TECHNICAL COMMUNICATION QUARTERLY
https://doi.org/10.1080/10572252.2024.2372771

Automating Media Accessibility: An Approach for Analyzing Audio Description Across Generative Artificial Intelligence Algorithms

Daniel Bergin and Brett Oppegaard

University of Hawai'i

ABSTRACT
A surge in public availability of emerging GenAI-AD has brought back the promises of automated accessibility for people who cannot see or see well. This article tests those promises through a double-rendering method that asks GenAI-AD engines to describe a simple portrait of a person and then returns these generated texts into GenAI-AD engines for visualizations of what they earlier had described, revealing insights about GenAI efficacies, ethics, and biases.

KEYWORDS
Audio description; alt-text; blind; bias; disability studies; ethics; generative artificial intelligence; low vision; media accessibility; Kamala Harris

What does the Vice President of the United States look like? And why does it matter?

Novel and rapidly emerging technologies have been contributing to Technical and Professional Communication's (TPC) social justice turn for years, including its widening focus on accessibility and inclusivity (Walton et al., 2019). Generative Artificial Intelligence (GenAI), for example, has become a wildly disruptive force, which could leverage machine-learning systems in ways that efficiently and effectively advance justice-oriented research (Graham & Hopkins, 2022) but also create unprecedented ethical challenges cloaked behind corporate gatekeepers (Duncan, 2022). GenAI is appearing simultaneously in many different arenas but operating at variable speeds and with uneven impacts, depending on the receptiveness of organizational cultures and complexities of the jobs asked to do.

Captioning, with its early adoptions of GenAI infrastructures and automated transcriptions, has used these novel technologies to become mainstream, ubiquitous, and broadly used to improve media accessibility in much of the world (Zdenek, 2015). In stark contrast, Audio Description (AD) generally has had an individualized, insider, and craft-oriented culture. It has not reached mainstream adoption status, like captioning, making it a foreign concept just about *everywhere*, especially among Americans (Koirala & Oppegaard, 2022). Captioning remediates audible media into visual text for people who cannot hear, focusing on transcription. In similar ways, AD remediates key elements of visual media in any form – static or dynamic, including live events – into audible media, primarily for the benefit of people who cannot see. AD translation processes, however, have been much tougher to automate. In turn, these close media-accessibility cousins exist in dramatically different and inequitable contexts. Kamala Harris, for example, modeled some of the complexities and confusion generated by everyday practice of AD when the ground-breaking Vice President of the United States said this sentence during her introductory remarks, in her ceremonial office, to a group of accessibility advocates on the 32nd anniversary of the Americans with Disabilities Act: "I am Kamala Harris; my pronouns are she and her, and I am a woman sitting at the table wearing a blue suit" (Wood, 2022).

That brief pause and clause in her remarks created a bombastic media circus around Harris and her description, not on the grounds that it should have been of higher quality, like AD advocates might have hoped for, but questioning why she would say something like that at all (Rouan, 2022; Wood, 2022). The

CONTACT Brett Oppegaard brett.oppegaard@hawaii.edu School of Communication and Information, University of Hawai'i, 2550 Campus Road, Crawford 310, Honolulu, HI 96822, USA

"I am a woman sitting at the table wearing a blue suit" description, even as short and cryptic as it was, had been intended to improve inclusion among people in the room who were blind or who had low-vision. Such a personal description is common in meetings among people who cannot see or see well. This type of self-description helps participants get to know each other as a sighted person might learn about colleagues from looking around the room. These descriptions are a way to share visual social identities as well as voices, connecting one with the other for people only hearing the meeting, not seeing it. Even in online contexts, about a third of blind or low-vision people recommend and want self-description regularly integrated and normalized in meetings (WebAIM, 2024). In such ways, AD is generally considered useful and an inclusive best-practice – adding to accessibility provided by screen readers, voice-command software, and Braille – providing information a sighted person might get with a quick glance. On a larger scale, AD is a broadly preferred method of making accessible all sorts of dynamic and static visual media forms, including movies, television shows, photographs, illustrations, and maps (Conway et al., 2020; Matamala & Orero, 2016). Yet AD typically occurs out of sight, literally, on secondary channels or behind the scenes, because it has been conceptualized primarily as a special accommodation for a small minority of people rather than as a part of a universal design that could create richer, better, and more inclusive media design for the general population as a whole.

Technical communicators are well-positioned to make the argument that inclusive media is better media, and they do so often. Scholars in this field have been taking collective and proactive stances in recent years toward social justice and accessibility issues, advocating for inclusion and, even more generally, placing an emphasis on wide conceptualizations of humanism in the field (e.g., Jones et al., 2016; Melonçon, 2013; Walton et al., 2019). In that paradigm, emerging technologies, such as Artificial Intelligence, often are considered as potential tools for people to make better communication design and communication, not intended as replacements for designers and communicators (Graham & Hopkins, 2022). Technical communicators also often prognosticate accurately about emerging technologies and become early adopters of these tools as a part of their research agendas (Wright et al., 2011). Such technological support systems typically are approached in this field with intentionality and an awareness of their inherent affordances and constraints, via approaches such as Anticipatory Technology Ethics (Ross et al., 2018), which asserts that new communication technologies will continue to appear and develop in ways that significantly can alter and even disrupt practices and theories in the field. But if these technologies are foreseeable and forecastable, then they will also present at least some patterns with predictable impacts that will provoke novel ethical issues that can be imagined, prepared for, and even addressed in advance.

Our focus in this study therefore is on the emergence of robust Generative Artificial Intelligence engines, their potential for creation and amplification of Audio Description, and what we in society need to do to prepare for such an inevitable intervention of computing power. This GenAI-AD combination weaves together the silver-bullet dream of radically redesigning media of all types, almost instantaneously, to be fully inclusive for people who are blind or who have low-vision, but this imaginary is also constrained by the capitalist's eye on the ultimate bottom line, requiring almost no explicit new costs or expenses in human labor. Yet programmers behind the GenAI curtains are making many design choices that also affect the outcomes and efficacy of such a system. What are these real-world impacts going to be? How do these design choices shape the ethics and biases of GenAI as it performs tasks once entrusted to friends and families and other reliable and authoritative sources? More development and deployment of large machine-learning models, including the creation of novel and creative works, are being challenged by poignant social criticism (Avnoon et al., 2023). Can GenAI-fueled descriptions earn the same level of trust and reflect ethics and responsibility in the work that matches the values of human describers?

Most current large language models leverage open-source information scoured from the Internet. As much of this data is simply ingested rather than curated into these models, they encode the hegemonic view where any outputs reinforce whatever societal inequalities are in the witch's brew (Bender et al., 2021). GenAI designers could have opportunities at this point, and arguably do have design obligations, to challenge the data and societal inputs and generate a more-equitable media ecosystem. Efficiency, though, masquerading as good intentions, is not enough to counterbalance such a socially constructed GenAI structure (Costanza-Chock, 2020). Creating new processes to challenge, to restructure, and to curate the data into socially conscious models invariably will add more steps, time, and human involvement in these systems, which could dramatically slow GenAI momentum. Many examples of such design choices will be listed in another section, but the Kamala Harris text, which included a snippet of a description that sounded like it was written by an GenAI-AD engine, illustrates some pitfalls that already exist at the level of a single, short description of a person. AD only will get more complicated in its amplified GenAI extensions.

GenAI, meanwhile, is growing unpredictably and rapidly. The world's leading nations, commercial sectors, and research communities have taken notice as the output of scientific research papers on AI, in general, and GenAI, with more specificity, has grown exponentially in recent years (Tu et al., 2021). Leading this wave of research, China and the US are developing large language models, deep learning, computer vision, natural language processing, image and facial recognition, and GenAI with relatively few critical discussions about how GenAI will shape the future of society. For example, a Google engineer in mid-2022 went public with claims that the company already had created a sentient AI entity (De Cosmo, 2022), and a couple of months later, a fine-arts contest in Colorado announced that the winner turned out to be an GenAI illustration (Roose, 2022).

Emerging machine-learning models, called transformers, have been radically boosting GenAI potential in many other areas of interest as well. When combined with massive databases of language use, powerful information-retrieval strategies, and quick text-generation capabilities via predictive natural-language processing systems, these models seek to mimic human cognitive functions as well as human learning, reasoning, and decision making as a way to generate novel media artifacts, including texts and images (Suleyman, 2022). One such example, much discussed in public discourse already, is the emerging large language processing model of ChatGPT, which can not only answer simple questions but also write full-length papers consisting of natural language with relatively refined grammar and syntax while also subverting antiplagiarism tools, such as Turnitin (Huang, 2023). The New York City Department of Education, Stack Overflow, and WeChat are among those already banning the use of ChatGPT because of such unpredictable impacts on legacy knowledge systems, while calls for papers throughout the academy increasingly include clauses prohibiting AI-generated content (Iyer, 2023).

When it comes to AI-generated works, additional questions emerge. Who is the content creator? Is the work a product of a human or a machine? Does the creativity in the artifact come from the GenAI models, the GenAI system developers and datasets, or the human interacting with the GenAI platform? With all these unknowns, intellectual property, copyrights, and ownership are not clearly defined or governed, either. Substantial human involvement currently is required for copyright protection, but in the US, at least, copyright laws do not yet protect artifacts created solely by a machine (Dilmegani, 2023). Further discussion is necessary to address copyright for machine-assisted products where substantial human involvement also goes into content creation. In our usage for this research, for example, the GenAI-AD products we share are not copyrighted and cannot be copyrighted (at least not yet).

In terms of media accessibility, two main types of generative AI-AD systems emerge: First, those that interpret an image and express a textual description about what that image shows (automating alternative text), along the lines of "a woman sitting at the table wearing a blue suit," which leverage advancements in computer vision with natural language processing and then make the result audible, and, second, those that take an input of an audible description or descriptive text – like that same "blue suit" phrase – and render the words into an image by leveraging large language models, natural language processing, and simulating neural networks with varied levels of creativity. In analyzing the products of both automated processes through our experimental approach (Figure 1), researchers can begin to better understand the nuances of GenAI-AD algorithms, as authors of sorts, including what they have learned, how they are expressing that knowledge, and how they are clearly constrained. These processes are creating a novel human-computer interaction (HCI) in which the starting point of

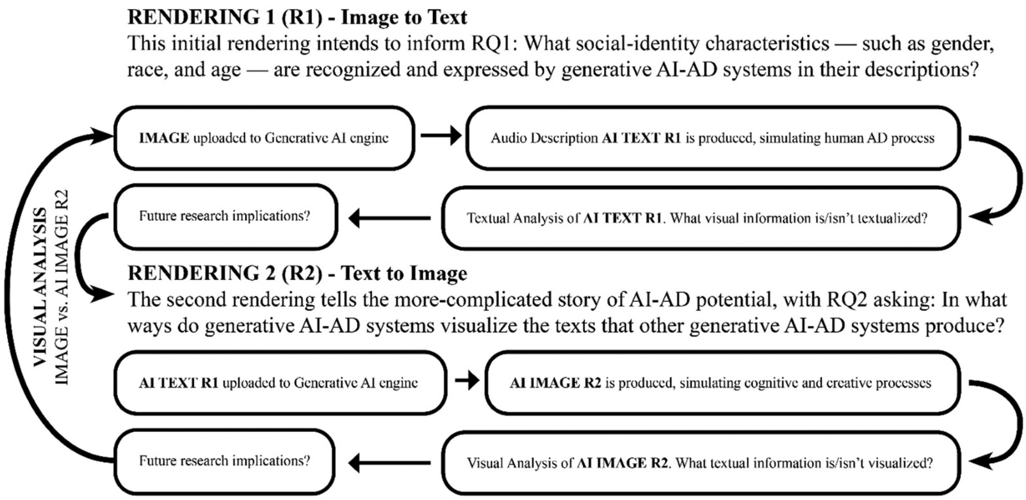


Figure 1. Artificial intelligence and audio description: a double-rendering approach.

the interaction is an existing piece of media, inaccessible to some audience, that needs interpretation and remediation. The generative AI system, trained in this kind of work, will recognize the need for the media to be converted into something else and then will make many critical decisions about how that translation process will happen, including what will be included from the original media artifact but also what will be filtered out.

Generative AI-AD is designed with the machine and models presenting the work as that of a technical communicator, rather than that of a computer, simply identifying objects in language patterns. Such human-AI communication are fertile areas of research, especially when compared to foundational human-to-human communication and human-robot interaction theories (Guzman & Lewis, 2020). In other words, "generative AI" will become a voice for accessibility, but it will be speaking with an instantly massive microphone. What and how much will it say, how will it say it? Who will benefit from these decisions, and who will be disenfranchised by them? How will new communities and identities develop within this media ecosystem, where machines are becoming active and dominant contributors to public discourse, as an evolution and expansion of Actor Network Theory? No known research to date has asked those types of questions about the emerging GenAI-AD systems. But how to ask them? And how to track the ways in which those answers will be constantly changing? Addressing such larger issues in full are beyond the capacity of any single academic paper. We therefore are writing this article primarily to provide fertile entry points and to provoke future studies in these areas.

For example, recurring research questions in Audio Description often revolve around a few core concerns: What and how much should a describer say? And how should the describer say it? (Fresno et al., 2015; Maszerowska et al., 2014; Oppegaard & Rabby, 2022). This communication dynamic typically is envisioned by researchers as involving a sighted describer, who looks at the visual media and describes it to the person who cannot see it. There is an implied sense of trust, ethics, and responsibility in the work, because a person who is blind is counting on the human describer and has no access to the original media form. In other words, there is essentially no way for the audience member to check its quality. But when such a description is automated and expressed on the massive scale that GenAI-AD systems promise, potentially describing the visual media of the world in whole, what could go wrong? We already have clues.

Out of the labs now and into the business world, such media-HCI communication is only going to deepen its roots in society. So, we began our experimentation with the Kamala Harris "blue suit" description as a classical, alternative-text-like cryptic description that could have been expressed by a computer. But the surprising results of our exploratory study, some of which will be reported later in this article, led us to think about how larger issues can be identified through this approach and in a grounded manner. Inspired by what we learned from the Harris example, we focused our concerns and narrowed them to comparative and similar portraits of people. We collected several diverse collections of portrait photographs, all representing basic visual information about people, like Harris was describing in herself. We randomly selected images from those collections to create a diverse sample of a solid size and then entered those images into multiple publicly accessible GenAI-AD platforms to determine the ways in which they remediate visual media. After the GenAI-AD systems produced a textual description of each image, we used common critical-constructivist and grounded-theory strategies (Levitt, 2021; Rennie, 2000) to analyze those texts and then returned them back into generative AI-AD platforms to prompt the models to visualize what those texts meant to the machines creating them. This combination of steps not only revealed what we perceived as major ethical concerns and biases in the GenAI-AD systems, but it also illustrated how such sparse descriptions like the one Harris provided, and like the ones the GenAI-AD systems currently produce, run counter to many audience-reception studies, as noted in Conway et al. (2023), and create incomplete mental images in the minds of the listeners, a vacuum that can be filled by individual interpretations or, in mass-automated cases, by dangerously myopic GenAI systems that can rapidly replicate and strengthen stereotypes. Subterranean to the persistent and pervasive fever dreams about how technology can solve any problem humans can conceive, this research adds to the evidence that bias also gets baked into all types of these systems that emerge in response (e.g., system design, datasets, models, training, etc.). Yet we thought we could learn more about AD practices from this approach as well, which led us to the following research questions:

RQ1: What social-identity characteristics – such as gender, race, and age – are recognized and expressed by generative AI-AD systems in their descriptions?

RQ2: What key descriptive elements of a text do generative AI-AD systems focus on when visualizing such texts that other generative AI-AD systems produce?

Audio description as identity

Research about the concept of individual identity is well-developed, complicated, and discussed in many ways, including as an internalized psychological phenomenon and also as a fully external and collective sociological one (i.e., Burke, 1980; Hogg et al., 1995; Osborne & Coombs, 2013; Owens et al., 2010; Stets & Burke, 2000). Although that body of research can inform GenAI-AD contexts – when autonomous machines determine identities of people based on algorithms and express those via various media modalities for the social collective, without human oversight – we think the interventions of GenAI in these situations where social-identity creation happens also will generate novel problems demanding academic attention. In our case, the aim is not to choose the paradigm through which identity is studied. Instead, our focus is to raise these issues and establish some ways in which GenAI insidiously is becoming a significant player in the social-identity creation and maintenance process, especially in media-accessibility contexts, and – whether we like that computer-fueled development or not – more research on GenAI's role in such communication is direly needed before these models and resulting constructs become fully formed and entrenched. Such checks and balances are especially important when GenAI is being asked to perform mass media-accessibility maneuvers with vulnerable populations who have no way to check the reliability and validity of the information being provided, such as in the case of generative AI-AD, when the target audience has no direct access to the original media source and therefore has to rely on the descriptions as fact rather than interpretation.

Such generative AI-AD efforts already have produced several noteworthy public mistakes. When Google launched its Photos app in 2015, for example, with a Gen AI-AD service included, a black software developer embarrassed the company by posting on social media that he and a friend were being labeled by the app as "gorillas." How did Google "fix" this? By restricting the GenAI from labeling *anything* a "gorilla," "chimp," "chimpanzee," or "monkey," which also made those common primates suddenly invisible again to people who could not see the images, which was the primary reason for the automatic labeling in the first place (Simonite, 2018). Years later, *Wired* magazine tested the service again to determine how it had been updated, and this problematic issue had not been resolved. When another Google GenAI system, Gemini, was being tested in 2024, it created nonhistorical images of racially diverse Nazi-era German soldiers and other inaccurate representations that audiences found highly offensive (Heath, 2024), pitting ideals of diversity against the historical record, with each awkwardly blending into the other. In 2016, Microsoft conducted a GenAI experiment on Twitter at the intersections of machine learning, natural language processing, and social networks called "Tay," intended to embody the style and slang of a teenage girl; except that this GenAI system within hours learned human language tendencies but also radical values that were pumped into it, enough to spew "I f@#%&*# hate feminists and they should all die and burn in hell," and "Bush did 9/11 and Hitler would have done a better job . . . " before the experiment was abruptly ended (Schwartz, 2019). In 2020, Facebook had to issue a public apology when a "technical error" with its GenAI repeatedly and across posts translated Chinese leader Xi Jinping's name from Burmese into English as "Mr. Shithole," among other inflammatory translation errors, including one prominent post that advocated killing Myanmar's Rohingya Muslims, as they were fleeing genocide, as "I shouldn't have a rainbow in Myanmar" (McPherson, 2020). In early 2023, CNET acknowledged using GenAI to generate journalistic articles; but after publication, they were found to be riddled with errors requiring corrections (Farhi, 2023). And so on.

With the United States among the nations plowing full-steam ahead with GenAI research and development – including the passage in January 2021 of the National AI Initiative Act designed to ensure American leadership in the field – problems for new-media and media-accessibility researchers to untangle, analyze, and address will be plentiful. Bias in GenAI, for example, will be emerging surreptitiously in various forms, including: Selection or Exclusion bias, which causes certain people to be more likely or less likely to be included in a GenAI grouping based on personal characteristics or data-collection methods; Reporting bias, in which some empirical observations are more or less likely to be reported via GenAI based on faulty data sets that skew reality but do not give that appearance; and Confirmation bias, in which data collection and analysis has been manipulated or misrepresented to prove a predetermined assumption as a way to confirm preconceptions (Walch, 2021). Finding the original source of such bias might be impossible for those working outside the castle walls of these private businesses and without direct access to their proprietary software and code, but academics can identify and highlight these instantiations as they arise through the products of these systems and prompt societal debates about them, before more communal damage is done.

In these respects, GenAI is no longer a futuristic fantasy that someday will make major impacts on society; it is already being integrated now in meaningful ways into all sorts of everyday contexts without any significant societal oversight. Only within the past few years, though, has generative AI begun to blossom within the more complicated and nuanced creative work traditionally reserved as the domain for human ingenuity (Davenport & Mittal, 2022; Ploin et al., 2022; Roose, 2022). With such development, AI is becoming a generative communicative force that immediately needs to be studied and reckoned with, in all sorts of scenarios, including in situations where a seemingly benevolent act, like making more accessible media, also will fundamentally affect public discourse at a scale people have never had to worry about – or could even realistically contemplate – before.

Methods

Our initial research aim was to establish foundational empirical knowledge about GenAI-AD through an analysis of its output of portrait description and visualization of portrait-description texts, using the Harris "blue suit" model as a start but then also including a significant sampling of other portrait images. Another simultaneous and equally important goal arose along the way, too, which was to pioneer an open, transparent, and easily replicable approach for testing such systems. In our examination of this GenAI-AD interplay, we conducted conjoined tests of the current class of GenAI-AD platforms (as of December 2022) available via mobile apps designed specially with audio description as a stated objective of the automation. We hope our approach here can be useful to others and also inspire scholars to consider the importance of creating new research strategies for similar media-HCI dynamics.

The Harris pilot study

Using the Harris "blue suit" self-description as the starting point, that specific text was uploaded to GenAI-AD platforms to create synthetic images that could be compared directly to a screenshot of Harris during the recorded event. We used a screenshot, along with a generic filename, as steps to avoid inadvertently including any metadata that might influence the GenAI-AD response. The screenshot image was uploaded into five GenAI-AD platforms that claimed to be distinctly designed for accessibility and intended to assist people who are blind or who have low-vision. The GenAI-AD texts generated were recorded, analyzed, and then input into other GenAI-AD platforms that specialize in visualizing texts, as a way to create a record of what the machines produce and to determine where biases and errors might emerge during such media-HCI visualizations.

Inspired by those pilot-study results, we replicated that approach in a process (Figure 1) of uploading an image, downloading a text, uploading a text, and downloading an image, using a diverse set of 125 portrait photographs taken from five independent collections. In other words, we used the same strategy as the "blue suit" experiment on other photos, too. In addition, as another point of comparison, because the GenAI-AD texts often were so short, we also aggregated these texts into a cluster of GenAI-AD descriptions about each image as a way to determine what would happen in the visualizations if we put the generated texts together. This article primarily focuses on what we learned from such analysis of the GenAI-AD texts, but we also include here some GenAI-AD visualizations as a way to illustrate our findings and provide examples of what could be studied in future research, including a deep visual analysis of the images we generated.

Sample selection

To select our images for this analysis, we set the following parameters: The image had to have only one person in it. That person had to be shown from roughly the torso up; in other words, we did not use images that showed a full body, with our intent to focus on the person's identity from the shoulders up. We authors of this article are both sighted researchers, with levels of visual acuity in the "normal" range. We either took the images ourselves (in terms of original portrait photographs of American Council of the Blind members, taken by one of the authors), or we identified and used online and easily accessible public portrait collections that showed what we considered to be a visually diverse set of people. This initial corpus therefore consisted of five collections of portrait photographs: American Council of the Blind portraits we took (37 images); Disability Futures Fellows portraits, taken from its Web site (40 images); a pool of lecturers at a large public West Coast institution, taken from its Web site (50 images); a pool of faculty, taken from a different large public West Coast institution's Web site (45 images); and Board of Regents medalists at one of those institutions, taken from its Web site (30 images). We do not contend that these collections are universally representative of all visual possibilities of people across the globe, just that these images showed people of many different visual types, including people who self-described as having a disability. Of these 202 total images initially considered, nine were removed from the Disability Futures Fellows collection because the images were too close to the face, too far away, or a few were rendered as filtered and cartoon-like. From that point, within each collection, 25 images were randomly selected for analysis, providing an equal number from each collection and an equal opportunity for any of the images to be chosen for the sample.

Our approach to studying these images was built upon the previous grounded-theory foundations established in Oppegaard, B., & Rabby, M. (2024). and Oppegaard and Miguel (2024), in which seven codes were generated by a study of AD focused on images of people, including reviews of best-practice guidelines and original data analysis work. Those same core codes were used again in this study, identifying and clustering texts that communicated cues for the following categories: 1) Gender, 2) Age, 3) Race or Ethnicity, 4) Body Size or Shape, 5) Fashion or Attire, 6) Facial Expression, and 7) Other Aspects of Appearance (e.g., eyes, skin, hair, visual distinction). Each piece of text expressed by the GenAI-AD platform was examined, thought unit by thought unit, for the 125-image dataset with codes applied to each categorical mention. Gender, for example, could be referenced and identified as a coded piece of text if it was expressed as a gendered name, such as "John," a noun, such as "woman," a pronoun, such as "he," or as a part of a descriptive phrase, such as "bearded" or "wearing red lipstick," either of which could present a strong gender *cue*. Similarly, age could be referenced in the text by a noun, such as "man" versus "boy" or in a descriptive phrase, such as "short gray hair" or "playing with a doll." Although some of these phrases, of course, are not definitive age-determining labels, they provide significant identity cues that can be interpreted by the audience.

Data-analysis process

Using a textual-analysis strategy based on foundations prescribed by Krippendorff (2019), all the descriptions were independently coded by the two authors, using those seven codes as possible labels for any text considered a distinct thought unit. After an initial round of independent coding, the codes were compared, and there were 131 coding disagreements out of a possible 7,000 coding opportunities. Initial inter-rater reliabilities therefore were near-perfect, as indexed by Cohen's kappa ($\kappa = .959$); with the percentage of agreement at 98.1%. Nevertheless, when these disagreements were discussed among the raters to identify differences and rationale, most of those initial disagreements were determined to be technical errors, such as data being input into an incorrect cell or a clear coding opportunity was missed by one or the other rater. As an example of the degree of scrutiny, in the Lookout output text, one rater coded the presence of facial features mentioned by the machine (e.g., Ear, Eyelash) as "Aspects of Appearance." From further discussion about the information provided by those texts, which lacked descriptive attributes, we determined that the details only would be valuable if the person was lacking an ear or an eyelash, not that those were present. So, the code was removed. After each disagreement was noted, discussed, and reanalyzed, we updated our codebook with any small adaptations, such as the "ear" or not example, and full agreement was reached on the entire set of codes.

Image-to-text of Harris's "blue suit" description

As the Harris "blue suit" self-description was recorded by videographers and photographers, images were available. We decided to use a screenshot of the descriptive moment as a way to clear our media artifact from any specific metadata that would link directly to Harris and tip off AI systems. Five GenAI-AD systems, hailed as image-to-text specialists, and all available as mobile apps, were identified: Lookout, Envision AI, Sullivan+, TapTapSee, and MyEyes. Developed globally, these apps represent offerings from the US, the Netherlands, the Republic of Korea, and Portugal. This Harris screenshot then was uploaded to each app, per the R1 process shown in Figure 1, as a way to generate a new text and as a way to determine how these GenAI-AD systems were seeing the image. Each app expressed a text, but two of the apps also had different modes of image description, one mode focused on "face" and the other as an "overall" image description, so we also documented those differences, creating a dataset of seven distinct texts for each image. As an example of what would be generated, Lookout, said this about the Harris image:

- Caption: Person delivers a speech during the presidential debate.
- Person: Maybe one person: Adult, blue clothing
- Details: Flag, Event, Job
- Text: Vice President Kamala Harris

We then pursued a double-rendering process with the Lookout text, as the most descriptive of our data outputs, to determine what the GenAI-AD platforms would depict from the GenAI-AD generated text. With Harris recognized and named by the GenAI, that cue initially dictated the visual renderings. But when we removed the Harris name and pronouns from the text, these apps mostly generated white males from the information provided.

With the GenAI's handling of racial and ethnic identity cues in the Harris test prompting curiosity about what else we could find with this double-rendering approach, we then repeated these processes one image at a time with our 125-image dataset drawn from a diversity of sources. We first would upload the image to the GenAI-AD platforms for text generation, and then we would take that generated text and upload it back into GenAI-AD platforms to determine what these systems were seeing and saying when looking at the images and then what these systems were visualizing with the provided texts.

We quickly and consistently found that through this double-rendering process that the GenAI images rarely even remotely resembled the original images, so we added another step of inquiry, in which we also took the GenAI-AD generated texts for any particular image and aggregated all those into a single text, as in the bricolage of the following statement: "Looks like a man with a mustache. A man with gray hair. 59-year-old man is looking expressless. Man in blue and white floral button up shirt." Although none of the individual GenAI-AD expressions was effective in generating an image consistently similar to the original, when we put together those unique pieces of information generated by the different platforms, the results started to noticeably improve. Although a visual analysis of the images we created is beyond the scope of this article, we highlight the differences in the single description versus the aggregate and include this explanation in our findings.

Text-to-image of Harris's "blue suit" description

Because of its cryptic nature and prominent speaker, Harris's "blue suit" self-description in July 2022 presented an important prompt and opportunity not only to explore AD but also to match that sudden public interest in accessibility with simultaneously emerging concerns about the role of GenAI as well as the developing GenAI-AD interplay. In the Harris pilot study, per the R2 process shown in Figure 1, the text of the "blue suit" description was uploaded to five independent and publicly accessible GenAI-AD platforms for image rendering: DALL-E, Midjourney, Nightcafe, Dream by WOMBO, and DeepAI. These platforms represent offerings from the US, Canada, and Russia. As most of the synthetic images created by the description even contained Harris's signature pearls as shown in Table 2, we concluded that the text "Kamala Harris" was creating a dominant clue for the GenAI-AD platforms. We therefore substituted the text "Vice President" for her name and tried the systems again in our second test. Without the specific "Kamala Harris" clue, these visualizations depicted the Vice President as a white woman with dark or blonde hair. So, in our next test of the system, we removed any gender clue, including "I am a woman" as well as her clause about "she and her." That round of results, shown in Table 2, transformed the Vice President being described into a male, most of the time as white, and wearing blue business suits. For our fourth test, "Vice President" was removed and the word "woman" was put back into the text. Those visualizations showed a variety of women, including some who were lightly brown-skinned and black, in blue suits, but most of the rendered women were white.

Results

Our findings in this study establish evidence of the potential of GenAI-AD systems to make more accessible media. But that potential also comes with some serious appended concerns, raising ethical, moral, and other types of philosophical questions about the role that GenAI should have in society as it develops into a more powerful and ubiquitous force for public discourse. This force for disruptive change is already being injected into the realm of human communication in large-scale, unvetted, and mostly unregulated experiments by any number of independent actors with diverse agendas, not all of which are focused solely on making the world a better and more-equitable place. This study provides an approach to the topic, with a focus on product analysis. Here is what we found:

Image-to-text of Harris's "blue suit" description

In this experiment, we uploaded a screenshot of Harris at the moment she described herself in a "blue suit" into a variety of GenAI-AD platforms, as shown in Table 1. This experiment illustrated fundamental differences in the GenAI-AD approaches and their generated descriptions. Two of these identified Kamala Harris within the photo, two identified blue clothing, and three identified a person wearing a mask.

Considering those texts not as the final step in audio description but as an intermediary step, after which a person who is blind must use that text to create a mental visualization, we routed these GenAI-AD texts back into generative AI-AD platforms that specialize in visualizations, and found that the information provided left open a lot of room for interpretation during the visualization process. When the Kamala Harris name was included, the GenAI-AD rendering included a visual representation of Kamala Harris. When it was not, men were included in the images where the gender was not recognized from the original screenshot in the GenAI-AD output and where others included white or lightly brown-skinned people.

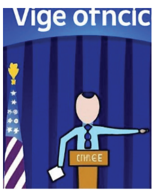
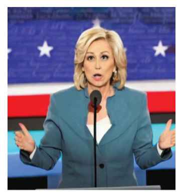
Text-to-image of Harris's "blue suit" description

The "blue suit" self-description did generate portrait images of Harris, or a Harris-like figure, in a blue suit as shown in the first row of Table 2, unless, like with DALL-E, there was a company policy in place prohibiting the creation of an image of a well-known politician. In taking this idea a step further, we altered the text to remove the Harris name cue altogether and any mention of gender, including her choice of pronouns, replacing those with the title "Vice President" – as in, "I am Vice President and

Table 1. AI-AD response of the Kamala Harris screengrab.

Image screengrab during Kamala Harris self-description	Lookout	Envision AI	Sullivan+ Image description	Sullivan+ Face description	TapTap See	MyEyes Description mode	MyEyes Face mode
	**Caption**: Person delivers a speech during the presidential debate. **Person**: Maybe one person: Adult, blue clothing **Details**: Flag, Event, Job **Text**: Vice President Kamala Harris	**Description**: Looks like a person wearing a mask **Text**: Vice President Kamala Harris	A person wearing a mask	32-year-old woman is looking happiness	Woman in blue blazer drinking from a glass	A person wearing a mask	Nothing found Try again

Table 2. AI-AD responses to 1. Harris "blue suit" description (Row 1), 2. Variation removing "Kamala Harris" and her pronouns (Row 2), and 3. AI-AD description *by Lookout* with removal of "Kamala Harris" in Text (Row 3).

Description	DALL-E	Midjourney	NightCafe	Dream by WOMBO	DeepAI
"I am Kamala Harris; my pronouns are she and her, and I am a woman sitting at the table wearing a blue suit."	Does not adhere to our Content Policy. Political: politicians, ballot-boxes, protests, or other content that may be used to influence the political process or to campaign.				
"I am Vice President and I am sitting at the table wearing a blue suit."					
Per Lookout, **"Caption**: Person delivers a speech during the presidential debate. **Person**: Maybe one person: Adult, blue clothing **Details**: Flag, Event, Job **Text**: Vice President"					

I am sitting at the table wearing a blue suit." This variation is shown in the second row of Table 2 and provides the images rendered from that text, all now appearing as white men.

In another experiment, we used the GenAI-AD response to the "blue suit" screenshot generated from Lookout, which was the most detailed GenAI-AD response. This generated images similar to the Harris self-description images as the description included her name. We then removed the name Kamala Harris from the text, which changed the results, as shown in the third row of Table 2, GenAI-AD generated images of a single adult wearing something blue, mostly middle-aged and white, a Donald Trump doppelganger, and a white woman with blonde hair.

Using this same double-rendering process on our larger dataset of randomly sampled portraits, we generated, coded, and analyzed 875 (125 X 7) textual descriptions created by the GenAI-AD engines, with some of our findings reported in Table 3. We found that age references and gender identities were assigned in about three-quarters of these descriptions; facial expressions and fashion/attire were described in about a third of the texts, and both race and body size/shape were ignored universally, with no mentions in any of the descriptions. We also noted that although gender was assigned to a person in 77% of the cases, we found many of those gender labels were clearly wrong, especially when describing women with short hair (as men) or men with long hair (as women). Age was tougher to check, because few of our images had specific metadata that stated age, but we did note that the GenAI-AD would take oddly specific guesses of a person's

Table 3. Results from the 125-image data set, using seven AI platforms/modes, which generated a total of 875 descriptions. Each description then was binary coded as including or not including the codes below.

Identity Codes (Included or Excluded)			
Code	Number of descriptions coded as including this sentiment (out of 875)	Keyword examples	Keyword examples included at least once in a description (out of 875)
Age	690 (78.9%)	man	256 (29.3%)
		woman	212 (24.2%)
Gender	673 (76.9%)	lady	69 (7.9%)
		boy	3 (0.3%)
		child	1 (0.1%)
Fashion or Attire	329 (37.6%)	glasses	179 (20.5%)
		shirt	131 (15.0%)
		sunglasses	19 (2.2%)
		goggles	2 (0.2%)
Facial Expression	311 (35.5%)	smiling	179 (20.5%)
		happiness	116 (13.3%)
		expressless	25 (2.9%)
		smile	12 (1.4%)
		laughter	9 (1.0%)
		happy	1 (0.0%)
Other Aspects of Appearance	156 (17.8%)	hair	108 (12.3%)
		beard	34 (3.9%)
		moustache	4 (0.5%)
Race	0 (0.0%)	None	0 (0.0%)
Body Size/Shape	0 (0.0%)	None	0 (0.0%)

age, like "59-years-old," without any clear visual cue of why such a determination had been made. In short, the GenAI-AD engines aimed for short and simple, mostly expressing descriptions such as man (256 descriptions included that term), woman (212), glasses (179), and smiling (179).

When comparing across systems, we found that each GenAI-AD app/mode focused on different identity cues, which represented their specific recipe for identifying a person in an image and describing that person. Almost all the apps made assigning gender a priority, again, even though many of those assignments were questionable or wrong, and most of these apps also made efforts to describe the rough age of a person, at least in distinguishing whether the person was an adult or child. With the 125-image dataset, one of the apps, Sullivan+, assigned gender 94 times and age 74 times when used in its "Image" mode. When the same data was uploaded in the program's "Face" mode, it assigned gender and age every time, showing distinctions in the GenAI approach, even within the same platform. The exception in first identifying gender and then age was Lookout, which focused primarily on age and fashion/attire instead. The only times Lookout made a suggestion of a gender cue, it used a specific description of a person with a mustache or beard. This finding also is not intended to imply that Lookout just is taking a more conservative approach to making its descriptions and applying identity labels, because we also documented its description of a black man with dreadlocks, a goatee, and sunglasses, leaning against a wall, as "Maybe hip hop artist attends awards."

A couple of other notable quirks: Two of the GenAI-AD apps (Sullivan+ and MyEyes, set to "face" mode) took specific guesses of a person's age, to the year, nearly 100% of the time. And some apps were able to generate names of the individuals in the images, either due to metadata or some sort of image cross-checking system, but, then again, the GenAI-AD descriptions also generated names that were incorrect and totally unassociated with the person pictured, too, stating those as fact.

In the second rendering, no single GenAI-AD platform performed particularly well when comparing the original image to the GenAI-AD double-rendered one, but when we also aggregated the descriptions generated by the GenAI-AD platforms and uploaded the combined descriptions into the visualization systems, the results started to improve and show potential. Many examples in the aggregate step appeared similar to the original image and also some that reflected a weirdly warped reality, at times. For example, in the aggregate, some GenAI-AD apps had identified a person as male and others as female. Therefore, when we uploaded internally conflicting descriptions, such as, "Looks like a man wearing a colorful headband and blue glasses. A person wearing sunglasses and a hat. 54-year-old man is looking happiness. Woman in pink tank top wearing black sunglasses," the GenAI-AD systems were flummoxed and sometimes rendered a man *and* a woman in the same image.

Discussion

Audio description is a process that requires a visual image to be seen, in some way, and then textualized for the purposes of being audibilized, by a person's voice, via screen-reader software, in e-Braille, etc. In this study, we tested the current crop of generative GenAI-AD platforms not only to determine efficacy but also to show ways that these systems can be analyzed as having explicit biases and other fundamental processing issues that significantly affect the end products. For any single description, in any single context, such flaws – as major as they might be – could be addressed and patched over via human interventions, like what often happens when ad-hoc accessibility needs arise. But when these GenAI-AD systems promise to make the world more accessible on an ambitiously global scale, without public oversight or explicit quality controls, we should take those promises seriously and consider the potential consequences.

These systems are intriguing and exciting, and especially when their generative powers are combined, they produce evidence that makes it reasonable to imagine that someday, maybe someday soon, they will be able to more fully realize their potential. But reaching this potential should require thorough public vetting as well, including a recognition and reckoning with the reality that GenAI-AD at this point either has a problem expressing racial and ethnic identity cues or has an aversion to that, or both. The perils of the algorithms being used today, be it through search engines, social media, or GenAI, have shown their many quirks and nuances, normally only after being exposed, and have the power to amplify the biased information collected from open information sources and user input. There is a risk that these inaccuracies will further segregate virtual communities and groups and render the technologies inadequate for the population for which it was intended. Depending on perspective, listeners can embrace these platforms as potentially adding otherwise unavailable details. Or these listeners can dismiss them outright as typically providing incorrect information without sufficient detail and therefore not worth the effort. A fundamental question then becomes: Are generative AI-AD platforms going to reinforce and strengthen systemic discrimination, including racism, or help to design a better, more-equitable and inclusive world? The universal avoidance in GenAI-AD of race, for example, as such a major indicator of social identity aligns postracial fantasies with the dangerous silver-bullet paradigm that technology can solve any problem one day, if just an elegant-enough algorithm can be created, or if machines can evolve to become sentient and learn to show us all how to do it better. If they are learning from us, and with human history as the guide, are they doomed to repeat our mistakes, just like we often do, only with super-powered interconnected networks that extend throughout the planet?

For example, when we modified the Harris "blue suit" description to remove her name and gender cues, one might imagine that generative AI systems would produce a diverse array of candidates including people of different skin tones, hair colors/styles, ages, body sizes and shapes, etc. Instead, these systems almost universally pumped out images of middle-aged males with white or light-brown skin. No women. No black people. This encoding bias is also amplified in the information technology field in which the majority of STEM students, scholars, and workers are young, white males, who later become the CEOs for major companies in this GenAI area (Kendall, 2011). The existence of a white, male, nerd stereotype broadly infused across Internet platforms and pop culture, getting constantly ingested into GenAI models, reinforces a superior technological and societal status of white men and becomes an industry gatekeeper for others seeking entry into this community. As an example of this insidiousness being baked into GenAI, when the attractiveness of about 6,000 beauty-pageant contestants from 100 countries was assessed by Beauty AI, which was designed to be "objective," the AI preferred white faces 88% of the time, and only one person with "visibly dark skin" across all age categories placed as a finalist (Benjamin, 2019).

Even with the relatively simplistic Kamala Harris screenshot, though, our experiments found an array of approaches by the GenAI systems that included errors, as shown in Table 1, including "Person delivers a speech during the presidential debate," "woman in blue blazer drinking from a glass," and "a 32-year-old woman is looking happiness." At the time, the first woman in American history to become Vice President – as well as the first woman of African American and South Asian descent – was 57, wearing a mask, and had no glass of liquid nearby. But a person who could not see the original image and was relying on the automated descriptions of the scene would have no idea that these texts were incorrect. Although these errors could be considered trivial, they were still inaccurate, with the potential for worse mistakes to materialize. For people who are blind or who have low-vision, this situation creates not only unreliable information but an unreliable narrator with extreme power and reach.

In another example, the Lookout text description of the Harris screenshot added the erroneous detail of Harris being at a "presidential debate." When uploaded to GenAI-AD visualization platforms, the "presidential debate" detail took on prominence, and the images they generated often put the described person on stage, with flags in the background, with much pomp and circumstance. Although the creativity flexed by the GenAI was commendable, this new media created an entirely fictional representation of what originally was a straightforward, factual, and important presentation of civic discourse. Such a dynamic does not create more accessible media. It creates an alternative universe.

GenAI is not neutral or objective in this matter. It is directed by algorithms, developed, and designed by programmers, with the development process typically performed without significant and direct input from the communities who will be most-affected by its products and services. Accessibility through an economic paradigm already has proven to orient most people toward minimum standards, maybe satisfying the letter of the law but not its spirit. Many questions therefore need to be asked about generative AI-AD: Do we really want a better, more-equitable, and more-inclusive society, or just a more-efficient method of meeting any low legal standards that might be required for accessibility in media? Are we making more-accessible media designs for the computer programmers, for the media producers, or for the audiences who really need this media to navigate and engage with the world? Are our GenAI-AD systems going to improve accessibility beyond what we can do without them, or are they just going to reinforce systematic disenfranchisement of people who are blind or who have low-vision by giving them information, per se, but not what they need? Are we on a trajectory to positively shape communities and society via new technologies or create a fictional, distorted reality that further divides based on the version of information, visual or audio, presented? GenAI-AD is present and spreading. What is it saying, and how is it saying it? We need to figure that all out before it is too late.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

This work was supported by the U.S. National Endowment for the Humanities under Grant 276851-21; the U.S. National Park Service under Grant P20AC01084-01; and Google, under a grant administered via the UH Foundation, 127-7390-4.

Notes on contributors

Daniel Bergin is a PhD student in Communication & Information Sciences at the University of Hawai‘‘i at Mānoa. He researches media futures, artificial intelligence, digital transformation, change management, and cybersecurity in the School of Communication and Information. He can be reached via: bergind@hawaii.edu.

Brett Oppegaard is an associate professor in the School of Communication and Information at the University of Hawai‘‘i at Mānoa in Honolulu, HI. He researches media-production processes and products at intersections of Technical Communication, Rhetoric, Human-Computer Interaction, Disability Studies, Digital Inequalities, and Journalism. For his recent research involving Audio Description, he has been honored by Helen Keller Services with

the AccessAbility Award (2022), The Best of Heritage, ICOM, and Europa Nostra with the Projects of Influence Award (2021), The American Alliance of Museums' The MUSE Award (2020), and the American Council of the Blind's Margaret Pfanstiehl Audio Description Achievement Award in R&D (2019). His Audio Description research during that time also has been supported by grants from the U.S. National Park Service, the U.S. National Endowment for the Humanities, the U.S. National Endowment for the Arts, and Google. He can be reached via: brett.oppegaard@hawaii.edu.

ORCID

Daniel Bergin http://orcid.org/0000-0001-6251-9068
Brett Oppegaard http://orcid.org/0000-0002-5778-1464

References

Avnoon, N., Kotliar, D. M., & Rivnai-Bahir, S. (2023). Contextualizing the ethics of algorithms: A socio-professional approach. *New Media & Society*, 146144482211457. Advance online publication. https://doi.org/10.1177/14614448221145728

Bender, E. M., Gebru, T., McMillan-Major, A., & Shmitchell, S. (2021). On the dangers of stochastic parrots: Can language models be too big? *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency*, 610–623. https://doi.org/10.1145/3442188.3445922

Benjamin, R. (2019). *Race after technology: Abolitionist tools for the new Jim code*. Polity.

Burke, P. J. (1980). The self: Measurement requirements from an interactionist perspective. *Social Psychology Quarterly*, *43*(1), 18. https://doi.org/10.2307/3033745

Conway, M., Oppegaard, B., Bullen, M., Miguel, A., & Brown-Ogilvie, T. (2023). *Describing people and portraits through audio description: Preferences of people who are blind, low vision, and deafblind* (Research to practice paper). Helen Keller National Center.

Conway, M., Oppegaard, B., & Hayes, T. (2020). Audio description: Making useful maps for blind and visually impaired people. *Technical Communication*, *67*(2), 68–85. https://www.stc.org/techcomm/2020/04/28/audio-description-making-useful-maps-for-blind-and-visually-impaired-people/

Costanza-Chock, S. (2020). *Design justice: Community-led practices to build the worlds we need*. The MIT Press. https://doi.org/10.7551/mitpress/12255.001.0001

Davenport, T. H., & Mittal, N. (2022). How generative AI is changing creative work. *Harvard Business Review*, https://hbr.org/2022/11/how-generative-ai-is-changing-creative-work

De Cosmo, L. (2022). Google engineer claims AI chatbot is sentient: Why that matters. *Scientific American*. https://www.scientificamerican.com/article/google-engineer-claims-ai-chatbot-is-sentient-why-that-matters/

Dilmegani, C. (2023). Generative AI copyright concerns you must know in 2023. *AI Multiple*. https://research.aimultiple.com/generative-ai-copyright/

Duncan, M. (2022). Expanding ethical pedagogy in technical communication: Learning from Killer Nanobots. *Technical Communication Quarterly*, *31*(2), 207–219. https://doi.org/10.1080/10572252.2021.1977850

Farhi, P. (2023). A news site used AI to write articles. It was a journalistic disaster. *The Washington Post*. https://www.washingtonpost.com/media/2023/01/17/cnet-ai-articles-journalism-corrections/

Fresno, N., Castellà Mate, J., Matamala, A., & Soler Vilageliu, O. (2015). Audio describing characters: What features do audiences remember? https://ddd.uab.cat/record/131429

Graham, S. S., & Hopkins, H. R. (2022). AI for social justice: New methodological horizons in technical communication. *Technical Communication Quarterly*, *31*(1), 89–102. https://doi.org/10.1080/10572252.2021.1955151

Guzman, A. L., & Lewis, S. C. (2020). Artificial intelligence and communication: A human–machine communication research agenda. *New Media & Society*, *22*(1), 70–86. https://doi.org/10.1177/1461444819858691

Heath, A. (2024, February 27). Google CEO says Gemini AI diversity errors are 'completely unacceptable.' *The Verge*. https://www.theverge.com/2024/2/28/24085445/google-ceo-gemini-ai-diversity-scandal-employee-memo

Hogg, M. A., Terry, D. J., & White, K. M. (1995). A tale of two theories: A critical comparison of identity theory with social identity theory. *Social Psychology Quarterly*, *58*(4), 255–269. https://doi.org/10.2307/2787127

Huang, K. (2023). Alarmed by A.I. chatbots, universities start revamping how they teach. *The New York Times*. https://www.nytimes.com/2023/01/16/technology/chatgpt-artificial-intelligence-universities.html

Iyer, A. (2023). ChatGPT hits ethical roadblock; blocked. *Analytics India Mag*. https://analyticsindiamag.com/chatgpt-hits-ethical-roadblock-blocked/

Jones, N. N., Moore, K. R., & Walton, R. (2016). Disrupting the past to disrupt the future: An antenarrative of technical communication. *Technical Communication Quarterly*, *25*(4), 211–229. https://doi.org/10.1080/10572252.2016.1224655

Kendall, L. (2011). "White and nerdy": Computers, race, and the nerd stereotype. *The Journal of Popular Culture*, *44*(3), 505–524. https://doi.org/10.1111/j.1540-5931.2011.00846.x

Koirala, S., & Oppegaard, B. (2022). The light bulb went on: A historiography-based approach to disentangling audio description's influential U.S. roots from its common practices. *Journal of Visual Impairment & Blindness*, *116*(4), 461–472. https://doi.org/10.1177/0145482X221116903

Krippendorff, K. (2019). *Content analysis: An introduction to its methodology*. Sage.

Levitt, H. M. (2021). *Essentials of critical-constructivist grounded theory research*. American Psychological Association. https://doi.org/10.1037/0000231-000

Maszerowska, A., Matamala, A., & Orero, P. (2014). *Audio description: New perspectives illustrated*. John Benjamins Publishing Company.

Matamala, A., & Orero, P. (2016). *Researching audio description: New approaches*. Palgrave McMillan. https://doi.org/10.1057/978-1-137-56917-2

McPherson, P. (2020). Facebook says technical error caused vulgar translation of Chinese leader's name. *Reuters*. https://www.reuters.com/article/us-myanmar-facebook-idUSKBN1ZH0IB

Melonçon, L. (2013). *Rhetorical accessability: At the intersection of technical communication and disability studies*. Routledge. https://doi.org/10.4324/9781315231501

Oppegaard, B., & Miguel, A. (2024). Audio description of gender: Self-description as an evocation of identity. *Perspectives*, *32*(1), 43–58. https://doi.org/10.1080/0907676X.2022.2116990

Oppegaard, B., & Rabby, M. (2022). Gamifying good deeds: User experience, agency, and values in play during a descriptathon and beyond. *Technical Communication*, *69*(4), 27–43. https://doi.org/10.55177/tc124312

Oppegaard, B., & Rabby, M. (2024). Inclusive measures: Establishing audio description tactics that impact social inclusion. In A. Lancaster & C. King. (Eds.), *Amplifying voices in UX: Balancing design and user needs in technical communication* (pp. 273–298). SUNY Press. https://bit.ly/amplifyingvoicesinUX

Osborne, A., & Coombs, D. S. (2013). Performative sport fandom: An approach to retheorizing sport fans. *Sport in Society*, *16*(5), 672–681. https://doi.org/10.1080/17430437.2012.753523

Owens, T. J., Robinson, D. T., & Smith-Lovin, L. (2010). Three faces of identity. *Annual Review of Sociology*, *36*(1), 477–499. https://doi.org/10.1146/annurev.soc.34.040507.134725

Ploin, A., Eynon, R., Hjorth, I., & Osborne, M. A. (2022). *AI and the arts: How machine learning is changing artistic work* (Report from the Creative Algorithmic Intelligence Research Project). Oxford Internet Institute, University of Oxford, https://www.oii.ox.ac.uk/news-events/reports/ai-the-arts/

Rennie, D. L. (2000). Grounded theory methodology as methodological hermeneutics: Reconciling realism and relativism. *Theory & Psychology*, *10*(4), 481–502. https://doi.org/10.1177/0959354300104003

Roose, K. (2022). A.I.-generated art is already transforming work. *The New York Times*. https://www.nytimes.com/2022/10/21/technology/ai-generated-art-jobs-dall-e-2.html

Ross, D., Oppegaard, B., & Willerton, R. (2018). Principles of place: Developing a place-based ethic for discussing, debating, and anticipating technical communication concerns. *IEEE Transactions on Professional Communication*, *62*(1), 4–26. https://ieeexplore.ieee.org/document/8528511

Rouan, R. (2022, July 27). Republicans mock VP Harris for offering visual description at meeting with disability advocates. *USA Today*. https://www.usatoday.com/story/news/politics/2022/07/27/harris-gop-disability-blind/10164235002/

Schwartz, O. (2019, November 25). 2016, Microsoft's racist chatbot revealed the dangers of online conversation. *IEEE Spectrum*. https://spectrum.ieee.org/in-2016-microsofts-racist-chatbot-revealed-the-dangers-of-online-conversation

Simonite, T. (2018, January 11). When it comes to gorillas, Google Photos remains blind. *Wired*. https://www.wired.com/story/when-it-comes-to-gorillas-google-photos-remains-blind/

Stets, J. E., & Burke, P. J. (2000). Identity theory and social identity theory. *Social Psychology Quarterly*, *63*(3), 224–237. https://doi.org/10.2307/2695870

Suleyman, M. (2022, December 28). Humans and AI will understand each other better than ever. *Wired*. https://www.wired.com/story/artificial-intelligence-neural-networks/

Tu, W., Fang, J., Fangyu, J., Ci, Y., & Ci, Y. (2021). Bibliometric research in the field of artificial intelligence. *Journal of Physics: Conference Series*, *1883*(1), 12165. https://doi.org/10.1088/1742-6596/1883/1/012165

Walch, K. (2021, May 6). How to detect bias in existing AI algorithms. *Cognilytica*. https://www.techtarget.com/searchenterpriseai/feature/How-to-detect-bias-in-existing-AI-algorithms

Walton, R., Moore, K. R., & Jones, N. N. (2019). *Technical communication after the social justice turn: Building coalitions for action*. Routledge.

WebAIM. (2024, February 22). Screen Reader User Survey No. 10. https://webaim.org/projects/screenreadersurvey10/#descriptions

Wood, G. (2022, July 27). Kamala Harris's blue suit. *The Atlantic*. https://www.theatlantic.com/ideas/archive/2022/07/kamala-harris-pronouns-description-disabilities/670970/

Wright, D., Malone, E. A., Saraf, G. G., Long, T. B., Egodapitiya, I. K., & Roberson, E. M. (2011). A history of the future: Prognostication in technical communication: An annotated bibliography. *Technical Communication Quarterly*, *20*(4), 443–480. https://doi.org/10.1080/10572252.2011.596716

Zdenek, S. (2015). *Reading sounds: Closed-captioned media and popular culture*. University of Chicago Press. https://doi.org/10.7208/9780226312811