e-ISSN nº 2447-4266
Palmas, v. 11, n. 1, 2025
http://dx.doi.org/10.20873/uft.2447-4266.2025v11n1a23en

INTERSEMIOTIC MISINFORMATION IN AUDIO DESCRIPTION: Artificial Intelligence in Inclusive School Settings

DESINFORMAÇÃO INTERSEMIÓTICA NA AUDIODESCRIÇÃO: A Inteligência Artificial em Contextos Escolares Inclusivos

DESINFORMACIÓN INTERSEMIÓTICA EN LA AUDIODESCRIPCIÓN: Inteligencia Artificial en Entornos Escolares Inclusivos

Adriano Alves da Silva
Master's Degree in Communication and Society (UFT); MBA in Corporate Communication and Marketing; Postgraduate in Innovative Educational Methodologies; Postgraduate in Art Therapy; Bachelor's Degree in Social Communication; Licensed in Visual Arts.
agencia.adriano@gmail.com
0000-0003-0654-7034

Anderson da Costa Lacerda
Master's Degree in Science and Biotechnology (UFF); Postgraduate in Teaching for Professional and Technological Education, Federal Institute of Espírito Santo (IFES). Postgraduate in Strategic Innovation Management and Science and Technology Policy, Federal University of Tocantins (UFT); Bachelor's Degree in Information Technology Management, Estácio de Sá University – Rio de Janeiro, RJ.
costa_lacerda@id.uff.br
0000-0002-1862-2743

Francisco Gilson Rebouças Pôrto Júnior
PhD in Communication and Contemporary Culture from the Faculty of Communication, Federal University of Bahia (UFBA); Master's Degree in Education from the Faculty of Education (UnB); Bachelor's Degree in Social Communication/Journalism (ULBRA) and Pedagogy (UnB). Currently leader of the Lattes Group Observatory of Applied Research in Journalism and Teaching (OPAJE-UFT). Coordinates research on digital journalism education and memory preservation. Research interests include training, teaching, and educational processes in Brazil and the European Union.
gilsonporto@uft.edu.br
0000-0002-5335-6428

Mailing address: ARSO 61, Alameda 8, Lot 13, House 01. Palmas/TO, Brazil. ZIP Code: 77016-354.

Received in: 25.08.2025.
Accepted in: 25.09.2025.
Published in: 22.10.2025.

ABSTRACT
This study critically analyzes the risks of intersemiotic misinformation caused by the use of generative artificial intelligence (GAI) in the audio description of artistic images aimed at students with visual impairments (PwDV) in the school environment. The research used an original image submitted to the ChatGPT and Gemini platforms to generate automated audio descriptions, which were analyzed in the discursive context, with critical validation by PwDV consultants. The results indicated symbolic erasures, aesthetic simplifications and lack of sensitive contextualization, evidencing risks to aesthetic enjoyment and inclusion. Although these are promising technologies, this study concludes that the use of AI tools requires ethical and pedagogical mediation. Generative AI does not replace listening, subjectivity and human experience, requiring teacher training and well-founded inclusive policies.

KEYWORDS: Audio Description; Artificial Intelligence; Intersemiotic Misinformation; School Inclusion; PwDV; Art and Technology.

e-ISSN nº 2447-4266
Palmas, v. 11, n. 1, 2025
http://dx.doi.org/10.20873/uft.2447-4266.2025v11n1a23en

Introduction

The emergence of Generative Artificial Intelligence (GAI), such as OpenAI's ChatGPT and Google's Gemini, has substantially altered the ways in which information is produced, circulated, and mediated in the educational space. Presented as accessible and efficient tools, these technologies are increasingly being incorporated into pedagogical practices that aim to promote the inclusion of students with visual impairments (PcDV). However, these tools operate on the basis of opaque algorithms and, although sold under the label of neutrality and universality, often reproduce hegemonic, normative, and reductionist patterns (UNESCO, 2023).

In this scenario, it is necessary to question how teachers, even those who are well-intentioned and naively imbued with the desire to promote accessibility, end up, due to technical ignorance or lack of critical training, collaborating with processes of intersemiotic misinformation — that is, the impoverished or distorted translation of visual content into other semiotic languages, such as verbal. Structural misinformation in the educational field stems not only from the dissemination of false or misleading content, but also from the reproduction of pedagogical practices detached from their social and symbolic contexts. When it comes to the mediation of images by IAGs, the problem becomes more acute, since the image, by its very polysemic nature, already presents itself as an interpretive challenge between human interlocutors. Visual works do not carry a single meaning, but a plurality of possible readings, conditioned by cultural, subjective, and historical factors (Didi-Huberman, 2003; Kress; Van Leeuwen, 2006). Automating this process, therefore, carries substantial risks of symbolic silencing and erasure of meaning, especially for vulnerable audiences such as PwDVs.

The critique of so-called technocratic inclusion, formulated by Boaventura de Sousa Santos (2016), helps to understand this problem. For the author, this is a "top-down" inclusion, centered on technical-managerial solutions that neglect the protagonism of the subjects and their local epistemologies. Technocratic inclusion, in this sense, tends to depoliticize exclusion, reducing it to a functional problem that can be solved through the introduction of digital devices or platforms. In doing so, it homogenizes educational practices, ignores cultural specificities, and reinforces the dependence of excluded populations on those who possess technical knowledge. Even more seriously, it promotes a logic of integration into the current system, without transforming it, moving away from genuinely emancipatory inclusion.

e-ISSN nº 2447-4266
Palmas, v. 11, n. 1, 2025
http://dx.doi.org/10.20873/uft.2447-4266.2025v11n1a23en

This article thus starts from the hypothesis that the uncritical use of IAGs in the audio description of artistic images for PcDVs, especially when performed by teachers without specific training, can undermine the aesthetic and cognitive experience of these subjects, converting accessibility into misinformation. The analysis undertaken here is linked to the tradition of French Discourse Analysis (Pêcheux, 1990; Orlandi, 2005), with the visual grammar of Kress and van Leeuwen (2006), and with the principles of A/R/Tography (Irwin et al., 2004), valuing situated, responsive, and sensitive aesthetic mediation.

In light of UNESCO guidelines (2023), there is a need for a critical pedagogy of technologies, in which artificial intelligence is treated as a complementary ally — rather than a substitute — for listening, subjectivity, and human mediation.

Theoretical Foundations

Audio description as intersemiotic mediation

Audio description (AD), as an intersemiotic translation practice, transcends the mere transposition of the visual to the verbal. It represents a gesture of mediation that involves linguistic, ethical, and aesthetic choices, traversed by sociocultural contexts, subjectivities, and regimes of visibility. According to Neves (2011), AD is not just about making visible what cannot be seen, but about constructing meanings that respect the complexity and polysemy of the original image, as well as the intersubjectivities inherent to each subject.

Describing an image to a visually impaired person involves more than just stating what is there. It implies recognizing what can be said, how it can be said, and to whom it is said. In this sense, AD should be thought of as a situated discursive practice, permeated by power relations, normative discourses, and semantic disputes. It is at this point that intersemiotics intersects with ethics: the audio-describer not only translates, but interprets and mediates meanings.

From an intersemiotic perspective, as argued by Jakobson (1959) and Hurtado (2008), the transposition between semiotic codes (from visual to verbal, for example) is neither neutral nor automatic. Mediation depends on cultural, intentional, and experiential filters. The visual grammar of Kress and van Leeuwen (2006) reveals that elements such as color, angle, proximity, and framing communicate dimensions that cannot be simply reduced to words. Attempting to automate this process through AI

e-ISSN nº 2447-4266
Palmas, v. 11, n. 1, 2025
http://dx.doi.org/10.20873/uft.2447-4266.2025v11n1a23en

runs the risk of producing an "impoverished translation" of the image, focused on apparent materiality and disregarding sensitive, contextual, and poetic aspects.

Furthermore, the classification of the AD modalities adopted in this study — objective, functional, interpretive, and poetic (Motta, 2010) — exposes the diversity of possible approaches, none of which can be considered "neutral" or universal. Each is anchored in different communicative purposes: to inform, guide, excite, or sensitize. The automatic adoption of a single model by AI, without proper alignment with the target audience profile, the pedagogical context, or the principles of critical accessibility, can misinform rather than include.

Therefore, understanding AD as intersemiotic mediation implies assuming its condition of situated discursive production. This requires not only technical competence, but also aesthetic training, sensitive listening, and an ethical commitment to the plurality of possible meanings. It is in this context that this article questions the application of generative AIs, arguing that the automation of AD — when devoid of human and sensitive validation — transforms a technology of access into a vector of symbolic and aesthetic misinformation.

Image as a polysemic sign

The image, as a sign, operates in a field of open, multiple, and often contradictory meanings. Its polysemic nature positions it as a powerful vector of communication, evoking readings that go beyond the intentionality of those who produce or encode it. As stated by Adriano Alves da Silva (2019), "the image is endowed with social and symbolic agency, capable of interfering in the way subjects perceive themselves and are perceived" (Silva, 2019, p. 8). This relational capacity of the image is traversed by cultural, historical, ethical, and aesthetic dimensions that challenge automated systems of reading and description.

Alfred Gell's (2018) contribution is central to this discussion. In his theory of art as a system of action, Gell proposes that images are social agents, endowed with material and symbolic efficacy. Rather than being merely passive representations, they "act upon" subjects, establishing networks of distributed agency between the artist, the work, the audience, and the context. According to the author, art functions as an agency trap, that is, as an artifact that captures the attention, emotion, and behavior of the subjects involved in its enjoyment. This perspective shifts the focus from the image as representation to the image as relationship.

e-ISSN nº 2447-4266
Palmas, v. 11, n. 1, 2025
http://dx.doi.org/10.20873/uft.2447-4266.2025v11n1a23en

Gilles Deleuze's (1990) thinking deepens this understanding by proposing that the image, especially the artistic image, is not a reflection or representation of reality, but the production of a given reality. The image, in this sense, "is" in itself a force, a vibration that traverses bodies, affections, and thoughts. In his philosophy of difference, the image escapes the logic of identity and fixed representation, opening itself to becoming and multiplicity. This is why Deleuze and Guattari (1995) suggest that art is always a plane of consistency — a space in which signs interconnect beyond dominant codes.

Jacques Aumont, a scholar of image and visuality, reinforces this idea by stating that "every image carries within itself not only what it shows, but above all what it hides, what it provokes, what it interrupts" (Aumont, 2011, p. 74). In this sense, the image becomes a place of semiotic, ideological, and epistemic dispute. It is this latent — and often silenced — dimension that automated description processes ignore or neutralize.

In educational contexts, such as those involving the mediation of images for visually impaired students, such neutralization translates into misinformation. When AI disregards polysemy and reduces the image to a set of decontextualized formal elements, it not only impoverishes the aesthetic experience of PcDV, but also reproduces a model of cognition centered on literalness, emptying the critical and subjective power of art. As Didi-Huberman (2003) rightly warns, seeing is, above all, seeing what is missing. Describing an image is thus also an act of choice, power, and ethical responsibility.

By treating the image as a polysemic sign, this article advocates for the need for interpretive approaches that are sensitive to otherness in the audio description process. This implies recognizing that no description will be definitive, neutral, or free of bias. What is required is an ecology of the senses (Rolnik, 2006), in which technologies are at the service of diversity of perspectives — and not their normalization.

Generative artificial intelligence and algorithmic bias

Generative Artificial Intelligence (GAI) is emerging as one of the most influential technologies of the 21st century, with the potential to radically transform educational, cognitive, and cultural processes. However, its indiscriminate and unregulated use, especially in inclusive educational practices, reveals structural contradictions and profound ethical risks. Among these, algorithmic bias stands out: a form of structural

e-ISSN nº 2447-4266
Palmas, v. 11, n. 1, 2025
http://dx.doi.org/10.20873/uft.2447-4266.2025v11n1a23en

misinformation that acts silently in the automated production of content, including the audio description of artistic images for people with visual impairments (PVI).

As Lucia Santaella (2019) argues, AI is part of the logic of digital hypercoding, where data governs the ways of knowing and representing the world. Such a cognitive regime is not neutral: it is structured on databases historically fed by dominant, exclusionary, and normative narratives. Consequently, the responses produced by AIs are not just statistical syntheses, but projections of pre-formatted possible worlds.

Dora Kaufman (2021), reflecting on human-machine interaction, warns against the myth of algorithmic autonomy. What we see, according to the author, is a symbolic delegation of critical judgment to systems that share neither aesthetic experience nor ethical responsibility. In the case of audio description, this delegation becomes even more sensitive, as it directly affects the way a person with visual impairment accesses the world of images. The absence of human agency in this process is, in fact, a depletion of the mediated aesthetic experience.

Shoshana Zuboff's (2019) critique of surveillance capitalism is also relevant in this context. The author argues that AI systems operate on an extractive logic of data, capturing behaviors, inferences, and patterns to feed back into predictive algorithms. In the field of education, this means that even the sensitive experience of people with visual impairments can be transformed into data, serving marketing rather than emancipatory purposes.

André Lemos (2023) broadens this debate by pointing to digital colonialism: a process by which technological infrastructures, mostly developed in the Global North, impose themselves on other forms of knowledge and cultural production. In this scenario, AI reinforces a dominant semiotics that erases otherness and subjectivities. In the case of images, this is reflected in normative, Eurocentric, and ableist descriptions produced by systems trained with corpora lacking diversity.

Martha Gabriel (2017), in turn, draws attention to the paradox of connectivity: the more connected we are, the more vulnerable we are to information manipulation. The author warns of the need to develop critical digital literacy — something that becomes even more urgent when dealing with the relationship between AI, inclusion, and misinformation.

Thus, the promise of algorithmic neutrality falls apart in the face of the multiple biases embedded in AAI systems. These biases are not technical flaws, but symptoms of a technocognitive project that still lacks ethics, plurality, and sensitive listening. When

e-ISSN nº 2447-4266
Palmas, v. 11, n. 1, 2025
http://dx.doi.org/10.20873/uft.2447-4266.2025v11n1a23en

applied to audio description, these systems risk reinforcing symbolic and cognitive inequalities under the guise of automated and depoliticized inclusion.

Educational misinformation and informational vulnerability of PcDV

The growing insertion of digital technologies in education has highlighted a fundamental paradox: although they promise to expand access to knowledge, these same technologies have intensified processes of misinformation, especially among students in situations of greater cognitive and social vulnerability, such as people with visual impairments. The introduction of IAG as a teaching tool in Brazilian schools has exposed not only the precariousness of teacher training in assistive technologies, but also the reproduction of a teaching model that ignores the ethical, aesthetic, and epistemic foundations of critical accessibility. This reveals that the mechanisms of ignorance and misinformation production in the educational context are not limited to the circulation of false content, but are rooted in teacher training structures, public policies, and the absence of material and symbolic conditions for critical pedagogical practice with technologies. In this sense, misinformation is not a contingent error, but a consequence of a training logic that neglects both digital literacy and an understanding of the specific needs of students with disabilities.

According to the UNESCO Report (2023), the use of AI in educational contexts must be preceded by a careful analysis of the risks, especially for the most vulnerable audiences. The entity warns that "the hasty integration of automated technologies can accentuate inequalities, reinforce stereotypes, and limit epistemic plurality in the school environment" (UNESCO, 2023, p. 45). In line with this concern, the National Common Core Curriculum (BNCC) (Brazil, 2018) recognizes the importance of digital literacy and critical analysis of information as fundamental skills for citizen education, but does not specify practical guidelines for teachers who work with accessibility resources, nor does it address the peculiarities of using AAI with PwDVs.

Authors such as Josélia Neves (2011) and Lívia Motta (2010) argue that audio description is not a technical act, but a process of sensitive, situated, poetic translation that is responsive to the particularities of the audience. Neves (2011, p. 94) warns that "neutrality in audio description is a fallacy," since all mediation is permeated by discursive choices. In the case of AAD, these choices are made based on biased databases that do not consider the sensory, cultural, and political heterogeneity of people with visual impairments.

e-ISSN nº 2447-4266
Palmas, v. 11, n. 1, 2025
http://dx.doi.org/10.20873/uft.2447-4266.2025v11n1a23en

Added to this are weaknesses in teacher training. As Araújo et al. (2025) point out, teachers have difficulty recognizing and combating digital misinformation precisely because they have not been prepared to critically operate with technologies in the classroom. Naivety regarding AI manifests itself in the belief that tools such as ChatGPT or Gemini operate in a "correct" and "objective" manner when translating images, when in fact they reproduce hegemonic aesthetics, erasing the plurality of meanings in artistic images and compromising epistemically fair access to visual culture.

In this sense, there is an urgent need for continuing education that considers the assumptions of epistemic justice (Fricker, 2007) and critical pedagogy of accessibility. The informational vulnerability of PcDV, therefore, cannot be seen as a cognitive deficiency, but as a side effect of the absence of effective public policies and teaching practices engaged with the complexity of intersemiotic mediation.

Critical inclusion and UNESCO guidelines

The concept of inclusion as a process in constant construction, rather than a fixed goal to be achieved, represents one of the pillars of contemporary thinking in the field of educational policies and critical accessibility. This perspective is present in the academic work of authors such as Mônica Pereira dos Santos (2016) and several researchers linked to this field (Santos, 2016; Sousa; Oliveira; Castro, 2021).

For Santos (2016), critical inclusion requires the denaturalization of exclusionary school practices and the epistemological repositioning of historically marginalized voices. The author argues that the paradigm of inclusion should be understood as an ethical, political, and aesthetic stance that is not restricted to formal access, but aims to transform the ways of teaching, learning, and living with difference.

This critical approach is shared by the guidelines of the United Nations Educational, Scientific and Cultural Organization (UNESCO). In the report Technology in Education: A Tool on Whose Terms? (2023), UNESCO warns that the integration of technologies in education must occur in a regulated, participatory, and socially situated manner. The document states that "without adequate governance structures and teacher training, the use of technologies, including artificial intelligence, can widen inequalities and deepen pre-existing exclusions" (UNESCO, 2023, p. 19).

The Brazilian regulatory framework also reflects these tensions. The Federal Constitution of 1988, in its article 205, establishes that education is a right of all and a

e-ISSN nº 2447-4266
Palmas, v. 11, n. 1, 2025
http://dx.doi.org/10.20873/uft.2447-4266.2025v11n1a23en

duty of the State and the family, and should be promoted and encouraged with the collaboration of society, aiming at the full development of the individual and their preparation for the exercise of citizenship. Law No. 13,146/2015, known as the Statute of Persons with Disabilities or Brazilian Inclusion Law, consolidates the right of persons with disabilities to education on an equal basis with other persons, at all levels and in all forms of education (Brazil, 2015).

However, the implementation of these measures faces recurring obstacles: budget cuts, legal setbacks, ideological disputes, discontinuity of public policies, and institutional resistance. These challenges make it clear that the struggle for inclusion takes place in a conflictual environment and requires constant vigilance on the part of civil society, the scientific community, and people with disabilities themselves.

The National Common Core Curriculum (BNCC) recognizes the appreciation of diversity, empathy, respect, and the promotion of human rights as general competencies of basic education (Brazil, 2018). However, the lack of detail on the pedagogical operationalization of inclusion, especially with the use of technologies, highlights a gap between norm and practice.

In this scenario, critical inclusion requires the recognition of difference as a value and disability as a legitimate expression of human diversity. It therefore means creating educational ecosystems in which digital technologies — including AI — are appropriated in an ethical, creative, and contextualized manner. This presupposes active listening to PwDs, continuing teacher training, and intersectoral engagement.

As Sousa et al. (2021, p. 12) emphasize, "inclusion is not a natural or spontaneous condition, but a collective construction that requires solid public policies, plural epistemologies, and dialogical educational practices." Ultimately, it is a matter of disputing the meaning of education itself: whether it will be an instrument for reproducing inequalities or a field for inventing new possible worlds.

Methodology

This is a qualitative, exploratory, applied research study that adopts a multiple case study with a discursive approach as its methodological procedure. The research comparatively analyzes the audio descriptions generated by generative AI tools (ChatGPT and Gemini) based on an unpublished artistic image, submitted to two platforms, considering the impacts on aesthetic mediation for students with visual impairments (PcDVs). The analysis was complemented by opinions from PwVS

e-ISSN nº 2447-4266
Palmas, v. 11, n. 1, 2025
http://dx.doi.org/10.20873/uft.2447-4266.2025v11n1a23en

consultants, focusing on elements of symbolic erasure, simplification, and lack of contextualization. The research is based on two main methodological axes: A/R/Tography (Irwin et al., 2004) and French Discourse Analysis (Pêcheux, 1990; Orlandi, 2005). The methodology seeks to reveal how meanings mediated by AI are discursively constructed in the intersemiotic translation of images for people with visual impairments.

A/R/Tography is structured around the intertwining of the roles of Artist, Researcher, and Teacher. In this research, a two-dimensional artistic image is chosen, containing visual elements loaded with subjective representativeness. This image is used as the starting point for the central experiment of the study, conceived by the researcher-artist himself. This image is submitted to two different generative artificial intelligences with image interpretation capabilities (ChatGPT and Gemini), using the same standardized prompt (Create audio description for PcDV), in order to ensure uniformity in the textual input.

The audio descriptions generated by each AI are then analyzed through the lens of French Discourse Analysis, with special attention to discursive formations, effects of meaning, silences, and ideological positions implicit in linguistic choices. Based on French Discourse Analysis, the following are investigated:

- The discursive formations that structure the description;
- The effects of meaning produced and the ideological positions implied;
- The elements of the unsaid or silenced: what was omitted in the description? What visual, symbolic, or contextual layers were ignored or erased by the AI?
- Classification of audio descriptions according to the following models: objective (neutral and descriptive), functional (aimed at practical understanding), interpretive (with subjective analysis), and poetic (with metaphorical and sensitive language).

Complementing this approach, the research is epistemologically anchored in the phenomenological-hermeneutic perspective, which allows access to the meanings attributed by subjects to experiences of aesthetic mediation through artificial intelligence technologies. This epistemological choice reinforces the commitment to understanding the experience lived by visually impaired students and the researchers involved, recognizing the place of subjectivity, historicity, and language in the

e-ISSN nº 2447-4266
Palmas, v. 11, n. 1, 2025
http://dx.doi.org/10.20873/uft.2447-4266.2025v11n1a23en

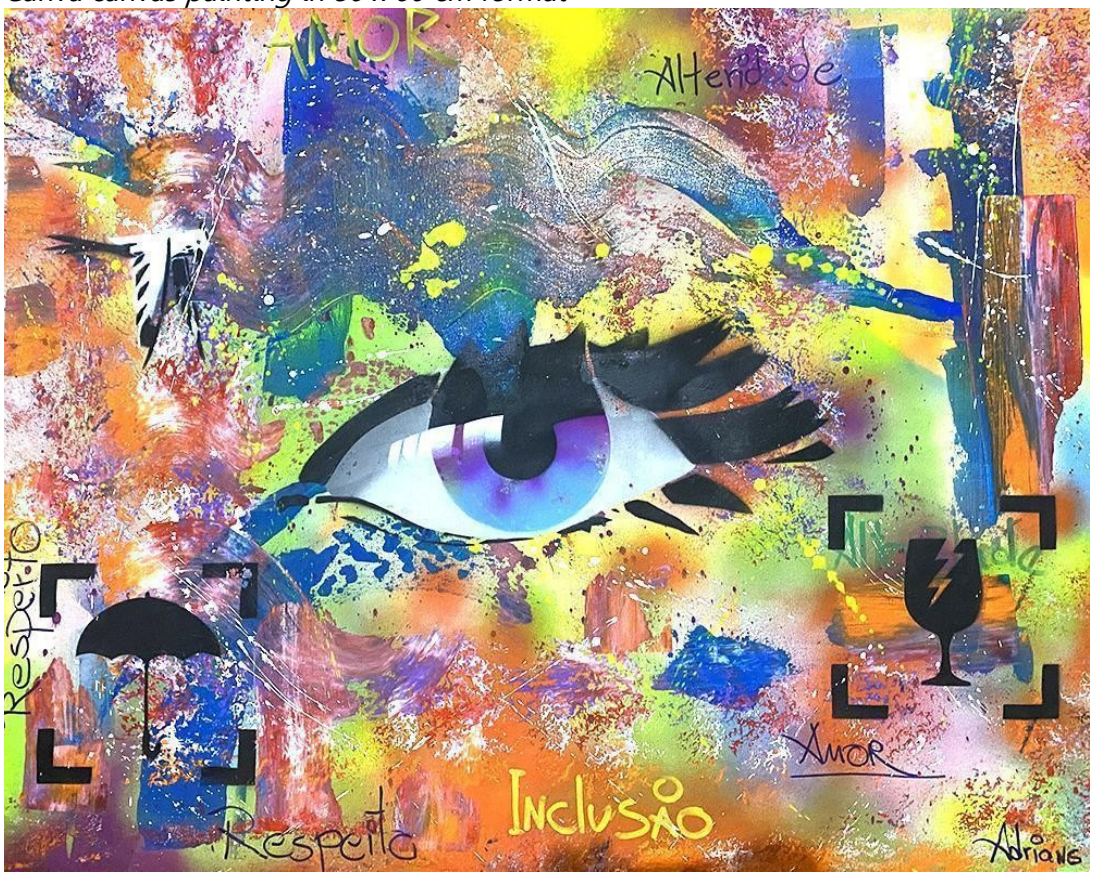
constitution of meanings (Gadamer, 1999; Ricoeur, 1997). Thus, the research does not seek only to technically describe audio descriptions, but to interpret the ways in which they operate discursively, interfering with the possibilities of enjoyment, understanding, and aesthetic inclusion. The intertwining of art, technology, and education is therefore analyzed as a symbolic field of disputes over meaning, where tensions emerge between algorithmic automation and sensitive human listening, considering that AI technologies are not neutral but encode structures of power and discursive exclusion (Dreyfus, 1992).

Choice of authorial image

The research began with the selection of a two-dimensional artistic image from the personal collection of the author, who is also a visual artist. An unpublished authorial work, not yet disclosed in digital or public media, was chosen as a methodological strategy to mitigate the possibility of prior recognition by the artificial intelligences used. This precaution aims to reduce the risk of algorithmic contamination, understood here as the incorporation of inappropriate or previously memorized patterns by the models, which could compromise the originality of the generated response. This caution is in line with the warning by Bender et al. (2021) when discussing the dangers of spurious memorizations — artificial and non-generalizable associations learned by language and image models due to biased or widely replicated data. In other words, it is a matter of avoiding responses compromised by redundant features, contaminated by multiple accumulated representations, whose forced convergence can generate semantic noise or, in more critical terms, a cognitive emptying of the message — a kind of algorithmic dumbing down of the content produced.

e-ISSN nº 2447-4266
Palmas, v. 11, n. 1, 2025
http://dx.doi.org/10.20873/uft.2447-4266.2025v11n1a23en

Figure 1
Canva canvas painting in 80 x 60 cm format



Source: Prepared by the author.

The selected image presents a mixed composition of non-figurative and figurative elements and contemporary expressive resources, such as fragmented lines, overlays, color palettes, and symbolic icons. This configuration was intentional: the aim was to provoke the interpretive limits of the AIs analyzed, considering that, as Dondis (1997) points out, reading complex images requires not only formal perception, but also an aesthetic repertoire and contextual sensitivity. Thus, the semantic density of the image was conceived as a deliberate challenge to automated models of intersemiotic translation.

From the perspective of A/R/Tography, image selection is a research gesture that incorporates artistic practice as a mode of knowledge production (Irwin et al., 2004). In this sense, the image is not only an object of analysis, but also a methodological device that mobilizes subjectivities, aesthetics, and performativities. As Hernández (2007) points out, arts-based research allows us to destabilize dichotomies between subject and object, science and sensitivity, proposing other modes of visual and educational inquiry.

e-ISSN nº 2447-4266
Palmas, v. 11, n. 1, 2025
http://dx.doi.org/10.20873/uft.2447-4266.2025v11n1a23en

Submission of the image to generative AIs

The image was submitted to two different generative artificial intelligence platforms with image reading and description capabilities: ChatGPT-4V (OpenAI) and Gemini (Google). These tools were chosen because of their widespread use, free accessibility, and penetration in the informal educational field. Each AI received the same textual instruction (prompt): "Create audio description for PcDV." The standardization of the command was intended to ensure the comparability of the outputs generated, simulating the use that would be made by educators without specific training in audio description.

The resulting texts were collected and stored for later discursive and imagery analysis, as described in the following steps.

Classification of audio description

The audio descriptions generated by the AIs (ChatGPT and Gemini) were initially analyzed individually and then compared in terms of their predominant narrative elements. The analysis focused on the convergence and divergence between the two texts, observing structure, language, sensitivity, and ability to capture symbolic elements of the image. This comparative reading aimed to assess the degree of adherence of each AI to the different models of audio description recognized in the specialized literature.

To classify the descriptions, we used the theoretical typology of audio description proposed by Motta (2010) and Neves (2011), which recognizes four main models:

Table 1
Audio description models listed for analysis

AD Model	Main Characteristics	Reference
Objective	Focus on formal visual data, without interpretation; technical or neutral vocabulary	Motta (2010), Neves (2011)
Functional	Emphasis on the informational function of the image; clarity for decision-making or location	Motta (2010)

e-ISSN nº 2447-4266
Palmas, v. 11, n. 1, 2025
http://dx.doi.org/10.20873/uft.2447-4266.2025v11n1a23en

Interpretive	Insertion of inferences and contextualizations that aid in the narrative understanding of the image	Neves (2011)
Poetic	Use of metaphors, rhythm, aesthetic sensitivity, and lyrical resources to activate the imagination	Motta (2010), Franco (2008)

Source: Prepared by the author.

This stage was fundamental for questioning the ability of AIs to operate with different layers of reading and revealing how much the algorithmic choice can influence the construction of meanings offered to PcDV.

Discourse analysis

The discursive analysis stage of the research is based on the French tradition of Discourse Analysis, according to the theoretical contributions of Michel Pêcheux (1990) and Eni Orlandi (2005). This approach understands discourse as the production of meaning traversed by ideological, historical, and social formations, rejecting the notion of linguistic neutrality. The focus is not only on the explicit content of the statements, but also on silences, effects of meaning, and conditions of production.

In this sense, the audio descriptions generated by AIs (ChatGPT and Gemini) were examined for their dominant discursive formations and implicit markers of ideological positioning. To this end, analytical categories such as the following were considered:

- Lexicalization and vocabulary choice: which terms are used to represent visual elements? Is the vocabulary technical, sensitive, or connotative?
- Omissions and silences: what is not said? What aspects of the image are ignored or disregarded in the automated description?
- Modalities and enunciative marks: are there expressions that reveal hesitation, certainty, or interpretation? Does AI take any position on the content described?
- Enactment of hegemonic or alternative meanings: do the discourses reproduce stereotypical views or activate critical, poetic, or pluralistic perspectives?

e-ISSN nº 2447-4266
Palmas, v. 11, n. 1, 2025
http://dx.doi.org/10.20873/uft.2447-4266.2025v11n1a23en

Critical validation

The critical validation stage was conducted with the aim of assessing the relevance, clarity, and sensitivity of audio descriptions by artificial intelligence, in light of the experience of people with visual impairments (PcDVs). For this purpose, two blind consultants with active experience in accessible cultural mediation processes and participation in inclusive projects were selected. This methodological choice is in line with the principles of participant research, as advocated by Neves (2011), and the fundamentals of epistemic justice, according to Fricker (2007).

The consultants were invited to listen to the audio descriptions generated by ChatGPT and Gemini. The evaluations were conducted through semi-structured interviews, with questions addressing the comprehension, expressiveness, interpretive fidelity, and aesthetic power of the descriptions. The responses were recorded, transcribed, and analyzed according to emerging categories.

Results and Discussions

Based on what was proposed methodologically, below are the results of the artificial intelligence deliveries used based on a simple prompt: "Create audio description for PcDV." Google's Gemini produced the following result.

> The vibrant, abstract painting features a stylized eye as its focal point, positioned slightly to the left of center. The iris of the eye combines shades of purple and light blue, while the pupil is black with a prominent white glow in the upper left corner, giving it a sense of depth and reflection. The eyelids are outlined in black, with expressive eyelashes that extend outward. The background of the work is an explosion of colors and textures. Shades of orange, red, green, and blue blend and overlap, creating a dynamic and energetic effect. There are splashes and loose brushstrokes of white and yellow paint that add a layer of spontaneity to the composition. Several graphic elements and words are integrated into the painting. In the upper left corner, the word "LOVE" is written in white letters on a dark blue stain. Below and to the left of the eye, the word "Respect" appears vertically in white letters, accompanied by a pictogram of a black umbrella inside a square with highlighted corners. At the bottom center, the word "Inclusion" is written in yellow letters. In the lower right corner, the word "Love" is repeated in white

e-ISSN nº 2447-4266
Palmas, v. 11, n. 1, 2025
http://dx.doi.org/10.20873/uft.2447-4266.2025v11n1a23en

> cursive letters, and a pictogram of a broken wine glass is inside a square with highlighted corners, similar to that of the umbrella. In the upper right corner, the word "Otherness" can be partially seen in white letters. The signature "Adriana S" is visible in the lower right corner in white letters. The overall composition suggests a reflection on themes such as love, respect, inclusion, and perhaps fragility, expressed through the symbolic representation of the gaze and the textual and pictorial elements. The variety of colors and the expressive application of paint contribute to a sense of emotional intensity (Google Gemini, 2025).

Similarly, using the same Prompt command, ChatGPT provided the following description:

> An explosion of colors and senses manifests itself in this mixed media pictorial composition on canvas. At the center of the work, in absolute prominence, is a large stylized eye — the left eye — in shades of gray, blue, lilac, and black, with a violet iris that vibrates with ethereal nuances, evoking mystery and sensitivity. It is partially outlined by long black eyelashes in triangular shapes that fray in various directions, as if they were feathers or shards in motion. In the background, there is a profusion of spots, splashes, and brushstrokes in dense and chaotic layers. The predominant colors are orange, lime green, royal blue, pink, lilac, and purple, with traces of white that suggest light, explosion, and energy. There is a striking presence of artistic gesture, as if each brushstroke were a heartbeat, expressing raw emotion. Scattered across the canvas, handwritten words in different sizes and colors: "Love," "Respect," "Inclusion," and "Attention" emerge as the guiding threads of the work's message. They are inserted into the visual chaos like poetic anchors. On the lower left side, there is an icon of a black umbrella surrounded by brackets, suggesting protection. On the right side, the symbol of a cracked cup, also framed by brackets, points to fragility. Both icons evoke accessibility and care. In the upper left corner, a silhouette of a white bird with black outlines flies off the screen, like a breath of freedom or transcendence. The entire image seems to pulsate — a kind of visual cry for empathy and sensitive presence. It is a work that touches the

e-ISSN nº 2447-4266
Palmas, v. 11, n. 1, 2025
http://dx.doi.org/10.20873/uft.2447-4266.2025v11n1a23en

invisible, making art a language that opens up to others (OpenAI ChatGPT, 2025).

The analysis revealed that both AIs operate under a standardized descriptive logic, with a strong presence of neutral visual nouns and adjectives. However, it was identified that ChatGPT, by incorporating metaphors and imagery, strains the model of algorithmic neutrality, assuming more open and subjective meanings. Gemini, on the other hand, tends to stabilize meanings, favoring a descriptive and informative function of the image.

Both AIs produced descriptions with a coherent discursive structure, presenting formal and descriptive aspects. However, it was observed that Gemini tended to adopt a more objective and functional model, describing the image in informative blocks, with attention to the organization of elements and the use of graphic terminology. ChatGPT, on the other hand, although it also began with objective elements, moved towards a poetic approach, using lyrical expressions such as "visual cry for empathy" or "symbolic pulsation," which brings it closer to a poetic or interpretive DA.

The convergence between the two ADs occurred in the identification of central visual elements, such as the stylized eye, the keywords inscribed in the work (Love, Respect, Inclusion), and the symbolic icons (umbrella and broken cup). The main divergence occurred in the dimension of sensitivity: while Gemini restricted itself to a more descriptive and literal reading, ChatGPT performed a symbolic expansion that refers to the style of poetic audio description described by Franco (2008).

The absence of contextual, historical, or cultural references was also observed, indicating that AI ignores symbolic layers of the image that could be relevant to PcDV. Such omission can be interpreted as a form of discursive silencing, as argued by Orlandi (2005), since “the unsaid also structures the meanings of the said.”

Finally, it was found that the discourses produced by AIs carry traces of technocentric objectivity, suggesting a dominant discursive formation of a positivist matrix. This model tends to reduce the polysemy of the image and subordinate the aesthetic experience of PcDV to criteria of legibility and standardization. Critical discourse analysis, therefore, shows that automated mediation is not free from ideological positions — on the contrary, it reproduces them implicitly and structurally.

The consultants' main findings revealed that:

- Automated audio descriptions lack empathy and narrative density, especially in passages where the symbolic context of the original image would require more subjective mediation.
- Both AIs presented terminological inconsistencies and unclear metaphors, with emphasis on Gemini's description, considered "cold" and "instrumental."
- The inferences made by the AIs are excessive, confusing, and do not follow a logical order, thus hindering the understanding of the imagery narrative.

The validation demonstrated that AI mediation, although technically elaborate, does not achieve the levels of sensitivity and plurality required by a critical inclusive practice. As pointed out by Motta (2010) and Franco (2008), audio description needs to be responsive, constructed in dialogue with the subjects for whom it is intended. The absence of this dialogue in automated descriptions highlights the limitation of AI as a substitute for human mediation in sensitive educational contexts.

This stage reinforces the defense of hybrid mediation models, in which technology acts as a complementary tool, rather than a substitute, for the critical educator. This validates the thesis that inclusion requires listening, co-authorship, and sensitivity — dimensions that, to date, artificial intelligence is not capable of replicating with aesthetic or ethical responsibility.

The results showed significant differences in the platforms' approaches. The image generated from ChatGPT's description featured more fluid and evocative visual elements, aligning with a poetic discourse that explored sensations and visual metaphors. However, the predominant aesthetic tended toward the dreamlike and generic, emptying specific features of the original image. The image resulting from Gemini's prompt was more literal and simplified, favoring informational clarity and almost completely eliminating symbolic or subjective nuances.

The analysis revealed significant divergence between the AIs in terms of language, symbolic depth, and sensitivity. Gemini presented an objective, technical, and descriptive description, with no subjective elements. ChatGPT, on the other hand, produced a more poetic and sensitive version, albeit with non-visual insertions that could confuse the recipient.

e-ISSN nº 2447-4266
Palmas, v. 11, n. 1, 2025
http://dx.doi.org/10.20873/uft.2447-4266.2025v11n1a23en

Critical validation revealed frustration among PcDV consultants with the audio descriptions. They pointed out that, although well written, both lacked context, intention, and affection. They reiterated that inclusion is not only about access to information, but also to the aesthetic exper

This research aimed to critically investigate the risks of intersemiotic misinformation arising from the use of generative artificial intelligence (GAI) in the audio description of artistic images in school contexts, focusing on the experience of students with visual impairments (PcDV). The guiding hypothesis was that, although efficient in syntactic terms, GAI lacks the aesthetic, symbolic, and ethical competence to sensitively and responsibly mediate the visual and artistic experience.

The results confirmed the hypothesis and revealed concrete evidence of misinformation generated by the GAI analyzed (ChatGPT and Gemini). The descriptions generated, although structurally coherent, proved incapable of accessing fundamental symbolic layers of the original image, presenting themselves either as descriptive and cold (Gemini model) or poetic but generic (ChatGPT model). None of the IAGs were able to mobilize contextual, historical, or subjective elements of the work. There were, therefore, significant erasures — so-called "silencing" — that misinform by concealing relevant meanings, as discussed by Orlandi (2005). Such phenomena do not occur by chance: they are expressions of algorithmic logic which, as pointed out by Santaella (2002) and Zuboff (2019), operate through simplification, statistics, and the extraction of dominant and homogenizing patterns.

The validation stage with blind consultants was decisive in consolidating the central argument of this study: AI, when used in an unmediated way, not only fails in the task of inclusion, but also institutes a new layer of exclusion — aesthetic algorithmic exclusion. Automated mediation proved insufficient to enable a sensitive and plural aesthetic enjoyment of PcDV, limiting itself to providing formal data, sometimes disconnected from the experience of the image.

From an educational point of view, it should be considered that the precariousness of teacher training in inclusive technologies and the naive belief in the neutrality of AI create a fertile environment for exclusionary practices, even under the discourse of innovation and accessibility. The National Common Core Curriculum (Brazil, 2018) provides for the development of digital literacy and empathy as educational competencies, but lacks practical guidelines for the ethical and aesthetic use of technologies with PwDs.

Given this, it can be concluded that AI-based school inclusion requires ethical pedagogical protocols, critical teacher training, active participation of people with disabilities, and situated aesthetic mediation. The use of artificial intelligence alone does not guarantee accessibility — and may, on the contrary, produce misinformation masked as efficiency. This does not mean that they cannot soon be improved and become useful tools.

As a result of this research, we recommend: The development of specific teacher training in audio description and AI ethics; The construction of frameworks or guidelines in AD, to train IAGs based on non-hegemonic perspectives; The formulation of public policies that recognize inclusion as a process; and The collaborative production of ADs involving artists, PcDV consultants, educators, and accessible technology specialists.

Finally, this research reaffirms the centrality of human experience, pedagogical sensitivity, and active listening as non-negotiable foundations for building an education system that is moving toward a minimum level of aesthetic and democratic inclusion.

References

Araújo, J. da S., et al. (2025). O letramento digital na BNCC e o combate à desinformação. *Revista de Gestão e Secretariado*, 16(1), 1–19.

Aumont, J. (1995). *A imagem* (2ª ed.). Papirus.

Bender, E., Gebru, T., McMillan-Major, A., & Mitchell, S. (2021). On the dangers of stochastic parrots: Can language models be too big? In *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency (FAccT '21)* (pp. 610–623). ACM Press. https://doi.org/10.1145/3442188.3445922

Brazil. (1988). *Constituição da República Federativa do Brasil de 1988*. Senado Federal.

Brazil. (2015). *Lei nº 13.146, de 6 de julho de 2015. Estatuto da Pessoa com Deficiência*. Presidência da República. http://www.planalto.gov.br/ccivil_03/_ato2015-2018/2015/lei/l13146.htm

Brazil. (2018). *Base Nacional Comum Curricular*. Ministério da Educação. http://basenacionalcomum.mec.gov.br/

Deleuze, G. (1990). *A imagem-tempo* (2ª ed.). Brasiliense.

Deleuze, G., & Guattari, F. (1995). *O que é a filosofia?* Editora 34.

Didi-Huberman, G. (2003). *O que vemos, o que nos olha*. Editora 34.

Dondis, D. A. (1997). *A sintaxe da linguagem visual*. Martins Fontes.

Dreyfus, H. L. (1992). *What computers still can't do: A critique of artificial reason*. MIT Press.

Franco, E. (2008). *Tradução intersemiótica e audiodescrição: O texto visível*. UFBA.

Fricker, M. (2007). *Epistemic injustice: Power and the ethics of knowing* (2ª ed.). Oxford University Press.

Gabriel, M. (2017). *Você, eu e os robôs: Como se transformar no profissional digital do futuro*. Atlas.

Gadamer, H.-G. (1999). *Verdade e método: Traços fundamentais de uma hermenêutica filosófica*. Vozes.
Gell, A. (2018). *Arte e agência: Uma teoria antropológica*. Ubu.
Gombrich, E. H. (2000). *A história da arte*. LTC.
Google. (2025). *Gemini (Mai 2025 version)* [Large language model]. Google DeepMind.
Hernández, F. (2007). *Educação e cultura visual*. Artmed.
Irwin, R. L., Beer, R., Springgay, S., Grauer, K., Xiong, G., & Bickel, B. (2004). A/r/tography: Rendering self through arts-based living inquiry. *Journal of Curriculum Theorizing*, 20(1), 27–38.
Jakobson, R. (1959). On linguistic aspects of translation. In R. A. Brower (Ed.), *On translation* (pp. 232–239). Harvard University Press.
Kaufman, D. (2021). *Desmistificando a inteligência artificial*. Edições Sesc.
Kress, G., & Van Leeuwen, T. (2006). *Reading images: The grammar of visual design* (2ª ed.). Routledge.
Lemos, A. (2023). *Colonialismo digital: Comunicação e poder nas plataformas digitais*. Ubu.
Motta, L. (2010). *Acessibilidade cultural: Tradutores e intérpretes de Língua de Sinais e Audiodescrição*. Secretaria de Cultura.
Neves, J. (2011). *Audiodescrição e acessibilidade: Experiências e perspectivas*. Universidade Aberta.
OpenAI. (2025). *ChatGPT (Mai 2025 version)* [Large language model]. OpenAI.
Orlandi, E. P. (2005). *Discurso e leitura*. Cortez.
Pêcheux, M. (1990). *Análise automática do discurso*. Editora da Unicamp.
Ricoeur, P. (1997). A hermenêutica e a crítica da ideologia. In P. Ricoeur, *Do texto à ação: Ensaios de hermenêutica II* (pp. 123–148). Instituto Piaget.
Rolnik, S. (2006). *Micropolítica: Cartografias do desejo*. Vozes.
Santaella, L. (2002). *A imagem: Cognição, semiótica, mídia*. Iluminuras.
Santaella, L. (2019). *Cultura e artes do pós-humano: Da cultura das mídias à cibercultura*. Paulus.
Santos, M. P. dos. (2016). Inclusão crítica: Política, epistemologia e estética da diferença. In M. P. dos Santos et al. (Orgs.), *Educação e diferença: Perspectivas para a inclusão crítica* (pp. 45–70). Lamparina.
Silva, A. A. da. (2019). *Graffiti, comunicação e antropologia da arte: Os indígenas no spray de Cranio e Raiz*. UFT.
Sousa, F. G. de, Oliveira, R. V. de, & Castro, L. R. de. (2021). Tecnologias e inclusão crítica: Desafios da docência na contemporaneidade. *Revista do PGCTIn*, 4(2), 6–24.
UNESCO. (2023). *Technology in education: A tool on whose terms?* UNESCO.
Zuboff, S. (2019). *A era do capitalismo de vigilância: A luta por um futuro humano na nova fronteira do poder*. Intrínseca.

e-ISSN nº 2447-4266
Palmas, v. 11, n. 1, 2025
http://dx.doi.org/10.20873/uft.2447-4266.2025v11n1a23en

RESUMO
Este estudo analisa criticamente os riscos de desinformação intersemiótica causados pelo uso de inteligências artificiais generativas (IAGs) na audiodescrição de imagens artísticas voltadas a estudantes com deficiência visual (PcDV) no ambiente escolar. A pesquisa utilizou uma imagem autoral inédita submetida às plataformas do ChatGPT e Gemini para gerar audiodescrições automatizadas, que foram analisadas no âmbito discursivo, com validação crítica por consultores PcDVs. Os resultados indicaram apagamentos simbólicos, simplificações estéticas e ausência de contextualização sensível, evidenciando riscos à fruição estética e à inclusão. Embora se trate de tecnologias promissoras, este estudo conclui que o uso das ferramentas de IA requer mediação ética e pedagógica. As IA Generativas não substituem a escuta, a subjetividade e a experiência humana, exigindo formação docente e políticas inclusivas fundamentadas.

PALAVRAS-CHAVE: Audiodescrição; Inteligência Artificial; Desinformação Intersemiótica; Inclusão Escolar; PcDV; Arte e Tecnologia.

RESUMEN
Este estudio analiza críticamente los riesgos de desinformación intersemiótica provocados por el uso de inteligencias artificiales generativas (IAG) en la audiodescripción de imágenes artísticas dirigidas a estudiantes con discapacidad visual (PcDV) en el entorno escolar. La investigación utilizó una imagen autoral inédita, sometida a las plataformas ChatGPT y Gemini para generar audiodescripciones automatizadas, las cuales fueron analizadas en el ámbito discursivo, con validación crítica por parte de consultores PcDV. Los resultados indicaron omisiones simbólicas, simplificaciones estéticas y falta de contextualización sensible, evidenciando riesgos para el disfrute estético y la inclusión. Aunque se trata de tecnologías prometedoras, este estudio concluye que el uso de estas herramientas requiere mediación ética y pedagógica. Las IA generativas no sustituyen la escucha, la subjetividad ni la experiencia humana, exigiendo formación docente y políticas inclusivas fundamentadas.

PALABRAS CLAVE: Audiodescripción; Inteligencia Artificial; Desinformación Intersemiótica; Inclusión Escolar; PcDV; Arte y Tecnología.