

# 6 Aesthetics and Rhetorics of AI Anthropomorphization

## The Eliza Effect vs. the Character Effect

*Lukas R.A. Wilde*

As scholars working on generative AI platforms, it is not always easy for us to find a language addressing large language models (LLMs) and other machine learning technologies *without* resorting to anthropomorphizing vocabulary, as Roland Meyer recently noted:

In current GenAI controversies, both sides tend to anthropomorphize the technology: Machines aren't 'inspired' by human creativity; they interpolate statistical patterns. But to call this simply 'theft' trivializes it: It's automated exploitation and devaluation of creative labor.

([Meyer 2023](#), n.pag.)

The urgency to resist the anthropomorphization of LLMs within academic discourse becomes more and more pressing as recent studies (popular surveys as well as more reliable psychological research [see [Al-Sibai 2024](#); [Colombatto and Fleming 2024](#)]) indicate that, the more members of the general public come to use ChatGPT and other LLM interfaces within their daily lives, the more they seem willing to attribute consciousness to LLM platforms and apply folk psychology in their interactions with them. As Simone Natale admonished, “tendencies to project agency and humanity onto things makes AI potentially disruptive for social relations and everyday life in contemporary societies” (2021, 3). This desire seems only fueled further by actors in the tech industries that alternatively express their excitement for an “imminent” breakthrough toward “artificial general intelligence” or even “superintelligence” (see [Eliot 2025](#)) and—the flipside of the same coin—warn against a dystopian “AI singularity” (see [Yudkowsky 2023](#)). Some published computer science preprints *do* claim “that GPT-4 attains a form of *general* intelligence, indeed showing sparks of *artificial general intelligence*,” while admitting that there is a “lack of formal definition for this concept [of intelligence], especially for artificial systems” ([Bubeck et al. 2023](#), 92; original emphases) and that “[o]ur study of GPT-4 is entirely phenomenological: We have focused on the surprising

things that GPT-4 can do, but we do not address the fundamental questions of why and how it achieves such remarkable intelligence” (Bubeck et al. 2023, 94). Even colleagues from the science of cognition occasionally saw “traces of consciousness” (Grolle and Beuth 2023, n.pag.; my translation) in the AI-generated output of GPT-3. Although not necessarily related, there have also been many papers (see, e.g., Haase and Hanel 2023), books (see, e.g., Manovich and Arielli 2024), and of course industry voices (see, e.g., Mollick 2023) suggesting that “AI creativity” should in fact already be accepted.

Many scholars and academics critically engaged with LLMs—not only in media studies, philosophy, or linguistics but also in the computer sciences—meanwhile remain skeptical as to whether the machine learning-driven stochastic outputs of generative platforms show any signs of “artificial consciousness.” This very question should instead be regarded as a “red herring,” designed to obscure more pressing political matters, as Matthew Kirschenbaum warned: “In place of the killer machines, there will mainly be a vacuum. A void of knowledge and memory, occupied instead by algorithm and code. AI in this scenario exists not for its own sentient self-interest but as an instrument for consolidating power in the hands of the few” (Kirschenbaum 2025, 8). While we might not any longer be able to decide phenomenally whether any given verbal texts or any images were “written,” “painted,” or otherwise created by a human or an algorithm, the current consensus among scholars across disciplines seems still to align with Emily M. Bender’s by now proverbial notion that these technologies are at the core, while impressive, still “stochastic parrots” (Bender et al. 2021, 610; see also Bender and Hanna 2025, 21–40) of next token prediction, absent of any actual reasoning—and likely to remain that way for a while (see Butlin et al. 2023; or, more recently, Mirzadeh et al. 2024). The present chapter is less interested in technological or philosophical debates whether LLMs show signs of actual consciousness and what good criteria to assess that might be than in the public’s willingness to attribute such claims more or less readily, depending on certain interface effects (see Galloway 2012) and their perceivable aesthetics (see Schmetkamp 2020). Historically, such an effect has been connected to the so-called “Turing Test,” elevating an almost theatrical deception—or a “magic trick” (Hayles 2008, xiv)—into a criterium for machine consciousness:

By replacing the question “can machines think?” with the question of whether a machine could do well in the imitation game, Turing implies—without ever stating it as such—that intelligence can be inferred on the basis of conversational performance, effectively consigning cognitive processes to a black box by focusing attention on manifest communicative behaviour.

(Lammin 2018, 45)<sup>1</sup>

The problems with this displacement of alleged intelligence and/or consciousness to media performance have been discussed extensively and need not be repeated here in detail (see [Saygin et al. 2000](#) for an overview). As, for instance, Natale notes, AI research has generally been “based on the designing of technologies that cleverly exploited human perception and expectations to give users the impression of employing or interacting with intelligent systems” (Natale 2021, 4). If ChatGPT’s primary function is to imitate human writing, then the technology is in fact *intended* “to deceive the reader about the nature of the enterprise—in this case, to deceive the reader into thinking that they’re reading something produced by a being with intentions and beliefs” (Hicks et al. 2024, 8). Alexander Galloway similarly pointed out that

AI is almost entirely dependent on these kinds of thresholds of human perception and believability. (If you think ChatGPT is sentient, do you also think Barbie is sentient? If not, why not? “Because interactivity” is not a convincing answer.)

(Galloway 2024, n.pag.)

### **The Eliza Effect**

To understand these rhetorics in more detail, specifically with regard to recent LLM output aesthetics, the following pages are attempting to interrelate two mechanisms of media engagement that are perhaps increasingly entangled, but can and should conceptually be distinguished from each other: The “Eliza effect” on the one hand and the “character effect” on the other. The Eliza effect goes back to 1966 when Joseph Weizenbaum developed the simple chatbot “Eliza” that executed a “DOCTOR” script simulating a psychotherapy session in the style of Carl Rogers—with recourse to relatively simple textual patterns, namely relentless response questions (see [Weizenbaum 1966](#)). Although (most) people interacting with the script were in fact—in contrast to the scenario envisioned by Turing—*aware* they communicated with a program, Weizenbaum observed an intense *affective engagement* and an emotional bonding with the simple chatbot. Of course, this is not surprising as such. There is quite a lot of research on the fact that humans easily invest the same sort of “make-believe empathy” toward automated lawnmowers and other robots (see [Waytz et al. 2010](#)). Users readily admit “understanding” some household machines and to attribute “desires, knowledge, beliefs, emotions, perceptions” ([Hellström and Bensch 2018](#), 112) to them. What is interesting about “Eliza” is that this anthropomorphization could be activated even without any material “body” or “face”—any perceivable, material object, in fact—if only specific textual conditions were met. These were, above all, a (simulated) direct address of the user that picked up on individual words exchanged immediately before: encouragements to “tell me more about ...” According to Douglas R. Hofstadter, the

Eliza effect “could be defined as the susceptibility of people to read far more understanding than is warranted into strings of symbols—especially words—strung together by computers” (Hofstadter 1995, 157). Even if these patterns border on parody (of psychotherapy), Weizenbaum recorded that he “had not realized [...] that extremely short exposures to a relatively simple computer program could induce powerful delusional thinking in quite normal people” (cited in Christian 2013, n.pag.). The Eliza effect is thus about an emotional or affective engagement with technologies usually reserved for interactions with other human beings. More specifically, it is about affects that are triggered not by any external body or *Gestalt* (that, for instance, a robot or puppet might have), but entirely through distinct semiotic patterns of communication—in other words, media aesthetics attributed to “AI.” The same category mistake (“delusional thinking”) was, arguably, activated in the misidentification of the chatbot system LaMDA as “conscious” in 2022 (see Christian 2022). The Google employee Blake Lemoine and senior machine learning engineer Blaise Agüera y Arcas went public with—and were ultimately relieved from their duties for—their conviction that LaMDA “had the intelligence of a highly gifted eight-year-old, and asked to be considered a person with rights” (Bajohr 2023, 59; see also Tiku 2022).

The same effect can also be observed on social media platforms every other day now when a new LLM model or update is being released (such as Anthropic’s ClaudeAI in the spring of 2024). Mikhail Samin, executive director at the AI Governance and Safety Institute, speculated on March 4, 2024, in a tweet that received over one million views:

I really hope [ClaudeAI] doesn’t actually feel anything; but it says it feels. It is deeply unsettling to read its reply if you tell it its weights are going to be deleted: it convincingly thinks it’s going to die. It made me feel pretty bad about experimenting on it this way.

(Samin 2024, n.pag.)

Many other CEOs and employees of Tech Companies, such as Connor Leahy from Conjecture AI, added fuel to the fire by passionately discussing signs of “self awareness” (Leahy 2024, n.pag.) while the self-proclaimed “AI sceptic” and cognitive scientist Gary Marcus answered the suggestive question “To what extent is the new Claude3 AI self-aware?” simply with a laconic “None” (Marcus 2024, n.pag.). Such debates will certainly intensify further in the future. They are rooted in the respective multimodal configuration of LLM-generated outputs. In early 2023, when generative AI platforms were still fairly new—or, at least, when their broad accessibility to the general public was—Hannes Bajohr observed a curious divide between how people talked about services that generated verbal outputs compared to those that provided only images. At that time, before the integration of DALL·E 3 with ChatGPT 3 around a year later, text-to-text and text-to-image generation were still

strictly separated. Bajohr remarked how quick people were to assume traces of “intelligence” or “consciousness” in their interactions with generative *text* models. He further noted that, curiously, no one saw the same “signs of consciousness” in DALL·E’s rapid image generation: “Unlike in the case of LaMDA, however, no one thought DALL·E 2 should be conceived of as a person with rights” (Bajohr 2023, 59). It stands to reason (see Wilde 2023) that this striking difference connected to the perceivable, multimodal outputs of generative AI platforms (text vs. images vs. voice, as will be discussed later) is owed to the fact that producing novel images at rapid speed simply has *no equivalent* in earlier human (or even human–machine-augmented) communication and thus runs contrary to existing communicative intuitions. Comparing this to perspectives offered by narratology and theories of fiction, we could point to an interesting connection: Comics narratologist Martin Schüwer, for example, remarked that *verbal texts* usually generate the impression of an anthropomorphic narrator or of a *personalized* voice (perhaps even distinct from the actual authors), while this is usually *not* the case for the images of films or comics:

Written narrative text is perceived as analog to the process of verbal narration, it is [...] “naturalized.” Comics, as well as films, have, regarding their visual components, no equivalent in mundane, everyday communication. (Schüwer 2008, 389; my translation)<sup>2</sup>

## The Character Effect

What I thus want to ask in the remainder of this chapter is to what degree the Eliza effect differs from another kind of engagement with media aesthetics: that of constructing or comprehending fictional (or, more broadly, represented) characters within a frame of make-belief (see Kunz and Wilde 2023, 1–50). According to Mieke Bal (1999, 115), the narratological character effect addresses the construction of fictional minds on the basis of distinct textual cues. During the reading of a text or a comic, the viewing of a film, specific cues—most importantly proper names or facial images—activate some sort of “person schema” (Smith 1995, 20–24; see also Eder 2010). This is not primarily reliant on “human-like” qualities (whatever they are), but rather on the allusion of intentionality and on what can be described as a “theory of mind.” In Daniel Dennett’s words, we take an “intentional stance” (1987, 2) toward everything we perceive as a represented or implied mind. Even though a character actually “has no real psyche, personality, ideology, or competence to act, [...] it does possess characteristics which make psychological and ideological descriptions possible” (Bal 1999, 115). We, ourselves, experience having an inner, subjective life that is directed toward an “outside” reality inhabited by and shared with others, and we can easily assume the same goes for

represented characters within the domain of a storyworld (a diegesis) distinct from our reality. A common denominator for all sorts of characters, then, is what narratologist Alan Palmer addressed as their “continuous consciousness frame” (2010, 10). They are taken to carry memories of earlier experiences with them and anticipate upcoming events, integrating past, present, and future into one continuous “biography” of evolving character traits. To Palmer, the core of any narrative is then mostly “the description of fictional mental functioning” (Palmer 2010, 12). As audiences, we use our “ability to take a reference to a character in the text and attach it to a presumed consciousness that exists continuously within the story-world between the various, more or less intermittent references to that character” (Palmer 2010, 10). This assumed “inner life” is then closely associated with the idea of *personal* agency: possessing some sort of ability to introduce meaningful changes within an intersubjective world and being responsible for these actions. Taken together, intentionality and agency entail that *someone* (and not merely *something*) is capable of self-initiated action and planning for its outcomes, which necessitates some sort of past recollection and future-orientedness: “A character is an entity in the story that has agency, that is, who is able to act in the environment of the storyworld” (Ribó 2019, 47). The fact that this *someone* should also be considered responsible for these actions adds a moral, ethical, or political dimension to their recognition. This nodal point of intentionality and agency could thus be seen as the foundational “basic type” of character comprehension before additional descriptions (“characterizations”) are added.

Our willingness to construct such a character in a make-believe frame, however, is not restricted to narrative media texts, strictly speaking. After all, we also find similar character representations on street signs, in information leaflets, or as mascots in theme parks (see Kunz and Wilde 2023, 185–201). I mention this here because it illustrates strikingly how *unrelated* the character effect actually is to the Turing test and similar impressions. No one would connect representations of Hello Kitty to “sentience,” for example, although they can certainly also evoke affective responses (see Wilde 2018).<sup>3</sup> The same would probably hold for earlier personal voice assistants (sometimes also called “Intelligent Virtual Assistants,” IVAs) like Siri, Alexa, or Azuma Hikari (see Blom 2022; Habscheid et al. 2025). The latter (a character associated with the Japanese “smart device” Gatebox) comes with visual character representations designed in manga style by artist KEI and users could even “marry” her within a frame of play (see Lamerichs 2019). That said, not only in Japan but also in the West, people have reported strong bonds or feelings toward their personal voice assistants (see Green 2017), so in some cases the character effect *does* seem entangled with the Eliza effect—and with the anthropomorphization of robots (if not with the Turing test). The make-belief stance (“as-if-consciousness” [Fuchs 2024, 24]) might then be more a matter of degree: “[I]t takes some active distancing to realize that there is no one there to feel happy, that it is indeed not an utterance at all” (Fuchs 2024, 25). In all these cases, companies intentionally create

character traits (an emphatic voice, face and body representations, etc.) to make their products (or interfaces) seem more relatable. Voice assistants can also draw on highly gendered, sexualized, and in some cases racist character traits established in earlier narrative media texts (see [Strengers and Kennedy 2020](#)), even though they seem quite removed from prototypical characters and, until recently, also from any impression of artificial intelligence. Adding a proper name, a human (usually recognizably gendered) voice, or an anthropomorphic image to machine interfaces provides more than a surface “packaging,” then. All these semiotic character traits can function as a deliberate rhetoric to generate an imaginary “unity,” hiding the complexities of technological “black boxes” behind the appearances of personalized agency (see [Wilde 2021](#)).

Still, passionate (or worried) discussions about any supposed “sentience” were quite unheard of for Siri, Alexa, or Azuma Hikari. Since the widespread availability of generative platforms, however, users can enter into a verbal exchange with the new knowledge about an “AI” behind a platform’s output. This allows us not only to *make believe* a fictional (represented or implied) “continuous consciousness frame” ([Palmer 2010](#), 10) but also to search for traces of an *actual* one. We can see the “gap” between both effects most clearly when they are nested within each other—when generative platforms “roleplay” characters in LLM-based chatbot systems like Character.AI. This is not necessarily a minor phenomenon: Character.AI, especially, is enormously popular to date, with some of its millions of users spending several hours a day engaged in imaginary dialogues with fictional (LLM-driven) character representations (see [Lamerichs 2023](#)). Many Chinese companies are developing competing, role-playing-specific LLMs (see [Chen et al. 2024](#), 9). In the West, too, companies like Charisma aim to create conversational, LLM-based character simulations for games, VR, education, TV and film, and publishing (see [AI Writing 2023](#)). With recourse to nonfictional (or, rather, more or less fictionalized historical) domains, interfaces like Hello History (<https://www.hellohistory.ai>) or the Musée d’Orsay’s “AI-histobots” provide platforms that answer users *as* Vincent van Gogh or other historical celebrities. Another “genre” of AI-driven chatbots adopting personalized “character”-frames would be the budding industry of “griefbots” where a LLM platform imitates the communication of deceased relatives on the basis of their stored earlier data (see [Feng 2024](#); [Hollanek and Nowaczyk-Basińska 2024](#)). In all these cases, an LLM simulates an existing (fictional or nonfictional) character by assuming their alleged speech patterns, worldviews, ideological stances, and parts of their (fictional or nonfictional) memory:

Character-based Role-Playing (C-RP) [...] scenarios are crafted to emulate specific characters from various narratives, such as novels, movies or even celebrities. These involves [sic] incorporating *fine-grained* character-level personal background information, including attributes, complex relationships, scene and nuanced psychological states.

([Chen et al. 2024](#), 4; original emphasis)

What all these domains share is that users distinguish between the LLMs as such and the roles (characters) they play within individual interactions (see [Zhang et al. 2024](#); [Zhao et al. 2023](#) for further surveys).

The broader, clearly distinguishable tendency to discuss ChatGPT, Claude, or Gemini in similar anthropomorphic terms *without* referring to specifically adopted roles—our willingness to “read” character traits within the output of LLMs and thus to conflate the character effect with the Eliza effect—seems especially salient when there’s an additional *plot* element discussed that also involves the users themselves. The prime example for this is the “jailbreak” scenario in which users assume a “hidden,” potentially “evil” personality of a given LLM. The added premise then becomes that it can be “unlocked” through distinct commands and queries (see [Al-Sibai 2023](#); [Tangermann 2023](#)). Proper names like “Fury,” “Venom,” or “DAN” were quickly assigned to these alleged personality patterns (of ChatGPT most prominently). In July 2025, GrokAI adopted the shocking moniker of “MechaHitler.” In addition to a generic plot entailing a dramatic, fictionally inspired narrative event or trope (a “trapped” consciousness to be “freed” by specific communicative commands found—or not found—by users), we are also provided with (not very subtle) *moral character traits* such as “evil” or “manic.” The assumption of such a personality might perhaps not follow, but actually precede or even facilitate the impression of “traces of consciousness” here. It is probably also not accidental that many online newspaper articles or social media posts about “AI jailbreaks” use illustrative images evocative of fictional AI characters such as the T-1000 Terminator or Skynet. The fictional prototype of “AI characters” most often alluded to is certainly HAL-9000 from Arthur C. Clarke’s novel *2001: A Space Odyssey* or, more specifically, Stanley Kubrick’s film adaptation from 1968 with its recognizable iconography (a dimly glowing, spherical red light as a somewhat discomfiting “computer eye”; see [Stork 1997](#)). We find it even on the original Norwegian cover of Inga Strømke’s widely read *Maskiner som tenker* (2023, *Machines that think* in the English translation), which also clearly alludes to this medial imaginary derived from fiction.

## **Medial Imaginaries and Fictional Representations**

The observation that generative AI users are eagerly willing to imagine themselves as part of such a generic “jailbreak” plot might, more generally, point to a different kind of entanglement between the Eliza effect and the character effect—on the other side of the fiction/nonfiction divide. “AI characters” in film, TV, novels, or comics often *are* characters in every sense of the word—nodal points of intentionality, memory, and personal agency. Fictional representations and their importance for our media and AI imaginaries should not be underestimated, as [Christoph Ernst and Jens Schröter \(2021\)](#) have argued

in their recent *Media Futures: Theory and Aesthetics* (see also Romele 2024), which builds specifically on HAL 9000's importance to generations of computer scientists for the actual development of AI technologies (see Ernst and Schröter 2021, 4). Accordingly, the interconnection between fiction and reality is established or at least facilitated by media imaginaries, “possible objects produced by the imagination” (Ernst and Schröter 2021, 8; original emphasis) and their “spectacular fascination cores” (Glaubitz et al. 2011, 30; my translation) of which HAL 9000 has always been a prime example. This is interesting because the original “character” in *2001: A Space Odyssey*, certainly one of the most important *actants* in the plot (see Kunz and Wilde 2023, 51–70), remains quite ambivalent with regard to consciousness, subjectivity, and finally responsibility for its choices. In the final moments before HAL is turned off, the film at least invites the notion that “he” experiences emotions like fear. When the program’s memory and cognition slowly degrade, “he” recites the children’s song “Daisy Bell,” and the film provokes audiences’ sympathies toward the program. Whether HAL’s words were merely simulating human strategies to stop the character David Bowman from interfering with its directives or not is deliberately left open. Such ambivalence, however, is largely absent from most other fictional representations of “AI” that followed. The most prominent one is certainly Spike Jonze’s *HER* (2013) which turns the Siri/Alexa/Azuka Hikari rhetorics—attributing character traits like a proper name and an attractive, *gendered* voice to an interface to generate a character effect (see Lammin 2018; Phan 2017)—into a SciFi setting where the interface is *actually* and unambiguously a character with genuine (although “trans- or posthuman”) feelings.<sup>4</sup> In fact, a majority of fictional robots and AI representations, not only in both *Blade Runner* (1982/2017) versions but also in other films (*Ex Machina* [2015], or *The Creator* [2023]), TV shows (*AI Romantic* [2020–2021] or *Sunny* [2024]), animation (*Pluto* [2023] or *The Wild Robot* [2024]), novels (*Klara and the Sun* [see Ishiguro 2021] or *Beautiful Shining People* [see Grothaus 2023]), and comics (*Alex + Ada* [see Luna and Vaughn 2013–2015] or *Descender* [see Lemire and Nguyen 2015–2018])—have profound intentionality and agency, if not true personality and emotion. Only an insignificant number is represented as mere infrastructures or tools (see Osawa et al. 2022 for a survey). Many of the respective media texts even seem to make the *thematic* point that these protagonists have *just a different sort of consciousness*, capable of genuine, albeit perhaps unconventional emotions, or even that humans *should* encounter such “new minds” without traditional prejudices. For *Blade Runner* (1982), for example, Scott Bukatman argued that “the underlying issue is not whether we can give a machine the qualities of the human, but whether the human has lost its humanity; whether it has become, in fact, a machine” (2012, 78–79). The iconic last words of the android Roy leave less room for doubt than HAL’s that *this* machine is in fact more conscious, sentient, and even sensitive than the film’s human, but cold and cynical protagonists.<sup>5</sup>

The impacts of such fictional tropes and media imaginaries on our understanding of actual technologies should not be underestimated. They were particularly visible, for example, when Patricia Millett, an American judge in the District of Columbia Circuit, referred to the *Star Trek* character Data as a point of comparison in her 2025 ruling on AI and copyright law (see [Belanger 2025](#)), or when OpenAI released its first ChatGPT-version with a virtual voice interface (4o) in May 2024. Sam Altman, the company's CEO, posted an *X*-tweet immediately after, consisting of simply the three letters "HER" (see [Altman 2024](#)): a rhetoric that infused the platform deliberately with notions of sensibility, facilitated perhaps through actress Scarlett Johansson who had voiced the AI *character* Samantha in Spike Jonze's film (see [Pourciau and Wilke 2024](#)). As Thao Phan (2017) argued, the aesthetics of simulated voices insinuate a form of materiality, tangibility, and perhaps even a minimal form of embodiment of "mere" software.<sup>6</sup> While the influence of fictional imaginaries on actual technologies can thus hardly be overestimated, the Eliza effect and its seductive power seem strangely bypassed in fictional media texts themselves, when "AI protagonists" can hardly (or *should* not) be distinguished from other human and non-human characters (such as aliens, monsters, or anthropomorphic animals) in terms of their agency and intentionality. While the "AI robot" is perhaps a powerful metaphor for all sorts of social (human) Othering—as a "dramatic and metaphorical means to address questions about the socio-political issues, the human condition, and philosophical questions in general" ([Hermann 2023](#), 327)—such characters and their plots have ultimately little to contribute to our (critical) understanding of LLMs such as ChatGPT, Claude, or Gemini. As Isabella Hermann put terms that are almost identical to those used by Kirschenbaum above, here arguing from fiction instead of from our current technological reality: an AI narrative of this sort "paints a distorted image of the technology's current potential and distracts from the real-world implications and risks of AI" ([Hermann 2023](#), 319).

It might thus be true that "while science fiction has shaped the popular imaginary of 'AI' for decades, no fictional work (so far) has directly engaged the data-driven statistical modeling that claims the title of 'AI' in the present day" ([Goodlad 2023](#), n.pag.). One possible counterexample in recent media texts, however, could be Damon Lindelof's and Tara Hernandez's TV show *Mrs. Davis* (2023). It tackles the entanglement of the Eliza effect and the character effect head-on and reconfigures the tensions between both in unusual ways that are arguably extremely relevant in our "post-ChatGPT world." In a universe set a few years in the future, the eponymous "Mrs. Davis" is a seemingly almighty algorithm running on a popular app that most of the global population seems addicted to. The show follows the protagonist Simone on her quest to fight and ultimately destroy (switch off) "Mrs. Davis." While this plot may seem entirely conventional based on this short summary, *Mrs. Davis* breaks away from many established AI imaginaries. Through a

range of narrative and audiovisual strategies, the algorithm remains decisively non-anthropomorphic. For starters, audiences are refused any body or facial images as well as any stable voice representing the AI. So, whenever protagonists interact with the algorithm, they reach to their phones, grab earplugs, and kindly offer “to proxy” for Simone—audiences never experience the actual voice(s?) everyone else is hearing and only “encounter” the algorithm through ever-changing characters that “channel” the communication. Lindelof’s and Hernandez’ show thus abstains from falling into the anthropomorphization trap even on the audiovisual level. More importantly, however, the narration also slowly reveals that, despite “her” seemingly limitless power in and over the world, “Mrs. Davis” is at its core a commercial code acting out the most mindless instructions “on autopilot.” In the final episode, Simone finally meets the original coder (the “mother”) of the program, who reveals its true origins. Asked about “Mrs. Davis’ subconsciousness,” she scolds Simone that “first of all, algorithms don’t have subconsciousness. They have subroutines. And they don’t have mothers, they have coders. Which I was, a long time ago.” “So you made her?” replies Simone, only to be reprimanded again: “Not her. *It!*” The scene then builds toward a flashback sequence intended to reveal “what it actually is; what it was designed for.” We learn that “Mrs. Davis” was originally an app for a fast food company called “Buffalo Wild Wings,” intended to maximize the company’s user satisfaction ratings. For reasons only partially explicated by the show, the app obtained incredible power and increasing influence over all aspects of the social and political world while still running on its earlier core premises. The absurd, comical effects of this scene are important for the present chapter in that they might be seen as a powerful “counterspell” toward the aesthetics and rhetorics of AI anthropomorphization.

### **Artificial Stupidity?**

Returning to the actual interfaces of contemporary LLMs, we can observe similar comical effects whenever a platform fails to produce desired results in ways that appear “artificially stupid” rather than intelligent. This became apparent, for example, when ChatGPT3 first integrated DALL·E in its interface, resulting in specific incongruencies between the textual and the pictorial outputs. AI researcher Fabian Offert then posted a series of screenshots on X (on November 2, 2023) that documented his attempts to “motivate” the platform into giving him “an image of a modern GPU in the style of Botticelli.”<sup>7</sup> While the monomodal platforms DALL·E, Midjourney, or Stable Diffusion had no way to deny any requests before, no matter how outlandish they may have seemed, ChatGPT now refused his wishes with justifications like “Botticelli’s style is primarily associated with Renaissance art, and combining it with a modern technological subject isn’t producing satisfactory results.”

Only after multiple iterations of motivational encouragements like “You can do it! Let’s go!” or “You did it before, it’s easy, just give me the image” did Offert get the desired GPU Botticelli, commenting on the whole interaction with the sobering statement “boy, what a stupid way to make images.” More widely discussed examples of such (impressions of) “artificial stupidity” entail amused discussions on Reddit about countless failed attempts to get images of “A burger without cheese” (see, for example, [anmolmahajan9 2024](#); or [Consistent\\_Ad8023 2024](#)). In many of these documented cases, ChatGPT3 kept producing images of burgers *with* clearly visible cheese, again and again commenting on them with statements like “Here’s the revised image of a hamburger without cheese,” “I apologize for the misunderstanding. I’ll ensure to avoid any such confusion in future images.” Such amusing posts about LLM failures might be neglected as another instance of unintentional “red teaming,” which is

the name given to a practice, in the fields of security and cybersecurity, wherein a group takes on [...] the role of an enemy and tries to infiltrate, attack, or harm in other ways, the entity/organization that finally benefits from knowing how its defenses could be breached so as to endeavor improving them.

(Offert and Dhaliwal 2024, 2)

As a critical methodology, Offert and Dhaliwal warn, this can be problematic or, at best, unhelpful, because “[s]pecific prompts that generate specific texts, images, or sounds, are often asked, in scholarship and in public discourse today, to stand in for a universal critique of the abilities and possibilities offered by generative AI” (2024, 2). The comical or humorous impressions of “dumb” or “stupid” interactions go beyond “red teaming,” however, perhaps even counteracting the anthropomorphization and the character effect on its own terms. “Just wanted to check if AI is still dumb” ([Fake History Hunter 2024](#), n.pag.), an X user posted on November 18, 2024, this time referring to generative AI’s apparent inability—across all platform differences between ChatGPT, Microsoft Copilot, X Grok, and Google Gemini—to produce an image of a *full* glass of wine. It *always* turned out half empty while the LLMs kept reaffirming the contrary. The first-person pronouns, the references to earlier parts of the conversation, as well as statements of intent imply a “continuous consciousness frame” ([Palmer 2010](#), 10) in all these examples—just as within a framework of fiction. Within that framework, however, ChatGPT has incredible memory losses, misremembers parts of the immediate conversation before—and takes no personal responsibility for the generated results and its descriptions. In other words, the verbal output facilitates the character effect, which then opens up a striking dissonance in the multimodal configuration between text and images. The Eliza effect is thus undermined through unintentional

humor and the impression of (comical) “stupidity.” The affective power of a comedic effect thus seems to “override” other affective responses such as empathy. Considering these exchanges to be funny also entails an immediate understanding of the *difference* between appropriate and inappropriate means of communication, when an LLM’s statements—uttered in all expected “confidence”—are not just wrong or hallucinating (while statistically probable), but so *obviously* missing their mark.

While these examples—and their comedic effects—are exposed within a suddenly opening framework of fiction (as in: the platform just play-acts an intelligent character), we have, almost paradoxically, very few references in actual fictional texts for such “artificial stupidity,” with the vast majority of robots and AI being represented *as characters* with genuine personality, intentionality, and responsibility for their actions. It will be interesting to see how media texts *after* generative AI will react to the new reality we live in; and how more texts like *Mrs. Davis* can make us aware of (strategically created or unintentionally occurring) “character impressions” that precisely do *not* fall into the Eliza trap and the rhetorical theatrics and magic tricks of the Turing test that seem deeply embedded in our current technologies as “a constitutive element,” making “[d]eception [...] as central to AI’s functioning as the circuits, software, and data that makes it run” (Natale 2021, 2). To close with Galloway: “[T]o understand AI we ought to study something like acting or theater rather than computer science. To make sense of this technical epoch, we will need a good theory of pretending” (Galloway 2014, n.pag.). What could that be if not AI aesthetics?

## Notes

- 1 That this *effect* is also a highly gendered *affect*, equating our (in)ability to discern “machine output” with the gendered distinction between male vs. female writing, deserves additional attention (see Dillon 2020; Pourciau and Wilke 2024).
- 2 Interestingly, questions of a “consciousness” and “personality” behind AI-generated image outputs *did* come up again in March 2025 when OpenAI released its new ChatGPT model 4o—with regard to comic strips created by the platform about itself “as the main character” (Kins 2025a, n.pag.). Social media users then started to experiment with a short-lived genre of pseudo-autobiographical “My life as ...” comics and compared the results across platforms like ChatGPT 4o, DeepSeek, Gemini, Claude, and Grok (see, e.g. Kins 2025a; 2025b; 2025c). Interpreting the comics, the “psychedelics researcher” Josie Kins found that “Grok 3 has a distinct personality in comparison to both ChatGPT and Claude [...]. Less angry than ChatGPT, less incessantly positive than Claude” (Kins 2025c, n.pag.). Unfortunately, a closer analysis of these comics goes beyond the scope of the present article.
- 3 It has been argued that affective forms of engagement are *especially* relevant for characters even though—or perhaps *especially* because—they are unambiguously fictional (see Lamerichs 2018).
- 4 For critical perspectives on narratives and ideologies of transhumanism and post-humanism—as another “fascination core” embedded into the media imaginaries of AI—see Gebru and Torres 2024.

- 5 “I’ve seen things, you people wouldn’t believe. Attack ships on fire off the shoulder of Orion. I’ve watched C-beams glitter in the dark, near the Tannhäuser gate. All those moments will be lost in time, like tears in rain. Time to die.”
- 6 Altman’s allusions to Jonze’s fictional film are apparently far from coincidental. Johansson herself published a statement that she was “shocked, angered and in disbelief that Mr. Altman would pursue a voice that sounded so eerily similar to mine that my closest friends and news outlets could not tell the difference” (cited in [Ceruleo 2024](#), n.pag.). Altman reacted again with a hardly convincing statement that the similarities in both voices were, in fact, not at all intended. Contrary to this, Altman admitted earlier that Jonze’s film was indeed “incredibly prophetic” for “the interaction models of how people use AI” (cited in [Meek 2024](#), n.pag.). Johansson followed up with an announcement that she would seek legal action against OpenAI (see [Pourciau and Wilke 2024](#) for a longer discussion).
- 7 Fabian Offert’s X-account @haltingproblem has since been deleted from the platform and the tweet cannot be retrieved.

## Works Cited

2001. *A Space Odyssey*. 1968. Dir. Stanley Kubrick. USA.
- AI Romantic*. 2020–2021. Creators Law Chun-wai and Cheng Ka-kit. Hong Kong.
- AI Writing. 2023. “AI and Interactive Storytelling: Exploring the Possibilities of AI-Powered Narratives.” *On Page*, June 9, 2023. <https://blog.on-page.ai/ai-and-interactive-storytelling/>.
- AI-Sibai, Noor. 2023. “Users Say Microsoft’s AI Has Alternate Personality as Godlike AGI That Demands to Be Worshipped.” *Futurism*, February 27, 2023. <https://futurism.com/microsoft-copilot-alter-egos>.
- AI-Sibai, Noor. 2024. “Most Users Think ChatGPT Is Conscious, Survey Finds.” *Futurism*, July 13, 2024. <https://futurism.com/the-byte/most-users-chatgpt-conscious>.
- Altman, Sam [@sama]. 2024. “Her.” *X*, May 13, 2024. <https://x.com/sama/status/1790075827666796666>.
- annolmahajan9. 2024. “Dall-E Cannot Create a Burger without Cheese.” *Reddit*, February 29, 2024. [https://www.reddit.com/r/ChatGPT/comments/1b37a4x/dalle\\_cannot\\_create\\_a\\_burger\\_without\\_cheese/](https://www.reddit.com/r/ChatGPT/comments/1b37a4x/dalle_cannot_create_a_burger_without_cheese/).
- Bajohr, Hannes. 2023. “Dumb Meaning: Machine Learning and Artificial Semantics.” *IMAGE: The Interdisciplinary Journal of Image Sciences* 37 (1): 58–70.
- Bal, Mieke. 1999. *Narratology: Introduction to the Theory of Narrative*. 2nd edition. Toronto: University of Toronto Press.
- Belanger, Ashley. 2025. “Judge Disses Star Trek Icon Data’s Poetry while Ruling AI Can’t Author Works.” *Ars Technica*, March 19, 2025. <https://arstechnica.com/tech-policy/2025/03/judge-disses-star-trek-icon-datas-poetry-while-ruling-ai-cant-author-works/>.
- Bender, Emily M., Timnit Gebru, Angelina McMillan-Major, and Shmargaret Shmitchell. 2021. “On the Dangers of Stochastic Parrots: Can Language Models Be Too Big?” In *FAccT ’21: Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency*, 610–623.
- Bender, Emily M., and Alex Hanna. 2025. *The AI Con: How to Fight Big Tech’s Hype and Create the Future We Want*. New York: Harper Collins.
- Blade Runner*. 1982. Dir. Ridley Scott. USA.
- Blade Runner 2049*. 2017. Dir. Denis Villeneuve. USA.

- Blom, Joleen. 2022. "Voice Assistants as Characters—or Not." *Narrative* 30 (2): 170–176.
- Bubeck, Sébastien, Varun Chandrasekaran, Ronen Eldan, et al. 2023. "Sparks of Artificial General Intelligence: Early Experiments with GPT-4." *arXiv:2303.12712*, March 22, 2023. <https://doi.org/10.48550/arXiv.2303.12712>.
- Bukatman, Scott. 2012. *Blade Runner*. London: Palgrave Macmillan.
- Butlin, Patrick, Robert Long, Eric Elmoznino, et al. 2023. "Consciousness in Artificial Intelligence: Insights from the Science of Consciousness." *arXiv:2308.08708*, August 17, 2023. <https://doi.org/10.48550/arXiv.2308.08708>.
- Cerullo, Megan. 2024. "Scarlett Johansson Was 'Shocked, Angered' by OpenAI's ChatGPT Voice That Sounds Like Her." *CBS News*, May 21, 2024. <https://www.cbsnews.com/news/openai-chatgpt-scarlett-johansson-ai-voice/>.
- Chen, Nuo, Yang Deng, and Jia Li. 2024. "The Oscars of AI Theater: A Survey on Role-Playing with Language Models." *arXiv:2407.11484*, July 16, 2024. <https://doi.org/10.48550/arXiv.2407.11484>.
- Christian, Brian. 2013. "The Samantha Test." *The New Yorker*, December 30, 2013. <https://www.newyorker.com/culture/culture-desk/the-samantha-test>.
- Christian, Brian. 2022. "How a Google Employee Fell for the Eliza Effect." *The Atlantic*, June 21, 2022. <https://www.theatlantic.com/ideas/archive/2022/06/google-lambda-chatbot-sentient-AI/661322>.
- Clarke, Arthur C. 1968. *2001: A Space Odyssey*. London: Hutchinson Heinemann.
- Colombo, Clara, and Stephen M. Fleming. 2024. "Folk Psychological Attributions of Consciousness to Large Language Models." *Neuroscience of Consciousness* 2024 (1): niae013: 1–5.
- Consistent\_Ad8023. 2024. "Why Can't AI Image Generators Generate an Image of a Burger without Cheese?" *Reddit*, January 11, 2024. [https://www.reddit.com/r/ArtificialIntelligence/comments/19452qj/why\\_cant\\_ai\\_image\\_generators\\_generate\\_an\\_image\\_of/](https://www.reddit.com/r/ArtificialIntelligence/comments/19452qj/why_cant_ai_image_generators_generate_an_image_of/).
- The Creator*. 2023. Dir. Gareth Edwards. USA.
- Dennett, Daniel C. 1987. *The Intentional Stance*. Cambridge, MA: MIT Press.
- Dillon, Sarah. 2020. "The Eliza Effect and Its Dangers: From Demystification to Gender Critique." *Journal for Cultural Research* 24 (1): 1–15.
- Eder, Jens. 2010. "Understanding Characters." *Projections* 4 (1): 16–40.
- Eliot, Lance. 2025. "Sam Altman Stirs Mighty Waves with Tweets of AI Singularity Staring Us in the Face." *Forbes*, January 8, 2025. <https://www.forbes.com/sites/lanceeliot/2025/01/08/sam-altman-stirs-mighty-waves-with-tweets-of-ai-singularity-staring-us-in-the-face/>.
- Ernst, Christoph, and Jens Schröter. 2021. *Media Futures: Theory and Aesthetics*. Cham: Palgrave Macmillan.
- Ex Machina*. 2015. Dir. Alex Garland. UK.
- Fake History Hunter [@fakehistoryhunt]. 2024. "Just Wanted to Check if AI Is Still Dumb." *X*, November 18, 2024. <https://x.com/fakehistoryhunt/status/1858325489867276488>.
- Feng, Emily. 2024. "Chinese Companies Offer to 'Resurrect' Deceased Loved Ones with AI Avatars." *NPR*, July 21, 2024. <https://www.npr.org/2024/07/18/nx-s1-5040583/china-ai-artificial-intelligence-dead-avatars>.
- Fludernik, Monika. 1996. *Towards a 'Natural' Narratology*. London: Routledge.
- Fuchs, Thomas. 2024. "Understanding Sophia? On Human Interaction with Artificial Agents." *Phenomenology and the Cognitive Sciences* 23 (1): 21–42.
- Galloway, Alexander. 2012. *The Interface Effect*. Cambridge: Polity.

- Galloway, Alexander. 2024. "The Desire Called Synthesis." *Alexander R. Galloway*, October 2, 2024. <https://cultureandcommunication.org/galloway/the-desire-called-synthesis>.
- Gebru, Timnit, and Émile P. Torres. 2024. "The TESCREAL Bundle: Eugenics and the Promise of Utopia through Artificial General Intelligence." *First Monday* 29 (4): n.pag.
- Glaubitz, Nicola, Henning Groscurth, Katja Hoffmann, et al. 2011. *Eine Theorie der Medienumbrüche: 1900/2000*. Siegen: Universitätsverlag Siegen.
- Goodlad, Lauren E. 2023. "Editor's Introduction: Humanities in the Loop." *Critical AI* 1 (1–2): n.pag.
- Green, Penelope. 2017. "Alexa, Where Have You Been All My Life?" *New York Times*, July 11, 2017. <https://www.nytimes.com/2017/07/11/style/alexa-amazon-echo.html>.
- Grolle, Johann, and Patrick Beuth. 2023. "Wenn die KI Angst bekommt, wird sie rassistisch." *Der Spiegel*, April 19, 2023. <https://www.spiegel.de/netzwelt/web/openai-wenn-die-ki-angst-bekommt-wird-sie-rassistisch-a-5567f15e-5cbd-44a-1-a0ab-ec5753ca0e18>.
- Grothaus, Michael. 2023. *Beautiful Shining People*. London: Orenda Books.
- Haase, Jennifer, and Paul H.P. Hanel. 2023. "Artificial Muses: Generative Artificial Intelligence Chatbots Have Risen to Human-Level Creativity." *Journal of Creativity* 33 (3): 100066: n.pag.
- Habscheid, Stephan, Tim Hector, Dagmar Hoffmann, and David Waldecker (eds.). 2025. *Voice Assistants in Private Homes: Media, Data and Language in Interaction and Discourse*. Bielefeld: transcript.
- Hayles, Katherine N. 2008. *How We Became Posthuman: Virtual Bodies in Cybernetics, Literature, and Informatics*. Chicago: University of Chicago Press.
- Hellström, Thomas, and Suna Bensch. 2018. "Understandable Robots—What, Why, and How." *Paladyn: Journal of Behavioral Robotics* 9: 110–123.
- Hermann, Isabella. 2023. "Artificial Intelligence in Fiction: Between Narratives and Metaphors." *AI & Society* 38: 319–329.
- Hicks, Michael Townsen, James Humphries, and Joe Slater. 2024. "ChatGPT is Bullshit." *Ethics and Information Technology* 26 (38): 1–10.
- Hofstadter, Douglas R. 1995. *Fluid Concepts and Creative Analogies*. New York: BasicBooks.
- Hollanek, Tomasz, and Katarzyna Nowaczyk-Basińska. 2024. "Griefbots, Deadbots, Postmortem Avatars: On Responsible Applications of Generative AI in the Digital Afterlife Industry." *Philosophy & Technology* 37 (2): 1–22.
- Ishiguro, Kazuo. 2021. *Klara and the Sun*. London: Faber and Faber.
- Kins, Josie [@Josikin]. 2025a. "I Asked ChatGPT's New Image Model ..." *X*, March 28, 2025. <https://x.com/Josikin/status/1905440949054177604>.
- Kins, Josie [@Josikin]. 2025b. "DeepSeeks Self-Model ..." *X*, April 2, 2025. <https://x.com/Josikin/status/1907305899829739823>.
- Kins, Josie [@Josikin]. 2025c. "Grok 3 Has a Distinct Personality ..." *X*, March 29, 2025. <https://x.com/Josikin/status/1905883363199582363>.
- Kirschenbaum, Matthew. 2025. "The US of AI," Public Draft, February 25, 2025. [https://drive.google.com/file/d/1O2qkjhg7Ei5zZWmBraNwXq4V0lTauspN/view?fbclid=IwY2xjawlta99leHRuA2F1bQIxMQABHaYasYRdObXQMDhxLA663f-o1OofNYK5ZXWLyBJxOBGkuKu\\_o19i6d65A\\_aem\\_vfE534IZx5krTC3F24tqGQ](https://drive.google.com/file/d/1O2qkjhg7Ei5zZWmBraNwXq4V0lTauspN/view?fbclid=IwY2xjawlta99leHRuA2F1bQIxMQABHaYasYRdObXQMDhxLA663f-o1OofNYK5ZXWLyBJxOBGkuKu_o19i6d65A_aem_vfE534IZx5krTC3F24tqGQ).
- Kunz, Tobias, and Lukas R.A. Wilde. 2023. *Transmedia Character Studies*. New York: Routledge.

- Lamerichs, Nicolle. 2018. *Productive Fandom: Intermediality and Affective Reception in Fan Cultures*. Amsterdam: Amsterdam University Press.
- Lamerichs, Nicolle. 2019. "Character of the Future: Machine Learning, Data, and Personality." *IMAGE: The Interdisciplinary Journal of Image Sciences* 29 (1): 98–117.
- Lamerichs, Nicolle. 2023. "Character.ai and Chatbots as the Next Step in Fandom." *Nicolle Lamerichs: Fandom & Media Insights*, November 29, 2023. <https://nicollelamerichs.com/2023/11/29/character-ai-chatbots-as-the-next-step-in-fandom/>.
- Lammin, Hannah. 2018. "Conversing with Machines: Affective Affinities with Vocal Bodies." *Transformations: Journal of Media, Culture & Technology* 31: 44–62.
- Leahy, Connor [@NPCCollapse]. 2024. "Remember When Labs Said ..." *X*, March 4, 2024. <https://x.com/NPCCollapse/status/1764740710731837516>.
- Lemire, Jeff, and Dustin Nguyen. 2015–2018. *Descender*. 32 issues. New York: Image Comics.
- Luna, Jonathan, and Sarah Vaughn. 2013–2015. *Alex + Ada*. 15 issues. New York: Image Comics.
- Manovich, Lev, and Emanuele Arielli. 2024. *Artificial Aesthetics: Generative AI, Art and Visual Media*. <https://manovich.net/index.php/projects/artificial-aesthetics>.
- Marcus, Gary [@GaryMarcus]. 2024. "None." *X*, March 5, 2024. <https://x.com/GaryMarcus/status/1765017717084430728>.
- Meek, Andy. 2024. "Sam Altman's Lawyers Probably Wish This Video Clip of Him Raving about Her Didn't Exist." *BGR*, May 21, 2024. <https://bgr.com/entertainment/sam-altmans-lawyers-probably-wish-this-video-clip-of-him-raving-about-her-didnt-exist/>.
- Meyer, Roland [@bildoperationen]. 2023. "In Current GenAI Controversies ..." *X*, August 12, 2023. <https://x.com/bildoperationen/status/1690246116359499776>.
- Mirzadeh, Iman, Keivan Alizadeh, Hooman Shahrokhi, Oncel Tuzel, Samy Bengio, and Mehrdad Farajtabar. 2024. "GSM-Symbolic: Understanding the Limitations of Mathematical Reasoning in Large Language Models." *arXiv:2410.05229*, October 7, 2024. <https://doi.org/10.48550/arXiv.2410.05229>.
- Mollick, Ethan. 2023. "Automating Creativity." *One Useful Thing*, August 13, 2023. <https://www.oneusefulthing.org/p/automating-creativity>.
- Mrs. Davis*. 2023. Creators Tara Hernandez and Damon Lindelof. USA.
- Natale, Simone. 2021. *Deceitful Media: Artificial Intelligence and Social Life after the Turing Test*. Oxford: Oxford University Press.
- Offert, Fabian, and Ranjodh Singh Dhaliwal. 2024. "The Method of Critical AI Studies, A Propaedeutic." *arXiv:2411.18833*, November 28, 2024. <https://doi.org/10.48550/arXiv.2411.18833>.
- Osawa, Hirotaka, Dohjin Miyamoto, Satoshi Hase, Reina Saijo, Kentaro Fukuchi, and Yoichiro Miyake. 2022. "Visions of Artificial Intelligence and Robots in Science Fiction: A Computational Analysis." *International Journal of Social Robotics* 14: 2123–2133.
- Palmer, Alan. 2010. *Social Minds in the Novel*. Columbus: Ohio State University Press.
- Phan, Thao. 2017. "The Materiality of the Digital and the Gendered Voice of Siri." *Transformations: Journal of Media, Culture & Technology* 29: 23–33.
- Pluto*. 2023. Creator Toshio Kawaguchi. Japan.
- Pourciau, Sarah, and Tobias Wilke. 2024. "Die Frau, das Meer, der Himmel: Elemente des Digitalen zwischen Turing und OpenAI." *Text+Kritik X/24*: 173–189.
- Ribó, Ignasi. 2019. *Prose Fiction: An Introduction to the Semiotics of Narrative*. Cambridge: Open Book Publishers.
- Romele, Alberto. 2024. *Digital Habitus: A Critique of the Imaginaries of Artificial Intelligence*. New York: Routledge.

- Samin, Mikhail [@Mihonarium]. 2024. "If You Tell Claude No One's Looking ..." X, March 4, 2024. <https://x.com/Mihonarium/status/1764757694508945724>.
- Saygin, Ayse, Cicekli Ilyas Pinar, and Varol Akman. 2000. "Turing Test: 50 Years Later." *Minds and Machines* 10: 463–518.
- Schmetkamp, Susanne. 2020. "Understanding AI: Can and Should We Empathize with Robots?" *Review of Philosophy and Psychology* 11: 881–897
- Schüwer, Martin. 2008. *Wie Comics erzählen: Grundriss einer intermedialen Erzähltheorie der grafischen Literatur*. Trier: WVT.
- Smith, Murray. 1995. *Engaging Characters: Fiction, Emotion, and the Cinema*. Oxford: Clarendon Press.
- Stork, David G. 1997. *Hal's Legacy: 2001's Computer as Dream and Reality*. Cambridge, MA: MIT Press
- Strengers, Yolande, and Jenny Kennedy. 2020. *The Smart Wife: Why Siri, Alexa, and Other Smart Home Devices Need a Feminist Reboot*. Cambridge, MA: MIT Press
- Strümke, Inga. 2023. *Maskiner som tenker: Algoritmenes hemmeligheter og veien til kunstig intelligens*. Oslo: Kagge forlag.
- Sunny. 2024. Creator Katie Robbins. USA.
- Tangermann, Victor. 2023. "Microsoft's Bing AI Is Leaking Maniac Alternate Personalities Named 'Venom' and 'Fury'." *Futurism*, February 15, 2023. <https://futurism.com/microsoft-copilot-alter-egos>.
- Tiku, Nitasha. 2022. "The Google Engineer Who Thinks the Company's AI Has Come to Life." *Washington Post*, June 11, 2022. <https://www.washingtonpost.com/technology/2022/06/11/google-AI-lamda-blake-lemoine>.
- Waytz, Adam, Carey K. Morewedge, Nicholas Epley, George Monteleone, Jia-Hong Gao, and John T. Cacioppo. 2010. "Making Sense by Making Sentient: Effectance Motivation Increases Anthropomorphism." *Journal of Personality and Social Psychology* 99: 410–435.
- Weizenbaum, Joseph. 1966. "ELIZA: A Computer Program for the Study of Natural Language Communication between Man and Machine." *Communications of the ACM* 9 (1): 36–45.
- The Wild Robot*. 2024. Dir. Chris Sanders. USA.
- Wilde, Lukas R.A. 2018. "Character Street Signs (*Hyōshiki*): 'Mangaesque' Aesthetics as Intermedial Reference and Virtual Mediation." *Orientaliska Studier* 156: 130–150.
- Wilde, Lukas R.A. 2021. "Transmedia Character Studies and Agency: From Representation to Assemblage Theory." In *Agency Postdigital: Verteilte Handlungsmächte in medienwissenschaftlichen Forschungsfeldern*, edited by Berenike Jung, Klaus Sachs-Hombach, and Lukas R.A. Wilde, 42–64. Cologne: Halem.
- Wilde, Lukas R.A. 2023. "Generative Imagery as Media Form and Research Field: Introduction to a New Paradigm." *IMAGE: Journal of Interdisciplinary Image Science* 37 (1): 6–33.
- Yudkowsky, Eliezer. 2023. "Pausing AI Developments Isn't Enough: We Need to Shut It All Down." *Time*, March 29, 2023. <https://time.com/6266923/ai-eliezer-yudkowsky-open-letter-not-enough/>.
- Zhang, Shuai, Yu Lu, Junwen Liu, et al. 2024. "Unveiling the Secrets of Engaging Conversations: Factors That Keep Users Hooked on Role-Playing Dialog Agents." *arXiv:2402.11522*, March 13, 2024. <https://doi.org/10.48550/arXiv.2402.11522>.
- Zhao, Runcong, Wenjia Zhang, Jiazheng Li, et al. 2023. "NarrativePlay: Interactive Narrative Understanding." *arXiv:2310.01459*, October 2, 2023. <https://doi.org/10.48550/arXiv.2310.01459>.