SEVEN
publicações acadêmicas

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

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

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

https://doi.org/10.56238/sevened2025.029-091

Márcio Rodrigo Elias Carvalho[1], Lucyneid Barros Carvalho[2], Sérgio Ferreira da Silva[3], José Dirnece Paes Tavares[4], Antônio Ribas Reis[5], Flavia Chaves Valentim Rodrigues[6], Vera Lucia da Silva Farias[7], Rogério dos Santos Morais[8]

ABSTRACT

The rise of Generative Artificial Intelligence (GAI) has dramatically reshaped the information and communication ecosystem, establishing a central paradox: the same technology that enables the creation and dissemination of disinformation at 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, the potential of AI as a tool for automated fact-checking and analysis of disinformation patterns is highlighted. 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 demands a multifaceted approach that integrates robust technological regulation, ethical adaptation of communication practices, and a massive investment in media and algorithmic literacy for society.

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

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

[1] Artificial Intelligence Specialist. FASUL Educacional. E-mail: marcior@gmail.com
[2] Specialist in Neuropsychopedagogy. UNINASSAU Aracaju. E-mail: lucyneidbarros@gmail.com
[3] Master in Psychoanalysis. FASUL Educacional. E-mail: admsergioferreira@gmail.com
[4] Master in Regional Management and Development. FASUL Educacional. E-mail: dirnece@gmail.com
[5] Dr. in Dissemination of Knowledge. FASUL Educacional. E-mail: amigoribas@gmail.com
[6] Master in Regional Development. FASUL Educacional. E-mail: profa.flaviavalentim@gmail.com
[7] Dr. in Soil Sciences – Agronomy. FASUL Educacional. E-mail: professoraveralucia@fasuleducacional.edu.br
[8] Dr. in Engineering. FASUL Educacional. E-mail: roger.dos.santos.morais@gmail.com

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.

RESUMEN

El auge de la Inteligencia Artificial Generativa (IAG) ha transformado drásticamente el ecosistema de la información y la comunicación, 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, se 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 exige 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 INTRODUCTION

The contemporary communication landscape is defined by a profound and continuous transformation, driven by the ubiquity of networked digital technologies (FOLETTO, 2024). The consolidation of social media platforms as the main arenas of interaction and informational exchange has irreversibly altered traditional communication paradigms, reconfiguring the production, distribution, and consumption of content on a global scale (RECUERO et al., 2025). This new ecosystem, characterized by speed and accessibility, has opened doors to new forms of interaction, but has also introduced complex challenges, notably the dissemination of inaccurate or deliberately false information (RECUERO et al., 2025; MENDES; MATTOS, 2025).

In this already volatile context, the advent and popularization of Generative Artificial Intelligence (AGI) represents a critical inflection point. Tools such as ChatGPT, DALL-E, Midjourney and others, which are based on Large Language Models (LLMs) and Generative Adversarial Networks (GANs), have migrated from restricted research environments to the public domain, triggering a disruption in multiple sectors, with a particularly acute impact on the field of Communication and Information (GOMES; OLIVEIRA, 2024; SPINAK, 2023; BLIKSTEIN; FERNANDES; COUTINHO, 2024). The ability of these technologies to generate synthetic texts, images, and audio with an increasing degree of realism introduces a new layer of complexity to the informational environment.

This article starts from the premise that AGI operates as a "double helix" in the communication ecosystem, a technological paradox with profound implications for society. The first helix represents IAG's ability to degrade the public sphere, functioning as a potent catalyst for the production and dissemination of disinformation. Technology allows the mass creation of false, sophisticated and hyper-personalized content, from texts that mimic journalistic style to audiovisual *deepfakes* almost indistinguishable from reality, constituting an unprecedented threat to the integrity of democratic processes, trust in institutions and the social fabric itself (GOLDSTEIN; LOHN, 2024; CETAS, 2024; FRONTIERS, 2025).

On the other hand, the second helix represents the potential for safeguarding. The same technological foundation — machine learning and natural language processing — offers the most advanced and scalable tools to detect, analyze, and mitigate disinformation flows. AI systems can identify patterns, verify claims, and trace the origin of false narratives with an efficiency that transcends human capacity, becoming crucial allies for journalists, fact-

checkers, and researchers (BONTRIDDER; POULLET, 2021; CHORAŚ et al., 2021; GRAVES, 2018; YANG; MENCZER, 2023).

In view of this paradoxical scenario, this article aims to carry out a critical and systematic review of the literature to analyze the duality of AGI as a vector and, simultaneously, an antidote to disinformation, with a specific focus on the Brazilian context. The justification for this study lies in the urgency of the debate, fueled by the processing of fundamental regulatory frameworks, such as Bill No. 2338/2023, and by the concrete challenges already observed in events of great national relevance, such as the recent electoral processes (GOMES; OLIVEIRA, 2024; DATA PRIVACY BRASIL, 2024).

To achieve this objective, the article is structured as follows: the Methodology section details the research design as a literature review. The Results section explores the two facets of the algorithmic paradox, first dissecting the technical mechanisms of synthetic disinformation and its impacts, and then analyzing the potential of AI as a mitigation tool, culminating with an analysis of the Brazilian regulatory landscape. The Discussion section delves into the ethical implications for journalism and critically evaluates legislative responses. Finally, the Conclusion summarizes the main arguments and proposes an agenda for future research.

2 METHODOLOGY

This study is characterized as a systematic literature review article, a research modality that aims to synthesize and critically analyze the accumulated knowledge on a given topic from secondary sources. The objective is to map the state of the art, identify trends, gaps and controversies in the field of study of the intersection between Generative Artificial Intelligence, disinformation and communication, with special attention to the Brazilian context.

The corpus of analysis for this review was constituted by a multifaceted set of documents, selected to provide a comprehensive and in-depth view of the phenomenon. The sources were grouped into the following categories:

1. **Academic and Scientific Articles:** Publications from journals indexed in databases such as SciELO and annals of relevant congresses in the area of Communication in Brazil, such as those of the National Association of Graduate Programs in Communication (Compós) and the Brazilian Association of Researchers in Journalism (SBPJor), were analyzed. These documents provide the theoretical and empirical basis on research trends, the impacts of platformization, and epistemological

discussions in the field (FOLETTO, 2024; RECUERO et al., 2025; MENDES; MATTOS, 2025; TELLES; MONTARDO, 2025; HJARVARD, 2014).

2. **Technical Reports and Expert Analysis:** Reports and publications from *think* tanks, research institutes, and non-governmental organizations (national and international) that are dedicated to monitoring the impact of AI on society were included. These materials offer in-depth analysis on the mechanisms of disinformation campaigns and case studies on the use of AI in electoral contexts (CETAS, 2024; ADA LOVELACE INSTITUTE, 2024).
3. **Legislative Documents and Legal Analyses:** The corpus includes the text of Bill No. 2338/2023, which aims to regulate AI in Brazil, and the European Union's AI Act, as well as legal analyses and news covering the legislative process of both. These sources are fundamental for the comparative analysis of regulatory frameworks and for understanding the political and legal challenges involved (BRASIL, 2023; BARROS, 2024; TOZZINIFREIRE, 2024; CALDERONIO, 2024; DATAGUIDANCE, 2025).

The data treatment was carried out through Thematic Content Analysis, as proposed by Bardin (1977), a methodology already used in studies on the subject in Brazil (TELLES; MONTARDO, 2025). The process involved the floating reading of all the material for impregnation of the content, followed by the identification and coding of units of meaning. Based on this codification, four central thematic categories were established that structure the presentation of the results and the discussion: (1) Mechanisms for the production of disinformation via AGI; (2) AI applications for disinformation mitigation; (3) Socio-political impacts and ethical challenges; and (4) Regulatory responses and governance.

It is recognized as a limitation of this study the fact that the review is limited to the scope of the pre-selected research corpus. Although comprehensive, it is not intended to exhaust the totality of academic production on the subject, but rather to offer an in-depth and focused analysis that synthesizes the main strands of the current debate, supporting the thesis of the "double helix" of AGI.

3 RESULTS: THE ALGORITHMIC PARADOX IN THE PUBLIC SPHERE

The analysis of the research corpus reveals the paradoxical nature of Generative Artificial Intelligence, which acts simultaneously as a powerful vector for the degradation of the informational environment and as a promising tool for its safeguarding. This section

unfolds this duality, presenting the results of the literature review in three axes: the mechanization of falsehood, the potential of the algorithmic antidote, and the regulatory and ethical challenges in the Brazilian context.

3.1 THE MECHANIZATION OF FALSEHOOD: GENERATIVE AI AS A VECTOR OF DISINFORMATION

IAG's ability to amplify disinformation lies in two main technological pillars: Generative Adversarial Networks (GANs), responsible for creating synthetic audiovisual content, and Large Language Models (LLMs), which automate large-scale text production.

GANs operate through a deep learning architecture composed of two neural networks that compete with each other: the generator and the discriminator. The generator creates data samples (such as images or sounds) from random noise, while the discriminator evaluates these samples, trying to distinguish those that are fake from those that are real (coming from a training dataset). The generator is trained to "deceive" the discriminator, and the discriminator is trained not to be deceived. This continuous adversarial process results in a generator capable of producing synthetic content that is increasingly realistic and convincing (GOODFELLOW et al., 2014; CVISIONLAB, 2023; IBM, 2024). This is the technology underlying *deepfakes*, which allow the manipulation or creation of fake videos and audios of public figures, with immense disruptive potential for the manipulation of public opinion and defamation (BATISTA; SANTAELLA, 2024; FONTÃO; DIAS, 2022).

In parallel, LLMs, such as those in the GPT (Generative Pre-trained Transformer) series, are trained on vast volumes of internet text to learn patterns, grammar, facts, and writing styles. Their primary function is to predict the next most likely word in a sequence, which allows them to generate texts that are coherent, contextually relevant, and mimic human writing (AWS, 2024; IBM, 2024). This capability enables the mass automation of disinformation campaigns, the creation of fake news articles, the generation of fraudulent comments on social networks, and the personalization of misleading narratives for specific audiences, all at a cost and speed previously unimaginable (GOLDSTEIN; LOHN, 2024; MATZ et al., 2024).

The impact of these technologies is already observable in multiple contexts. In electoral processes, the threat is particularly acute. Reports on the 2024 municipal elections in Brazil, although they indicate an impact that is not yet decisive, serve as a warning of the potential for disruption in larger-scale elections, such as presidential elections (DATA

PRIVACY BRASIL, 2024; JORNAL DA CULTURA, 2024). Concrete cases have already been identified, including the use of AI to generate *campaign jingles*, the manipulation of images of female candidates with sexual content, and the creation of satirical videos to attack public figures, blurring the lines between criticism and disinformation (DATA PRIVACY BRASIL, 2024; JORNAL DA CULTURA, 2024; DISINFORMANT, 2024). The concern is that such tools can be used to create and disseminate *deepfakes* and other false content to influence the electorate, especially in an already polarized political environment (BATISTA; SANTAELLA, 2024; CONECTAS, 2024).

In addition to politics, IAG threatens the integrity of other domains. The ability to generate fake scientific articles, for example, poses a risk to the credibility of science and the peer review process (SPINAK, 2023). In the corporate sphere, the generation of harmful content can be used to attack the reputation of brands and companies, requiring new communication and crisis management strategies (BLIKSTEIN; FERNANDES; COUTINHO, 2024). The main consequence of this proliferation is the drastic reduction in the cost and technical complexity of producing high-quality disinformation, effectively "democratizing" access to manipulation tools that were previously exclusive to state actors or large organizations (CETAS, 2024).

However, the most insidious impact of AGI may transcend the mere production of falsehoods. The growing difficulty of the general public to distinguish authentic content from synthetic content (GOLDSTEIN; LOHN, 2024; NIGHTINGALE; FARID, 2022) generates a dangerous side effect. The widespread awareness that any video, audio, or image can be a *deepfake* creates an environment of universal skepticism. In this scenario, malicious actors can exploit this uncertainty to discredit genuine, compromising evidence by simply claiming that it is false. This phenomenon, known as the "liar*'s dividend*" (CHESNEY; CITRON, 2019), means that the greatest harm of AGI is not only the proliferation of the false, but the systemic erosion of trust in any form of digital evidence. This undermines society's ability to discern fact from fiction and benefits those who wish to operate without accountability, making truth a matter of belief rather than evidence.

3.2 THE ALGORITHMIC ANTIDOTE: THE POTENTIAL OF AI IN COMBATING DISINFORMATION

In contrast to its destructive potential, Artificial Intelligence also emerges as the main tool to combat disinformation at scale. Systems based on machine learning and Natural

Language Processing (NLP) are being developed to automate and enhance the *fact-checking* process, overcoming the speed and volume limitations of human verification (BONTRIDDER; POULLET, 2021; GRAVES, 2018).

These mitigation tools operate on multiple fronts. They can analyze large volumes of data from social networks and the web to identify the spread of false narratives in real time. Using NLP, algorithms can analyze the textual structure, sentiment, and context of a claim to detect patterns associated with misinformation (YANG; MENCZER, 2023; ZHOU; ZAFARANI, 2020). In addition, they are able to automatically cross-reference information with databases from reliable sources, such as news agencies, scientific articles, and government records, to verify the veracity of a statement (YANG; MENCZER, 2023). Research has demonstrated the effectiveness of algorithms trained to differentiate factual news from rumors with high accuracy, analyzing the frequency and combination of words (RODRIGUES; MATTOS, 2024).

However, the application of AI to combat disinformation is not without its challenges. One of the main limitations is the difficulty of algorithms in interpreting nuances of human language, such as sarcasm, irony, and satire, which can lead to incorrect classifications (ZHOU; ZAFARANI, 2020). Another significant challenge is the reliance on the quality of training data. If the data used to train the models contains biases, AI can perpetuate or even amplify these biases in its analyses (YANG; MENCZER, 2023). The constant evolution of disinformation techniques also requires that detection models be continuously updated so as not to become obsolete. For these reasons, human oversight remains indispensable. The most effective approach is human-machine collaboration, in which AI acts as a tool to assist and scale the work of journalists and fact-checkers, who apply final critical and contextual judgment (THE FIX MEDIA, 2024).

In Brazil, academic research on the topic reflects this global trend, with a particular focus on the defensive application of AI. A survey of the papers presented at the congresses of the Brazilian Association of Journalism Researchers (SBPJor) between 2015 and 2022 revealed that, although the field is still incipient, the predominant line of research is the use of AI for the detection of *fake news* by *fact-checking agencies*. Other emerging themes include the analysis of the impact of AI on the credibility of journalism and on the new attributions of the profession (TELLES; MONTARDO, 2025). This indicates that the Brazilian academic community is actively engaged in exploring the potential of AI as an "algorithmic antidote", even if research on the creation of disinformation by AGI is less explored.

3.3 THE BRAZILIAN SCENARIO IN FOCUS: REGULATORY AND ETHICAL CHALLENGES

The dual nature of AGI poses a significant challenge for policymakers around the world, who seek to balance promoting innovation with mitigating risk. In Brazil, this debate materializes in Bill No. 2338/2023, a regulatory framework that aims to establish rights, duties, and a governance structure for the development and use of AI in the country (BRASIL, 2023). The Brazilian proposal, approved by the Federal Senate in December 2024, dialogues intensely with the European Union's AI Act, which came into force in August 2024, adopting a similar risk-based approach (BARROS, 2024; SENADO NOTÍCIAS, 2024).

The risk-based approach classifies AI systems according to the potential danger they pose to health, safety, and fundamental rights. Unacceptable risk practices, such as social scoring systems (*social scoring*) by governments or subliminal manipulation techniques, are prohibited. High-risk systems, which include applications in critical areas such as health, justice, public safety, and electoral processes, are subject to strict governance, transparency, risk management, and human oversight obligations (BRASIL, 2023; TOZZINIFREIRE, 2024).

The processing of PL 2338/2023 in the Senate was marked by intense debates, especially around the inclusion of explicit provisions to ensure the integrity of information and combat disinformation. Opposition to these passages, on the grounds of protecting freedom of expression, led to negotiations and the removal of some mentions, although the principle of "integrity of information" was maintained as one of the foundations of the law (AGÊNCIA BRASIL, 2024a; AGÊNCIA BRASIL, 2024b). This dispute highlights the central tension in AI regulation: how to curb the malicious use of technology without curtailing fundamental rights.

To contextualize and deepen the analysis of the Brazilian regulatory effort, Table 1 presents a detailed comparison between the main points of the European *AI Act* and PL 2338/2023.

Table 1

Comparison between the European Union's AI Act and PL 2338/2023 (version approved by the Senate)

Characteristic	EU AI Act	PL 2338/2023 (version approved in the Senate)	Comparative analysis
Main Approach	Risk-based (unacceptable, high, limited, minimal).	Risk-based (excessive, high).	Both take a risk-based approach, aligning with a global trend. The Brazilian

			categorization is more simplified (BRASIL, 2023; DATAGUIDANCE, 2025).
Prohibited Practices (Unacceptable/Excessive Risk)	It prohibits *government social scoring*, vulnerability exploitation, subliminal manipulation, and real-time remote biometric identification (with exceptions).	It prohibits systems that exploit vulnerabilities, subliminal manipulation, and *social scoring* with unfair effects.	The bans are similar, reflecting a consensus on the most dangerous uses of AI. The Brazilian Bill is aligned with European protection principles (BRASIL, 2023; TOZZINIFREIRE, 2024).
High-Risk Systems (Sectors)	Explicit list of sectors: critical infrastructure, education, employment, essential services, law enforcement, migration, administration of justice, etc.	Similar list: health, justice, security, credit, employment, education, critical infrastructure, autonomous vehicles, biometric identification.	The coverage of high-impact sectors is quite convergent, showing that both texts identify the same points of social vulnerability (BRASIL, 2023; GT LAWYERS, 2025).
High Risk Bonds	Strict requirements for risk management, data governance, technical documentation, transparency for users, human oversight, and cybersecurity.	It requires governance measures, algorithmic impact assessment, safety testing, and human oversight mechanisms.	The obligations are conceptually aligned, focusing on governance and transparency as pillars for the reliability of the systems (BRASIL, 2023; DATA PRIVACY BRASIL, 2024).
Citizens' Rights	Indirect, focused on the obligations of providers. More explicit rights come from other legislation (e.g., GDPR).	Explicit chapter on rights, including information, explanation of automated decisions, challenge, human review, and non-discrimination.	The Brazilian Bill is more explicit in codifying the rights of individuals affected by AI systems, a notable difference and an advance in relation to the European text (DATAGUIDANCE, 2025; DATA PRIVACY BRASIL, 2024).
Governance and Oversight	Establishment of a European AI Board to coordinate national supervisory authorities.	Creation of the National AI Regulation and Governance System (SIA), with the National Data Protection Authority	Both create a centralized governance structure, but Brazil assigns the function to its existing data protection authority, which can

		(ANPD) as the competent authority.	generate synergies but also overload (BRASIL, 2023; DATA PRIVACY BRASIL, 2024).
Sanctions	Fines of up to 35 million euros or 7% of annual global turnover, whichever is greater.	Fines of up to R$ 50 million per infraction or 2% of the group's revenue in Brazil.	Sanctions are robust in both cases, using billing as a basis to ensure that penalties are significant even for large technology companies (GT LAWYERS, 2025; ARTIFICIAL INTELLIGENCE ACT, 2024).

Source: Prepared by the authors (2025), based on Brazil (2023); Dataguidance (2025); GT Lawyers (2025).

The comparative analysis reveals that Brazil is aligning itself with the best international practices, but with important adaptations, such as the explicit emphasis on citizens' rights. However, the effectiveness of this legislation will depend crucially on its implementation and the State's enforcement capacity.

4 DISCUSSION: NAVIGATING THE HYBRID INFOSPHERE

The presentation of the results shows that contemporary society is entering a "hybrid infosphere", where the distinction between human-generated and machine-generated content becomes increasingly blurred. Navigating this new environment requires more than just technological advancements; requires a thorough reassessment of professional practices, ethical frameworks, and regulatory strategies.

The implications for journalism and communication are profound and multifaceted. The era of AGI forces a redefinition of skills and one's own professional identity. The ability to ascertain facts and write texts, although still fundamental, is no longer enough. The need for a new form of expertise emerges: the ability to critically interact with AI systems. This includes knowing how to formulate the right questions to extract useful information, understanding the limitations and biases inherent in algorithmic models, and, above all, exercising responsible curation over content generated or aided by machines (DE-LIMA-SANTOS; SALAVERRIA, 2021; PARANHOS NETO, 2024). The role of the journalist shifts from a mere content producer to that of a validator and contextualizer in an ecosystem saturated with synthetic information.

This transformation demands an urgent update of the profession's codes of ethics. Traditional principles such as truthfulness and impartiality need to be reinterpreted in light of the new challenges. Studies point to the need to incorporate new ethical guidelines, such as radical transparency about the use of AI in news production, ensuring meaningful human oversight at all stages of the editorial process, and implementing robust mechanisms to audit and mitigate algorithmic biases that could lead to discrimination or distortion of reality (ESTRELLA TUTIVÉN; GARDE CANO, 2024; DE-LIMA-SANTOS; SALAVERRIA, 2021). Public trust, an already fragile asset, will depend on the ability of news organizations to adopt these practices proactively and transparently.

In this context, the regulatory response of the State becomes a fundamental pillar. The comparative analysis between PL 2338/2023 and the *European AI Act* shows that Brazil is building a solid legislative base, inspired by an internationally recognized model. The risk-based approach and entitlement enumeration are steps in the right direction. However, the effectiveness of any legislation lies not only in its formulation, but in its ability to be implemented and enforced. Herein lies a critical risk for the Brazilian context: that of a "symbolic regulation".

The process to reach robust legislation on paper is only the first step. Overseeing high-risk AI systems, conducting algorithmic audits, and investigating AI-caused harm require technical capacity and highly specialised financial and human resources (YANG; MENCZER, 2023; ZHOU; ZAFARANI, 2020). The attribution of this responsibility to the ANPD, although logical, imposes a monumental challenge on an agency that is already facing difficulties in fully overseeing the General Data Protection Law (LGPD). In addition, the political disputes that sought to soften the obligations related to the fight against disinformation during the passage of the bill in the Senate (AGÊNCIA BRASIL, 2024b) signal strong pressure from sectors of the technology industry and political groups that may continue to act to weaken the application of the law. Therefore, there is a possibility that Brazil has legislation that is advanced in theory, but with a limited practical application due to implementation bottlenecks, lack of resources and regulatory capture, making it ineffective to deal with the complex challenges of the hybrid infosphere.

This leads us to the conclusion that regulation, by itself, is an incomplete solution. The most resilient and sustainable line of defense against disinformation in the age of AGI is a critically informed citizenry. It is imperative to invest in large-scale media and algorithmic literacy programs. This education must go beyond the traditional teaching of how to identify

fake news. It is necessary to empower citizens to understand the basic principles of operation of recommendation algorithms and AGI models, fostering a healthy skepticism and an active posture in verifying the information consumed (CHESNEY; CITRON, 2019; NIGHTINGALE; FARID, 2022). Without a population able to critically navigate this new environment, even the best of regulations will have a limited impact.

5 CONCLUSION

This review article set out to analyze the paradox of Generative Artificial Intelligence in the communication ecosystem, articulating the argument that the technology operates as a "double helix": on the one hand, a powerful engine for the creation and dissemination of sophisticated disinformation; on the other, a set of essential tools for its detection and mitigation. The analysis, focused on the Brazilian context, reveals that the country is at a crucial moment, seeking to build a regulatory framework while already facing the practical impacts of this disruptive technology.

The synthesis of the literature confirms that AGI, through technologies such as GANs and LLMs, has cheapened and scaled the production of false content, representing a real threat to informational integrity, especially in sensitive contexts such as electoral processes. The phenomenon of the "liar's dividend" emerges as one of the most damaging consequences, where the erosion of trust becomes widespread, affecting the credibility of even authentic content. By contrast, the same technological foundation offers promising solutions for *fact-checking* automation, although these still rely on human oversight and face significant technical challenges.

The final balance points out that there is no single or simple solution to the challenges imposed by IAG. The balance between fostering innovation and protecting society from the risks of algorithmic disinformation depends on a synergistic and multifaceted approach, which should be based on three interdependent pillars:

1. **Robust Technological Regulation:** PL 2338/2023 represents a fundamental step, but its effectiveness will depend on rigorous implementation and technically capable inspection, overcoming the risk of becoming a "symbolic regulation".
2. **Professional Ethical and Technical Adaptation:** Communication professionals, especially journalists, must incorporate new skills and update their codes of ethics to address the curation of synthetic content and transparency in the use of AI tools.

3. **Digitally Empowered Citizenship:** Massive and continuous investment in media literacy and algorithmic programs that prepare citizens to critically consume information in the new hybrid infosphere is crucial.

In view of the above, and recognizing the gaps in current knowledge, an agenda for future research in the area of Communication and Information in Brazil is proposed, which includes:

- **Reception and Impact Studies:** Longitudinal investigations on the cognitive and behavioral effects of continuous exposure to synthetic content on the perception of reality and the formation of public opinion.
- **Effectiveness of Media Literacy:** Applied research to evaluate the effectiveness of different pedagogical approaches in algorithmic literacy, adapted to the diverse sociocultural and educational contexts of Brazil.
- **Regulatory Implementation Analysis:** Studies that critically monitor and analyze the implementation of the AI legal framework, focusing on the operational challenges of the ANPD and other bodies of the governance system, as well as the practical results of its oversight.
- **Development of Interdisciplinary Governance Models:** Fostering research that unites Communication, Law, Computer Science, and Social Sciences to develop AI governance models that are not only technically sound and legally compliant, but also socially fair and culturally appropriate to the Brazilian reality.

The journey to navigate the era of IAG is just beginning. Building a safer and more reliable digital future will depend on the ability of academia, the state, the private sector, and civil society to collaborate in building these three pillars.

REFERENCES

Agência Brasil. (2024a, December 3). Debate sobre desinformação adia votação de projeto que regula IA. Agência Brasil. https://agenciabrasil.ebc.com.br/politica/noticia/2024-12/debate-sobre-desinformacao-adia-votacao-de-projeto-que-regula-ia

Agência Brasil. (2024b, December 5). Acordo sobre desinformação permite aprovação de regulação da IA. Agência Brasil. https://agenciabrasil.ebc.com.br/politica/noticia/2024-12/acordo-sobre-desinformacao-permite-aprovacao-de-regulacao-da-ia

Artificial Intelligence Act. (2024). Brazil AI Act. https://artificialintelligenceact.com/brazil-ai-act/

Amazon Web Services. (2024). O que é um grande modelo de linguagem (LLM)? Amazon Web Services. https://aws.amazon.com/pt/what-is/large-language-model/

Bardin, L. (1977). Análise de conteúdo. Edições 70.

Barros, G. O. (2024, August 2). Uma análise do AI Act. Gabinete de Estratégia e Estudos do Ministério da Economia. https://www.gee.gov.pt/pt/documentos/estudos-e-seminarios/artigos/10243-gee-em-analise-regulamento-da-inteligencia-artificial/file

Batista, A. R. F., & Santaella, L. (2024). Prognósticos das deepfakes na política eleitoral. Organicom, 21(42), 1–15. https://doi.org/10.11606/issn.2238-2593.organicom.2024.208109

Blikstein, I., Fernandes, M., & Coutinho, M. (2024, June 11). Estudo analisa o impacto da inteligência artificial na comunicação corporativa nas redes sociais. FGV Notícias. https://portal.fgv.br/noticias/estudo-analisa-impacto-inteligencia-artificial-comunicacao-corporativa-redes-sociais

Bontridder, N., & Poullet, Y. (2021). AI for counter-disinformation. European Parliament.

Brasil. Senado Federal. (2023). Projeto de Lei nº 2338, de 2023. Dispõe sobre o uso da Inteligência Artificial. https://www25.senado.leg.br/web/atividade/materias/-/materia/157668

Calderonio, V. (2024). The state of the art of the Brazilian Bill on Artificial Intelligence with a focus on civil responsibility. Rivista Italiana di Informatica e Diritto, 2. https://doi.org/10.32091/RIID0087

Centre for Emerging Technology and Security. (2024). AI-enabled influence operations: Safeguarding future elections. The Alan Turing Institute. https://cetas.turing.ac.uk/publications/ai-enabled-influence-operations-safeguarding-future-elections

Chesney, R., & Citron, D. (2019, June 21). Deep fakes: A looming challenge for privacy, democracy, and national security. Lawfare. https://www.lawfaremedia.org/article/deep-fakes-looming-challenge-privacy-democracy-and-national-security

Choraś, M., Demestichas, K., Giełczyk, A., Herrero, Á., Ksieniewicz, P., Remoundou, K., & Woźniak, M. (2021). A survey on machine learning for fake news detection. Information Fusion, 72, 1–22. https://doi.org/10.1016/j.inffus.2021.02.008

Conectas. (2024, August 28). Eleições 2024: Desinformação causa danos concretos na democracia e na vida das pessoas. Conectas Direitos Humanos. https://conectas.org/noticias/eleicoes-2024-desinformacao-causa-danos-concretos-na-democracia-e-na-vida-das-pessoas

CVisionLab. (2023). Deepfake and GANs: How generative adversarial networks work. https://www.cvisionlab.com/cases/deepfake-gan/

Data Privacy Brasil. (2024). IA no primeiro turno: O que vimos até aqui? https://www.dataprivacybr.org/documentos/ia-no-primeiro-turno-o-que-vimos-ate-aqui/

DataGuidance. (2025). Brazil: Similarities and key differences between the Brazilian AI Bill and the EU AI Act. https://www.dataguidance.com/opinion/brazil-similarities-and-key-differences-between

de-Lima-Santos, M. F., & Salaverría, R. (2021). Do jornalismo de dados à inteligência artificial: Desafios enfrentados pelo La Nación na implementação da visão computacional para a produção de notícias. Palabra Clave, 24(3), e2437. https://doi.org/10.5294/pacla.2021.24.3.7

Desinformante. (2024). Inteligência artificial e desinformação | IAí? https://desinformante.com.br/observatorio-ia/

Estrella Tutivén, I. V., & Garde Cano, C. (2024). Ética jornalística na era da inteligência artificial: Rumo a uma atualização dos códigos deontológicos no contexto ibero-americano. Comunicação e Sociedade, 45. https://doi.org/10.17231/comsoc.45(2024).4822

Foletto, L. (2024). Pesquisa em comunicação: Tendências emergentes e desafios no cenário digital. FGV Mídia. https://mestrado-doutorado.fgv.br/blog/noticia/pesquisa-em-comunicacao-tendencias-emergentes-e-desafios-no-cenario-digital

Fontão, F. F., & Dias, J. A. (2022). Bots, fake news, fake faces, deepfakes e sua eventual influência no processo eleitoral democrático. Revista da ANPAL, 1(1). https://doi.org/10.29327/238711.1.1-4

Frontiers. (2025). AI-driven disinformation: Policy recommendations for democratic resilience. Frontiers in Artificial Intelligence, 8. https://doi.org/10.3389/frai.2025.1335867

Goldstein, J. A., & Lohn, A. (2024). Deepfakes, elections, and shrinking the liar's dividend. RAND Corporation. https://www.rand.org/pubs/articles/2024/deepfakes-elections-and-shrinking-the-liars-dividend.html

Gomes, P. A. N., & Oliveira, M. L. P. (2024). Inteligência artificial generativa e a desinformação no Brasil. Siti, 6, e155. https://doi.org/10.20873/uft.2447-4266.2024v6n2e155

Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A., & Bengio, Y. (2014). Generative adversarial nets. Advances in Neural Information Processing Systems, 27. https://papers.nips.cc/paper/5423-generative-adversarial-nets

Graves, L. (2018). Understanding the promise and limits of automated fact-checking. Reuters Institute. https://reutersinstitute.politics.ox.ac.uk/our-research/understanding-promise-and-limits-automated-fact-checking

GT Lawyers. (2025). Marco legal da inteligência artificial no Brasil: O PL 2338/2023 em foco. https://www.gtlawyers.com.br/en/artigos/marco-legal-da-inteligencia-artificial-no-brasil-o-pl-2338-2023-em-foco/

Hjarvard, S. (2014). A midiatização da cultura e da sociedade. Unisinos.

IBM. (2024). What are generative adversarial networks (GANs)? https://www.ibm.com/think/topics/generative-adversarial-networks

Jornal da Cultura. (2024, October 16). Relatório aponta que uso da IA nas eleições não causou impacto negativo [Video]. YouTube. https://www.youtube.com/watch?v=dmjWy8GQbJM

Matz, S. C., Kosinski, M., Nave, G., & Stillwell, D. (2024). AI-driven personalization can be used to influence and manipulate. PNAS, 121(9), e2311891121. https://doi.org/10.1073/pnas.2311891121

Mendes, C. M., & Mattos, M. Â. (2025). Limites e potencialidades de conceitos de desinformação e congêneres acionados em artigos brasileiros sobre covid-19. E-Compós, 28. https://doi.org/10.30962/ec.2858

Nightingale, S. J., & Farid, H. (2022). AI-synthesized faces are indistinguishable from real faces and more trustworthy. PNAS, 119(8), e2120481119. https://doi.org/10.1073/pnas.2120481119

Paranhos Neto, O. S. (2024). A inteligência artificial e seus impactos na produção jornalística: Uma análise sobre a qualidade, veracidade e relevância da informação. Siti, 7, e188. https://doi.org/10.20873/uft.2447-4266.2024v7n3e188

Recuero, R., Gruzd, A., Soares, F. B., Zago, G., & Silva, R. (2025). A desinformação narrada pela checagem: Estudo das eleições de 2018 e 2022. E-Compós, 28. https://doi.org/10.30962/ec.2860

Rodrigues, N., & Mattos, D. (2024, July 8). Estudo mostra uso de inteligência artificial na detecção de fake news. Agência Brasil. https://sintepe.org.br/estudo-mostra-uso-de-inteligencia-artificial-na-deteccao-de-fake-news/

Senado Notícias. (2024, December 10). Senado aprova regulamentação da inteligência artificial; texto vai à Câmara. Agência Senado. https://www12.senado.leg.br/noticias/materias/2024/12/10/senado-aprova-regulamentacao-da-inteligencia-artificial-texto-vai-a-camara

Spinak, E. (2023, August 30). Inteligência artificial e a comunicação da pesquisa. SciELO em Perspectiva. https://blog.scielo.org/blog/2023/08/30/inteligencia-artificial-e-a-comunicacao-da-pesquisa/

Telles, M. K., & Montardo, S. P. (2025). Inteligência artificial e desinformação: Uma análise dos trabalhos apresentados em congressos da SBPJOR de 2015 a 2022. Conexão - Comunicação e Cultura, 22(1). https://doi.org/10.22409/conexaov22n1a56323

The Fix Media. (2024, November 7). AI course: Using AI for fact-checking in journalism. https://thefix.media/2024/11/07/ai-course-using-ai-for-fact-checking-in-journalism/

TozziniFreire. (2024). Lei de Inteligência Artificial da União Europeia entra em vigor. https://tozzinifreire.com.br/boletins/lei-de-inteligencia-artificial-da-uniao-europeia-entra-em-vigor

Yang, K.-C., & Menczer, F. (2023). Cutting through the noise: A machine learning approach to news trustworthiness. Science Advances, 9(48), eadi7638. https://doi.org/10.1126/sciadv.adi7638

Zhou, X., & Zafarani, R. (2020). A survey of fake news: Fundamental theories, detection methods, and opportunities. ACM Computing Surveys, 53(5), Article 109. https://doi.org/10.1145/3395046