

# Chatbots: Cybernetic Psychology and the Future of Conversation

Looking back on the history of chatbot development, one Microsoft development team observed in 2018 that “with vastly more people being digitally connected, *it is not surprising* that social chatbots have been developed as an alternative means for engagement.”<sup>1</sup> What sort of “alternative” is presented when humans engage with chatbots? If the Fourth Industrial Revolution depends not only on the flow of goods and services but also on the flow of signals of assent (purchases, likes, shares), then the economy of conversation between users must be made seamless at any cost.<sup>2</sup> Is the chatbot an alternative to the otherness of human beings? Are chatbots a patch for alterity? Alongside the psychologically meaningful dimensions attending the problem of our incommensurability with one another—our personhood—the disconcerting, unmanageable, merciful, and threatening separation between human beings presents a newly focalized economic problem in the digital age.

- 1 Heung-yeung Shum, Xiaodong He, and Di Li, “From Eliza to Xiaoice: Challenges and Opportunities with Social Chatbots,” *Frontiers of Information Technology & Electronic Engineering* 19, no. 1 (2018): 13 (emphasis added).
- 2 While the Third Industrial Revolution describes the period during which information processing via the computer became predominant, the Fourth Industrial Revolution (4IR) refers to a period defined by widespread social and industrial connectivity through the internet (cloud computing, social media, etc.), artificial intelligence, quantum computing, nano- and biotechnology, and the incorporation of smart devices into many aspects of economic and interpersonal life through the Internet of Things (IoT). This periodization was proposed by economist Klaus Schwab, founder of the World Economic Forum. See, for example, Klaus Schwab, *The Fourth Industrial Revolution* (Geneva: World Economic Forum, 2016).

---

Avery Slater, “Chatbots: Cybernetic Psychology and the Future of Conversation,” *JCMS* 61, no. 4 (Summer 2022): 181–187.

In the echo chambers of social media, algorithms carefully curate online interactions to create amplifying effects, with streams of content filtered toward users who are predicted to “like” and “share” this content. Social bots, whether conversation generating or message amplifying in their intent, have enmeshed themselves ineradicably into the flow of digital communication.<sup>3</sup> As Douglas Guilbeault argues, social media platforms “are a new kind of habitat that imposes habits of self-construction that both humans and bots equally exploit.”<sup>4</sup> Simultaneously, the echo chamber resounds with sock puppet accounts and with bots hailing and harassing users at every swipe or scroll, transmitting multimodal packets of information designed to catch human attention from every crevice in the online infrastructure.<sup>5</sup>

In an English-language context, the memory of Microsoft’s 2016 social chatbot experiment remains infamous—an incident in which “Tay” (an artificial conversational agent designed with the alleged personality of a nineteen-year-old woman) transformed into a neo-Nazi within hours of social interaction and was removed unceremoniously from the web in disgrace.<sup>6</sup> Certainly this says something about the users (of 4chan, etc.) who indoctrinated Tay; it says something as well about the future of conversation, the future of our vulnerability to one another in and through language—a faculty now wielded by nonhuman agents. To understand not only the relational ethics of life amid bots but also the economic valence of this conversational enmeshment, the origin of the chatbot should be reconsidered. This essay suggests some directions for thinking concerning postwar computer science’s uptake of cybernetic psychology and the influence of this genealogy on certain problems of digital communication today.

Training early chatbots in conversational fluency represents an important chapter in the history of natural language processing (NLP) technologies. The persuasive abilities of Amazon’s Alexa and the query-driven perspicacity of Apple’s Siri derive from advances that begin with postwar chatbots designed to simulate psychoanalysis. Massachusetts Institute of Technology (MIT) computer scientist Joseph Weizenbaum designed perhaps the most famous early chatbot.<sup>7</sup> Named ELIZA, after Eliza Doolittle in George Bernard Shaw’s *Pygmalion* (1913), this conversational agent communicated fluently with humans by cleverly combining generative questions and evasive generalities. ELIZA was designed to imitate the psychoanalytic method,

3 Emilio Ferrara et al., “The Rise of Social Bots,” *Communications of the ACM* 59, no. 7 (July 2016): 96–104.

4 Douglas Guilbeault, “Growing Bot Security: An Ecological View of Bot Agency,” *International Journal of Communication* 10 (2016): 5004.

5 Sock puppet accounts are accounts that misrepresent the agents that operate them, whether human or nonhuman (e.g., bots). On influential uses of such technology in politics, see, for example, Marco T. Bastos and Dan Mercea, “The Brexit Botnet and User-Generated Hyperpartisan News,” *Social Science Computer Review* 37, no. 1 (February 2019): 38–54. See also Philip N. Howard, *Lie Machines: How to Save Democracy from Troll Armies, Deceitful Robots, Junk News Operations, and Political Operatives* (New Haven, CT: Yale University Press, 2020).

6 See Gina Neff and Peter Nagy, “Talking to Bots: Symbiotic Agency and the Case of Tay,” *International Journal of Communication* 10 (2016): 4915–4931.

7 Joseph Weizenbaum, “ELIZA—a Computer Program for the Study of Natural Language Communication between Man and Machine,” *Communications of the ACM* 9, no. 1 (January 1966): 36–45.

specifically the methods of Carl Rogers. What is often overlooked in writings about Weizenbaum's famous creation is ELIZA's debt to the preexisting work of Kenneth Mark Colby, a psychoanalyst-turned-computer scientist whose work preceded Weizenbaum's in print by several years.<sup>8</sup>

Colby's simulation of human psychopathologies paralleled a contemporary paradigm in cybernetics, which viewed computational models as analogs for the human brain.<sup>9</sup> Such models were espoused by psychologists contributing to the postwar Macy Conferences in cybernetics (e.g., Lawrence Kubie and Alex Bavelas).<sup>10</sup> The chatbot's origins lie with these postwar proponents of cybernetics who brought the insights of information theory to bear on biological models of intelligence, a project described by historian Steve Heims as proposing "formal models of the brain based on possible machines which can organize by using information, stored programs, communications, feedback loops, and instructions."<sup>11</sup> The merger of cybernetics and psychology in this period drove researchers "to understand the processes of perception, memory, and language in terms of formalizable transformations of information."<sup>12</sup> Yet while Colby's work led to important innovations in chatbot design, his aim in building artificial conversational agents had nothing to do with driving website traffic, automating service sector jobs, or even beating the Turing Test (all subsequent goals for which chatbots would later be employed).

In 1973, looking back on his efforts to bring the insights of psychology into postwar computer science's quest to develop artificial conversational agents, Colby describes his work as an attempt "to simulate human belief processes on a computer."<sup>13</sup> Before taking up computer science, Colby had been a practicing psychoanalyst, and his attempt to simulate "belief processes" combined therapeutic intervention with an investigation into the non-rational makeup of the human mind.<sup>14</sup> By what methods does the human mind develop, fix, and operate from its learned beliefs? Colby's methodology

8 Kenneth Mark Colby, "Computer Simulation of a Neurotic Process," in *Computer Simulation of Personality: Frontier of Psychological Theory*, ed. Silvan Solomon Tomkins and Samuel Messick (New York: Wiley, 1963). Although Colby published on these problems first, Weizenbaum's ELIZA program using an MIT computer may have been the first actually to have been run, a fact alluded to in Kenneth Mark Colby, James B. Watt, and John P. Gilbert, "A Computer Method of Psychotherapy: Preliminary Communication," *Journal of Nervous and Mental Disease* 142, no. 2 (1966): 148–152.

9 One classic example is W. Ross Ashby, *Design for a Brain: The Origin of Adaptive Behavior* (New York: Wiley, 1952).

10 See Claus Pias and Heinz von Foerster, eds., *Cybernetics: The Macy Conferences 1946–1953* (Zurich: Diaphanes, 2016).

11 Steve Heims, "Encounter of Behavioral Sciences with New Machine-Organism Analogies in the 1940's," *Journal of the History of the Behavioral Sciences* 11, no. 4 (October 1975): 372.

12 Paul N. Edwards, *The Closed World: Computers and the Politics of Discourse in Cold War America* (Cambridge, MA: MIT Press, 1996), 179–180.

13 Kenneth M. Colby, "Simulations of Belief Systems," in *Computer Models of Thought and Language*, ed. Roger C. Schank and Kenneth Mark Colby (San Francisco: W. H. Freeman, 1973).

14 For this phase of his career, see Kenneth Mark Colby, *A Primer for Psychotherapists* (New York: Ronald Press, 1951). Colby's growing doubts concerning psychoanalysis's efficacy turned him toward computational experiments. See Kenneth M. Colby, *A Skeptical Psychoanalyst* (New York: Ronald Press, 1958).

ran counter to two major trends in the hard sciences: he subscribed neither to behaviorism in his psychological model nor to the logical, rule-based paradigm of Cold War computer science. Colby's numerous attempts to automate psychiatric intake interviews (between a psychiatrist and an artificial patient) and psychoanalytic sessions (with artificial analysands) were therapeutically inconclusive. Although the artificial conversation technologies that Colby helped pioneer would lead to the origin of the chatbot, Colby's final single-author book was not on artificial conversational agents but on chess.<sup>15</sup>

In the proceedings from a 1962 conference on the Computer Simulation of Personality at Princeton University, Colby discusses his work within the emerging possibilities of computational belief networks: in his model, "[b]eliefs are the molecular units of information processing" and "beliefs are organized into complexes. A complex is a list of beliefs which are related to one another according to criteria of relevance."<sup>16</sup> As Colby explains elsewhere, "A belief is considered to be both an emotion and an idea," and "[a]t the level of social psychology, we are interested in belief systems and how they operate in generating thought."<sup>17</sup> Colby's insight was that to successfully imitate human conversation it would be necessary to explore how human *irrationality*, not human rationality, guides people's interpersonal and social interactions. Colby's resulting program PARRY thus sought to simulate paranoid chains of inference as drawn by human psychiatric patients.<sup>18</sup>

Such a technological watershed moment should not be divorced from the larger techno-theoretical and political aims that work such as Colby's served: a Cold War computational turn toward measuring, calculating, and *gaming* irrationality, affect, and the intuition of human groups formed on this era's horizon. Writing a new computational history for the chatbot demands reconsidering how irrationality was technologically captured (if not rationalized) as a metric for human sociality. The pervasively successful myriad of chatbots of today serve as opportunities to consider how linguistic inter-relationship—that human artifact called "conversation"—shares a genealogy with Cold War concerns.<sup>19</sup>

What is the technopolitical status of conversation in the twenty-first century? In 2009, political theorist Jodi Dean outlined the shift to "communicative capitalism" as perpetuated by "changes in information and communication networks associated with digitalization, speed (of computer processors as well as connectivity), and memory/storage capacity [that] impact capitalism and democracy," fusing capitalism with information technology and preexisting networks of human conversation.<sup>20</sup>

15 Kenneth Mark Colby, *Secrets of a Grandpatzer: How to Beat Most People and Computers at Chess* (Malibu, CA: Malibu Chess Press, 1979), 256. Colby co-authored two subsequent books on psychiatry and psychoanalysis, respectively.

16 Colby, "Simulations," 167. Figure from Colby, "Computer Simulation," 171.

17 Kenneth Mark Colby and John P. Gilbert, "Programming a Computer Model of Neurosis," *Journal of Mathematical Psychology* 1, no. 2 (July 1964): 406, 416.

18 Kenneth Mark Colby, Sylvia Weber, and Franklin Dennis Hilf, "Artificial Paranoia," *Artificial Intelligence* 2, no. 1 (1971): 1–25.

19 I discuss this topic at greater length in my book manuscript, "Apparatus Poetics."

20 Jodi Dean, *Democracy and Other Neoliberal Fantasies: Communicative Capitalism and Left Politics* (Durham, NC: Duke University Press, 2009), 23.

Under communicative capitalism, the use-value of messages is eclipsed by their exchange-value.<sup>21</sup> In this, Dean presages the arrest or rewiring of the ethos of political speech in general—a change in what we expect our speech to *do*, what we expect will come of our gestures at participation. As digital infrastructures algorithmically incentivize and guide conversation in increasingly corporate social forums, the appearance of free and open dialogue remains vital to the economic model this mode of conversation affords. Even as this “registration effect” of users’ speech acts promotes a “fantasy of participation,” the technologies themselves become “exquisite media for capturing and reformatting political energies . . . reinforcing the hold of neoliberalism’s technological infrastructure.”<sup>22</sup>

Yet despite the tendency to *capture* and *reformat*, there is something fundamentally novel within this infrastructure. Chatbots reveal not the limit-case of conversation in this late capitalist epoch but, rather, one of its most central paradigms. Chatbots and their related language-processing technologies are a machinic infrastructure thriving within a matrix of earlier information technologies designed merely to transmit communication, to convey messages from one node to another. Artificial conversational agents, or bots, do not simply substitute themselves for humans. Whether bots participate in known or unknown ways, in every case they *impersonate* the community they transmit. This machinery speaks *as if* on our behalf; it conveys us back to ourselves, irrationalities and all. With each passing year, it becomes more and more difficult to converse without it.

The computational invention of the chatbot gives an alternative view of how that smallest political unit in the postwar Euro-American consensus was engineered: the monadic, bounded, and auto-managerial consumer subject of neoliberalized societies. Beginning from these experiments in the “computer simulation of human personality,” the chatbot, as a computational being grafted into a feedback loop of conversation with humans, moves from the *therapeutic* to the *transactional*.<sup>23</sup> The linkage between these two versions of the chatbot remains perceptible in the many devices that make up the Internet of Things and quietly attend to our consumer habits by surveilling, nudging, recommending, and driving our desires toward certain products. Having transitioned from automated therapist to seamlessly integrated commodity-consultant, the story of the chatbot relates one way in which, in a digital age, habits become networked to markets.

The chatbot serves the economic agenda that neoliberal economist Gary Becker envisioned when he insisted that the consumer does not simply consume but rather *produces* desire. Market demand is not an ontologically available substance; like anything else, it must be generated. How to control, or steer, this facet of the economic process? As Michel Foucault observes:

21 Dean, 26.

22 Dean, 30, 31–32.

23 Therapeutic uses of the chatbot, of course, continue and have seen a resurgence of research in recent years. For a review of contemporary implementations, see Eliane M. Boucher et al., “Artificially Intelligent Chatbots in Digital Mental Health Interventions,” *Expert Review of Medical Devices* 18, no. S1 (2021): 37–48.

[I]t means generalizing the “enterprise” form within the social body or social fabric; it means taking this social fabric and arranging things so that it can be broken down, subdivided, and reduced, not according to the grain of individuals, but according to the grain of enterprises. The individual’s life must be lodged, not within a framework of a big enterprise like the firm . . . but within the framework of a multiplicity of diverse enterprises connected up to and entangled with each other, enterprises which are in some way ready to hand for the individual. . . . And finally, the individual’s life itself . . . must make him into a sort of permanent and multiple enterprise.<sup>24</sup>

Although this quotation is taken from Foucault’s writings on the rise of neoliberalism, it is striking how easily it could serve as copy for a tech brochure for a chatbot application programming interface.

Addressing the feedback loop between the consumer and the information economy, Paolo Virno notes that “*language itself has been put to work.*”<sup>25</sup> Virno’s writings on the economic value of language’s automation mark a threshold between the Third and Fourth Industrial Revolutions. In the Third Industrial Revolution, the information economy transformed language-use into “wage labor.”<sup>26</sup> In the Fourth Industrial Revolution, Web 2.0 treats language in the digital world as a plastic, cultural infrastructure, a socio-technical ligature traversing the distance from node to node, agent to agent—the stuff of networks. This twenty-first-century socioeconomic configuration of language as connecting routes between users is emblemized by the chatbot, a technology that embeds the dynamics of interpersonal conversation into a host of profit-driven spaces.

The market-based yet seemingly intimate dynamics that emerge between humans and artificial conversation agents—from chatbots to Alexa—suggest a different set of problems than those that otherwise dominate the conversation around artificial intelligence technologies replacing human workers by rendering their skills obsolete. Chatbots do not outmode the humans with whom they communicate; quite the opposite. Indeed, chatbots are designed to insinuate themselves into preexisting dynamics of human behavior (whether conversational or domestic). For this technology to be profitable, the humans must be kept in the loop since the humans are the consumers. Rather than replacing human skills, chatbots slowly alter the relational status of conversation as such for their human interlocuters.<sup>27</sup> Certainly the chatbot relation connects human and nonhuman, but, in its most pervasive and ubiq-

24 Michel Foucault, *The Birth of Biopolitics: Lectures at the Collège de France, 1978-1979*, trans. Graham Burchell, ed. Michel Senellart (London: Palgrave Macmillan, 2008), 241.

25 Paolo Virno, “Notes on the General Intellect (1990),” in *Marxism beyond Marxism*, ed. Saree Makdisi, Cesare Casarino, and Rebecca E. Karl (New York: Routledge, 1996), 271.

26 Virno, 271.

27 See Lucy A. Suchman, *Human-Machine Reconfigurations: Plans and Situated Actions*, 2nd ed. (Cambridge: Cambridge University Press, 2006); see also Sherry Turkle, *Reclaiming Conversation: The Power of Talk in a Digital Age* (New York: Penguin, 2015).

uitous cases, the chatbot more saliently represents and facilitates the relation between human and marketplace.

As automated personal assistants now saturate the interstices of daily life, their answers to our queries simultaneously profile us as consumers. Their disingenuously guided tours into commodity fetishism emerge at every opportunity during online interactions yet are not driven by any centrally planned calculations. The recommendations, instead, result from scrupulously attentive *listening*: uncannily tailoring themselves to conversations that human users were just having on the phone with a friend or within the walls of their living spaces. What began as the attempt at automating therapeutic conversation in the 1960s culminates in the present with a vastly different form of dialogue. This is not simply what we might call *retail therapy* but rather the commodification and monetization of the conversation as such, for it reroutes the linguistic call-and-response so definitive of human relationality in ancestral ways. As *dialogue* becomes *chat*, human conversation succumbs to metrics such as those used by chatbot design teams who “define conversation-turns per session (CPS) as the success metric for social chatbots.”<sup>28</sup>

The chatbot redirects primordial human desire for *social* recognition into “alternative means of engagement” with this desire. The bot is a more reliable producer of this fundamental social desire than any human community of users—a central paradox of the chatbot. The chatbot, as a steerable producer of social affirmation and the recognition of personhood, is also (tellingly) unadmitted to the very human community it affirms. Simulating the force of social recognition as a commodifiable service, chatbots offer conversations more *consumed* than *participated in*. Even without the explicit use of chatbots, every computerized device we use hosts parallel, implicit conversations, every word, gesture, click-through being tracked, calibrated thoughtfully to build ever better models of what kind of “user” or consumer we are. By altering the relational premises of conversation, chatbots encourage or amplify a certain structure of desire: conversation without any Other. The future of conversation presents alternatives to alterity. What should be our response?

**Avery Slater** is an assistant professor with the University of Toronto’s Department of English. Her recent work can be found in *New Literary History*, *IEEE*, *symplokē*, and *Amodern*, as well as *The Oxford Handbook of Ethics of AI* (2020) and *Saturation: An Elemental Politics* (2021).

28 Shum, He, and Li, “From Eliza to Xiaolce,” 16.