LUMEN ET VIRTUS
ISSN: 2177-2789

A DUPLA HÉLICE DA INTELIGÊNCIA ARTIFICIAL GENERATIVA: VETOR E ANTÍDOTO PARA A DESINFORMAÇÃO NO BRASIL

THE DOUBLE HELIX OF GENERATIVE ARTIFICIAL INTELLIGENCE: VECTOR AND ANTIDOTE TO DISINFORMATION IN BRAZIL

LA DOBLE HÉLICE DE LA INTELIGENCIA ARTIFICIAL GENERATIVA: VECTOR Y ANTÍDOTO CONTRA LA DESINFORMACIÓN EN BRASIL

https://doi.org/10.56238/levv16n52-044

Data de submissão: 16/08/2025 **Data de publicação:** 16/09/2025

Márcio Rodrigo Elias Carvalho
Especialista em Inteligência Artificial
Instituição: FASUL Educacional
E-mail: marcior@gmail.com

Lucyneid Barros Carvalho
Especialista em Neuropsicopedagogia
Instituição: UNINASSAU Aracaju
E-mail: lucyneidbarros@gmail.com

Sérgio Ferreira da Silva
Mestre em Psicanálise
Instituição: FASUL Educacional
E-mail: admsergioferreira@gmail.com

José Dirnece Paes Tavares
Mestre em Gestão e Desenvolvimento Regional
Instituição: FASUL Educacional
E-mail: dirnece@gmail.com

Antônio Ribas Reis
Doutor em Difusão do Conhecimento
Instituição: FASUL Educacional
E-mail: amigoribas@gmail.com

Flavia Chaves Valentim Rodrigues
Mestre em Desenvolvimento Regional
Instituição: FASUL Educacional
E-mail: profa.flaviavalentim@gmail.com

Vera Lucia da Silva Farias
Doutora em Ciências do Solo – Agronomia
Instituição: FASUL Educacional
E-mail: professoraveralucia@fasuleducacional.edu.br

Rogério dos Santos Morais
Doutor em Engenharia
Instituição: FASUL Educacional
E-mail: roger.dos.santos.morais@gmail.com

RESUMO

A ascensão da Inteligência Artificial Generativa (IAG) reconfigurou drasticamente o ecossistema de comunicação e informação, instaurando um paradoxo central: a mesma tecnologia que permite a criação e disseminação de desinformação em escala e sofisticação sem precedentes também oferece as ferramentas mais promissoras para sua detecção e mitigação. Este artigo de revisão analisa criticamente esta dualidade, que se assemelha a uma dupla hélice, com foco no contexto brasileiro. A metodologia adotada foi a revisão sistemática de literatura, abrangendo artigos acadêmicos, relatórios técnicos e documentos legislativos para mapear as tendências e desafios atuais. Os resultados demonstram, por um lado, como tecnologias como Redes Adversariais Generativas (GANs) e Grandes Modelos de Linguagem (LLMs) são instrumentalizadas para a produção de deepfakes e narrativas falsas, impactando processos democráticos e a confiança pública. Por outro lado, evidencia-se o potencial da IA como ferramenta de fact-checking automatizado e análise de padrões de desinformação. A análise aprofunda-se no debate regulatório brasileiro, por meio de um estudo comparativo entre o Projeto de Lei nº 2338/2023 e o AI Act da União Europeia, revelando convergências e desafios de implementação. Conclui-se que a navegação segura na nova infosfera híbrida demanda uma abordagem multifacetada, que integre uma regulação tecnológica robusta, a adaptação ética das práticas comunicacionais e um investimento massivo em literacia midiática e algorítmica para a sociedade.

Palavras-chave: Inteligência Artificial Generativa. Desinformação. Comunicação Digital. Regulamentação. Jornalismo.

ABSTRACT

The rise of Generative Artificial Intelligence (GAI) has drastically reconfigured the communication and information ecosystem, establishing a central paradox: the same technology that enables the creation and dissemination of disinformation on an unprecedented scale and sophistication also offers the most promising tools for its detection and mitigation. This review article critically analyzes this duality, which resembles a double helix, focusing on the Brazilian context. The methodology adopted was a systematic literature review, encompassing academic articles, technical reports, and legislative documents, to map current trends and challenges. The results demonstrate, on the one hand, how technologies such as Generative Adversarial Networks (GANs) and Large Language Models (LLMs) are instrumentalized to produce deepfakes and false narratives, impacting democratic processes and public trust. On the other hand, it highlights the potential of AI as a tool for automated fact-checking and analysis of disinformation patterns. The analysis delves into the Brazilian regulatory debate through a comparative study between Bill No. 2338/2023 and the European Union's AI Act, revealing similarities and implementation challenges. The conclusion is that safe navigation in the new hybrid infosphere requires a multifaceted approach that integrates robust technological regulation, ethical adaptation of communication practices, and massive investment in media and algorithmic literacy for society.

Keywords: Generative Artificial Intelligence. Disinformation. Digital Communication. Regulation. Journalism.

RESUMEN

El auge de la Inteligencia Artificial Generativa (IAG) ha reconfigurado drásticamente el ecosistema de la comunicación y la información, estableciendo una paradoja central: la misma tecnología que permite la creación y difusión de desinformación a una escala y sofisticación sin precedentes también ofrece

las herramientas más prometedoras para su detección y mitigación. Este artículo de revisión analiza críticamente esta dualidad, que se asemeja a una doble hélice, centrándose en el contexto brasileño. La metodología adoptada fue una revisión sistemática de la literatura, que abarcó artículos académicos, informes técnicos y documentos legislativos, para mapear las tendencias y los desafíos actuales. Los resultados demuestran, por un lado, cómo tecnologías como las Redes Generativas Antagónicas (GAN) y los Grandes Modelos de Lenguaje (LLM) se instrumentalizan para producir deepfakes y narrativas falsas, lo que impacta los procesos democráticos y la confianza pública. Por otro lado, destaca el potencial de la IA como herramienta para la verificación automatizada de hechos y el análisis de patrones de desinformación. El análisis profundiza en el debate regulatorio brasileño mediante un estudio comparativo entre el Proyecto de Ley n.º 2338/2023 y la Ley de IA de la Unión Europea, revelando similitudes y desafíos de implementación. La conclusión es que la navegación segura en la nueva infoesfera híbrida requiere un enfoque multifacético que integre una sólida regulación tecnológica, la adaptación ética de las prácticas de comunicación y una inversión masiva en alfabetización mediática y algorítmica para la sociedad.

Palabras clave: Inteligencia Artificial Generativa. Desinformación. Comunicación Digital. Regulación. Periodismo.

1 INTRODUÇÃO

O cenário comunicacional contemporâneo é definido por uma profunda e contínua transformação, impulsionada pela onipresença das tecnologias digitais em rede (FOLETTO, 2024). A consolidação das plataformas de mídia social como principais arenas de interação e troca informacional alterou de maneira irreversível os paradigmas tradicionais de comunicação, reconfigurando a produção, a distribuição e o consumo de conteúdo em escala global (RECUERO et al., 2025). Este novo ecossistema, caracterizado pela velocidade e acessibilidade, abriu portas para novas formas de interação, mas também introduziu desafios complexos, notadamente a disseminação de informações imprecisas ou deliberadamente falsas (RECUERO et al., 2025; MENDES; MATTOS, 2025).

Nesse contexto já volátil, o advento e a popularização da Inteligência Artificial Generativa (IAG) representam um ponto de inflexão crítico. Ferramentas como ChatGPT, DALL-E, Midjourney e outras, que se baseiam em Grandes Modelos de Linguagem (LLMs) e Redes Adversariais Generativas (GANs), migraram de ambientes de pesquisa restritos para o domínio público, desencadeando uma disrupção em múltiplos setores, com um impacto particularmente agudo no campo da Comunicação e Informação (GOMES; OLIVEIRA, 2024; SPINAK, 2023; BLIKSTEIN; FERNANDES; COUTINHO, 2024). A capacidade dessas tecnologias de gerar textos, imagens e áudios sintéticos com um grau de realismo cada vez maior introduz uma nova camada de complexidade ao ambiente informacional.

Este artigo parte da premissa de que a IAG opera como uma "dupla hélice" no ecossistema comunicacional, um paradoxo tecnológico com implicações profundas para a sociedade. A primeira hélice representa a capacidade da IAG de degradar a esfera pública, funcionando como um potente catalisador para a produção e disseminação de desinformação. A tecnologia permite a criação em massa de conteúdo falso, sofisticado e hiper-personalizado, desde textos que mimetizam o estilo jornalístico até *deepfakes* audiovisuais quase indistinguíveis da realidade, constituindo uma ameaça sem precedentes à integridade dos processos democráticos, à confiança nas instituições e ao próprio tecido social (GOLDSTEIN; LOHN, 2024; CETAS, 2024; FRONTIERS, 2025).

Em contrapartida, a segunda hélice representa o potencial de salvaguarda. A mesma base tecnológica — o aprendizado de máquina e o processamento de linguagem natural — oferece as ferramentas mais avançadas e escaláveis para detectar, analisar e mitigar os fluxos de desinformação. Sistemas de IA podem identificar padrões, verificar alegações e rastrear a origem de narrativas falsas com uma eficiência que transcende a capacidade humana, tornando-se aliados cruciais para jornalistas, checadores de fatos e pesquisadores (BONTRIDDER; POULLET, 2021; CHORAŚ et al., 2021; GRAVES, 2018; YANG; MENCZER, 2023).

Diante deste cenário paradoxal, o presente artigo tem como objetivo realizar uma revisão crítica e sistemática da literatura para analisar a dualidade da IAG como vetor e, simultaneamente, antídoto para a desinformação, com um foco específico no contexto brasileiro. A justificativa para este estudo reside na urgência do debate, alimentada pela tramitação de marcos regulatórios fundamentais, como o Projeto de Lei nº 2338/2023, e pelos desafios concretos já observados em eventos de grande relevância nacional, como os processos eleitorais recentes (GOMES; OLIVEIRA, 2024; DATA PRIVACY BRASIL, 2024).

Para atingir tal objetivo, o artigo está estruturado da seguinte forma: a seção de Metodologia detalha o delineamento da pesquisa como uma revisão de literatura. A seção de Resultados explora as duas facetas do paradoxo algorítmico, primeiro dissecando os mecanismos técnicos da desinformação sintética e seus impactos, e depois analisando o potencial da IA como ferramenta de mitigação, culminando com uma análise do cenário regulatório brasileiro. A seção de Discussão aprofunda as implicações éticas para o jornalismo e avalia criticamente as respostas legislativas. Por fim, a Conclusão sintetiza os principais argumentos e propõe uma agenda para pesquisas futuras.

2 METODOLOGIA

Este estudo caracteriza-se como um artigo de revisão sistemática de literatura, modalidade de pesquisa que visa sintetizar e analisar criticamente o conhecimento acumulado sobre um determinado tema a partir de fontes secundárias. O objetivo é mapear o estado da arte, identificar tendências, lacunas e controvérsias no campo de estudo da interseção entre Inteligência Artificial Generativa, desinformação e comunicação, com especial atenção ao contexto brasileiro.

O corpus de análise para esta revisão foi constituído por um conjunto multifacetado de documentos, selecionados para proporcionar uma visão abrangente e aprofundada do fenômeno. As fontes foram agrupadas nas seguintes categorias:

1. Artigos Acadêmicos e Científicos: Foram analisadas publicações de periódicos indexados em bases de dados como SciELO e anais de congressos relevantes da área de Comunicação no Brasil, como os da Associação Nacional dos Programas de Pós-Graduação em Comunicação (Compós) e da Associação Brasileira de Pesquisadores em Jornalismo (SBPJor). Esses documentos fornecem a base teórica e empírica sobre as tendências de pesquisa, os impactos da plataformização e as discussões epistemológicas no campo (FOLETTO, 2024; RECUERO et al., 2025; MENDES; MATTOS, 2025; TELLES; MONTARDO, 2025; HJARVARD, 2014).

2. Relatórios Técnicos e Análises de Especialistas: Foram incluídos relatórios e publicações de *think tanks*, institutos de pesquisa e organizações não governamentais (nacionais e internacionais) que se dedicam ao monitoramento do impacto da IA na

sociedade. Esses materiais oferecem análises aprofundadas sobre os mecanismos das campanhas de desinformação e estudos de caso sobre o uso de IA em contextos eleitorais (CETAS, 2024; ADA LOVELACE INSTITUTE, 2024).

3. Documentos Legislativos e Análises Jurídicas: O corpus inclui o texto do Projeto de Lei nº 2338/2023, que visa regulamentar a IA no Brasil, e do *AI Act* da União Europeia, bem como análises jurídicas e notícias que cobrem o processo legislativo de ambos. Essas fontes são fundamentais para a análise comparativa dos marcos regulatórios e para a compreensão dos desafios políticos e legais envolvidos (BRASIL, 2023; BARROS, 2024; TOZZINIFREIRE, 2024; CALDERONIO, 2024; DATAGUIDANCE, 2025).

O tratamento dos dados foi realizado por meio da Análise de Conteúdo Temática, conforme proposto por Bardin (1977), uma metodologia já empregada em estudos sobre o tema no Brasil (TELLES; MONTARDO, 2025). O processo envolveu a leitura flutuante de todo o material para impregnação do conteúdo, seguida da identificação e codificação de unidades de significado. A partir dessa codificação, foram estabelecidas quatro categorias temáticas centrais que estruturam a apresentação dos resultados e a discussão: (1) Mecanismos de produção de desinformação via IAG; (2) Aplicações de IA para mitigação da desinformação; (3) Impactos sociopolíticos e desafios éticos; e (4) Respostas regulatórias e governança.

Reconhece-se como limitação deste estudo o fato de a revisão se ater ao escopo do corpus de pesquisa pré-selecionado. Embora abrangente, não se pretende esgotar a totalidade da produção acadêmica sobre o tema, mas sim oferecer uma análise aprofundada e focada que sintetiza as principais vertentes do debate atual, fundamentando a tese da "dupla hélice" da IAG.

3 RESULTADOS: O PARADOXO ALGORÍTMICO NA ESFERA PÚBLICA

A análise do corpus de pesquisa revela a natureza paradoxal da Inteligência Artificial Generativa, que atua simultaneamente como um poderoso vetor para a degradação do ambiente informacional e como uma promissora ferramenta para sua salvaguarda. Esta seção desdobra essa dualidade, apresentando os resultados da revisão de literatura em três eixos: a mecanização da falsidade, o potencial do antídoto algorítmico e os desafios regulatórios e éticos no contexto brasileiro.

3.1 A MECANIZAÇÃO DA FALSIDADE: IA GENERATIVA COMO VETOR DE DESINFORMAÇÃO

A capacidade da IAG de amplificar a desinformação reside em dois pilares tecnológicos principais: as Redes Adversariais Generativas (GANs), responsáveis pela criação de conteúdo audiovisual sintético, e os Grandes Modelos de Linguagem (LLMs), que automatizam a produção de texto em larga escala.

As GANs operam por meio de uma arquitetura de aprendizado profundo composta por duas redes neurais que competem entre si: o gerador e o discriminador. O gerador cria amostras de dados (como imagens ou sons) a partir de um ruído aleatório, enquanto o discriminador avalia essas amostras, tentando distinguir as que são falsas das que são reais (provenientes de um conjunto de dados de treinamento). O gerador é treinado para "enganar" o discriminador, e o discriminador é treinado para não ser enganado. Esse processo adversarial contínuo resulta em um gerador capaz de produzir conteúdo sintético cada vez mais realista e convincente (GOODFELLOW et al., 2014; CVISIONLAB, 2023; IBM, 2024). Essa é a tecnologia subjacente aos *deepfakes*, que permitem a manipulação ou criação de vídeos e áudios falsos de figuras públicas, com um potencial disruptivo imenso para a manipulação da opinião pública e a difamação (BATISTA; SANTAELLA, 2024; FONTÃO; DIAS, 2022).

Paralelamente, os LLMs, como os da série GPT (Generative Pre-trained Transformer), são treinados em vastos volumes de texto da internet para aprender padrões, gramática, fatos e estilos de escrita. Sua função principal é prever a próxima palavra mais provável em uma sequência, o que lhes permite gerar textos coerentes, contextualmente relevantes e que mimetizam a escrita humana (AWS, 2024; IBM, 2024). Essa capacidade viabiliza a automação em massa de campanhas de desinformação, a criação de artigos de notícias falsas, a geração de comentários fraudulentos em redes sociais e a personalização de narrativas enganosas para públicos específicos, tudo a um custo e velocidade anteriormente inimagináveis (GOLDSTEIN; LOHN, 2024; MATZ et al., 2024).

O impacto dessas tecnologias já é observável em múltiplos contextos. Em processos eleitorais, a ameaça é particularmente aguda. Relatórios sobre as eleições municipais de 2024 no Brasil, embora indiquem um impacto ainda não decisivo, servem como um alerta para o potencial de disrupção em pleitos de maior escala, como as eleições presidenciais (DATA PRIVACY BRASIL, 2024; JORNAL DA CULTURA, 2024). Casos concretos já foram identificados, incluindo o uso de IA para gerar *jingles* de campanha, a manipulação de imagens de candidatas com teor sexual e a criação de vídeos satíricos para atacar figuras públicas, borrando as linhas entre crítica e desinformação (DATA PRIVACY BRASIL, 2024; JORNAL DA CULTURA, 2024; DESINFORMANTE, 2024). A preocupação é que tais ferramentas possam ser usadas para criar e disseminar *deepfakes* e outros conteúdos falsos para influenciar o eleitorado, especialmente em um ambiente político já polarizado (BATISTA; SANTAELLA, 2024; CONECTAS, 2024).

Além da política, a IAG ameaça a integridade de outros domínios. A capacidade de gerar artigos científicos falsos, por exemplo, representa um risco à credibilidade da ciência e ao processo de revisão por pares (SPINAK, 2023). No âmbito corporativo, a geração de conteúdo prejudicial pode ser usada para atacar a reputação de marcas e empresas, exigindo novas estratégias de comunicação e gestão de crise (BLIKSTEIN; FERNANDES; COUTINHO, 2024). A principal consequência dessa proliferação

é a drástica redução do custo e da complexidade técnica para produzir desinformação de alta qualidade, efetivamente "democratizando" o acesso a ferramentas de manipulação que antes eram exclusividade de atores estatais ou grandes organizações (CETAS, 2024).

Contudo, o impacto mais insidioso da IAG pode transcender a mera produção de falsidades. A crescente dificuldade do público em geral para distinguir conteúdo autêntico de conteúdo sintético (GOLDSTEIN; LOHN, 2024; NIGHTINGALE; FARID, 2022) gera um efeito colateral perigoso. A consciência generalizada de que qualquer vídeo, áudio ou imagem pode ser um *deepfake* cria um ambiente de ceticismo universal. Nesse cenário, atores mal-intencionados podem explorar essa incerteza para desacreditar evidências genuínas e comprometedoras, simplesmente alegando que são falsas. Este fenômeno, conhecido como o "dividendo do mentiroso" (*liar's dividend*) (CHESNEY; CITRON, 2019), significa que o maior dano da IAG não é apenas a proliferação do falso, mas a erosão sistêmica da confiança em qualquer forma de evidência digital. Isso mina a capacidade da sociedade de discernir fatos de ficção e beneficia aqueles que desejam operar sem accountability, tornando a verdade uma questão de crença e não de evidência.

3.2 O ANTÍDOTO ALGORÍTMICO: O POTENCIAL DA IA NO COMBATE À DESINFORMAÇÃO

Em contraposição ao seu potencial destrutivo, a Inteligência Artificial também emerge como a principal ferramenta para combater a desinformação em escala. Sistemas baseados em aprendizado de máquina e Processamento de Linguagem Natural (PLN) estão sendo desenvolvidos para automatizar e aprimorar o processo de *fact-checking*, superando as limitações de velocidade e volume da verificação humana (BONTRIDDER; POULLET, 2021; GRAVES, 2018).

Essas ferramentas de mitigação operam em várias frentes. Elas podem analisar grandes volumes de dados de redes sociais e da web para identificar a disseminação de narrativas falsas em tempo real. Utilizando PLN, os algoritmos podem analisar a estrutura textual, o sentimento e o contexto de uma alegação para detectar padrões associados à desinformação (YANG; MENCZER, 2023; ZHOU; ZAFARANI, 2020). Além disso, são capazes de cruzar informações automaticamente com bancos de dados de fontes confiáveis, como agências de notícias, artigos científicos e registros governamentais, para verificar a veracidade de uma afirmação (YANG; MENCZER, 2023). Pesquisas têm demonstrado a eficácia de algoritmos treinados para diferenciar notícias factuais de boatos com alta precisão, analisando a frequência e a combinação de palavras (RODRIGUES; MATTOS, 2024).

No entanto, a aplicação de IA para o combate à desinformação não é isenta de desafios. Uma das principais limitações é a dificuldade dos algoritmos em interpretar nuances da linguagem humana, como sarcasmo, ironia e sátira, o que pode levar a classificações incorretas (ZHOU; ZAFARANI, 2020). Outro desafio significativo é a dependência da qualidade dos dados de treinamento. Se os dados utilizados para treinar os modelos contiverem vieses, a IA pode perpetuar ou até amplificar esses

preconceitos em suas análises (YANG; MENCZER, 2023). A constante evolução das técnicas de desinformação também exige que os modelos de detecção sejam continuamente atualizados para não se tornarem obsoletos. Por essas razões, a supervisão humana permanece indispensável. A abordagem mais eficaz é a colaboração homem-máquina, na qual a IA atua como uma ferramenta para auxiliar e escalar o trabalho de jornalistas e checadores de fatos, que aplicam o julgamento crítico e contextual final (THE FIX MEDIA, 2024).

No Brasil, a pesquisa acadêmica sobre o tema reflete essa tendência global, com um foco particular na aplicação defensiva da IA. Um levantamento dos trabalhos apresentados nos congressos da Associação Brasileira de Pesquisadores em Jornalismo (SBPJor) entre 2015 e 2022 revelou que, embora o campo seja ainda incipiente, a linha de pesquisa predominante é o uso de IA para a detecção de *fake news* por agências de *fact-checking*. Outros temas emergentes incluem a análise do impacto da IA na credibilidade do jornalismo e nas novas atribuições da profissão (TELLES; MONTARDO, 2025). Isso indica que a comunidade acadêmica brasileira está ativamente engajada em explorar o potencial da IA como um "antídoto algorítmico", ainda que a pesquisa sobre a criação de desinformação por IAG seja menos explorada.

3.3 O CENÁRIO BRASILEIRO EM FOCO: DESAFIOS REGULATÓRIOS E ÉTICOS

A dupla natureza da IAG impõe um desafio significativo para legisladores em todo o mundo, que buscam equilibrar a promoção da inovação com a mitigação de riscos. No Brasil, esse debate se materializa no Projeto de Lei nº 2338/2023, um marco regulatório que visa estabelecer direitos, deveres e uma estrutura de governança para o desenvolvimento e uso da IA no país (BRASIL, 2023). A proposta brasileira, aprovada no Senado Federal em dezembro de 2024, dialoga intensamente com o *AI Act* da União Europeia, que entrou em vigor em agosto de 2024, adotando uma abordagem semelhante baseada em risco (BARROS, 2024; SENADO NOTÍCIAS, 2024).

A abordagem baseada em risco classifica os sistemas de IA de acordo com o perigo potencial que representam para a saúde, a segurança e os direitos fundamentais. Práticas de risco inaceitável, como sistemas de pontuação social (*social scoring*) por governos ou técnicas de manipulação subliminar, são proibidas. Sistemas de alto risco, que incluem aplicações em áreas críticas como saúde, justiça, segurança pública e processos eleitorais, estão sujeitos a obrigações rigorosas de governança, transparência, gestão de risco e supervisão humana (BRASIL, 2023; TOZZINIFREIRE, 2024).

A tramitação do PL 2338/2023 no Senado foi marcada por intensos debates, especialmente em torno da inclusão de dispositivos explícitos para garantir a integridade da informação e combater a desinformação. A oposição a esses trechos, sob o argumento de proteger a liberdade de expressão, levou a negociações e à retirada de algumas menções, embora o princípio da "integridade da informação" tenha sido mantido como um dos fundamentos da lei (AGÊNCIA BRASIL, 2024a;

AGÊNCIA BRASIL, 2024b). Essa disputa evidencia a tensão central na regulação da IA: como coibir o uso malicioso da tecnologia sem cercear direitos fundamentais.

Para contextualizar e aprofundar a análise do esforço regulatório brasileiro, a Tabela 1 apresenta uma comparação detalhada entre os principais pontos do *AI Act* europeu e do PL 2338/2023.

Tabela 1 – Comparação entre o AI Act da União Europeia e o PL 2338/2023 (versão aprovada no Senado).

Característica	**EU AI Act**	**PL 2338/2023 (versão aprovada no Senado)**	**Análise Comparativa**
Abordagem Principal	Baseada em risco (inaceitável, alto, limitado, mínimo).	Baseada em risco (excessivo, alto).	Ambos adotam uma abordagem baseada em risco, alinhando-se a uma tendência global. A categorização brasileira é mais simplificada (BRASIL, 2023; DATAGUIDANCE, 2025).
Práticas Proibidas (Risco Inaceitável/Excessivo)	Proíbe *social scoring* governamental, exploração de vulnerabilidades, manipulação subliminar e identificação biométrica remota em tempo real (com exceções).	Proíbe sistemas que explorem vulnerabilidades, manipulação subliminar e *social scoring* com efeitos injustos.	As proibições são semelhantes, refletindo um consenso sobre os usos mais perigosos da IA. O PL brasileiro é alinhado com os princípios europeus de proteção (BRASIL, 2023; TOZZINIFREIRE, 2024).
Sistemas de Alto Risco (Setores)	Lista explícita de setores: infraestrutura crítica, educação, emprego, serviços essenciais, aplicação da lei, migração, administração da justiça, etc.	Lista similar: saúde, justiça, segurança, crédito, emprego, educação, infraestrutura crítica, veículos autônomos, identificação biométrica.	A cobertura de setores de alto impacto é bastante convergente, mostrando que ambos os textos identificam os mesmos pontos de vulnerabilidade social (BRASIL, 2023; GT LAWYERS, 2025).
Obrigações para Alto Risco	Requisitos rigorosos de gestão de risco, governança de dados, documentação técnica, transparência para usuários, supervisão humana e cibersegurança.	Exige medidas de governança, avaliação de impacto algorítmico, testes de segurança e mecanismos de supervisão humana.	As obrigações são conceitualmente alinhadas, focando em governança e transparência como pilares para a confiabilidade dos sistemas (BRASIL, 2023; DATA PRIVACY BRASIL, 2024).
Direitos dos Cidadãos	Indiretos, focados nas obrigações dos provedores. Direitos mais explícitos vêm de outras legislações (e.g., GDPR).	Capítulo explícito sobre direitos, incluindo informação, explicação sobre decisões automatizadas, contestação, revisão	O PL brasileiro é mais explícito na codificação dos direitos dos indivíduos afetados por sistemas de IA, uma diferença notável e um

		humana e não discriminação.	avanço em relação ao texto europeu (DATAGUIDANCE, 2025; DATA PRIVACY BRASIL, 2024).
Governança e Fiscalização	Criação de um Comitê Europeu de IA (*AI Board*) para coordenar as autoridades nacionais de supervisão.	Criação do Sistema Nacional de Regulação e Governança de IA (SIA), com a Autoridade Nacional de Proteção de Dados (ANPD) como autoridade competente.	Ambos criam uma estrutura de governança centralizada, mas o Brasil atribui a função à sua autoridade de proteção de dados já existente, o que pode gerar sinergias, mas também sobrecarga (BRASIL, 2023; DATA PRIVACY BRASIL, 2024).
Sanções	Multas de até 35 milhões de euros ou 7% do faturamento global anual, o que for maior.	Multas de até R$ 50 milhões por infração ou 2% do faturamento do grupo no Brasil.	As sanções são robustas em ambos os casos, utilizando o faturamento como base para garantir que as penalidades sejam significativas mesmo para grandes empresas de tecnologia (GT LAWYERS, 2025; ARTIFICIAL INTELLIGENCE ACT, 2024).

Fonte: Elaborado pelos autores (2025), a partir de Brasil (2023); Dataguidance (2025); GT Lawyers (2025).

A análise comparativa revela que o Brasil está se alinhando às melhores práticas internacionais, mas com adaptações importantes, como a ênfase explícita nos direitos dos cidadãos. No entanto, a eficácia dessa legislação dependerá crucialmente de sua implementação e da capacidade de fiscalização do Estado.

4 DISCUSSÃO: NAVEGANDO A INFOSFERA HÍBRIDA

A apresentação dos resultados evidencia que a sociedade contemporânea está ingressando em uma "infosfera híbrida", onde a distinção entre conteúdo gerado por humanos e por máquinas se torna cada vez mais tênue. Navegar neste novo ambiente exige mais do que apenas avanços tecnológicos; requer uma profunda reavaliação das práticas profissionais, dos marcos éticos e das estratégias regulatórias.

As implicações para o jornalismo e a comunicação são profundas e multifacetadas. A era da IAG força uma redefinição das competências e da própria identidade profissional. A habilidade de apurar fatos e redigir textos, embora ainda fundamental, já não é suficiente. Emerge a necessidade de uma nova forma de expertise: a capacidade de interagir criticamente com sistemas de IA. Isso inclui saber formular as perguntas corretas para extrair informações úteis, compreender as limitações e os

vieses inerentes aos modelos algorítmicos e, acima de tudo, exercer uma curadoria responsável sobre o conteúdo gerado ou auxiliado por máquinas (DE-LIMA-SANTOS; SALAVERRIA, 2021; PARANHOS NETO, 2024). O papel do jornalista se desloca de mero produtor de conteúdo para o de um validador e contextualizador em um ecossistema saturado de informações sintéticas.

Essa transformação demanda uma atualização urgente dos códigos deontológicos da profissão. Princípios tradicionais como a veracidade e a imparcialidade precisam ser reinterpretados à luz dos novos desafios. Estudos apontam para a necessidade de incorporar novas diretrizes éticas, como a transparência radical sobre o uso de IA na produção de notícias, a garantia de supervisão humana significativa em todas as etapas do processo editorial e a implementação de mecanismos robustos para auditar e mitigar vieses algorítmicos que possam levar à discriminação ou à distorção da realidade (ESTRELLA TUTIVÉN; GARDE CANO, 2024; DE-LIMA-SANTOS; SALAVERRIA, 2021). A confiança do público, um ativo já fragilizado, dependerá da capacidade das organizações de notícias de adotarem essas práticas de forma proativa e transparente.

Nesse contexto, a resposta regulatória do Estado torna-se um pilar fundamental. A análise comparativa entre o PL 2338/2023 e o *AI Act* europeu mostra que o Brasil está construindo uma base legislativa sólida, inspirada em um modelo internacionalmente reconhecido. A abordagem baseada em risco e a enumeração de direitos são passos na direção correta. No entanto, a eficácia de qualquer legislação não reside apenas em sua formulação, mas em sua capacidade de ser implementada e fiscalizada. Aqui reside um risco crítico para o contexto brasileiro: o de uma "regulação simbólica".

O processo para se chegar a uma legislação robusta no papel é apenas o primeiro passo. A fiscalização de sistemas de IA de alto risco, a realização de auditorias algorítmicas e a investigação de danos causados por IA exigem uma capacidade técnica e recursos financeiros e humanos altamente especializados (YANG; MENCZER, 2023; ZHOU; ZAFARANI, 2020). A atribuição dessa responsabilidade à ANPD, embora lógica, impõe um desafio monumental a uma agência que já enfrenta dificuldades para fiscalizar plenamente a Lei Geral de Proteção de Dados (LGPD). Adicionalmente, as disputas políticas que buscaram suavizar as obrigações relacionadas ao combate à desinformação durante a tramitação do projeto no Senado (AGÊNCIA BRASIL, 2024b) sinalizam uma forte pressão de setores da indústria de tecnologia e de grupos políticos que podem continuar a atuar para enfraquecer a aplicação da lei. Portanto, existe a possibilidade de o Brasil possuir uma legislação avançada em teoria, mas com uma aplicação prática limitada por gargalos de implementação, falta de recursos e captura regulatória, tornando-a ineficaz para lidar com os desafios complexos da infosfera híbrida.

Isso nos leva à constatação de que a regulação, por si só, é uma solução incompleta. A linha de defesa mais resiliente e sustentável contra a desinformação na era da IAG é uma cidadania criticamente informada. É imperativo investir em programas de literacia midiática e algorítmica em larga escala.

Essa educação deve ir além do ensino tradicional de como identificar *fake news*. É preciso capacitar os cidadãos a compreenderem os princípios básicos de funcionamento dos algoritmos de recomendação e dos modelos de IAG, fomentando um ceticismo saudável e uma postura ativa na verificação das informações consumidas (CHESNEY; CITRON, 2019; NIGHTINGALE; FARID, 2022). Sem uma população apta a navegar criticamente neste novo ambiente, mesmo a melhor das regulações terá um impacto limitado.

5 CONCLUSÃO

Este artigo de revisão se propôs a analisar o paradoxo da Inteligência Artificial Generativa no ecossistema comunicacional, articulando o argumento de que a tecnologia opera como uma "dupla hélice": por um lado, um poderoso motor para a criação e disseminação de desinformação sofisticada; por outro, um conjunto de ferramentas essenciais para a sua detecção e mitigação. A análise, focada no contexto brasileiro, revela que o país se encontra em um momento crucial, buscando construir um arcabouço regulatório enquanto já enfrenta os impactos práticos dessa tecnologia disruptiva.

A síntese da literatura confirma que a IAG, por meio de tecnologias como GANs e LLMs, barateou e escalou a produção de conteúdo falso, representando uma ameaça real à integridade informacional, especialmente em contextos sensíveis como os processos eleitorais. O fenômeno do "dividendo do mentiroso" emerge como uma das consequências mais danosas, onde a erosão da confiança se generaliza, afetando a credibilidade até mesmo de conteúdos autênticos. Em contrapartida, a mesma base tecnológica oferece soluções promissoras para a automação do *fact-checking*, embora estas ainda dependam de supervisão humana e enfrentem desafios técnicos significativos.

O balanço final aponta que não há uma solução única ou simples para os desafios impostos pela IAG. O equilíbrio entre o fomento à inovação e a proteção da sociedade contra os riscos da desinformação algorítmica depende de uma abordagem sinérgica e multifacetada, que deve assentar em três pilares interdependentes:

1. **Regulação Tecnológica Robusta:** O PL 2338/2023 representa um passo fundamental, mas sua eficácia dependerá de uma implementação rigorosa e de uma fiscalização tecnicamente capacitada, superando o risco de se tornar uma "regulação simbólica".
2. **Adaptação Ética e Técnica Profissional:** Profissionais da comunicação, especialmente jornalistas, devem incorporar novas competências e atualizar seus códigos deontológicos para lidar com a curadoria de conteúdo sintético e a transparência no uso de ferramentas de IA.
3. **Cidadania Digitalmente Capacitada:** É crucial um investimento massivo e contínuo em programas de literacia midiática e algorítmica que preparem os cidadãos para consumir informação

de forma crítica na nova infosfera híbrida.

Diante do exposto, e reconhecendo as lacunas no conhecimento atual, propõe-se uma agenda para pesquisas futuras na área de Comunicação e Informação no Brasil, que inclua:

- **Estudos de Recepção e Impacto:** Investigações longitudinais sobre os efeitos cognitivos e comportamentais da exposição contínua a conteúdo sintético na percepção da realidade e na formação da opinião pública.
- **Eficácia da Literacia Midiática:** Pesquisas aplicadas para avaliar a eficácia de diferentes abordagens pedagógicas em literacia algorítmica, adaptadas aos diversos contextos socioculturais e educacionais do Brasil.
- **Análise da Implementação Regulatória:** Estudos que monitorem e analisem criticamente a implementação do marco legal da IA, focando nos desafios operacionais da ANPD e de outros órgãos do sistema de governança, bem como nos resultados práticos de sua fiscalização.
- **Desenvolvimento de Modelos de Governança Interdisciplinares:** Fomento a pesquisas que unam Comunicação, Direito, Ciência da Computação e Ciências Sociais para desenvolver modelos de governança de IA que sejam não apenas tecnicamente sólidos e juridicamente conformes, mas também socialmente justos e culturalmente adequados à realidade brasileira.

A jornada para navegar a era da IAG está apenas começando. A construção de um futuro digital mais seguro e confiável dependerá da capacidade da academia, do Estado, do setor privado e da sociedade civil de colaborarem na edificação desses três pilares.

REFERÊNCIAS

AGÊNCIA BRASIL. Debate sobre desinformação adia votação de projeto que regula IA. *Agência Brasil*, 3 dez. 2024a. Disponível em: https://agenciabrasil.ebc.com.br/politica/noticia/2024-12/debate-sobre-desinformacao-adia-votacao-de-projeto-que-regula-ia. Acesso em: 10 set. 2025.

AGÊNCIA BRASIL. Acordo sobre desinformação permite aprovação de regulação da IA. *Agência Brasil*, 5 dez. 2024b. Disponível em: https://agenciabrasil.ebc.com.br/politica/noticia/2024-12/acordo-sobre-desinformacao-permite-aprovacao-de-regulacao-da-ia. Acesso em: 10 set. 2025.

ARTIFICIAL INTELLIGENCE ACT. *Brazil AI Act*. 2024. Disponível em: https://artificialintelligenceact.com/brazil-ai-act/. Acesso em: 10 set. 2025.

AWS. *O que é um grande modelo de linguagem (LLM)?* Amazon Web Services, 2024. Disponível em: https://aws.amazon.com/pt/what-is/large-language-model/. Acesso em: 10 set. 2025.

BARDIN, Laurence. *Análise de Conteúdo*. Lisboa: Edições 70, 1977.

BARROS, Gabriel Osório de. Uma análise do AI Act. *Gabinete de Estratégia e Estudos do Ministério da Economia*, Lisboa, 2 ago. 2024. Disponível em: https://www.gee.gov.pt/pt/documentos/estudos-e-seminarios/artigos/10243-gee-em-analise-regulamento-da-inteligencia-artificial/file. Acesso em: 10 set. 2025.

BATISTA, Anderson Röhe Fontão; SANTAELLA, Lucia. Prognósticos das deepfakes na política eleitoral. *Organicom*, São Paulo, v. 21, n. 42, p. 1-15, mai./ago. 2024.

BLIKSTEIN, Izidoro; FERNANDES, Manoel; COUTINHO, Marcelo. Estudo analisa o impacto da inteligência artificial na comunicação corporativa nas redes sociais. *FGV Notícias*, 11 jun. 2024. Disponível em: https://portal.fgv.br/noticias/estudo-analisa-impacto-inteligencia-artificial-comunicacao-corporativa-redes-sociais. Acesso em: 10 set. 2025.

BONTRIDDER, Nicolas; POULLET, Yves. *AI for counter-disinformation*. Brussels: European Parliament, 2021.

BRASIL. Congresso. Senado Federal. *Projeto de Lei nº 2338, de 2023*. Dispõe sobre o uso da Inteligência Artificial. Brasília, DF: Senado Federal, 2023.

CALDERONIO, Vincenzo. The state of the art of the Brazilian Bill on Artificial Intelligence with a focus on civil responsibility. *Rivista Italiana di Informatica e Diritto*, v. 2, 2024.

CETAS. *AI-Enabled Influence Operations: Safeguarding Future Elections*. The Alan Turing Institute, 2024. Disponível em: https://cetas.turing.ac.uk/publications/ai-enabled-influence-operations-safeguarding-future-elections. Acesso em: 10 set. 2025.

CHESNEY, Robert; CITRON, Danielle. Deep Fakes: A Looming Challenge for Privacy, Democracy, and National Security. *Lawfare*, 21 jun. 2019.

CHORAŚ, Michał et al. A survey on machine learning for fake news detection. *Information Fusion*, v. 72, p. 1-22, 2021.

CONECTAS. Eleições 2024: desinformação causa danos concretos na democracia e na vida das pessoas. *Conectas Direitos Humanos*, 28 ago. 2024. Disponível em:

https://conectas.org/noticias/eleicoes-2024-desinformacao-causa-danos-concretos-na-democracia-e-na-vida-das-pessoas. Acesso em: 10 set. 2025.

CVISIONLAB. *Deepfake and GANs: How Generative Adversarial Networks Work*. 2023. Disponível em: https://www.cvisionlab.com/cases/deepfake-gan/. Acesso em: 10 set. 2025.

DATA PRIVACY BRASIL. *IA no primeiro turno: o que vimos até aqui?* 2024. Disponível em: https://www.dataprivacybr.org/documentos/ia-no-primeiro-turno-o-que-vimos-ate-aqui/. Acesso em: 10 set. 2025.

DATAGUIDANCE. *Brazil: Similarities and key differences between the Brazilian AI Bill and the EU AI Act*. 2025. Disponível em: https://www.dataguidance.com/opinion/brazil-similarities-and-key-differences-between. Acesso em: 10 set. 2025.

DE-LIMA-SANTOS, Mathias Felipe; SALAVERRIA, Ramón. Do jornalismo de dados à inteligencia artificial: desafios enfrentados pelo La Nación na implementação da visão computacional para a produção de notícias. *Palabra Clave*, v. 24, n. 3, e2437, 2021.

DESINFORMANTE. *Inteligência Artificial e Desinformação | IAí?* 2024. Disponível em: https://desinformante.com.br/observatorio-ia/. Acesso em: 10 set. 2025.

ESTRELLA TUTIVÉN, Ingrid Viviana; GARDE CANO, Cristina. Ética Jornalística na Era da Inteligência Artificial: Rumo a uma Atualização dos Códigos Deontológicos no Contexto Ibero-Americano. *Comunicação e Sociedade*, v. 45, 2024.

FOLETTO, Leonardo. Pesquisa em Comunicação: Tendências emergentes e desafios no cenário digital. *FGV Mídia*, 2024. Disponível em: https://mestrado-doutorado.fgv.br/blog/noticia/pesquisa-em-comunicacao-tendencias-emergentes-e-desafios-no-cenario-digital. Acesso em: 10 set. 2025.

FONTÃO, Fabiano Fernando; DIAS, Jefferson Aparecido. BOTS, FAKE NEWS, FAKE FACES, DEEPFAKES E SUA EVENTUAL INFLUÊNCIA NO PROCESSO ELEITORAL DEMOCRÁTICO. *Revista da ANPAL*, v. 1, n. 1, 2022.

FRONTIERS. AI-driven disinformation: policy recommendations for democratic resilience. *Frontiers in Artificial Intelligence*, v. 8, 2025.

GOLDSTEIN, Josh A.; LOHN, Andrew. *Deepfakes, Elections, and Shrinking the Liar's Dividend*. RAND Corporation, 2024.

GOMES, Pollyany Annenberg Nascimento; OLIVEIRA, Maria Lívia Pachêco de. Inteligência artificial generativa e a desinformação no Brasil. *Siti*, Maceió, v. 6, e155, 2024.

GOODFELLOW, Ian et al. Generative Adversarial Nets. *Advances in Neural Information Processing Systems*, 27, 2014.

GRAVES, Lucas. *Understanding the Promise and Limits of Automated Fact-Checking*. Reuters Institute, 2018.

GT LAWYERS. *Marco Legal da Inteligência Artificial no Brasil: O PL 2338/2023 em Foco*. 2025. Disponível em: https://www.gtlawyers.com.br/en/artigos/marco-legal-da-inteligencia-artificial-no-brasil-o-pl-2338-2023-em-foco/. Acesso em: 10 set. 2025.

HJARVARD, Stig. A midiatização da cultura e da sociedade. *Unisinos*, 2014.

IBM. *What are generative adversarial networks (GANs)?* 2024. Disponível em: https://www.ibm.com/think/topics/generative-adversarial-networks. Acesso em: 10 set. 2025.

JORNAL DA CULTURA. Relatório aponta que uso da IA nas eleições não causou impacto negativo. *YouTube*, 16 out. 2024. Disponível em: https://www.youtube.com/watch?v=dmjWy8GQbJM. Acesso em: 10 set. 2025.

MATZ, S. C. et al. AI-driven personalization can be used to influence and manipulate. *PNAS*, v. 121, n. 9, 2024.

MENDES, Conrado Moreira; MATTOS, Maria Ângela. Limites e potencialidades de conceitos de desinformação e congêneres acionados em artigos brasileiros sobre covid-19. *E-Compós*, v. 28, 2025.

NIGHTINGALE, Sophie J.; FARID, Hany. AI-synthesized faces are indistinguishable from real faces and more trustworthy. *PNAS*, v. 119, n. 8, 2022.

PARANHOS NETO, Oldemburgo da Silva. A Inteligência Artificial e seus impactos na produção jornalística: uma análise sobre a qualidade, veracidade e relevância da informação. *Siti*, Maceió, v. 7, e188, 2024.

RECUERO, Raquel et al. A desinformação narrada pela checagem: estudo das eleições de 2018 e 2022. *E-Compós*, v. 28, 2025.

RODRIGUES, Nicollas; MATTOS, Diogo. Estudo mostra uso de inteligência artificial na detecção de fake news. *Agência Brasil*, 8 jul. 2024. Disponível em: https://sintepe.org.br/estudo-mostra-uso-de-inteligencia-artificial-na-deteccao-de-fake-news/. Acesso em: 10 set. 2025.

SENADO NOTÍCIAS. Senado aprova regulamentação da inteligência artificial; texto vai à Câmara. *Agência Senado*, 10 dez. 2024. Disponível em: https://www12.senado.leg.br/noticias/materias/2024/12/10/senado-aprova-regulamentacao-da-inteligencia-artificial-texto-vai-a-camara. Acesso em: 10 set. 2025.

SPINAK, Ernesto. Inteligência Artificial e a comunicação da pesquisa. *SciELO em Perspectiva*, 30 ago. 2023. Disponível em: https://blog.scielo.org/blog/2023/08/30/inteligencia-artificial-e-a-comunicacao-da-pesquisa/. Acesso em: 10 set. 2025.

TELLES, Marina Klein; MONTARDO, Sandra Portella. Inteligência artificial e desinformação: uma análise dos trabalhos apresentados em congressos da SBPJOR de 2015 a 2022. *Conexão - Comunicação e Cultura*, Caxias do Sul, v. 22, n. 01, 2025.

THE FIX MEDIA. *AI course: using AI for fact-checking in journalism*. 2024. Disponível em: https://thefix.media/2024/11/07/ai-course-using-ai-for-fact-checking-in-journalism/. Acesso em: 10 set. 2025.

TOZZINIFREIRE. *Lei de Inteligência Artificial da União Europeia entra em vigor*. 2024. Disponível em: https://tozzinifreire.com.br/boletins/lei-de-inteligencia-artificial-da-uniao-europeia-entra-em-vigor. Acesso em: 10 set. 2025.

YANG, Kai-Cheng; MENCZER, Filippo. Cutting through the noise: A machine learning approach to news trustworthiness. *Science Advances*, v. 9, n. 48, 2023.

ZHOU, Xinyi; ZAFARANI, Reza. A Survey of Fake News: Fundamental Theories, Detection Methods, and Opportunities. *ACM Computing Surveys*, v. 53, n. 5, 2020.