 2025. Disponível em: https://alziras.org.br. Acesso em: maio 2026.

INSTITUTO DEMOCRACIA EM XEQUE. Rotulando IA nas eleições 2026. São Paulo: Instituto Democracia em Xeque, 2026. Disponível em: https://institutodx.org/rotulandoia/. Acesso em: 16 maio 2026.

INSTITUTO DEMOCRACIA EM XEQUE. Semanal DX: relatório semanal de narrativas políticas e integridade democrática. São Paulo: Instituto Democracia em Xeque, 2026. Publicação semanal. Disponível em: https://institutodx.org/semanaldx. Acesso em: maio 2026.

INSTITUTO DEMOCRACIA EM XEQUE. Semanal DX (05.05.2026): relatório semanal de narrativas políticas e integridade democrática. Período de análise: 27 de abril a 04 de maio de 2026. São Paulo: Instituto Democracia em Xeque, 5 maio 2026. Disponível em: https://institutodx.org/semanaldx/05052026. Acesso em: maio 2026.

INSTITUTE FOR STRATEGIC DIALOGUE. Coordinated disinformation network uses AI, media impersonation to target German election, 2025. Disponível em: https://www.isdglobal.org/digital-dispatch/coordinated-disinformation-network-uses-ai-media-impersonation-to-target-german-election/. Acesso em: maio 2026.

INSTITUTO MARIELLE FRANCO; JUSTIÇA GLOBAL; TERRA DE DIREITOS. Regime de ameaça: violência política de gênero e raça no âmbito digital, 2025. Disponível em: https://violenciapolitica.org. Acesso em: maio 2026.

LACERDA, Elisa; KUCURUZA, Gabri. Fraudes: conhecendo o mundo da publicidade programática e o financiamento da desinformação. Sleeping Giants Brasil, 7 abr. 2026. Disponível em: https://sleepinggiantsbrasil.com/fraudes-conhecendo-o-mundo-da-publicidade-programatica-e-o-financiamento-da-desinformacao/. Acesso em: maio 2026.

LEGIS-ATIVO, Coletivo; MARUCI, Hannah. IA, eleições e o desafio da violência política de gênero e raça. Congresso em Foco, 15 jul. 2025. Disponível em: https://www.congressoemfoco.com.br/coluna/110225/ia-eleicoes-e-o-desafio-da-violencia-politica-de-genero-e-raca. Acesso em: maio 2026.

MARIALAB. Maria d'Ajuda: linha de ajuda em segurança digital para mulheres, pessoas não-binárias e LGBTQIAP+. Salvador: MariaLab, 2026. Serviço gratuito. Disponível em: https://mariadajuda.org. Contato: sos@mariadajuda.org. Acesso em: maio 2026.

MEIO & MENSAGEM. Aplicação de IA se dissemina com ferramentas das plataformas, 2026. Disponível em: https://www.meioemensagem.com.br/midia/aplicacao-de-ia-se-dissemina-com-ferramentas-das-plataformas. Acesso em: maio 2026.

MELLO, Patrícia Campos. A máquina do ódio: notas de uma repórter sobre fake news e violência digital. São Paulo: Companhia das Letras, 2020.

MULHERES NEGRAS DECIDEM. Produção de áudio e comunicação política, 2026. Disponível em: https://soundcloud.com/mulheresnegrasdecidem. Acesso em: maio 2026.

SANTINI, R. Marie; SALLES, Débora; BELIN, Luciane L; BELISÁRIO, Adriano; MATTOS, Bruno; MEDEIROS, Stéphanie G.; MELLO, Danielle; GRAEL, Felipe; SEADE, Renata; BORGES, Amanda; MURAKAMI, Lucas; CARDOSO, Rafael; DAU, Erick; LOUREIRO, Felipe;YONESHIGUE, Bernardo; CARMO, Vitor do; MAIA, Felipe. "Aprenda a evitar 'esse tipo' de mulher": estratégias discursivas e monetização da misoginia no YouTube. Rio de Janeiro: NetLab – Laboratório de Estudos de Internet e Redes Sociais, Universidade Federal do Rio de Janeiro (UFRJ). Publicado em Dezembro de 2024. Acesso em: maio 2026.

NETLAB UFRJ. Anúncios com IA usam imagem de políticos brasileiros para aplicar golpes. Rio de Janeiro: NetLab, Universidade Federal do Rio de Janeiro, 17 jun. 2024. Disponível em: https://netlab.eco.ufrj.br/post/anúncios-com-ia-usam-imagem-de-políticos-brasileiros-para-aplicar-golpes. Acesso em: 16 maio 2026.

NETLAB UFRJ; MINISTÉRIO DAS MULHERES. "Aprenda a evitar 'esse tipo' de mulher": estratégias discursivas e monetização da misoginia no YouTube. Rio de Janeiro: NetLab UFRJ, 2024. Atualização: mar. 2026. Disponível em: https://netlab.eco.ufrj.br. Acesso em: maio 2026.

NETLAB UFRJ; MINISTÉRIO DAS MULHERES. "Aprenda a evitar 'esse tipo' de mulher": OBSERVATÓRIO IA NAS ELEIÇÕES. Repositório de casos de uso de IA generativa nas eleições brasileiras. [S.l.]: Data Privacy Brasil; Alafia Lab; Desinformante, 2025–2026. Disponível em: https://observatorioianaseleicoes.com.br. Acesso em: maio 2026.

REIS, Sara. Inteligência artificial nas eleições: veja o que ficou decidido pelo TSE. Senado Verifica, Brasília, 6 mar. 2026. Disponível em: https://www12.senado.leg.br/verifica/materias-especiais/2026/inteligencia-artificial-nas-eleicoes-veja-o-que-ficou-decidido-pelo-tse. Acesso em: maio 2026.

PROJETO COMPROVA. Preta Gil não escreveu carta a Bolsonaro; vídeo foi gerado por inteligência artificial. Verificado por Correio Braziliense, Estadão e NSC Comunicação. 1 ago. 2025. Disponível em: https://projetocomprova.com.br/publica%C3%A7%C3%B5es/preta-gil-nao-escreveu-carta-a-bolsonaro-video-foi-gerado-por-inteligencia-artificial/. Acesso em: maio 2026.

Revista Fórum: Dona Maria: avatar de IA que ataca Lula já teve mais de 100 milhões de visualizações. Revista Fórum, abr. 2026. Disponível em: https://revistaforum.com.br/revista-forum/dona-maria-avatar-ia-tse. Acesso em: maio 2026.

SADI, Andréia. Caso Dona Maria: ministros do TSE avaliam que regras sobre IA nas eleições não são suficientes. G1 Blog da Andréia Sadi, 27 abr. 2026. Disponível em: https://g1.globo.com/politica/blog/andreia-sadi/post/2026/04/27/caso-dona-maria-ministros-do-tse-avaliam-que-regras-sobre-ia-nas-eleicoes-nao-sao-suficientes.ghtml. Acesso em: maio 2026.

SLEEPING GIANTS BRASIL. Saiba como foi a participação brasileira no maior evento da extrema direita global - CPAC 2026. Sleeping Giants Brasil, mar. 2026. Disponível em: https://sleepinggiantsbrasil.com. Acesso em: maio 2026.

SLEEPING GIANTS BRASIL. Fraudes: conhecendo o mundo da publicidade programática e o financiamento da desinformação. Sleeping Giants Brasil, 7 abr. 2026. Disponível em: http://sleepinggiantsbrasil.com/fraudes-conhecendo-o-mundo-da-publicidade-programatica-e-o-financiamento-da-desinformacao. Acesso em: maio 2026.

TECNOBLOG. Virou TikTok: Instagram muda interface e coloca feed de Reels na home, 2026. Disponível em: https://tecnoblog.net/noticias/virou-tiktok-instagram-muda-interface-e-coloca-feed-de-reels-na-home/. Acesso em: maio 2026.

THE ALAN TURING INSTITUTE — CETaS. From deepfake scams to poisoned chatbots: AI and election security in 2025. Autor: Sam Stockwell. CETaS Expert Analysis. Londres: Centre for Emerging Technology and Security (CETaS), 17 nov. 2025. Disponível em: https://cetas.turing.ac.uk/publications/deepfake-scams-poisoned-chatbots. Acesso em: maio 2026.

WE ARE SOCIAL; MELTWATER. Digital 2025 Global Overview Report, 2025. Disponível em: https://wearesocial.com/wp-content/uploads/2025/02/GDR-2025-v2.pdf. Acesso em: maio 2026.

WITNESS. Deepfake Rapid Response Force. Nova York: WITNESS, 2024. Disponível em: https://www.gen-ai.witness.org/deepfakes-rapid-response-force. Acesso em: maio 2026.

Imaginar é criar futuros

lampejo.digital

lampejo.digital