

Artificial intelligence, gender and race in the editorial guidelines of legacy media: omissions, challenges and opportunities for revising AI use policies in newsrooms

Inteligência artificial, gênero e raça nas diretrizes editoriais da mídia hegemônica: omissões, desafios e oportunidades de revisão das políticas de uso de IA nas redações

Eliseu Barreira Junior¹ 

Elisabeth Nicolau Saad Correa¹ 

ABSTRACT: *This article presents a critical analysis of the editorial guidelines of eight Brazilian and international media organizations regarding the use of Generative Artificial Intelligence, investigating whether—and how—issues of gender and race are incorporated into these documents. Drawing on the concepts of platformization, intersectionality, and data colonialism, we developed a qualitative study based on content analysis. Manuals and editorial guidelines published on the websites of the selected organizations were examined, focusing on three analytical categories: references to diversity, mitigation of algorithmic bias, and guidance regarding gender and racial stereotypes. The findings indicate that, although these organizations generally acknowledge the existence of biases associated with the use of AI, none incorporates a structured intersectional approach within their guidelines.*

Keywords: *journalism; platformization; intersectionality; data colonialism; artificial intelligence.*

¹Universidade de São Paulo – São Paulo (SP), Brazil.

Editors: Eliza Casadei  and Gabriela Almeida 

RESUMO: *Este artigo apresenta uma análise crítica das diretrizes editoriais de oito veículos da mídia brasileira e internacional sobre o uso de inteligência artificial (IA) generativa, investigando se (e como) questões de gênero e raça são incorporadas nesses documentos. Com base nos conceitos de plataformação, interseccionalidade e colonialismo de dados, realizamos um estudo qualitativo preliminar por meio de análise de conteúdo. Foram examinados manuais e diretrizes editoriais publicados nos websites dos veículos selecionados, com foco em três categorias analíticas: menções à diversidade, mitigação de vieses algorítmicos e orientações sobre estereótipos de gênero e raça. Os resultados demonstram que, embora os veículos reconheçam genericamente a existência de vieses associados ao uso de IA, nenhum incorpora uma abordagem interseccional estruturada em suas diretrizes.*

Palavras-chave: *jornalismo; plataformação; interseccionalidade; colonialismo de dados; inteligência artificial.*

Introduction

In recent decades, the media industry in general, and journalism in particular, have undergone successive waves of digital transformation that have profoundly reshaped the forms of content production, distribution and consumption (MALAR; SAAD, 2024). The digitalization promoted by the commercial internet, from the 1990s onward, challenged traditional models based on physical media (FERRARI, 2003). Subsequently, the rise of digital platforms and social networks, especially from the 2010s onward, radically altered the logic of distribution, amplifying the power of large platforms (POELL; NIEBORG; VAN DIJCK, 2020) and transforming the dynamics of journalistic monetization. Today, a new disruptive wave is underway, marked by the rise of Generative Artificial Intelligence (GenAI), whose depth and breadth bring even more structural impacts to journalistic practice.

The popularization of tools such as ChatGPT, launched in November 2022, has triggered a technological revolution that goes beyond the mere automation of tasks. Newsrooms around the world have been progressively integrating GenAI into supporting editorial processes, such as research, transcription of interviews, headline suggestions and automatic summaries. In Brazil, initiatives such as automated dubbing, editorial chatbots and the generation of news summaries illustrate the rapid incorporation of these tools into journalistic routines (MEIO & MENSAGEM, 2023; ROSA, 2024; SCIRE, 2025).

According to the Reuters Digital News Report 2025 (REUTERS INSTITUTE FOR THE STUDY OF JOURNALISM, 2025), audiences remain broadly skeptical of automated approaches to news production, but are more receptive to the use of AI as a support tool for journalists' work. In the Brazilian case, the report highlights that 9% of the population uses AI chatbots as a source of journalistic information, a percentage already close to the current reach of print journalism, which fell from 50% in 2013 to 10% in 2025.

In this context of the expansion and normalization of AI use in journalism, media groups have been publishing specific editorial guidelines on the technology, underlining that editorial applications of generative AI must always occur under direct human supervision. These documents have been incorporated into editorial style guides and codes of conduct, with varying degrees of detail.

Beyond these dimensions, the growing adoption of AI in journalistic practice also raises an urgent ethical debate about whether—and how—concerns regarding gender and race are incorporated into these normative frameworks. Notorious cases of algorithmic failures make evident the potential of AI to reproduce and even amplify historical prejudices. Examples include errors by large language models, such as Google Gemini, which generated distorted images when illustrating historical figures by ignoring racial contexts (TENBARGE, 2024), or biased responses from chatbots that reproduce gender and racial stereotypes, as documented by the United Nations Educational, Scientific and Cultural Organization (UNESCO) in recent studies (UNESCO; IRCAI, 2024).

In light of this scenario, this article critically analyzes the editorial guidelines of traditional media organizations in the face of the rise of generative AI, with a specific focus on whether—and how—issues of gender and race are recognized and operationalized in these orienting documents. Drawing on the critical frameworks that denounce the algorithmic codification of systemic racism (BENJAMIN, 2019) and on the proposal of re-existence (RICAURTE, 2023), which points to forms of resistance and reconstruction in the face of exclusionary technologies, the article seeks to investigate the extent to which current editorial norms recognize the risks of reproducing stereotypes, erasures and algorithmic biases that affect historically minoritized groups.

To this end, institutional reference documents from major international and Brazilian journalistic organizations will be examined, identifying the presence or absence of intersectional devices in the formulation of AI use policies, the level of sophistication of approaches to algorithmic bias, and the degree of awareness regarding ethical risks for marginalized

groups. From this analysis, the article aims to contribute critically to the contemporary debate on the ethical governance of AI in journalism from an intersectional perspective.

Theoretical-Epistemological Framework

Our epistemological approach articulates a set of interrelated concepts: the analytical devices of intersectionality; the algorithmic biases associated with GenAI; the journalistic narrative editorialized in these contexts; and the platformized environments in which such dynamics are produced and modulated.

When reflecting on issues of gender and race in journalistic organizations—increasingly impacted by GenAI systems—it becomes necessary to make explicit the conceptual correlations that guide our approach, considering that these systems operate through algorithmic devices that influence the production and circulation of journalistic narratives. Thus, when considering intersectional devices, we depart initially from the concept of the *dispositif* proposed by Agamben (2009), developed from the reflections of Foucault and Deleuze. For the author, this refers to “any mechanism, material or immaterial, capable of capturing, orienting, modeling and controlling the gestures, behaviors and opinions of living beings.” By tensioning relations of power and processes of subjectivation, Agamben (2009) reveals how objects, materialities and languages produce social positivities. This process is intensified in contemporary platformized environments, which operate through algorithmic systems and GenAI models.

As a complement to the *dispositif*, we speak of its operation in scenarios of intersectionalities, here treated in a broader and more transversal framing, beyond feminist approaches. In this way, intersectionality can be understood as a methodological instrument (SIRCAR, 2022) for the comprehension of multiple differences—ethnicity, gender, locality, cultures and social relations operated by systems of power that install themselves in public spaces. As detailed in the paragraphs that follow,

we indicate a more direct correlation to the field of intelligent systems through a critical theory applied to the intersectional methodology of their construction, enabling more inclusive and responsible outcomes with regard to local knowledge, specific figurations and imaginaries.

Drawing on the systematic literature review conducted by Poell, Nieborg and Van Dijck (2020), the impacts of the rise of artificial intelligence (AI) on media organizations are embedded in the context of platformization, in which large international organizations—such as Meta, Google and OpenAI—call into question models of content production, distribution and monetization.

In the emergence of a “platform society”, digital platforms have ceased to be mere intermediaries and have come to structure the economic, social and cultural dimensions of public and private life (VAN DIJCK; POELL; DE WAAL, 2018).

Second, in the context of digital journalism (FERRARI, 2003) and algorithmic racism (SILVA, 2022), the transparency of editorial processes becomes inseparable from an ethics of AI (SOTERO; COUTINHO; BEZERRA, 2025) that explicitly confronts its impacts, including gender biases and other intersectionalities, in consonance with the works of authors such as Crenshaw (1989), Lago, Kazan and Thamani (2018), and Collins and Bilge (2020).

Here, algorithmic bias is understood as the set of systematic distortions produced or amplified by computational systems on the basis of training data that reflect pre-existing historical and social inequalities. As Eubanks (2018) and Noble (2022) demonstrate, such biases are not accidental technical failures, but expressions of political and epistemological choices embedded in system design, which tend to reproduce and scale discriminations of race, gender and class.

In this scenario, journalistic newsrooms, while maintaining their internal ethical manuals and editorial principles, as they increasingly incorporate AI into their production routines, operate on technological infrastructure whose algorithmic dynamics are profoundly determined by infrastructural platforms (VAN DIJCK; POELL; DE WAAL, 2018)

dominated by big tech companies, whose own AI models may already contain codified biases.

As Pasquinelli and Joler (2020) argue, generative AI must be understood as part of a regime of “knowledge extractivism”. In this regime, enormous masses of cultural, social and behavioral data are collected, labeled and processed. This process reproduces geopolitical divisions of labor, with invisibilized and under-remunerated labor in peripheral countries. In this way, GenAI develops as part of a continuous regime of “data colonialism” (COULDRY; MEJIAS, 2019), in which social and cultural life itself is converted into an economic resource by hegemonic platforms.

If, on one hand, AI operates through a continuous extractivism of knowledge, on the other it reproduces gender and racial biases present in the databases that feed it. According to a Unesco document (UNESCO; IRCAI, 2024), it is possible to identify gender biases that create stereotypes in automatic translations and amplify the sexualization of women, while racialized groups are associated with criminality or violent behaviors. The study also observes that language models trained predominantly on data from the Global North marginalize voices, cultures and languages from the Global South.

The amplification of intersectional oppressions in the context of the rise of AI brings new dimensions to reflections on discrimination and algorithmic racism. The notion of intersectionality, developed by Crenshaw (1989; 2002) and deepened by Collins and Bilge (2020), demonstrates that forms of discrimination—organized along axes such as race, gender, class and nationality—do not operate in isolation, but articulate and mutually intensify one another, demanding equally integrated analytical and institutional responses.

Colonialism organizes social relations on the basis of the intersection of race and gender to sustain global hierarchies. This process was termed by Quijano (2005) the “coloniality of power”, described by Fanon (2008) as the inscription of colonial violence on racialized bodies, and denounced by Césaire (2006) as the dehumanizing logic of civilizatory

discourse. Current AI systems tend to reproduce, automate and scale these same patterns of domination in the production of knowledge and public discourse.

In the Brazilian and Latin American context, Silva (2022) demonstrates that algorithmic racism is a technological updating of structural racism historically sedimented in the social data that feed these models. Lippold and Faustino (2022) reinforce that the invisibilization or stigmatized representation of racialized populations in datasets reproduces racial hierarchies and global inequalities. To these dimensions, Ricaurte (2023) adds the concept of body-territory, according to which algorithmic violence is not abstract, but materially inscribed on the bodies of women, Black people and Indigenous peoples—shifting the debate on bias from a technical matter to a question of justice and human rights.

Objectives and Methodology

AI models applied to journalism are trained predominantly on databases originating from the Global North. This causes them to replicate what Lugones (2008) describes as the naturalization of the categories of gender and race as colonial fictions of domination. This logic articulates with what Grosfoguel (2011) identifies as the “geopolitics of knowledge”—that is, colonial hierarchies that determine which forms of knowledge are valid, universal and technologically operationalizable. In the Brazilian and Latin American context, Lélia Gonzalez (1984) had already pointed to the specificity of Brazilian-style racism and to the erasure of Black women from systems of representation, a dynamic that AI models tend to perpetuate and scale.

Thus, recommendation algorithms, automatic news generation and editorial prioritization systems may: render invisible narratives and voices from the Global South; reproduce gender and racial stereotypes; automate epistemological erasure.

Therefore, discussing the ethical use of AI in journalism requires going beyond traditional concerns with accuracy and technical transparency, incorporating intersectional reflections on:

- Q1. Who is being represented and who is being silenced?;
- Q2. Which epistemologies inform the training data?;
- Q3. How does the very architecture of the systems perpetuate or challenge the colonial hierarchies of gender and race?

It is in this sense that the critical analysis of documents and editorial principles of hegemonic newsrooms becomes the first indispensable step for thinking through these questions. Examining how Brazilian media organizations, and those situated in the Global North, approach (or neglect) aspects related to gender, race and their intersections allows mapping the degree of ethical and political awareness present in editorial practices in the face of the risks of the algorithmic reproduction of historical inequalities.

Beyond identifying the presence or absence of explicit norms, this analysis seeks to understand the extent to which legacy journalism recognizes (or naturalizes) the colonial and patriarchal patterns of knowledge production, and to what degree its guidelines manage (or fail) to challenge the power structures underlying the adoption of AI.

In view of the scenario outlined above, this article presents a preliminary and qualitative documentary analysis of manuals, codes of conduct and editorial guidelines on the use of AI in eight hegemonic national and international media organizations: G1/Grupo Globo, Folha de S. Paulo and O Estado de S. Paulo (Brazil); The New York Times (NYT) and Associated Press (USA); BBC and The Guardian (United Kingdom); and El País (Spain). The selection followed criteria of institutional relevance, international comparability and public availability of the documents, and was constructed intentionally (purposeful sampling), appropriate to exploratory qualitative studies.

Moreover, the inclusion of organizations headquartered in both the Global North and the Global South enables an analysis of similarities and dissimilarities that is sensitive to structural asymmetries of power, technological dependency and knowledge production, without any claim to statistical representativeness, but seeking to maximize the diversity of organizational contexts and editorial regimes.

For the analysis, articles and links publicly disclosed by each organization regarding guidelines on the use of AI tools by newsrooms and content producers were collected. The objective was to examine the presence, absence or treatment afforded to dimensions of gender, race, diversity, algorithmic justice and intersectionality in these institutional documents, given the growing academic and social debate on the risks of the perpetuation and amplification of structural inequalities by algorithmic systems.

Data collection was carried out through direct access to the institutional websites of the selected organizations, gathering publicly available manuals, codes of conduct and guidelines on the use of AI. The documents were submitted to qualitative content analysis (BARDIN, 2016), following these steps: pre-analysis of the material; thematic exploration and coding; and interpretive treatment of the results. The categories were defined a priori on the basis of the theoretical framework adopted: explicit mentions of diversity and representativeness; policies for the mitigation of algorithmic bias affecting minoritized groups; and guidelines regarding gender and racial stereotypes.

Analysis of the Media Organizations

The Global North media organizations—NYT, BBC, Guardian, El País and AP—operate in the contexts where AI tools are developed and where regulatory power over these technologies is concentrated, which affords them greater proximity to debates on algorithmic governance. The Brazilian organizations—G1, Folha and Estadão—occupy a position of dual dependency: technological, in adopting tools developed by foreign big tech companies; and epistemic, in adapting normative frameworks gestated in racial and social contexts distinct from their own. As Couldry and Mejias (2019) argue, data colonialism operates precisely through the naturalization of this asymmetry.

The document of G1/Grupo Globo¹ generically acknowledges the possibility of bias in the use of AI, stating that “journalists will adopt

1 Available at: <https://g1.globo.com/economia/midia-e-marketing/noticia/2024/06/27/grupo-globo-atualiza-principios-editoriais-para-incluir-orientacoes-sobre-inteligencia-artificial.ghml>. Accessed on: June 16, 2026.

strategies so that any errors and biases produced by artificial intelligence [...] do not result in errors or biases in journalistic coverage” (Section 2-a). However, this reference remains vague, in that it does not specify which forms of bias are at issue, nor does it present specific mitigation guidelines relating to social markers of difference, such as gender, race or class. The absence of an intersectional approach fails to recognize precisely the documented risks in the critical literature on algorithms of oppression (NOBLE, 2022), which warns of the way in which such technologies may reproduce pre-existing discriminations.

In the case of NYT, the guidelines² reaffirm central ethical principles of accuracy, accountability and transparency, directing that AI “serve the journalistic mission of seeking truth and expanding access to information,” with “human oversight and accountability at every stage of AI use.” Even so, when addressing the need to mitigate “bias,” there is no elaboration whatsoever on what types of biases are at issue, nor are there any mentions of gender, race, diversity or algorithmic justice. Such normative silence ignores the structural character of algorithmic biases already extensively discussed (RICAURTE, 2023), resulting in a scope limited to the technical dimensions of the problem.

The BBC’s guidelines³ go somewhat further in explicitly acknowledging a case of racial bias in a facial recognition system, mentioning that “an early use of facial recognition software in the UK passport renewal system found it struggled to identify some skin tones” (Section 4. Algorithmic bias and training data). Despite this isolated acknowledgment, the document treats bias as a technical failure, devoid of sociopolitical or racial contextualization. The structural implications of these failures for the reproduction of systemic inequalities are not elaborated upon. The omission of intersectional perspectives prevents the document from

2 Available at: <https://www.nytimes.com/press/principles-for-using-generative-ai-in-the-times-newsroom/>; <https://www.nytimes.com/editorial-standards/ethical-journalism.html>; <https://www.nytimes.com/2024/10/07/reader-center/how-new-york-times-uses-ai-journalism.html>. Accessed on: June 16, 2026.

3 Available at: <https://www.bbc.co.uk/editorialguidelines/guidance/use-of-artificial-intelligence>. Accessed on: June 16, 2026.

advancing toward an algorithmic governance more committed to social justice (RICAURTE, 2023).

The newspaper *The Guardian*, in its Editorial Code of Practice⁴ (p. 12), recognizes the risks of GenAI, including bias, intellectual property, plagiarism, inconsistency and inaccuracy. It establishes that the use of GenAI in the production of journalistic content must be rare, approved by senior editors and disclosed to the reader. The newspaper's code directs journalists to avoid "prejudiced or derogatory references to race, color, ethnicity, nationality, religion, age, sex, gender identity, sexual orientation, disability." However, this non-discrimination policy remains entirely disconnected from practices involving AI throughout the text. Thus, the code adopts a formal conception of editorial neutrality, but fails to recognize the risk of reproduction of algorithmic inequalities and proposes no specific safeguards in the face of the automation of editorial processes.

The guidelines of *El País*⁵ highlight some relevant practices, such as the mandatory human supervision of content generated with AI support, transparency with the public regarding the use of these tools, and the prohibition of deceptive synthetic content. It also generically acknowledges that there is "bias embedded within generative tools." However, it presents no guidelines on how to identify and mitigate intersectional biases, nor does it address the risk of differentiated impacts on marginalized social groups. Such absence denotes an instrumental vision of AI, focused on maintaining journalistic credibility, but disconnected from critical debates on algorithmic justice.

The document of *Folha de S.Paulo*⁶ clearly delimits that human judgment remains indispensable and that the primary objective of AI use is

4 Available at: https://uploads.guim.co.uk/2025/04/22/Editorial_Code.pdf; <https://www.theguardian.com/help/insideguardian/2023/jun/16/the-guardians-approach-to-generative-ai>. Accessed on: June 16, 2026.

5 Available at: <https://elpais.com/info/codigo-etico/>; <https://elpais.com/defensor-a-del-lector/2025-01-19/el-uso-de-la-inteligencia-artificial-en-el-pais.html>; <https://www.prisa.com/uploads/2025/01/libro-de-estilo-documento-sobre-el-uso-de-herramientas-de-ia.pdf>. Accessed on: June 16, 2026.

6 Available at: <https://temas.folha.uol.com.br/folha-projeto-editorial/manual-de-redacao-conduta/inteligencia-artificial.shtml>. Accessed on: June 16, 2026.

“to increase the efficiency” of operational processes. Some technical risks are acknowledged, such as “hallucinations, plagiarism, biases, technical limitations and copyright issues.” However, “biases” are treated as a broad and underspecified category, without any mention of social, racial or gender dimensions. Folha de S.Paulo thus reproduces a technocratic approach to risks, limiting the governance of AI to editorial review, without incorporating intersectional concerns.

The newspaper O Estado de S. Paulo⁷, in turn, presents one of the most detailed manuals among the Brazilian organizations analyzed, with the creation of an AI Committee and robust rules for review, validation and transparency. It acknowledges the risks of indiscriminate AI use, particularly in journalism, given its role in the democratic state governed by the rule of law. Even so, the document remains anchored in an essentially technical and juridical normative model. There is no reference whatsoever to race, gender, sexuality or social class in the established protocols, nor are there strategies for the mitigation of algorithmic discriminations or for the active promotion of diversity in the validation of the tools used.

Finally, the guidelines of the Associated Press (AP)⁸ reaffirm its commitment to accuracy, impartiality and speed, highlighting that AI may be used for translations, summaries and headline suggestions, provided it is supervised by journalists. It prohibits alterations to audiovisual content and requires rigorous source verification. However, despite presenting a set of editorial safeguards, the AP includes no mention whatsoever of race, gender, diversity or algorithmic justice. Risks are addressed from the perspective of disinformation, but not as an expression of structural inequalities encoded in the AI models themselves.

In general terms, the critical analysis of the documents reveals a common pattern among the organizations analyzed, according to Chart 1: AI use guidelines are strongly oriented by principles of editorial

7 Available at: <https://www.estadao.com.br/link/estadao-define-politica-de-uso-de-ferramentas-de-inteligencia-artificial-por-seus-jornalistas-veja/>. Accessed on: June 16, 2026.

8 Available at: <https://www.ap.org/wp-content/uploads/2024/02/ap-news-values-and-principles-1.pdf>. Accessed on: June 16, 2026.

Chart 1 – Summary of the results of the evaluation of artificial intelligence use policies of the selected journalistic organizations

Organization	Recognition of biases	Explicit mention of gender and race	Strategies for mitigating intersectional risks	Predominant approach
G1 / Grupo Globo (Brazil)	Generically acknowledges the possibility of biases	No explicit mention	Presents no specific strategies	Technical
The New York Times (USA)	Acknowledges the need to mitigate 'bias'	No mention of gender, race or diversity	No specific guidelines	Technical
BBC (United Kingdom)	Acknowledges a specific case of racial bias (facial recognition)	Indirect and isolated mention	No structural mitigation strategies	Technical
The Guardian (United Kingdom)	Acknowledges risks of bias and discrimination editorially	General non-discrimination policy without connection to AI	Proposes no specific safeguards linked to AI	Technical
El País (Spain)	Acknowledges 'bias embedded within generative tools'	No mention of gender, race or diversity	Presents no intersectional strategies	Technical
Folha de S.Paulo (Brazil)	Cites risks of 'biases' in broad terms	No specification of social markers	No intersectional guidelines	Technical
Estadão (Brazil)	Acknowledges risks and creates AI Committee	No mention of gender, race, class or sexuality	No specific diversity strategies	Technical
Associated Press (USA)	Points to risks of disinformation and requires human supervision	No mention of gender, race or algorithmic justice	Presents no intersectional strategies	Technical

Source: Elaborated by the authors.

accountability, human review and technical control of processes. However, there is a systematic silence regarding ethical risks related to the reproduction of intersectional inequalities of gender, race and class.

Ricaurte (2023) uses the concept of body-territory to argue that algorithmic violence is not abstract: it is materially inscribed on the bodies of women, Black people and Indigenous peoples, affecting their safety, dignity and opportunities. The strictly technical approach to bias, as adopted by the BBC and the NYT, ignores this dimension of violence—elevating the critique from a matter of best practices to a question of justice and human rights.

Furthermore, the emphasis on “human supervision”, as proposed by the organizations, functions more as a mechanism for safeguarding brand credibility than as a tool of social justice. An approach inspired by Ricaurte (2023) would require that “supervision” be, in fact, a form of community and intersectional governance.

Although the critical literature has already extensively documented phenomena such as algorithmic racism, such aspects remain absent from the institutional documents of the organizations of the so-called traditional media. As Silva (2022) argues, the advance of algorithmic technologies is anchored in social history itself to sustain the idea of a supposed AI. However, it ends up reproducing and updating oppressions such as structural racism.

The comparison between the Global North and the Global South reveals a paradoxical convergence. Despite the differences in context, all organizations adopt the same technical approach, without incorporating intersectional perspectives. This convergence, however, is not symmetrical. The Brazilian organizations fail doubly: by not questioning the models imported from the North and by ignoring local critical frameworks, such as Brazilian-style racism (GONZALEZ, 1984).

Conclusion

The results reveal a consistent pattern with respect to the three research questions. Regarding Q1, no organization incorporates normative

mechanisms for representativeness or safeguards directed at historically minoritized groups. Regarding Q2, the guidelines treat AI models as technically neutral, ignoring their anchorage in databases predominantly from the Global North. Regarding Q3, the emphasis on human supervision functions more as a safeguard for editorial credibility than as a mechanism of social justice.

The silence observed in the analyzed guidelines with respect to intersectional markers of gender and race is not merely an omission, but part of an epistemic regime that naturalizes inequality as if algorithmic models were neutral. As “The Noosphere Manifesto” synthesizes, AI compresses the world into a statistical model, simplifying social complexities and obliterating subalternized identities (PASQUINELLI; JOLER, 2020).

The absence of social justice principles in editorial guidelines allows the algorithmic curation of journalistic products themselves to reproduce dominant narratives, selectively amplifying or silencing minoritized groups.

A truly critical ethical governance of journalistic AI cannot be limited to technical risk reviews. It is necessary to incorporate effective principles of data sovereignty and historical reparation. Only in this way will it be possible to restore agency to the populations whose cultures and experiences are transformed into algorithmic substrate.

This study presents limitations that must be considered. The sample, composed of eight organizations, makes no claim to statistical representativeness, being appropriate for an exploratory and qualitative analysis. Moreover, the analysis was restricted to publicly available documents, without access to internal editorial practices or unpublished versions of the guidelines. Finally, the field is undergoing rapid transformation, so that new versions of the analyzed documents may have been published after the data were collected.

As an agenda for future research, the following are suggested: analyzing how these guidelines are effectively manifested in the productive routines of newsrooms, through ethnographic studies or interviews with journalists and editors; expanding the sample to include Global South

organizations underrepresented in this study; investigating the extent to which the AI tools licensed by journalistic organizations themselves have internal policies for the mitigation of intersectional biases; and longitudinally tracking the evolution of these guidelines as AI regulation advances in different national contexts.

Conflict of interest: none to declare.

Funding source: none.

Data availability statement: The entire dataset supporting the results of this study has been published within the article itself.

Authors' contribution: Barreira Junior, E.: conceptualization, methodology, formal analysis, investigation, writing of the original draft, review, and editing. Correa, E.N.S.: conceptualization, methodology, formal analysis, investigation, writing of the original draft, review, and editing.

References

- AGAMBEN, G. *O que é o contemporâneo?* Chapecó: Argos, 2009.
- BARDIN, L. *Análise de conteúdo*. São Paulo: Edições 70, 2016.
- BENJAMIN, R. *Race after technology*: abolitionist tools for the new Jim code. Medford: Polity, 2019.
- CÉSAIRE, A. *Discurso sobre el colonialismo*. Madrid: Akal, 2006.
- COLLINS, P. H.; BILGE, S. *Interseccionalidade*. São Paulo: Boitempo, 2020.
- COULDRY, N.; MEJIAS, U. *The costs of connection*: how data is colonizing human life and appropriating it for capitalism. Stanford: Stanford University Press, 2019.
- CRENSHAW, K. Demarginalizing the intersection of race and sex. *University of Chicago Legal Forum*, Chicago, v. 1989, n. 1, p. 139-167, 1989.
- CRENSHAW, K. Documento para o encontro de especialistas em aspectos da discriminação racial relativos ao gênero. *Revista Estudos Feministas*, Florianópolis, v. 10, n. 1, p. 171-188, 2002. <https://doi.org/10.1590/S0104-026X2002000100011>
- EUBANKS, V. *Automating inequality*: how high-tech tools profile, police, and punish the poor. New York: St. Martin's Press, 2018.
- FANON, F. *Pele negra, máscaras brancas*. Salvador: EDUFBA, 2008.
- FERRARI, P. *Jornalismo digital*. São Paulo: Contexto, 2003.
- GONZALEZ, L. Racismo e sexismo na cultura brasileira. *Ciências Sociais Hoje*, São Paulo, p. 223-244, 1984.
- GROSFUGUEL, R. Decolonizing post-colonial studies and paradigms of political economy: transmodernity, decolonial thinking, and global coloniality. *Transmodernity*:

Journal of Peripheral Cultural Production of the Luso-Hispanic World, v. 1, n. 1, 2011. <https://doi.org/10.5070/T411000004>

LAGO, C.; KAZAN, E.; THAMANI, M. Jornalismo e estudos de gênero: a interseccionalidade, onde está? In: CONGRESSO BRASILEIRO DE CIÊNCIAS DA COMUNICAÇÃO, 41., 2018, Joinville, SC. *Annals* [...]. São Paulo: Intercom, 2018.

LIPPOLD, W.; FAUSTINO, D. Colonialismo digital, racismo e acumulação primitiva de dados. *Germinal: Marxismo e Educação em Debate*, Salvador, v. 14, n. 2, p. 56-78, 2022. <https://doi.org/10.9771/gmed.v14i2.49760>

LUGONES, M. Colonialidad y género. *Tabula Rasa*, Bogotá, n. 9, p. 73-101, July/Dec. 2008.

MALAR, J. P.; SAAD, E. Transformação digital no jornalismo brasileiro: reflexões a partir de um panorama geral. In: SAAD, E. (org.). *Olhares híbridos sobre a comunicação contemporânea: a trajetória de 18 anos do COM+*. Cachoeirinha: Fi, 2024. p. 203-230.

MEIO & MENSAGEM. Folha cria plataforma de receitas com uso de inteligência artificial. *Meio & Mensagem*, 2023. Available at: <https://www.meioemensagem.com.br/midia/folha-cria-plataforma-de-receitas-com-uso-de-inteligencia-artificial>. Accessed on: June 22, 2025.

NOBLE, S. U. *Algoritmos da opressão: como os mecanismos de busca reforçam o racismo*. Santo André: Rua do Sabão, 2022.

ORGANIZAÇÃO DAS NAÇÕES UNIDAS PARA A EDUCAÇÃO, A CIÊNCIA E A CULTURA (UNESCO); INTERNATIONAL RESEARCH CENTRE ON ARTIFICIAL INTELLIGENCE (IRCAI). *Challenging systematic prejudices: an investigation into bias against women and girls in large language models*. Paris: Unesco, 2024.

PASQUINELLI, M.; JOLER, V. *O manifesto nooscópio: inteligência artificial como instrumento de extrativismo do conhecimento*. KIM Research Group; Share Lab, 2020. Available at: <https://nooscope.ai>. Accessed on: July 6, 2025.

POELL, T.; NIEBORG, D.; VAN DIJCK, J. Plataformização. *Fronteiras – Estudos Midiáticos*, v. 22, n. 1, p. 2-10, Jan./Apr. 2020. <https://doi.org/10.4013/fem.2020.221.01>

QUIJANO, A. Colonialidade do poder, eurocentrismo e América Latina. In: LANDER, E. (org.). *A colonialidade do saber: eurocentrismo e ciências sociais: perspectivas latino-americanas*. Buenos Aires: CLACSO, 2005. p. 107-130.

REUTERS INSTITUTE FOR THE STUDY OF JOURNALISM. *Digital News Report 2025*. Oxford: Reuters Institute, 2025.

RICAUURTE, P. Resistencia como reexistencia: la defensa del cuerpo-territorio en la sociedad algorítmica. *Pléyade*, n. 32, p. 65-92, 2023. <https://doi.org/10.4067/S0719-36962023000200065>

ROSA, V. Documentário da Globoplay tem dublagem realizada por IA; veja impressões. Estadão, 2024. Available at: <https://www.estadao.com.br/cultura/series/documentario-da-globoplay-tem-dublagem-realizada-por-ia-veja-impressoes-nprec/>. Accessed on: June 22, 2025.

SCIRE, S. *Let's get to the point: three newsrooms on generating AI summaries for news*. Nieman Lab, 2025. Available at: <https://www.niemanlab.org/2025/06/lets-get-to-the-point-three-newsrooms-on-generating-ai-summaries-for-news/>. Accessed on: June 22, 2025.

SILVA, T. *Racismo algorítmico*: inteligência artificial e discriminação nas redes digitais. São Paulo: Sesc SP, 2022.

SIRCAR, S. Emplacing intersectionality: autoethnographic reflections on intersectionality as geographic method. *Gender, Place & Culture*, v. 29, n. 7, p. 903-922, 2022. <https://doi.org/10.1080/0966369X.2021.1891865>

SOTERO, P. F. G.; COUTINHO, F. G. A.; BEZERRA, L. Os desafios éticos da inteligência artificial no jornalismo. *Comunicologia*, Brasília, v. 18, n. 1, p. 251-268, 2025.

TENBARGE, K. Google making changes after Gemini AI portrayed people of color inaccurately. *NBC News*, 2024. Available at: <https://www.nbcnews.com/tech/tech-news/google-making-changes-gemini-ai-portrayed-people-color-inaccurately-rcna140007>. Accessed on: June 22, 2025.

VAN DIJCK, J.; POELL, T.; DE WAAL, M. *The platform society*. Oxford: Oxford University Press, 2018.

About the authors

Eliseu Barreira Junior: Master's student in Communication Sciences at ECA-USP. Bachelor in Journalism from the same institution. Holds an MBA in Business Management from Fundação Dom Cabral (FDC). He is a member of the COM+ research group. E-mail: eliseu.barreira@usp.br

Elisabeth Nicolau Saad Correa: Senior full professor in the Department of Journalism and Publishing at ECA-USP. She is the coordinator of the COM+ research group, international speaker and digital strategist. E-mail: bethsaad@gmail.com

Received: 12/28/2025

Accepted: 04/22/2026