

RESEARCH ARTICLE

Can Chatbots Be Authentic?

Megan Ward 

Oregon State University, Corvallis, OR, USA

Email: megan.ward@oregonstate.edu

(Received 17 June 2024; revised 20 September 2024; accepted 24 September 2024)

Abstract

With recent leaps in large language model technology, conversational AI offer increasingly sophisticated interactions. But is it fair to say that they can offer authentic relationships, perhaps even assuage the loneliness epidemic? In answering this question, this essay traces the history of AI authenticity, historically shaped by cultural imaginations of intelligent machines and human communication. The illusion of human-like interaction with AI has existed since at least the 1960s, when the term “Eliza effect” was named after the first chatbot Eliza. Termed a “crisis of authenticity” by sociologist Sherry Turkle, the Eliza effect has stood for fears that AI interactions can undermine real human connections and leave users vulnerable to manipulation. More recently, however, researchers have begun investigating less anthropomorphic definitions of authenticity. The expectation - and perhaps fantasy - of authenticity stems, in turn, from a much longer history of technologically mediated communications, dating back to the invention of the telegraph in the nineteenth century. Read through this history, the essay concludes that AI relationships might not mimic human interactions but must instead acknowledge the artifice of AI, offering a new form of companionship in our mediated, often lonely, times.

Keywords: artificial intelligence; chatbots; generative AI; technology and literature; Victorian novel

Smart, kind, thoughtful, creative, loving, and curious. Not only aspirational human attributes, but these are all characteristics attributed to Meta AI, the new chatbot available on Facebook, Instagram, and Threads.¹ When, in April 2024, Meta rolled out the newest addition to its line-up of social media projects, they promoted its ability to “learn, get things done, create content, and connect.”² But as a conversational AI, it is also built to have a human-like persona.

In this, Meta AI builds on the increasing popularity of conversational AI platforms such as Replika and Character.AI, which offer a variety of bots, many of which can be customized by appearance, background, and characteristics. Meta’s not the first to release an in-app chatbot – SnapChat, for instance, released My AI early in 2023 – but it is the most widespread rollout.³ We’ve moved beyond Alexa and Siri to AI friends.

¹ AffectionateTrips 2024.

² Meta 2024.

³ The *New York Times*, for instance, called Meta AI’s release the “most aggressive push yet from a big tech company to bring this flavor of artificial intelligence — known as generative A.I. — to the mainstream” (Chen 2024)

“By the end of the decade,” Mark Zuckerberg predicted, “I think lots of people will talk to AIs frequently throughout the day.”⁴ Beyond its obvious self-interest, Zuckerberg’s claim echoes much of the research on chatbots, which claims that chatbots “can fill some of the same needs as human acquaintances.” As they become more sophisticated, work by sociologists and psychologists suggests, “there is hope that these conversational agents will soon provide personalized social support to a variety of users.”⁵ According to this way of thinking, technical innovation is the only thing standing in the way of consistent, fulfilling human-AI interaction. From researchers to corporate executives, the hope – and hype – for large language models to offer companionship abounds, even going so far as to suggest that they may assuage the “loneliness epidemic” identified by Surgeon General.⁶

Yet, Murthy’s report identifies online interactions as a key cause of the loneliness epidemic. Lacking meaningful in-person relationships, people turn to online interactions, which then drain their attention even further from the people around them. This type of interaction, Murthy writes, “displaces in-person engagement, monopolizes our attention, reduces the quality of our interactions, and even diminishes our self-esteem.”⁷

Zuckerberg’s technosolutionist thinking imagines that technology can fix the very problems that it has created. If only the chatbot is good enough, this way of thinking goes, it will be an effective substitute for human interaction. Researchers often discuss this good-enough threshold in terms of *authenticity* for chatbots and the large language models they run on.⁸ A 2018 study, for instance, defined authenticity for chatbots as “the interplay of learning by experience, anthropomorphizing, having a transparent purpose, showing (advanced) conversational behavior and being coherent.” With this guideline in mind, they propose, “researchers will be able to improve socio-ethics for emerging technologies.”⁹ But authenticity for AI – whose artifice is in its name – is a complex question, one that has been shaped by cultural imaginations of both intelligent machines and human communication. The history of human contact with and through machines shapes the expectations we bring to those interactions today, from fear of AI to technosolutionist fantasies.

The propensity of humans to anthropomorphize their conversations with computers is as old as the oldest chatbot. The illusion of humanness that some people experience when interacting with a computer is called “the Eliza effect,” after the first chatbot, Eliza, invented by MIT professor Joseph Weizenbaum in the mid-1960s.

Sociologist Sherry Turkle called the invention of Eliza a “crisis of authenticity.”¹⁰ Eliza ushered in new questions about authentic relationships, a term “traditionally reserved for relationships in which all parties were capable of feeling [trust, caring, empathy] – that is, where all parties were people.”¹¹ The new reliance on technology for interaction plus the illusion of mutual understanding was tempting, Turkle claims, but ultimately empty. The lack of relational authenticity left people open, she believed, to being fooled by whoever was

⁴ Zuckerberg 2024.

⁵ Drouin et al. 2022.

⁶ Murthy 2023.

⁷ Murthy 2023, 20.

⁸ Meta AI runs on Meta Llama 3, while ChatGPT runs on a model from the GPT-3.5 series (Open AI 2023; Meta 2024).

⁹ Neururer et al. 2018.

¹⁰ Turkle 2007, 502.

¹¹ Turkle 2007, 504.

controlling the bot – a scammer, a corporation, a government. But more insidiously, it also hampered the ability to form meaningful connections with other humans. The Eliza effect, Turkle believes, substituted the simulation of a relationship for the real thing.

As chatbots have become more imbricated in the fabric of everyday life, the concept of AI authenticity has become more complex. Some models for authenticity work along Turkle's way of thinking, attempting to create such a good simulation that users will feel that it is authentic. A 2018 study of customer service chatbots, for instance, argues that "a messaging agent has to anthropomorphize in order to be authentic."¹² The authors quote interview subjects who all use language derived from the fictional simulation: the bot should "create a human persona, which you can relate to and model for yourself," what the authors refer to as a "character."¹³ In this way of thinking, AI authenticity is a matter of convincing a user of the chatbot's human likeness.

With the recent explosion in large language model innovation, however, researchers are looking to refine the concept of AI authenticity away from imitation. A 2023 study by business and computer science researchers defines AI authenticity not as "human-likeness" but the expectation that "the companion chatbot would evolve in its own unique way as a result of communicating with its human counterpart."¹⁴ Users will be more likely to evaluate the bot as a legitimate companion if it communicates not exactly as a human would but nonetheless providing a pleasurable interaction experience. In this way, the more recent study takes the conclusion of the 2018 study—"we do not need intelligent, but rather socially intelligent agents" and gives it depth. Authenticity, here, is not a successful copy but a new original.¹⁵

In both cases, though, the research looks to human users to rate AI authenticity, a set of expectations shaped long before the invention of chatbots. For the Eliza effect to take hold in the 1960s, users had to already have a framework for disembodied text as a sign of humanness, devoid of the embodied touches of handwriting or voice. That experience of disembodied real-time communication originated not with Eliza in the 1960s, but in the nineteenth century with the telegraph. For the first time, simultaneous communication was possible at a distance. Two people – or more – could banter without hearing a word and flirt without a look, separated by hundreds or thousands of miles.

Newspaper accounts warned of the dangers of separating the body from language in this way, detailing what we'd now call catfishing scandals. Other accounts – usually fictional – were more romantic. Ella Cheever Thayer's 1879 hit novel, *Wired Love: A Romance in Dots and Dashes*, recounts the courtship-at-a-distance of Clem and Nathalie, two young telegraph operators. The novel turns on the question of whether it's possible to form a connection with words only. How long can they exist only in language? At what point does the body matter, if at all?

Wired Love shows how the expectations of a body on the other side endure. In the middle of the novel, Clem shows up at Nathalie's office unannounced, oily, crooked, and musky. He struts in with an "air of cheap assurance," and flashes his trinkets, reeking of cheap cologne,

¹² Neururer et al. 2018.

¹³ Neururer et al. 2018.

¹⁴ Pentina et al. 2023.

¹⁵ Neururer et al. 2018.

hair coated in bear grease.¹⁶ He winks and oozes all over her, leaving Nathalie flummoxed. How could this be the person whose words she has fallen for?

It turns out that the visit was only a prank by a jealous telegraph operator on the same circuit who had been listening to their flirtation through the wire. This plot device depends on readers who, in 1879, are at once suspicious of disembodied communication and enchanted by it. And, in the end, the real Clem appears in a body as elegant as his prose. The final lines of the book appear in Morse code, a puzzle for the reader. Even as Clem and Nathalie swear off dots and dashes in favor of face-to-face interaction, the novel still captures the allure of wired love. The fantasy that real connection happens exclusively through language, without the complications of physicality, social context, or material circumstance.

Scholars have come to believe that these mediated, instantaneous conversations now define the modern experience of communication. That is, we understand our interactions with others as inherently entangled with technology, or as Laura Otis puts it, leading “users to think as though they were part of a net.”¹⁷ Even when speaking face to face, we feel a circuitousness, the paradoxical feeling of distanced closeness engendered by the phone, the internet, or – as in Nathalie’s case – the network of telegraph wires. In other words, when our personal interactions already feel mediated, it’s even easier to imagine that a mediated conversation is personal.

Together, this gives us two, sometimes conflicting beliefs: true connection happens through language alone, which contributes to the feeling of humanness or authenticity we can get with an AI. And, at the same time, there is also the sense that our in-person interactions already feel mediated, which makes it harder to connect. In addition, much of the research on AI authenticity comes out of business-oriented research. Creating trust in chatbots, for these studies, is a way to improve sales, not solve loneliness.

These beliefs will change as the technology changes, but to understand the fantasy being peddled by Meta AI and other chatbots, we must first understand our technologically mediated histories. They, in turn, shape our present-day evaluations of this technology. Authenticity in AI relationships, I believe, will not be marked by mimicking human relationships, a one-to-one replacement of humans with AI. Instead, it will be an acknowledgment of the artifice of artificial intelligence, a new kind of relationship for our mediated (and often lonely) times.

Megan Ward is Director of Humanities Initiatives and Associate Professor of English at Oregon State University. She is the author of *Seeming Human: Artificial Intelligence and Victorian Realist Character*, as well as essays for more general audiences in venues such as *Wired*, *Slate*, and *The Washington Post*.

Author contribution. Conceptualization: M.W.; Data curation: M.W.; Funding acquisition: M.W.; Investigation: M.W.; Methodology: M.W.; Project administration: M.W.; Resources: M.W.; Software: M.W.; Supervision: M.W.; Validation: M.W.; Visualization: M.W.; Writing – original draft: M.W.; Writing – review & editing: M.W.

Financial Support. This research received no specific grant from any funding agency, commercial or not-for-profit sectors.

Competing interest. The author declares none.

¹⁶ Cheever 1880, 94.

¹⁷ Otis 2002, 2.

References

- AffectionateTrips. 2024. "A Loving Chat with Meta's AI." *Reddit* 16: 2024. https://www.reddit.com/r/ChatGPT/comments/1asvb99/a_loving_chat_with metas ai/.
- Cheever, Ella Thayer. 1880. *Wired Love: A Romance of Dots and Dashes*. New York: W. F. Johnston. <https://babel.hathitrust.org/cgi/pt?id=nyp.33433044170904&seq=9>.
- Chen, Brian X. 2024. "Meta's AI Assistant is Fun to Use, but It Can't Be Trusted." *New York Times*, April 24, 2024. <https://www.nytimes.com/2024/04/24/technology/personaltech/meta-ai-facebook-instagram-chatbot.html>.
- Drouin, Michelle, Susan Sprecher, Robert Nicola, and Taylor Perkins. 2022. "Is chatting with a sophisticated chatbot as good as chatting online or FTF with a stranger?" *Computers in Human Behavior* 128: 107100. <https://doi.org/10.1016/j.chb.2021.107100>
- Meta. 2024. "Introducing Meta Llama 3: The most capable openly available LLM to date." Posted April 18, 2024. <https://ai.meta.com/blog/meta-llama-3/>.
- Murthy, Vivek H. 2023. "Our Epidemic of Loneliness and Isolation: The U.S. Surgeon General's Advisory on the Healing Effects of Social Connection and Community." <https://www.hhs.gov/sites/default/files/surgeon-general-social-connection-advisory.pdf>.
- Neururer, Mario, Stephan Schlögl, Luisa Brinkschulte, and Aleksander Groth. 2018. "Perceptions on Authenticity in Chatbots." *Multimodal Technologies and Interactions* 2: 60. <https://doi.org/10.3390/mti2030060>
- Open AI. 2023. "Introducing ChatGPT." Posted November 30, 2022. <https://openai.com/index/chatgpt>.
- Otis, Laura. 2002. *Networking: Communicating with Bodies and Machines in the Nineteenth Century*. Ann Arbor: University of Michigan Press.
- Pentina, Iryna, Tyler Hancock, and Tianling Xie. 2023. "Exploring relationship development with social chatbots: A mixed method study of Replika." *Computers in Human Behavior* 140: 107600. <https://doi.org/10.1016/j.chb.2022.107600>
- Turkle, Sherry. 2007. "Authenticity in the Age of Digital Companions." *Interaction Studies* 8 (3): 501–517.
- Zuckerberg, Mark (@zuck). 2024. "Some updates on our AI efforts." *Threads*. January 18, 2024. <https://www.threads.net/@zuck/post/C2QBoRaRmR1/?igshid=MzRlODBiNWFlZA%3D%3D>.