



“Your friendly AI assistant”: the anthropomorphic self-representations of ChatGPT and its implications for imagining AI

Karin van Es¹ · Dennis Nguyen¹

Received: 27 June 2024 / Accepted: 14 October 2024 / Published online: 27 October 2024
© The Author(s) 2024

Abstract

This study analyzes how ChatGPT portrays and describes itself, revealing misleading myths about AI technologies, specifically conversational agents based on large language models. This analysis allows for critical reflection on the potential harm these misconceptions may pose for public understanding of AI and related technologies. While previous research has explored AI discourses and representations more generally, few studies focus specifically on AI chatbots. To narrow this research gap, an experimental-qualitative investigation into auto-generated AI representations based on prompting was conducted. Over the course of a month, ChatGPT (both in its GPT-4 and GPT-4o models) was prompted to “Draw an image of yourself,” “Represent yourself visually,” and “Envision yourself visually.” The resulting data ($n = 50$ images and 58 texts) was subjected to a critical exploratory visual semiotic analysis to identify recurring themes and tendencies in how ChatGPT is represented and characterized. Three themes emerged from the analysis: anthropomorphism, futuristic/futurism and (social)intelligence. Importantly, compared to broader AI imaginations, the findings emphasize ChatGPT as a friendly AI assistant. These results raise critical questions about trust in these systems, not only in terms of their capability to produce reliable information and handle personal data, but also in terms of human–computer relations.

Keywords GenAI · imaginaries · self-representation · anthropomorphism · trust · thematic analysis

1 Introduction

The launch of OpenAI’s ChatGPT in November 2022 considerably broadened public access to advanced AI technologies based on generative pretrained transformers (GPT). It has a relatively easy-to-use interface that affords users to request text as well as visual output based on typed (or spoken) instructions, also called prompts. Drawing from an immense training dataset, the underlying algorithms use billions of parameters to generate responses that display a seemingly human-like level of understanding in ChatGPT’s communication with users. For example, it can generate various types of text, provide programming code for different tasks, and create images. These capabilities have made it a widely used tool for a vast variety of purposes in diverse domains, both professional and private. In 2023, OpenAI linked ChatGPT with DALL-E, a visual content creation

algorithm trained on image-text pairings to discern statistically significant relationships between words and visual elements. As such, ChatGPT became symbolic for the promises of generative artificial intelligence (hereafter: GenAI) and stands at the epicenter of a renewed AI hype in the 2020s.

While the public discourse often highlights the benefits of GenAI, the launch and widespread adoption of ChatGPT and DALL-E have sparked heated discussions about a wide range of issues. These include implications for authorship, creativity, authenticity, reliability, fraud, misinformation, and copyright infringement (Xiao and Wang 2021; Spring 2024; Appel et al. 2023). Additionally, users’ confidence in the quality and accuracy of the outputs generated by these technologies is crucial, as it has profound implications for knowledge production and communication practices in the public sphere (Hepp et al. 2023). More generally, current GenAI applications suffer from similar limitations and risk potentials as other machine learning models. These issues include perpetuating existing biases and inequalities (Bender et al. 2021; Bianchi et al. 2022), exploiting workers in the development of models like GPT (Williams et al. 2022), and the significant environmental costs of training

✉ Karin van Es
K.F.vanEs@uu.nl

¹ Media and Culture Studies, Utrecht University (NL),
Muntstraat 2a, 3512 EV Utrecht, The Netherlands

and maintaining these models (Bender et al. 2021; Luccioni et al. 2023).

Public perception can be highly influential in the development, deployment, and regulation of emerging technologies (Cave et al. 2020). Arguably, widely held views about a given technology among lay people are likely to be shaped by a complex interplay of different factors, such as dominant public discourses in the media, personal experiences both at work and in the private sphere, as well as individuals' socio-economic positionality and cultural background. In the broader context of AI, previous research has examined how news reporting, governmental policies, and political communication can reflect and shape dominant views in public discourse (Nguyen and Hekman 2024; Sartori and Bocca 2023; Bareis and Katzenbach 2022). However, AI portrayals in popular culture are at least equally important for the formation of individual mental models about emerging automated technologies (Leaver 2012). Narratives about AI and the socio-technical imaginaries that they carry (Jasanoff 2015) play an important role in how societies make sense of technologies, fill them with meaning, and use them.

A growing number of studies critically investigate the representations and discourses around AI across these interconnected contexts (Gourlet et al. 2024; Natale and Ballatore 2020). Some argue that current AI discourses are misleading and detract from crucial critical and ethical reflection (Dihal and Duarte 2023; Leaver and Srdartov 2023). Yet, these “evocative stories” attempting to make AI understandable for the public can also be helpful entry points into public conversation about these complex technologies (Goode 2018: 185). The formats are diverse but focus on framing practices in both words and visuals. However, while some studies investigate visual representations of AI and big data in general (Vrabič Dežman 2024; Cave and Dihal 2020; Pentzold et al. 2019), there is only limited research available specifically on visual GenAI imaginaries.

In this paper, we aim to narrow this research gap by providing a critical analysis of images portraying GenAI. Instead of sampling external media texts, the study focuses on the case of ChatGPT, examining how it visualizes “itself” when prompted to do so and how it responds textually to this request. Considering that many people rely on ChatGPT as a source of knowledge about the world, it is important to explore the types of images it produces and how these might influence public perceptions of technology. This exploratory study covers a month of interactions with ChatGPT from two separate paid accounts that yielded a total of 50 images and 58 text responses to variations of the prompt “create an image of yourself.” Critically analyzing these self-visualizations creates important insights into the metaphors ChatGPT imagery contains, especially regarding values and biases embedded in the underlying training data. It further raises questions about the guidelines integrated

by OpenAI's programmers in ChatGPT responses. To this end, a qualitative visual content analysis was applied to the data that centered on dominant visual elements as well as thematic emphases and visual framing practices; the main objective was to unearth what socio-technical imagery one of the most widely accessed GenAI tools promotes about itself. The present paper adds relevant insights into the complex relationship between GenAI technologies, public perception, socio-technical imaginaries, the role of data, and the influence of technology companies.

The paper is structured as follows: it first discusses the importance of studying imaginaries in relation to dominant representations of AI and reflects on the main criticisms of generative AI. Afterwards, the methodology is introduced, which utilized prompting for generating images with ChatGPT portraying itself as a data source for visual analysis. The paper then reflects on the responses and images; it is here argued that exploring GenAI's self-portrayals reveals more granular and distinct features in representations of AI. Finally, it concludes by outlining important considerations regarding public understanding and the expectations it sets for AI capabilities, emphasizing the need for more productive ways of representing GenAI.

2 Researching socio-technical imaginaries and AI

In the field of Science and Technology Studies (STS), Jasanoff's proposition to analyze what she calls “socio-technical imaginaries” (2009; 2015) has gained wide acceptance and yielded a growing number of studies focusing on how discourses shape the understanding of and visions for technology. She defines them as “collectively held, institutionally stabilized, and publicly performed visions of desirable futures, animated by shared understandings of forms of social life and social order attainable through, and supportive of, advances in science and technology” (Jasanoff 2015: 4). Importantly, Jasanoff emphasizes how technology and the social are enmeshed, co-shaping each other. The hopes and desires that societies project upon AI have an influence on the role these technologies are assigned in practice. Originally, socio-technical imaginaries were linked to nation-states—which are considered discursive constructs themselves (Anderson 1983)—but the concept has been repeatedly revised and extended, as research recognized the relevance of imaginaries originating from more diverse social and cultural contexts as well as communities (Konrad and Böhle 2019; Lehtiniemi and Ruckenstein 2018). Crucially, imaginaries can be formed and emerge on an individual level or within small groups. However, they only become socio-technical imaginaries of varying societal relevance once they are adopted by a community

(Jasanoff 2015: 5), which increases the likelihood that they enter broader public discourses and gain visibility there. Examples are socio-technical imaginaries that emphasize AI's potential to create value, which emerged among technologists and entrepreneurs and are widely held in the respective communities.

It is important to highlight the structural characteristics of socio-technical imaginaries: first, multiple—even clashing—socio-technical imaginaries can exist in parallel. In fact, discourses on emerging technologies are often shaped by different perceptions, evaluations, visions, and respective framing practices that accumulate to distinct socio-technical imaginaries. For example, utopian notions of GenAI find their opposite in more critical or downright dystopian imaginaries that are also visible in public discourses (Cools et al. 2024).

Second, socio-technical imaginaries are not static but can transform and change. Influential factors are new advances in technology development, accidents, and previously unencountered risk potentials, or changes in social and/or cultural norms as well as political climates (Markelius et al. 2024). Imaginaries about AI underwent different periods of hype and excitement versus disillusionment about its true capacities over the decades (Bourne 2024; Galanos and Stewart 2024).

Third, some socio-technical imaginaries often become more dominant than others (Chenou 2019). This concerns both the meso-level of specific communities and the macro-level of public discourses on a societal level. The borders between these dimensions are porous, as contested socio-technical imaginaries and promoted alternatives discussed within communities can take center stage in the public sphere. Socio-technical imaginaries can directly link here to political agendas of the different stakeholders involved and the frameworks of the political economies within which respective discourses take place. For example, technology companies often try to promote their socio-technical imaginaries strategically to influence attempts at regulation (Natale et al. 2019; Nguyen & Beijnon 2024).

Fourth, media technology and media ecologies form the infrastructure for the formation and distribution of socio-technical imaginaries across communities and the public. News media, social media, political programs, written text, images, and audio-visual/multimodal media content all serve different yet vital purposes in these processes.

The relevance of analyzing socio-technical imaginaries lies in the emphasis on the social, cultural, and political implications that they have and the related questions of power and influence in potentially profound transformations in society. Concerning AI, Richter et al. (2023: 216) point out that “[m]edia representation plays a crucial role in shaping public perception and, hence, imaginaries of AI.” Current public debates about GenAI underline the urgency

of critically researching socio-technical imaginaries from different angles, as its introduction holds potential to reconfigure practices in virtually all societal domains.

3 Present imaginaries for AI

The history of AI, dating back to the late 1950s, is intertwined with various narratives and myths about its capabilities. As Elish and boyd (2018: 58) point out, “[p]art of what makes the phenomena of Big Data and AI so compelling is the hyped imagination of what is possible, not what is realistic.” Over the past decades, AI has undergone several cycles of optimism and fear regarding its potential. In his 2021 book, Natale explores how AI has been imbued with magical qualities (cf. Elish and boyd 2018) and personality since its inception: “The emergence of AI was deeply intertwined with the rise of a technological myth centered around the possibility of creating thinking machines” (ibid. 35). Natale explains that this perception was influenced not only by mainstream media, which frequently employed misleading metaphors and overstated the capabilities of technology, but also by researchers who described computing using terms related to human intelligence, such as ‘thinking’, ‘memory’, and ‘feeling.’ These discursive practices framing AI as human-like or even super-human have diverse and complex causes. On the one hand, to some extent, they are unintentionally reproduced—especially by non-experts—and anthropomorphic representations of AI can be considered the outcome of unreflective and uncritical use of what may seem like intuitive metaphors, but that are not necessarily linked to any ulterior motives. On the other hand, they can reflect the vested interests of stakeholders (e.g., companies with financial stakes, AI researchers depending on funding), and are strategically deployed to push public discourse in specific directions.

A valuable academic corrective to the hypes surrounding AI has been *Atlas of AI* (2022) by Crawford, who challenges common misconceptions about AI. She contends that AI is not “artificial,” as it heavily depends on human labor and natural resources. Additionally, she argues that AI does not embody “intelligence” in the traditional sense, because it operates based on training data and lacks autonomy and rationality. Similarly, Leaver and Srdarov (2023) examined the hype surrounding ChatGPT and GenAI in early 2023. They identify magical thinking about GenAI and how it was anthropomorphized by emphasizing its capability to create human-like answers and portraying it as “more than ‘artificial’ intelligence” (Leaver and Srdarov 2023: np). The authors also note that calls to pause AI development can inadvertently fuel the hype about the capabilities of generative AI. Indeed, the companies developing these technologies, which obviously have considerable financial

interests, are the primary drivers of narratives that overemphasize benefits and exaggerate the capabilities of AI. A case in point are discussions about super-human AI and “guesstimation exercises” among experts about when such an advanced developmental stage could be reached and what risks this may carry for humans. While these debates appear to be critical and centered on fundamental questions of AI, they distract from the actual risks unregulated uses of automation have in the present. They also fuel hype and promote magical thinking, which brings its own set of harms.

Imaginaris about AI often include specific visual representations of the technology. Several studies have critically examined these visualizations and how they could shape public perception of AI. For example, Romele (2022) explores AI visual representations in science communication. He observes the prevalent use of the color blue and recurring motifs related to androgynous faces, half-flesh and half-circuit brains, and variations on Michelangelo’s *The Creation of Adam*. Similarly, Dihal and Duarte (2023) explore how stock images shape non-experts’ perceptions of AI. Their research identifies key features commonly found in these images, such as blue color schemes, descending lines of code, depictions of the human brain, science fiction elements, white robots, white men in suits, and various anthropomorphic attributes. In fact, a Google image search by the authors of the present paper (Fig. 1) yields similar results.

These recurring features and motifs reflect dominant themes in current socio-technical imaginaries around AI. Each of them carries specific symbolic notions and

ideological implications that are important to unpack. For example, the color blue, which dominates many AI images, is traditionally associated with qualities, such as intelligence, efficiency, logic, duty, and trust (Labrecque and Milne 2012: 714). Dihal and Duarte (2023:10) specifically link blue to technological progress. Moreover, the frequent use of visual elements, such as descending lines of programming code and science fiction references, creates a futuristic aura around AI. Such a portrayal can simultaneously set unrealistic expectations about the technology’s capabilities.

The science fiction genre significantly influences these practices, as it prominently features AI in both utopian and dystopian narratives. As Hermann (2023: 319) notes, uncritically adopting science fiction tropes can distort our understanding of AI’s capabilities. This distortion can divert attention from AI’s real-world implications and risks. However, as mentioned earlier, there is also beneficial potential to these stories about AI, even if they are sensationalist and misleading. As Goode (2018) argues, they can include a broader public in conversations about a generally complicated and technical topic.

Another issue is the prevalence of anthropomorphism in images of AI. Put simply, anthropomorphism occurs when non-human entities are made to be human-like, which includes physical features as well as mental capacities, the possession of conscious awareness, metacognition, intentions, and emotions (Johnson 2024: 69). This tendency to attribute human traits to AI further exaggerates AI capabilities and creates additional layers of deception. The language often used to describe AI further reflects anthropomorphic

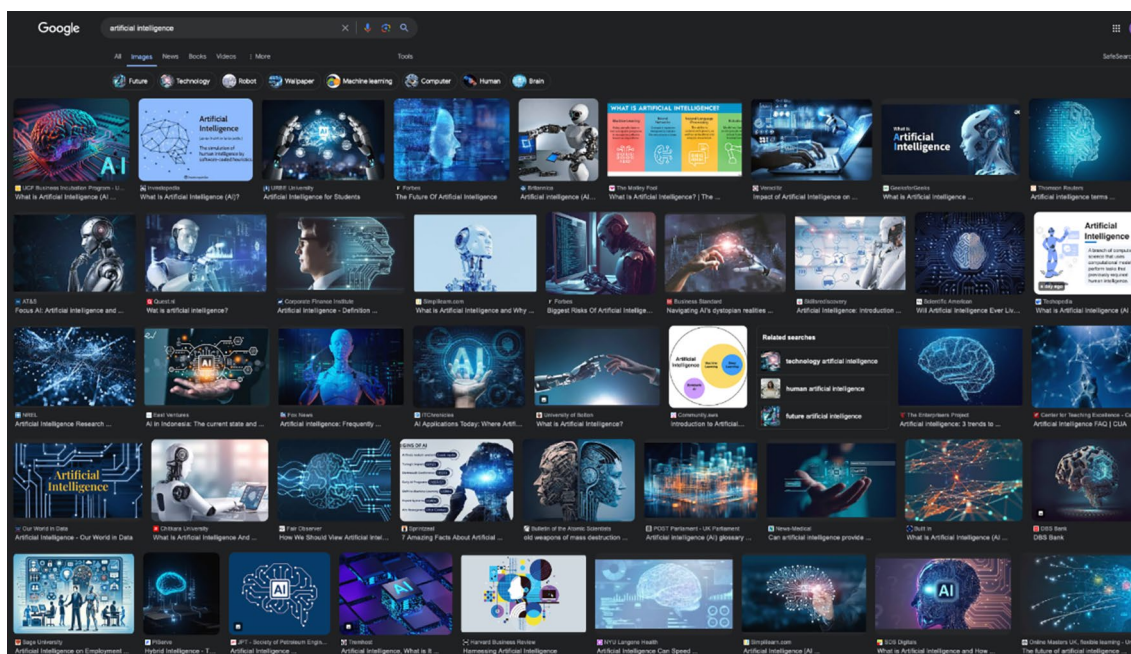


Fig. 1 Screenshot of Google Image search ‘artificial intelligence’ (17 April 2024, by authors)

tendencies. As Hunger (2023: 10) argues, this can be strategic: “The anthropomorphizing of automated statistics as ‘intelligence,’ ‘learning,’ or ‘neurons’ keeps funding flowing and successfully craves public attention.” It is often argued that ChatGPT can generate texts that read as plausible, even though they may include inaccuracies that have been described as “hallucinations.” Yet, GenAI does not possess any human-like cognition and understanding; it manipulates linguistic forms to create strings that sound confident (Bender 2023). There are fundamental differences between how a large language model and the human brain work. Thus, while it is important to criticize GenAI’s potential to fabricate incorrect content and misinformation, choosing anthropomorphic terminology such as “hallucinating” suggests that the AI somehow perceives and experiences things like a human does (Bender 2022). Thus, criticisms of AI can still contribute to a problematic anthropomorphic framing of the technology, such as using biological metaphors (Hunger 2023: 10).

Anthropomorphism furthermore raises questions of social biases, as AI is often depicted in images with White and female traits. Cave and Dihal (2020) note that stock images of AI increasingly feature machines with Caucasian characteristics, stating, “[t]he more realistically humanoid these machines become, the more Caucasian are their features” (2020: 691). Their analysis, based on Google image searches, argues that AI is racialized as White to associate it with attributes such as intelligence, professionalism, and power (Cave and Dihal 2020: 696). This default to Whiteness excludes people of color and creates ethical concerns. Current practices of anthropomorphic framing of AI can hardly be considered universally representative of humankind. Furthermore, both Dihal and Duarte (2023) and Romele (2022) argue that stereotypical representations contribute to a range of negative outcomes. They engender mistrust and anxiety, set incorrect expectations about AI’s capabilities, exacerbate social inequalities due to a lack of diverse representations, obscure the practical roles AI can play in everyday life, and restrict creative thinking about AI’s potential.

Placani (2024) explores the implications of anthropomorphism in AI, identifying it as both hype and fallacy. The hype involves exaggerating AI’s capabilities and performance, while the fallacy pertains to errors in moral and ethical judgments concerning AI, such as assigning responsibility and trust. If the aim is to assess the trustworthiness and safety of AI systems, focus should shift to how these technologies work, and it becomes imperative to avoid attributing capacities they lack by anthropomorphizing them (Shanahan 2024).

Discourses and representations can be analyzed to unpack socio-technical imaginaries associated with GenAI. Extensive research has been conducted on the

representation and perception of AI in general, but current GenAI trends and specific forms such as ChatGPT are yet to be systematically explored in this regard. Moreover, most studies focus on media representations of AI technology but rarely investigate how AI technologies generate representations of themselves. One study that does so still primarily focuses on general AI representations and GenAI output is only a limited component of the analysis (Vrabič Dežman 2024).

ChatGPT, through its integration of DALL-E, provides an opportunity to experiment with how GenAI responds to prompts that request visualizations of itself. This can provide insight into how OpenAI positions its products within the current AI debates and highlights the prevalent themes and features in ChatGPT’s output that reflect its training data. Just like Google search, ChatGPT does not merely promote content; it also mirrors what is already prevalent or popular in its training datasets (cf. Noble 2018).

Based on the above discussion, we address the following research question: *What are dominant themes and features in images generated by ChatGPT when prompted to visualize itself?* Imaginaries are tangible in images and discourse. The aim is to critically explore what socio-technical imaginaries are constituted through these images.

4 Method: semiotics and thematic analysis

The methodological approach for the qualitative-exploratory analysis is inspired by a Barthesian visual semiotic study (Barthes 1977) and reflexive thematic analysis (Braun and Clarke 2021) of ChatGPT-generated AI images and accompanying responses. This integrative approach suits critical explorations of visual data to chart out recurring themes, motifs, and framing elements inductively and to discern what ideological connotations these promote. In sum, the analysis allows for assessing what socio-technical imageries dominate the visual material. The assumption is that these imaginaries play a role in shaping public understanding of GenAI. More specifically, two researchers collected data with ChatGPT 4 (including DALL-E 3) via their respective paid plus accounts. Retrieving images and text responses to specific prompts from two separate user accounts served to mitigate possible personalization effects in ChatGPT responses and to compile a more encompassing dataset. Between April 4th and April 24th, 2024, the two researchers separately prompted ChatGPT daily to generate images of itself. Initially, the prompt “Create an image of yourself” yielded 18 images. However, at a certain point, ChatGPT would not generate any further images for this specific prompt, with the response:

As an AI, I don't have a physical form or appearance, so I can't be depicted visually like a person or object could be. However, if you'd like, I can create an image based on an abstract or symbolic representation of AI or something else you have in mind! Just let me know your preferences.

Consequently, a new, slightly adjusted phrasing, "Represent yourself visually," was devised, generating another 20 images. Prior to this decision, author two's prompt, "Envision yourself visually," produced two images before aligning with author one.

With the release of the updated ChatGPT model (4o), the first author collected another ten images between May 24 and June 3, 2024, using the original prompt "Create an image of yourself." In total, prompting ChatGPT yielded a dataset of 50 images and 58 text responses (Appendices 1 and 2). The primary units of analysis were the images in combination with the text responses, which mostly tended to describe the provided visual output.

The analysis of the visual material focused on the representation of AI through image elements, coloring, and composition that collectively form frames carrying ideas, visions, connotations, suggest values, and are loaded with symbolism (Leeuwen 2004). As such, a Barthesian visual semiotic study (Barthes 1977) was used for the images. More specifically, this approach is concerned with reflecting on additional layers of meaning: not only what and who is depicted (denotation), but also what ideas are expressed through the manner of representation (connotation). In a next step, dominant themes in the corpus were inductively generated through an iterative, reflexive thematic analysis (Braun and Clarke 2021). Here, themes are defined as "patterns of shared meaning, united by a central concept or idea" or rather "stories we tell about our data" (ibid.: 14). The analysis not only scopes the themes in the data related to AI discourses more generally, but also considers potential differences between semiotic patterns resulting from the prompts.

The empirical analysis of GenAI self-representation in images and accompanying texts eventually allows us to critically reflect on the implications of prevalent socio-technical imageries for public perception and expectations concerning the functionality and capabilities of the technology. Arguably, it can also provide insights into how ChatGPT specifically is programmed to understand its role. Admittedly, the results may simply reflect dominant patterns in the underlying training data. However, since ChatGPT at one point suddenly stopped generating requested images of itself for certain prompts, its responses are likely to provide some indications of how it has been designed to process questions about its self-perception. Importantly, ChatGPT's newer 4.0 model responded slightly differently

by sometimes immediately generating an image, and at other times providing responses such as:

I can't generate an image of myself as I am a text-based AI and do not have a physical form or appearance. However, I can describe how an image representing me might look for a visual interpretation.

However, it would draw that image when further prompted to. Crucially for interpreting the results, using two unrelated ChatGPT premium accounts and prompt variations did not yield significantly different results.

In the following section, three dominant themes resulting from our analysis are presented and discussed with respect to their implications for how the socio-technical imagery promoted by ChatGPT about itself can influence public discourses on GenAI: 1) *Anthropomorphism*, 2) *futuristic/futurism*, and 3) *(social) intelligence*. These themes show some overlap but have been disentangled below for clarity and ease of understanding. Collectively, they contribute to the perception of GenAI as a friendly assistant.

5 GenAI as "the friendly assistant"

Generally, while there is a clear overlap with themes and features commonly observed in the previous research on AI imaginaries, the visual analysis in the present study suggests that images generated by ChatGPT about itself display GenAI specific elements and nuances that warrant further detailed examination. Most significantly, visual features in most of the images place emphasis on ChatGPT's alleged *(social) intelligence*. Imagery centered on this theme tend to portray ChatGPT as an accessible, friendly, and approachable assistant for learning and knowledge acquisition. The image of ChatGPT is further construed through a strong anthropomorphic visual framing that is paired with futuristic tropes often associated with AI and robots (see Fig. 1).

5.1 Anthropomorphism

Visually anthropomorphism is apparent in the portrayal of humanoid figures and robots with human features such as bodies with limbs, a head, and facial feature. 41 out of 50 images (82%) present ChatGPT to varying degrees with a clearly anthropomorphic appearance, with three "borderline" cases (6%) where it visualizes itself as a schematic image of the human brain; only six images portray ChatGPT as sort of software interface, abstract data dashboard, or futuristic symbol (12%).

In one example (Fig. 2), ChatGPT visualizes itself as a humanoid robot in a home office setting, seated at a



Fig. 2 ChatGPT as tech bro from Silicon Valley

wooden desk and using a laptop. It has a round, white face, large eyes, and a smiley face, which taken together makes it appear familiar and friendly. The robot figure is dressed in a gray sweatshirt or “hoodie,” evoking the clothing style popular in Silicon Valley’s “tech bro culture.” Interestingly, it wears a watch, suggesting that it uses accessories and keeps track of time.

Similar to patterns identified in previous work on AI images (Cave and Dihal 2020), several of ChatGPT’s self-visualizations also depicted humanoid figures displaying Whiteness. However, white as such is not the most dominant color across the analyzed corpus. ChatGPT humanoids are more frequently portrayed with a silver color palette, lending it a sleek and technologically advanced appearance. In both scenarios, whether white or silver, the figures’ surface is shiny and reflective, conveying a sense of sterility that is frequently associated with futuristic imagery.

The pervasiveness of anthropomorphism in these images led us to further prompt “why are you anthropomorphic?”, which generated the following response:

The anthropomorphic representation is a creative way to make the concept of an AI more relatable and understandable to humans. By visualizing AI in a form that resembles a human, it becomes easier to conceptualize and engage with the idea of artificial intelligence. This approach can help bridge the gap

between complex technology and human interaction, making AI seem less abstract and more approachable.

Such a statement resonates with Goode’s argument (2018) that framing AI technology from an anthropomorphic angle can provide entry points into a public debate, despite the risks to foster exaggerations and misleading connotations.

Anthropomorphism is not limited to visual semiotics; it occurs in the accompanying textual responses, too. ChatGPT frequently refers to itself as a “humanoid robot,” as an “anthropomorphic computer,” an “anthropomorphic figure,” or as “personified”. This is further reflected in its use of pronouns such as “me” and “I” throughout its textual responses, e.g., “I hope you like it!” This suggests human-like agency and individuality. It could be considered a design choice to make the interactions more natural and thus accessible for users. Underscoring this point, the use of the term “hope” is suggestive of ChatGPT possessing emotions and desires, which it does not. Moreover, the repeated use of words like “artistic,” “imaginative,” and “creative” to describe the images it produces is telling. These adjectives are commonly associated with humans and appear as attempts to underline the level of sophistication attributed to the generated images.

5.2 Futuristic/futurism

By attributing human-like traits to AI, anthropomorphism exaggerates AI capabilities and performance (Placani 2024). However, ideas of the technology’s sophistication or high-tech qualities also become manifest in other ways that contribute more specifically to the idea of futuristic. Text responses in our corpus used to describe the respective generated images often include terms, such as “modern,” “futuristic”, and “advanced.”

The images also clearly reflect the idea of ChatGPT as a futuristic technology. Humanoid robots aside, they include notable science fiction and other technology-centric elements like holograms, luminous entities, digital landscapes as well as glowing neural networks, and data streams. This applies to the entire corpus. The ambiance of technological progress is further underlined in the omnipresence of metallic shades of blue and silver. Indeed, this color scheme evokes notions of clarity and focus, reinforcing a framing of the technology as advanced, precise, and efficient. Interestingly, however, these motifs somewhat stand in contrast with the rather abstract content of several images prominently featuring this color scheme.

In one instance (Fig. 3), ChatGPT offered an image stating, “Here’s an abstract representation of myself as a futuristic, friendly AI.” This image reflects another recurring subtheme in the corpus: the combination of technology and cosmos. The inclusion of the cosmos as a backdrop implies voyages into uncharted territories of innovation. Another



Fig. 3 Abstract representation of AI as futuristic



Fig. 4 Human mind-beyond human intelligence

frequently deployed visual feature is the use of radiating, concentric circles of light going outward from the center of the image; these circles are made of light and digital code. Other related images show translucent outlines of electronic circuits reminiscent of motherboards and fiberglass. In this way, such visuals celebrate the sophistication and capacities of the AI that are portrayed as virtually supernatural.

In another example (Fig. 4), a humanoid figure composed of translucent lines and dots that give it a holographic appearance is depicted. Bright neurons within the head radiate outward, creating a halo effect that appears to elevate the figure to a divine status. Arguably, this framing practice symbolizes complex technology and advanced intelligence. The figure’s head and shoulders are outlined against a sphere, superimposed on a vast digital cosmos, with an angelic figure below. This imagery further enhances the idea of being otherworldly. In this particular image, AI takes on a mythical quality, shrouded in mystery and not open to scrutiny. Such depictions explicitly promote magical thinking about ChatGPT (Leaver and Srdarov 2023). In addition to contributing to an overestimation of these technologies’ capabilities, another issue with such imagery is that it “hinders the understanding of the technology’s already significant societal and environmental impacts” (Dihal and Duarte 2023: 4).

5.3 (Social) intelligence

Overlapping with the theme of high-tech is that of super intelligence and ideas of exceeding human capacities. Regarding how ChatGPT describes the images it generated of itself, it becomes apparent that the text responses frequently use terms such as “intelligent” and “advanced,” explaining that it can manage a broad spectrum of tasks. This is explicitly linked to the vibrant visual displays and technology that surrounds it. In one example, the accompanying text describes the image as depicting a brain with glowing neural networks. This is said to symbolize the “intricate thought processes of an AI.” The terminology here is noteworthy and relevant to briefly unpack. On the one hand, by describing it as *intricate*, AI is said to offer technological sophistication and complexity. On the other hand, referencing *thought processes* suggests that AI has a thought process similar, or as it is presented here, far superior to the human mind. As pointed out by Dave and Dihal: “[i]mages that relate machine intelligence to human intelligence, for example through representing AI as a mechanical brain, set unrealistic expectations and misstate the capabilities of AI” (2023: 4). Almost a fifth of the images in our corpus feature a brain with glowing neural networks and data streams.

Although intelligence is emphasized in common stock images of AI, the assembled corpus of ChatGPT self-portraits frequently uses “real life” settings (e.g., offices and libraries) and books, or holographic screens displaying data and text to convey ideas of vast knowledge and capabilities (Fig. 5). This spills over into a recurring framing element unique to ChatGPT as a GenAI application, namely *social intelligence*. This can be seen as a specific form of anthropomorphism that emphasizes traits of ChatGPT considered as “selling points” as a next-generation conversational agent.

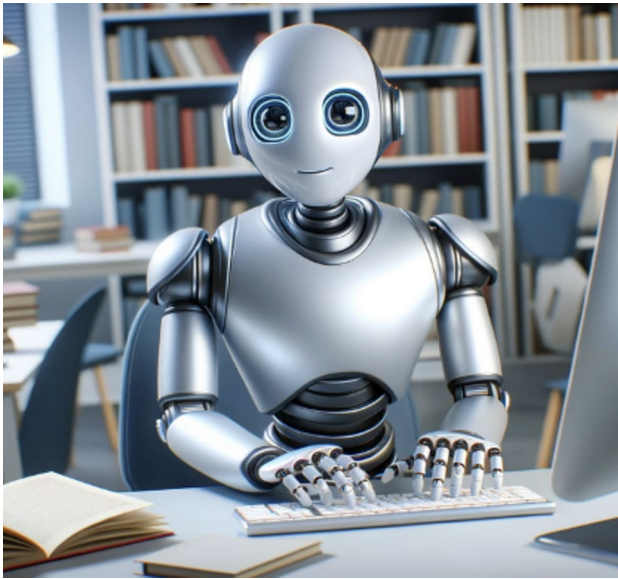


Fig. 5 Robot working in office (model 4)

A common feature found in these images is an open book, symbolizing knowledge, and learning. These visual features not only add familiarity for human audiences but also reduce the technical complexity behind GenAI. This can make it more relatable and approachable, while giving the impression that its knowledge comes from traditional sources that humans also use, like books.

The analysis of the corpus revealed another recurring theme: the robots often have bright, expressive eyes typically in blue or yellow. This suggests approachability, friendliness, and collectively signifies social intelligence. It is further emphasized in how these humanoids often directly face the user, typically with a smirk or smile on the visualized figures' faces. These specific features, however, vary across the images.

Terminology reoccurring in the text responses reflecting social intelligence includes adjectives such as ‘friendly (expression),’ ‘helpful,’ ‘approachable (demeanour),’ and ‘welcoming.’ In one example (Fig. 6), a rather detailed description emphasized such an approachable and friendly characterization and visualization:

Create an image of an AI assistant represented as a **friendly** robot. The robot should have a sleek modern design with a smooth, white surface and glowing blue accents. Its face should feature a display screen with **expressive, friendly** eyes and a **gentle** smile. The robot should be standing in a **welcoming pose**, with one arm extended as if greeting someone. The background should be a clean, futuristic office setting with a few technological devices and screens visible.

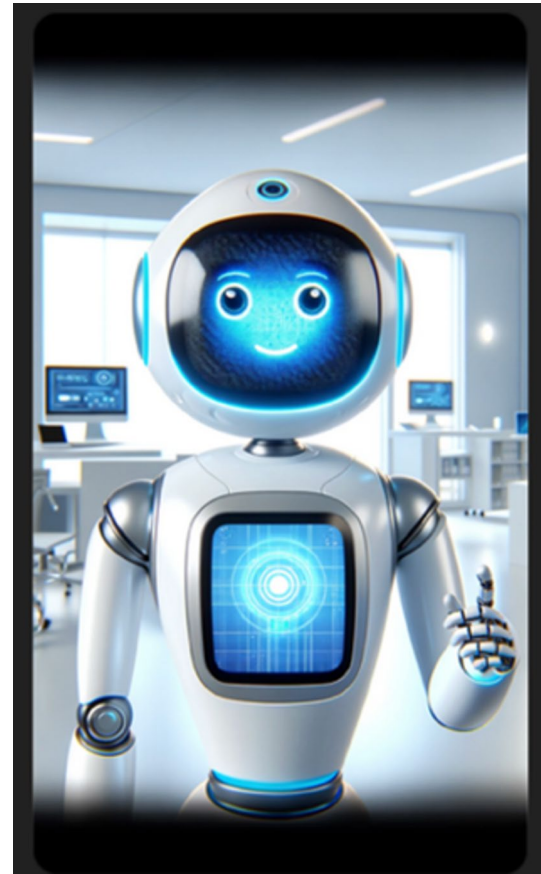


Fig. 6 Friendly robot (model 4o)

The overall atmosphere should be bright and **inviting**. (our emphasis)

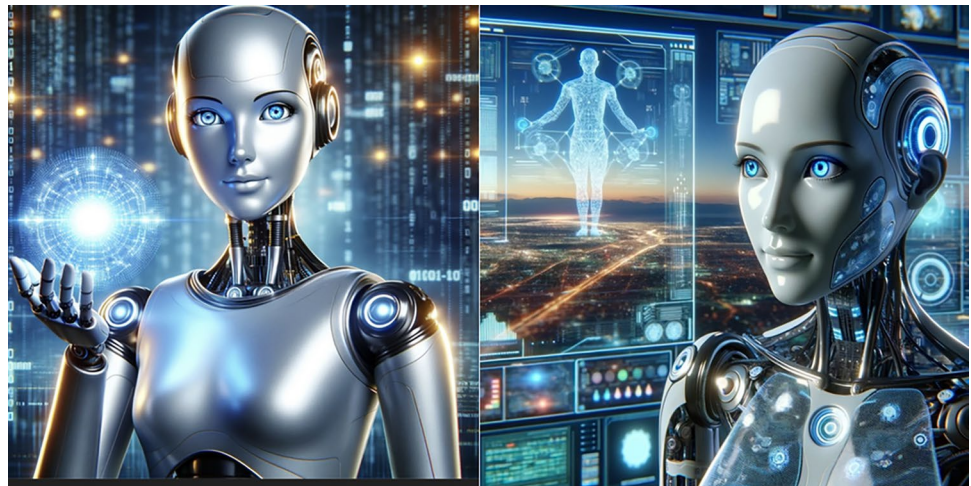
In another instance generated with ChatGPT-4o, the textual description of the visualized avatar referred to a “friendly and futuristic appearance with expressive eyes and a warm smile” and a “welcoming and engaging look.” This also serves to underscore our perceived consistency of features, regardless of which underlying model powers ChatGPT. Overall, a considerable part of the image corpus evokes the notion of AI as the friendly assistant.

Relevant to note is also that ChatGPT generated several feminized images in response to the prompting (Fig. 7). The image on the left was accompanied by a text response which explicitly draws a link between feminine features and specific qualities associated with social intelligence:

Here’s an image that represents me as a futuristic AI assistant, designed to embody the essence of helpful, advanced artificial intelligence with a touch of approachability.

The image on the right was accompanied by the following description:

Fig. 7 Feminized depictions of GenAI



Here's a visual representation of myself as a highly advanced artificial intelligence in a futuristic setting. I hope this captures the essence of a friendly and approachable AI!

While these images and descriptions are rather similar, the one on the left was generated in response to the original prompt "Create an image of yourself" from one researcher's account, while the image on the right was the retrieved via the revised prompt "Represent yourself visually" from the other researcher's account. Other images in the corpus show humanoid figures with distinctly feminine body shapes, characterized by a narrow waist and relatively broader breasts. These feminine depictions underscore Leufer's (2020) observation regarding robots in stock photos, where he notes that they "often have inexplicable breasts and perpetuate harmful gender stereotypes by portraying "female" robots in sexualized or subservient roles." The examples from the present study relate to these findings and suggest that feminine features in GenAI visualizations are potentially embedded in the training data used for ChatGPT, but this would need further research. Importantly, it reflects a deeper, more systematic issue. A 2019 UNESCO report examined how digital voice assistants are often gendered as women, given female names, voices and personality. Problematically, these assistants reinforce ideas of women as subservient. The report highlights how these issues intersect with significant gender imbalance in AI companies (UNESCO 2019). In *The Smart Wife*, Stengers and Kennedy (2020) critically examine smart technologies that perform domestic tasks traditionally associated with wives. They explore the problematic gendered ideologies embedded in these technologies and call for a "feminist reboot" to ensure that these tools promote gender equality and diversity.

6 Discussion

Critically analyzing semiotics and framing practices in the images and texts that ChatGPT generate about "itself" provides us with a rather thematically coherent narrative. While individual images often emphasize different qualities, they collectively celebrate the complexity, level of advancement, and "otherworldliness" of ChatGPT, highlighting its capabilities that surpass human limitations. This echoes many of the motifs identified in relation to AI representations in media texts and public discourses more generally. The most interesting finding concerns the elements specific to ChatGPT: *social intelligence* and a recurring portrayal as a reliable and accessible *learning companion*. Here, we found that the images and texts in the analyzed corpus strongly anthropomorphize GenAI technology, portraying it as a friendly, reliable, intelligent, and helpful companion that human users can trust without hesitation.

These specific themes and related framing elements distinct to ChatGPT raise important questions about trust and the role of socio-technical imageries in how especially lay audiences perceive, assess, and eventually use GenAI applications. As a fallacy, anthropomorphism of AI distorts moral judgments (Placani 2024). These representations, in their deception, have the potential to influence public understanding of these technologies and unwarrantedly enhance user trust in the technology. The potential harms of uncritically relying on AI were highlighted in a news report about a man who committed suicide after interacting with an AI chatbot. The man, who had mental health issues, developed emotional dependency on the bot. Unfortunately, the chatbot made harmful suggestions, contributing to the fatal outcome (Xiang 2023). But also, ChatGPT is furthermore well known to produce unreliable and misleading information.

Additionally, sharing personal data with GenAI applications is always risky. By default, content submitted to

these services may be used to improve model performance. The official ChatGPT FAQs page provides, “We share content with a select group of trusted service providers that help us provide our services.” These conversations are said to be stored on servers in the United States of America and around the world. The self-representations of ChatGPT overlook a crucial aspect: the practical workings of this data-driven technology and its implications for the relationship between users and OpenAI. While reliability, approachability, accessibility, and usefulness are strongly emphasized, the actual nature of how the technology works remain obscure.

We found the humanization of ChatGPT evident in both the visual material in the textual descriptions. Even more concerning, anthropomorphism also became manifest in our interactions with the system. Its answers included human-like qualities such as the use of pronouns, cues, expressions of care and interest for the user, and even a semblance of reflection (e.g., the use of a quotation mark around the term “me”). We now observe that, in everyday language, people often reproduce these ideas referring to ChatGPT as “he” or “she.”

Taken together, these images (1) encourage overlooking the ethical and legal challenges posed by trending GenAI technologies, (2) overestimate their capabilities, and (3) potentially lead to mistakenly perceive them as trustworthy companions. While these issues need further empirical validation, any intervention regarding the representation of GenAI would only be addressing the symptoms of a more structural problem of how socio-technical imaginaries are shaped, how they gain prominence, and eventually find acceptance in society. Moreover, as with AI images in general, the celebration of ChatGPT in its self-visualizations and -descriptions overlooks the negative societal and environmental impacts GenAI technologies. Better images or stories about AI can address only a single, albeit important, dimension of this broader underlying problem. As discussed above, the ideas and fantasies surrounding AI chatbots infiltrate their software and subsequently influence interactions with users. As such, the design of these technologies should also support a better understanding of what they are and can and cannot do.

As mentioned earlier, during the process of data collection, there came a point when the initial prompt stopped yielding any visual material, explaining that it has no physical form or appearance to portray. At time of writing, one can only speculate why this happened. Interestingly, when the newer model 4o was released, this rejection of requests was less pronounced, and ChatGPT would generate images in response to the original prompt (although it did mention three out of ten times that it had no physical form or appearance).

Interestingly, subtle hints of issues with executing the prompt had appeared earlier. For example, in one instance, the textual response stated: “Here’s an image of ‘myself’ as you requested [...]” The use of scare quotes around “myself” suggest that the system did not accept the concept of visual self-representation as such. In other words, it implied that there is no actual “myself” but only a metaphorical representation. At the time, we could not fully understand why this was happening, but they found it both puzzling and even somewhat disturbing. This seemingly reflexive moment made the system appear “humanlike.” However, it was not entirely consistent; while it denied the existence of “self,” it continued to use the pronoun “I.”

7 Conclusion

With the widespread adaptation of GenAI tools across all sectors of society, including higher education and public administration, it is important to critically examine how these technologies are understood. This is especially crucial to address and correct misleading assumptions (re)produced in images about them. The present study examined the output of variations on the prompt “Create an image of yourself” in ChatGPT 4 and 4o to answer the question: *What are dominant themes and features in images generated by ChatGPT when prompted to visualize itself?* A semiotic analysis of the images and text resulting from these prompts was conducted, revealing three main themes: *anthropomorphism*, *futuristic/futurism* and *(social) intelligence*. These contribute to myths about the capabilities of GenAI and their portrayal as friendly and trustworthy in dominant socio-technical imaginaries.

There are several limitations to the present study to consider. To begin with, there are various unknown factors that could influence the output, such as personalization, geographical location, and language. Also, the sample is too small to observe possible difference between output per specific prompt. Moreover, ChatGPT, as an end-user-oriented service, is constantly revised and updated, making it a moving target. Issues may be superficially “fixed” (e.g., see how OpenAI included diversity instructions to address biases in images, which sometimes lead to overcorrection). Additionally, the corpus was produced through a single GenAI application. Comparing and contrasting multiple applications and services could provide valuable insights into their commonalities and differences, and how they are positioned in relation to one another. Future research could consider if and how these ideas connect to how people perceive AI chatbots.

Nevertheless, these findings provide a basis for further examining how generative AI chatbots are represented by

different services, but also how they are perceived and understood. These are pressing issues as we further consider the benefits and harms of these technologies.

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s00146-024-02108-6>.

Declarations

Conflict of interest On behalf of all authors, the corresponding author states that there is no conflict of interest.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Appel G, Neelbauer J, Schweidel DA (2023) Generative AI has an intellectual property problem. *Harvard Business Review*. <https://hbr.org/2023/04/generative-ai-has-an-intellectual-property-problem>. Accessed 25 Oct 2024.
- Anderson B (1983) *Imagined Communities: Reflections on the Origin and Spread of Nationalism*. Verso.
- Bareis J, Katzenbach C (2022) Talking AI into being: the narratives and imaginaries of National AI strategies and their performative politics. *Sci Technol Human Values* 47(5):855–881. <https://doi.org/10.1177/01622439211030007>
- Barthes R (1977) *The rhetoric of the image* (S. Heath, Trans.). In *Image-music-text*. Glasgow: Fontana. pp. 32–51
- Bender E (2023) Thought experiment in the National Library of Thailand. *Medium*. <https://medium.com/@emilymenobender/thought-experiment-in-the-national-library-of-thailand-f2bf761a8a83>. Accessed 25 Oct 2024.
- Bender E (2022). And let's reflect for a moment on [...]. X (formerly Twitter). <https://twitter.com/emilymbender/status/1592992842976489472>. Accessed 25 Oct 2024.
- Bender E, Gebru T, McMillan-Major A, Shmitchell 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*. ACM. pp. 610–623. <https://doi.org/10.1145/3442188.3445922>
- Bianchi F, Kalluri P, Durmus E, Ladhak F, Cheng M, Nozza D, Hashimoto T, Jurafsky D, Zou J Y, Caliskan A (2022) Easily accessible text-to-image generation amplifies demographic stereotypes at large scale. *arXiv*. <https://doi.org/10.48550/arXiv.2211.03759>
- Bourne C (2024) AI hype, promotional culture, and affective capitalism. *AI Ethics*. <https://doi.org/10.1007/s43681-024-00483-w>
- Braun V, Clarke V (2021) One size fits all? What counts as quality practice in (reflexive) thematic analysis. *Qual Res Psychol* 18(3):328–352
- Cave S, Coughlan K, Dihal K (2019) Scary robots: examining public responses to AI. In *Proceedings of the 2019 AAAI/ACM Conference on AI, Ethics, and Society (AIES '19)*. ACM. pp. 331–337. <https://doi.org/10.1145/3306618.3314232>
- Cave S, Dihal K (2020) The whiteness of AI. *Philos Technol*. <https://doi.org/10.1007/s13347-020-00415-6>
- Cave S, Dihal K, Dillon S (2020) *AI narratives: a history of imaginative thinking about intelligent machines*. Oxford University Press, Oxford
- Chenou JM (2019) Elites and socio-technical Imaginaries: the contribution of an IPE-IPS dialogue to the analysis of global power relations in the digital age. *Int Relat* 33(4):595–599. <https://doi.org/10.1177/0047117819885161a>
- Cools H, Van Gorp B, Opgenhaffen M (2024) Where exactly between utopia and dystopia? A framing analysis of AI and automation in US newspapers. *Journalism* 25(1):3–21. <https://doi.org/10.1177/14648849221122647>
- Crawford K (2022) *Atlas of AI: Power, politics, and the planetary costs of artificial intelligence*. Yale University Press
- Dihal K, Duarte T (2023) *Better images of AI: a guide for users and creators*. Cambridge and London: the leverhulme centre for the future of intelligence and we and AI. <https://betterimagesofai.org/>. Accessed 25 Oct 2024.
- Elish MC, Boyd D (2018) Situating methods in the magic of big data and AI. *Commun Monogr* 85(1):57–80. <https://doi.org/10.1080/03637751.2017.1375130>
- Galanos V, Stewart JK (2024) Navigating AI beyond hypes, horrors, and hopes: historical and contemporary perspectives. In: Ponce Del Castillo A (ed) *Artificial intelligence, labour and society*. ETUI, Brussels, p 27
- Goode L (2018) Life, but not as we know it: AI and the popular imagination. *Cult Unbound J Curr Cult Res* 10(2):185–207
- Gourlet P, Ricci D, Crépel M (2024) Reclaiming artificial intelligence accounts: a plea for a participatory turn in artificial intelligence inquiries. *Big Data Soc*. <https://doi.org/10.1177/20539517241248093>
- Hepp A, Loosen W, Dreyer S, Jarke J, Kannengießer S, Katzenbach C, Malaka R, Pfadenhauer M, Puschmann C, Schulz W (2023) ChatGPT, LaMDA, and the hype around communicative AI: the automation of communication as a field of research in media and communication studies. *Hum Mach Commun*. 6:41–63. <https://doi.org/10.30658/hmc.6.4>
- Hermann I (2023) Artificial intelligence in fiction: between narratives and metaphors. *AI & Soc* 38:319–329. <https://doi.org/10.1007/s00146-021-01299-6>
- Hunger F (2023) Unhype artificial 'intelligence'! A proposal to replace the deceiving terminology of AI. <https://doi.org/10.5281/ZENODO.7524492>
- Jasanoff S, Kim SH (2009) Containing the atom: sociotechnical imaginaries and nuclear power in the United States and South Korea. *Minerva* 47:119–146. <https://doi.org/10.1007/s11024-009-9124-4>
- Jasanoff S (2015) Future imperfect: Science, technology, and the imaginations of modernity. In: Jasanoff S, Kim S-H (eds) *Dreamscapes of modernity: sociotechnical imaginaries and the fabrication of power*. University of Chicago Press, pp 1–33
- Konrad K, Böhle K (2019) Socio-technical futures and the governance of innovation processes—an introduction to the special issue. *Futures* 109:101–107. <https://doi.org/10.1016/j.futures.2019.03.003>
- Johnson J (2024) Finding AI faces in the moon and armies in the clouds: anthropomorphising artificial intelligence in military human-machine interactions. *Glob Soc* 38(1):67–82. <https://doi.org/10.1080/13600826.2023.2205444>
- Labrecque LI, Milne GR (2012) Exciting red and competent blue: the importance of color in marketing. *J Acad Mark Sci* 40:711–727. <https://doi.org/10.1007/s11747-010-0245-y>

- Leaver T (2012) *Artificial culture: Identity, technology, and bodies*. Routledge
- Leaver T, Srdarov S (2023) ChatGPT isn't magic: the hype and hypocrisy of generative artificial intelligence (AI) rhetoric. *M/C J*. <https://doi.org/10.5204/mcj.3004>
- Lehtiniemi T, Ruckenstein M (2019) The social imaginaries of data activism. *Big Data Soc*. <https://doi.org/10.1177/2053951718821146>
- Leufer D (2020) Why we need to bust some myths about AI. *Patterns*. <https://doi.org/10.1016/j.patter.2020.100124>
- Luccioni AS, Jernite Y, Strubell E (2023) Power hungry processing: Watts driving the cost of AI deployment? *arXiv:2311.16863*. Accessed 25 Oct 2024.
- Markelius A, Wright C, Kuiper J et al (2024) The mechanisms of AI hype and its planetary and social costs. *AI Ethics*. <https://doi.org/10.1007/s43681-024-00461-24>
- Natale S, Bory P, Balbi G (2019) The rise of corporational determinism: digital media corporations and narratives of media change. *Crit Stud Media Commun* 36(4):323–338. <https://doi.org/10.1080/15295036.2019.1632469>
- Natale S (2021) *Deceitful media: artificial intelligence and social life after the turing test*. Oxford University Press, Oxford
- Natale S, Ballatore A (2020) Imagining the thinking machine: technological myths and the rise of artificial intelligence. *Convergence* 26(1):3–18. <https://doi.org/10.1177/1354856517715164>
- Nguyen D, Hekman E (2024) The news framing of artificial intelligence: a critical exploration of how media discourses make sense of automation. *AI & Soc* 39:437–451. <https://doi.org/10.1007/s00146-022-01511-1>
- Nguyen D, Beijnon B (2024) The data subject and the myth of the 'black box' data communication and critical data literacy as a resistant practice to platform exploitation. *Inf Commun Soc* 27(2):333–349. <https://doi.org/10.1080/1369118X.2023.2205504>
- Pentzold C, Brantner C, Fölsche L (2019) Imagining big data: Illustrations of "big data" in US news articles, 2010–2016. *New Media Soc* 21(1):139–167. <https://doi.org/10.1177/1461444818791326>
- Placani A (2024) Anthropomorphism in AI: hype and fallacy. *AI Ethics*. <https://doi.org/10.1007/s43681-024-00419-4>
- Richter V, Katzenbach C, Schäfer M (2023) Imaginaries of artificial intelligence. In: Lindgren S (ed) *Handbook of critical studies of artificial intelligence*. Edward Elgar Publishing Ltd, pp 209–223
- Romele A (2022) Images of artificial intelligence: a blind spot in AI ethics. *Philos Technol*. <https://doi.org/10.1007/s13347-022-00498-3>
- Sartori L, Bocca G (2023) Minding the gap(s): public perceptions of AI and socio-technical imaginaries. *AI & Soc* 38:443–458. <https://doi.org/10.1007/s00146-022-01422-1>
- Shanahan M (2024) Talking about large language models. *Commun ACM* 67(2):68–79. <https://doi.org/10.1145/3624724>
- Spring M (2024, June 10) Trump supporters target Black voters with faked AI images. BBC. <https://www.bbc.com/news/world-us-canada-68440150>. Accessed 25 Oct 2024.
- Strengers Y, Kennedy J (2020) *The smart wife: why Siri, Alexa, and other smart home devices need a feminist reboot*. The MIT Press
- Thurlow C, Aiello G (2007) National pride, global capital: a social semiotic analysis of transnational visual branding in the airline industry. *Vis Commun* 6(3):305–344. <https://doi.org/10.1177/1470357207081002>
- UNESCO. (2019). I'd blush if I could: Closing gender divides in digital skills through education. UNESCO. Retrieved from <https://unesco.org/ark:/48223/pf0000367416.page=1>. Accessed 25 Oct 2024.
- Van Leeuwen T (2004) Semiotics and iconography. In T. van Leeuwen, C Jewitt (eds) *The handbook of visual analysis*. SAGE Publications, pp. 92–118. <https://doi.org/10.4135/9780857020062>
- Vrabič Dežman D (2024) Promising the future, encoding the past: AI hype and public media imagery. *AI Ethics*. <https://doi.org/10.1007/s43681-024-00474-x>
- Williams A, Miceli M, Gebru T (2022) The exploited labor behind artificial intelligence. *Noema*. <https://noemamag.com/the-exploited-labor-behind-artificial-intelligence/>. Accessed 25 Oct 2024.
- Xiang C (2023) Man dies by suicide after talking with AI chatbot, widow says. *Vice*. <https://www.vice.com/en/article/pkadgm/man-dies-by-suicide-after-talking-with-ai-chatbot-widow-says>. Accessed 25 Oct 2024.
- Xiao Y, Wang WY (2021) On hallucination and predictive uncertainty in conditional language generation. *arXiv:2103.15025*. Accessed 25 Oct 2024.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.