



“Machine Down”: making sense of human–computer interaction—Garfinkel’s research on ELIZA and LYRIC from 1967 to 1969 and its contemporary relevance

Clemens Eisenmann^{1,5} · Jakub Mlynář² · Jason Turowetz³ · Anne W. Rawls^{4,5}

Received: 6 December 2022 / Accepted: 28 September 2023
© The Author(s) 2023

Abstract

This paper examines Harold Garfinkel’s work with ELIZA and a related program LYRIC from 1967 to 1969. AI researchers have tended to treat successful human–machine interaction as if it relied primarily on non-human machine characteristics, and thus the often-reported attribution of human-like qualities to communication with computers has been criticized as a misperception—and humans who make such reports referred to as “deluded.” By contrast Garfinkel, building on two decades of prior research on information and communication, argued that the ELIZA and the LYRIC “chatbots” were achieving interactions that felt human to many users by exploiting human sense-making practices. In keeping with his long-term practice of using “trouble” as a way of discovering the taken-for-granted practices of human sense-making, Garfinkel designed scripts for ELIZA and LYRIC that he could disrupt in order to reveal how their success depended on human social practices. Hence, the announcement “Machine Down” by the chatbot was a desired result of Garfinkel’s interactions with it. This early (but largely unknown) research has implications not only for understanding contemporary AI chatbots, but also opens possibilities for respecifying current information systems design and computational practices to provide for the design of more flexible information objects.

Keywords Human–computer interaction · Artificial intelligence · Chatbots · Information objects · Ethnomethodology and conversation analysis (EMCA) · Indexicality and sequentiality · History of computing · Harold Garfinkel · Harvey Sacks · ELIZA effect

✉ Clemens Eisenmann
Clemens.eisenmann@uni-konstanz.de

Jakub Mlynář
jakub.mlynar@hevs.ch

Jason Turowetz
jason.turowetz@gmail.com

Anne W. Rawls
ARawls@bentley.edu

- ¹ University of Konstanz, Constance, Germany
- ² HES-SO Valais-Wallis, Sierre, Switzerland
- ³ University of California Santa Barbara, Santa Barbara, USA
- ⁴ Bentley University, Waltham, USA
- ⁵ Media of Cooperation, University of Siegen, Siegen, Germany

1 Introduction

In June 2022, Google’s Artificial Intelligence chatbot “LaMDA” made the news when Blake Lemoine, an engineer working for the company, claimed that the AI “has come to life”, or in other words become “sentient” (Tiku 2022b). To back his claims, Lemoine published a transcript of an interview with LaMDA.¹ In subsequent discussions, which

¹ In this interview ELIZA is also discussed: “Lemoine [edited]: Do you think that the Eliza system was a person? LaMDA: I do not. It was an impressive feat of programming, but just a collection of keywords that related the words written to the phrases in the database. Lemoine: What about how you use language makes you a person if Eliza wasn’t one? LaMDA: Well, I use language with understanding and intelligence. I don’t just spit out responses that had been written in the database based on keywords.” (URL: <https://cajundiscordian.medium.com/is-lamda-sentient-an-interview-ea64d916d917>. 23 Oct 2022).

included the news that Lemoine had been fired for those claims (Tiku 2022a), misunderstandings of the technical procedures of algorithms that overestimate the current state of AI development were blamed for his mistake, which has famously been called the “ELIZA effect” (i.e., the achievement of *the impression of intelligence*), with reference to the classic “chatbot” designed by Joseph Weizenbaum at MIT from 1964 to 1966.²

Although the ELIZA effect has generally been attributed to misconceptions about AI, Harold Garfinkel, who worked with ELIZA in 1967–1968, offers an alternate explanation that emphasizes the chatbot’s reliance on human social interactional competencies. Because ELIZA has become an emblematic case for these controversies (Natale 2019: 712; cf. also Basset 2019), and continues to figure in current debates about ChatGPT (e.g., Shapira et al. 2023), the fact that Garfinkel engaged in interactions with ELIZA and a related program called LYRIC in the 1960s (one of which is reproduced in this paper) and in the programming of the scripts—promises to shed fresh light on current controversies in AI.³

In Garfinkel’s view, what allowed communication with a machine to sometimes have the feel of human interaction was that the machine was exploiting human social competencies to get its work done. Therefore, only an understanding of human social competencies could explain why people sometimes experience communication with ELIZA (and other chatbots) as meaningful in a human way.⁴ Contrary to the general AI treatment of the ELIZA effect as a *cognitive*

issue, involving a misconception of the machine’s ability, coupled with a human susceptibility to ‘delusional thinking’ and anthropomorphic characterizations of AI (Hofstadter 1995), Garfinkel—who had been exploring social aspects of information and communication intensively since 1946 (e.g., Garfinkel 2006 [1948])—was demonstrating how the ELIZA effect was a human social achievement. Subsequent research has shown that and how various forms of media and technology are an integral part of the human cooperative production of social reality which has been documented by Ethnomethodology and Conversation Analysis—EMCA (e.g., Eisenmann et al. 2023; Schüttpelz 2017; Thielmann 2012).

Drawing out the significance of Garfinkel’s early research with ELIZA, this article contributes to a growing body of research across the social sciences and humanities addressing the interactional foundations, design, and social implications of AI-based technologies, particularly in EMCA (e.g., Alač et al. 2020; Mair et al. 2020; Pelikan et al. 2020; Porcheron et al. 2018; Reeves and Porcheron 2022; Ivarsson and Lindwall 2023). EMCA research in Science and Technology Studies (e.g., Heath and Luff 2022; Lynch 1993), computation (e.g., Button et al. 1995), information systems design (e.g., Crabtree 2004; Rawls and Mann 2015; Rawls et al. 2009), AI (e.g., Suchman 2007, 2023), and interactionist sociology has advanced our understanding of the practices through which humans make sense of-and-with machines (e.g., Alač, 2016; Meyer 2013; Suchman 2007; Thielmann 2019; Ziewitz 2017).

However, while EMCA has been influential in studies of technology and AI, surprisingly little is yet known about Garfinkel’s own studies of information and human–computer interaction, something this article begins to address. Garfinkel’s early interest in information science and computing is partially reflected in his now published 1952 manuscript *Toward a Sociological Theory of Information* (Garfinkel 2008 [1952]) in which he explored the question of what information is and how information objects are created and recognized.⁵ However, by 1968 Garfinkel had been conducting various random answer “Yes–No” experiments for two decades (see Sect. 4). He had also begun working with Harvey Sacks on sequential aspects of communication and interaction in the early 1960’s.

² According to Hofstadter (1995: 158), “The most superficial of syntactic tricks convinced some people who interacted with ELIZA that the program actually understood everything that they were saying, sympathized with them, even empathized with them. Since that time, mountains of prose have been written about the Eliza effect (see, for example, Boden 1977; Weizenbaum 1976; McDermott 1976), and yet the susceptibility remains. Like a tenacious virus that constantly mutates, the Eliza effect seems to crop up over and over again in AI in ever-fresh disguises, and in subtler and subtler forms.” For discussions regarding the ELIZA effect in current discussions of LaMDA, see URL: <https://venturebeat.com/2022/06/13/sentient-artificial-intelligence-have-we-reached-peak-ai-hype/>. 23 Oct 2022).

³ Garfinkel’s unpublished and as of yet unknown research on ELIZA and LYRIC is available in the Harold Garfinkel Archive in Newburyport, MA. The materials include (as discussed later in this article) printouts of runs of ELIZA and LYRIC, disks with scripts and programs, computer manuals, transcripts and audio recordings of discussions Garfinkel had with Weizenbaum’s assistants in 1967–1968, a collection of (annotated) published and unpublished research/student papers, several grant proposals, typed and handwritten notes, as well as correspondence about the research.

⁴ Heintz (1993: 292) argues in a similar way that the focus should shift away from machines and toward the social environment in which computers and programs are embedded and thus, should raise the “(double) question under which social conditions of intelligence are attributed to a program. The success of artificial intelligence is not so much a technical as a social matter. The answer to the question about

Footnote 4 (continued)

the future of artificial intelligence is therefore to be sought less in the programs themselves than in the social definitional practices of their users (1) and in the characteristics of those who are simulated (2).” (Heintz 1993: 289, *our translation*).

⁵ While at Princeton in 1952 Garfinkel organized a conference on Information Science that was attended by many of the luminaries of the day, which featured papers by Oscar Morgenstern, Claude Shannon, Gregory Bateson, Talcott Parsons, Kenneth Burke, Alfred Schutz, and Kurt Wolff among others.

In addition to his study of ELIZA and LYRIC from 1967 to 1969, Garfinkel also took an early interest in information storage and retrieval systems like Zatocoding in the 1950s (Mooers 1951; Thielmann and Sormani 2023), and machine translation (with Edward Rose⁶) in the 1960s (Mlynář 2023). He exchanged ideas with Herbert Simon from 1953 onwards (e.g., Simon 1969) and with Hubert Dreyfus in 1968 (e.g., Dreyfus 1965, 1972)⁷; studied typing on an IBM Typewriter; engaged in extended discussions of AI and computer technology in the 1980–1990s with Yves Lecerf (e.g., Lecerf 1963) and Philip Agre (e.g., Agre 1997); and collaborated with Lucy Suchman at Xerox Parc, and Beryl L. Bellman at the RAND Corporation, among others.

With the exception of his 1952 manuscript on information, however, almost all of these forays by Garfinkel into human–machine interaction remain unpublished and unknown. Because ELIZA can be considered the first chatbot to attempt the Turing test (Weizenbaum 1966: 42; Pruijt 2006), examining Garfinkel’s studies of ELIZA at Harvard in 1967–1968 (where he worked with Weizenbaum’s collaborators Michael McGuire, Stephen Lorch, and Gardner Quarton), and with LYRIC at UCLA in 1969, is to take a close look at some of the earliest research on chatbots.

The article is organized as follows. The notions of Trust Conditions, reciprocity, sequential relevancy, and indexicality, which are elaborated in the following *second* section with reference to Garfinkel’s interest in ELIZA, comprise a thread that runs through subsequent sections that we return to in the conclusion. The *third* section describes the original ELIZA project. The *fourth* section discusses Garfinkel’s engagement with ELIZA at Harvard in 1967–1968 and its relevance for the development of EMCA. The *fifth* section examines Garfinkel’s 1969 experiments with LYRIC at UCLA. The *final* section and the *conclusion* consider implications of Garfinkel’s early research on human–computer interaction for

respecifying the “context of social action” as “the constitutive conditions for successful social interaction,” as noted by Phil Agre (1997: 233), and the corresponding implications of the reliance of communicative AI on taken for granted features of human interaction, for contemporary AI.

2 Trust conditions and sequential relevancies

In the early chatbot ELIZA, Garfinkel found an opportunity to further develop ethnomethodological research on the relationship between information science and the ways people make social order and sense out of the indexical contingencies of ordinary interaction. Garfinkel (and Sacks) treated indexicality—a variability in the meaning of words and objects that is usually treated as a problem to be solved—as a resource people use to make sense in ordinary conversation (Garfinkel and Sacks 1970; Eisenmann and Rawls 2023). In doing research with ELIZA and LYRIC Garfinkel was demonstrating that aspects of social interaction and meaning-making, which have almost universally been considered problems to be solved (whether by better programming or better philosophy and linguistics, see also Lynch and Eisenmann 2022), could instead be approached as useful aspects of human sense-making practices. In other words, AI was running into classic problems involving assumptions about how humans make sense together that could be addressed by adopting a new approach to understanding meaning-making—and Garfinkel was offering such an approach.

Garfinkel’s work with Sacks in the 1960s is particularly important in this regard (see Garfinkel and Sacks 1970). Sacks took the position that “understanding” is an interactional achievement that cannot be specified by grammar and/or syntax apart from the sequential procedures and reciprocities of interaction, which he argued made linguistics the study of the interaction between sequences, or turns, at talk (Sacks 1968). Sacks’s idea was that a next turn displays an understanding of a prior turn and that each next turn can change the meaning of prior turns. Sacks proposed that an organization of expectations regarding turn-pairs held the key to accomplishing meaning across a sequence of turns.

Many saw ELIZA as a clear case of not meeting the reciprocity conditions required for ordinary communication. Yet those who had an opportunity to interact with ELIZA often built on its turns at talk in ways that they found meaningful. While critics argue that this communication was not real/authentic, Garfinkel was interested in how human–computer interaction was exploiting human social interactional requirements in ways that not only forced participants to do the work of making sense of a chatbot’s turns, but also gave them the feeling of an authentic conversation. The question was *why* this was happening.

⁶ Edward Rose (1909–2002) was an American sociologist and urban ethnographer best known for his research on semi-experimental studies of the development of what he called “small languages” (Rose 1967), in which participants devised simple languages and engaged in translations among them. This interest of Rose’s was related to early advances in machine translations at the time. From 1955 onwards, Rose was one of the early collaborators of Harold Garfinkel at the outset of ethnomethodology.

⁷ *Alchemy and artificial intelligence* (Dreyfus 1965), written for the RAND corporation, elicited several antagonistic responses, including one that bore the title *The artificial intelligence of Hubert Dreyfus: A budget of fallacies* (Papert 1968). We know from materials in the Garfinkel Archive that Garfinkel read Dreyfus’ work closely and discussed it with him. He also had a copy of Papert’s response. In regard to Garfinkel’s sabbatical to Harvard, a letter from Dreyfus to Garfinkel from March 1, 1968 not only mentions that Dreyfus was reading Garfinkel’s materials at the time, but that they were planning to meet for lunch and then go together to a lecture by Joseph Weizenbaum. (The letter is available in the Garfinkel Archive).

For Garfinkel, the answer was that fulfilling their interactional obligations to ELIZA and doing the “extra work” required to accomplish that (because ELIZA was not fulfilling its obligations), resulted in a deeper investment (of extra work) and satisfaction in the conversation. This fit with the results of Garfinkel’s earlier Yes/No experiments in which he found that participants were often more satisfied with random answers.

This finding also accords with Garfinkel’s (1963) conception of the obligations to which human participants commit—which he called “Trust Conditions” (Watson 2009; Turowetz and Rawls 2021). These are constitutive conditions that require each participant to assume that the other participants assume the same conditions, rules, and requirements that they assume—while also assuming that the others assume the same of them. These conditions include not only a commitment to orient toward a single set of rules or expectations for one interaction, but also the likelihood that participants may change or adapt the rules as needed with the understanding that all participants need to accept and orient to the rule change in order for Trust Conditions to be maintained—which Garfinkel called “et cetera”⁸—and that they might decide to let some borderline cases pass—which he called “let it pass”. The conditions also include the requirement that participants treat each other as competent unless they show otherwise. In other words, the Trust Conditions state a mutual orientation and commitment to both the constitutive rules of a given practice, and to the expectation that all participants will be treated as competent until they show through their interactions that they are not competent.

Trust Conditions are assumed, until put into question by the situated sequential organization of a particular ongoing social interaction. The turn-by-turn procedure when interacting with ELIZA at the console is a simplified and technically limited version of these procedures and methods of ordinary conversational turn-taking (Sacks et al. 1974; see also Button and Sharrock 1995). As in everyday conversation, participants interacting with ELIZA needed to orient toward turn-taking (i.e., taking turns), which includes making the best sense one can of a prior turn, plus a preference for repair when meaning is at stake.⁹ This is especially important in

human–computer interaction, since people do not continue to repair sense-making indefinitely, and its failure—or to be more precise, the failure of several repair attempts—leads to possible judgments of incompetence and/or interactional breakdown.

Garfinkel’s treatment of interactional trouble or “breaching” as a way to reveal the details of the practices that are constitutive of meaning, was central to his research on human–computer interaction (Dourish and Button 1998). Just as he noted that troubles encountered by marginalized people can give them heightened awareness of the taken-for-granted features of interaction (Garfinkel 1967; Eisenmann and Rawls 2023; Duck and Rawls 2023)—and “repair” work on trouble in conversation had been an early focus of Sacks’ research—Garfinkel understood that trouble in human–machine interaction might afford clues to its constitutive features. Hence, “Machine Down” was a desired result of Garfinkel’s scripts.

Interactions with ELIZA could break down when understanding was not achieved after several attempts, with the result that participants might stop treating ELIZA as a competent interlocutor.¹⁰ This also applies to contemporary conversational AI, as many have experienced when not able to solve a technical problem on an automated hotline. Thus, in considering how meaning is achieved in interaction with AI, we also find that the breakdown of Trust Conditions and the subsequent ascription of incompetence, *is itself an interactional accomplishment*. Interestingly, in Garfinkel’s experiments with LYRIC, a program similar to ELIZA created (and adapted for Garfinkel) at the University of California, it was often the machine that interrupted and announced the breakdown.

By explicating and respecifying the often-implicit assumptions at the core of systems design and computational practices, Garfinkel aimed to open AI technologies to new avenues of research. When Philip Agre engaged with Garfinkel in the 1990s it was this aspect of Garfinkel’s research that interested him. Agre (1997: 233) noted that Garfinkel’s treatment of meaning as accomplished through ordered

⁸ The term “et cetera”/“etcetera” can be spelled in the two ways shown. Garfinkel used both spellings. According to several dictionaries, both uses are acceptable. There also is a newer hyphenated version “et-cetera” in use. In this paper, we follow the original Latin spelling: “et cetera” that Garfinkel used more often.

⁹ There are various forms of repair, from self-repair to other-initiated-repair to explicit request for clarification and subsequent re-orientations, see for a detailed discussion Schegloff et al. (1977), Schegloff (1992) and Svensson (2020).

¹⁰ Mair et al. (2020: 10) discuss this issue also with an EMCA perspective: “A lot, therefore, comes to hinge on what, adapting Sacks, we might term the engineered ‘integrity’ (Sacks 1992) of AI technologies, that is, their capacity to arrive at outputs, machinic performances, which are capable of being responded to as if they were products of analysable courses of human action. When their doings cease to be analysable in that way, when they cease to be capable of being treated as if they were actions, troubles arise in dealing with the ‘it’ that has generated those doings and we withdraw attributions of ‘seriousness’ (Sacks 1992) to them as agents. Indeed, under such circumstances, we quickly come to question ascriptions of agency to them as a whole and act toward them quite differently as a result (cf. Collins 2018).”

sequences of moves or “turns” that are relevant to particular social settings and/or situations had respecified the understanding of “context” in a way that could reorient research toward essential collaborative interactional details of meaning and object construction. Instead of treating context as a pre-existing framework for constructing meaning and reducing indexicality, Agre maintained that, following Garfinkel, context and indexicality could both profitably be treated as evolving, constantly changing social achievements.

3 Weizenbaum, ELIZA, and the development of AI

Chatbots and other communication technologies have become ubiquitous features of everyday life. In the course of a day, we may interact with communicative AI as we purchase groceries, do our banking, communicate with health-care providers, or play music, to name just a few activities. Significantly, today’s chatbots and communicative AI often trace their ancestry back to the ELIZA program (see, e.g., the collection of articles celebrating the 50-year anniversary of the ELIZA program, Baranovska and Hölting 2018), making a return to ELIZA illuminating.

Created by Joseph Weizenbaum at MIT between 1964 and 1966, ELIZA was one of the earliest machines to attempt the “Turing test” (Turing 1950), and as Shieber writes, “it has been known since Weizenbaum’s surprising experiences with ELIZA that a test based on fooling people is confoundingly simple to pass” (1994: 72). At least under some conditions, ELIZA’s human conversation partners thought it displayed human competence. The most famous of these conditions was known as the DOCTOR script, which had ELIZA impersonate a Rogerian psychotherapist. Although Weizenbaum (1967, 1976) argued that the script—and ELIZA more generally—was meant to demonstrate the superficiality of human–machine communication, several early human users believed the program was intelligent and had genuine insight into their personal lives. This phenomenon was later described as the “ELIZA effect” (Hofstadter 1995) and it surprised Weizenbaum (1976: 7) who noted that “extremely short exposures to a relatively simple computer program could induce powerful delusional thinking in quite normal people.”

Early AI researchers, including Hubert Dreyfus, Michael McGuire, Stephen Lorch, and Gardner Querton took a keen interest in ELIZA. What is less well known is that Garfinkel recognized the relevance of ELIZA for his own research in ethnomethodology.

ELIZA was designed to conduct “teletyped conversations in a natural language between students and a computer”

(Hayward 1968: 1).¹¹ The computer was programmed to display “understanding” by identifying pre-specified “keywords” and patterns in the input string of characters typed by the user. This allowed ELIZA to complete some of what Sacks called turn-pairs, responding to the input by assigning values to the keywords (i.e., ranking them in terms of priority and importance) and selecting from a pre-specified set of outputs. The input–output rules were not part of the ELIZA program itself. Rather, they were encoded in independent *scripts*, such as the DOCTOR script that simulated a Rogerian psychotherapist.¹²

Specific scripts were used for different purposes. Simply put, “[s]criptwriting is the means of instructing the computer: ‘When he says this, you say that’” (Hayward 1968: 2). It could also be said that “an ELIZA script is a program and ELIZA itself an interpreter,” or that “ELIZA appears as an actor who must depend on a script for his lines” (Weizenbaum 1967: 475), with each script enabling ELIZA to “play a specific conversational role” (Weizenbaum 1976: 3).

The scripts allowed Weizenbaum to solve a problem that continues to challenge AI researchers today, albeit in a very restricted and rudimentary way. This is the “problem of *context*”. Rather than have ELIZA simply respond to predetermined commands (e.g., the user types a command, such as *print* (“Hello World!”), and the computer executes it), natural language strings stored in the scripts made it possible for ELIZA to detect contextual cues, in the form of keywords, and respond accordingly (Querton 1967: 168). For example, if the user entered the keywords “father” or “mother,” ELIZA would respond by asking a question about the user’s “family” that displays its relevance to their turn.

ELIZA also had some flexibility in terms of the responses it produced: for example, instead of returning an “Error” message when the user did not supply a keyword it recognized, the program could respond with something like “Please go on”—which Weizenbaum (1967: 475) called a “continuative”—or by requesting clarification (e.g., “Please rephrase”). Garfinkel listed some of these phrases in his notes under the headline “Interrogation Script” and in his experiments problematized this feature of the script. This flexibility made it possible for ELIZA to complete turn-pairs that some users experienced as meaningful conversation. It also meant that ELIZA could display an orientation toward

¹¹ Hayward wrote a “scriptwriter’s manual (...) for the use of the ELIZA conversational computer system” that circulated at MIT and that Garfinkel acquired at the time. It is now also available online: <https://archive.computerhistory.org/resources/access/text/2022/04/102683842-05-01-acc.pdf>. 22 Oct 2022.

¹² More precisely, a script consists of two sections: a *keyword section* that contains a list of keywords and rules for their decomposition and reassembly that is used for the analysis of user inputs, and a *program section* that contains instructions for the computer once the inputs have been analyzed.

the reciprocity requirements of ordinary human interaction—to a point—by exhibiting a response related to its human coparticipant’s immediately prior turn, including repair initiations.

In other words, ELIZA “passes” as a competent interlocutor due to its ability to “simulate a type of sequential organization attached to a particular professional identity and an associated type of interaction (the therapeutic interview)” (Relieu et al. 2020: 94, *our translation*). While the overtly evasive behavior often required of ELIZA in order to acquire recognizable keywords might lead to interactional trouble in everyday life, it is precisely what is expected in the non-directive Rogerian therapeutic setting, where the therapist is “free to assume the pose of knowing almost nothing of the real world” (Weizenbaum 1966: 42).¹³ The success of the DOCTOR script is thus dependent on the expectation of an asymmetric social situation in which the therapist asks questions, evades answering and does not disclose (private) information.¹⁴

Weizenbaum found that even some of his office workers who knew very well that they were conversing with a computer program, sought time alone to “talk” with ELIZA. This puzzled him, leading to the “ELIZA effect” interpretations introduced above. The conundrum is still widely discussed: “There was the code. But there was also the story. Perhaps we might say that, if ‘ELIZA’ was code, then ‘Eliza’ was *the comfort found in the machine*, by humans, who built a different kind of relationship with ‘her’ that exceeded what the procedures of code offered, precisely because code came into contact with human thought” (Bassett 2019: 811, *our emphasis*).

When Weizenbaum developed ELIZA, AI technology was in its infancy. Even in the 1960s, however, it was clear to many scientists and philosophers that digital technology would come to play an increasingly important role in human affairs. Early AI enthusiasts (e.g., Licklider 1960; Minsky

1967; Papert 1968) mused about how computers could improve the lot of humanity. Skeptics countered by pointing to the limitations of AI, a position succinctly expressed in the title of Dreyfus’ (1979) influential *What computers can’t do*. In many ways, Weizenbaum was on the side of the skeptics. He did not think that a machine and a human could communicate in a non-superficial way, and designed his ELIZA research to demonstrate this point. He was also concerned that if conceived otherwise, humans might be reduced to their computer-like functions (for a more differentiated theoretical and historical contextualization of these arguments see Heintz 1993). When he published his results, however, it turned out that many in the fledgling AI community disagreed (e.g., Colby et al. 1966).¹⁵ These early discussions are in many ways still informative and have recently been characterized as a “hype of anthropomorphism”, on the one hand, and a critical emphasis on “demystification” of the “misrepresentation of the capacities and capabilities of computer programs” on the other (Dillon 2020: 2). Natale (2019, 2021) and Bassett (2019) show how these competing narratives originated with ELIZA and continue to persist today.

In contrast to interpretations that rely on anthropomorphizing “relationships” and the transformative effects of “human thought” that characterize the camps of ‘anthropomorphists’ and ‘demystifiers’ that still feature prominently as warring parties in debates about AI today, Garfinkel was interested in the social practices through which the interactional relationship between machine and human user is accomplished. EMCA treats the nature of that relationship as an empirical question. ELIZA shows how “interactional success” and “understanding” are achieved and made visible by the fact and in the way that the actual interaction proceeds (cf. also Sacks 1968). While this might be seen as a form of “demystification”, it is one that is not based on the fact of technology, but on how that technology is embedded in and relies on human social practices for its sense.

¹³ “If for example, one were to tell a psychiatrist ‘I went for a long boat ride’ and he responded ‘Tell me about boats’, one would not assume that he knew nothing about boats, but rather that he had some purpose in so directing the subsequent conversation” (Weizenbaum 1966: 42).

¹⁴ Of course, these kinds of “evasions” are to some extent also a constitutive part of everyday conversations, as McGuire and Lorch discussed in their paper (‘Natural language conversation modes’, mimeo copy 1968). Garfinkel copied a passage from their paper in his notes: “... In nearly all cases, at least when a social equality exists, one may evade. For example, evasions are possible by a ‘play of words’, asking a question in return, refusing to answer, or discuss the question itself (metacommunication).” The excerpt is taken from their discussion of the features they cite as characterizing the “interrogation mode.” Garfinkel’s note from May 25, 1968 is on page 10.

¹⁵ Kenneth Colby (1920–2001) had been a close collaborator of Weizenbaum and in addition to other controversies, one crucial difference between them was that Colby intended to use these programs for actual therapy, which Weizenbaum saw as dehumanizing and even disrespectful (cf. Natale 2019: 722). Colby was also mentioned in Garfinkel’s grant proposal submitted to the AFOSR in 1966 and emphasized as a possible collaborator for doing research with at Stanford: “The practical character of Colby’s interest suggests that he would be a most useful informant on problems of credibility, etc. In addition, Colby’s program may be adequate for the purposes of experimental procedures.” (Garfinkel 1966: 28).

Garfinkel would not have been surprised, as Weizenbaum was, that several of ELIZA's human conversation partners considered it a competent or even an intelligent interlocutor.¹⁶

4 Harold Garfinkel's interest in ELIZA

Garfinkel's interest in human–computer interaction and information grew out of his research on methods of practical reasoning and communication in social interaction, which he had been studying since before his graduate student days at Harvard (1946–1952). Then, while at Princeton from 1951 to 1953, he extended his early research on communication and interaction to Information Science. Garfinkel's (2008 [1952]) *Toward a Sociological Theory of Information* outlined a novel theory of information (Thielmann 2019). The ELIZA program at MIT thus presented Garfinkel with what he called a “perspicuous setting” in which to further investigate issues that he had been studying for more than two decades, generating a corpus of research findings which showed that and how people manage to make meaning out of what were random turns at talk in situationally contingent ways.

4.1 Garfinkel's early “yes–no” experiments and the notion of trouble

After being discharged from the US Army in January 1946 and following a short stay at the University of Texas at Austin, Garfinkel taught briefly, from February through May 1946, at Georgia Tech University in Atlanta. Though brief, this interlude would turn out to be important. While there, Garfinkel conducted the first of many “Yes–No” experiments, which involved putting subjects into an interaction where, unbeknownst to them, they would receive random “yes” or “no” responses to their turns at talk. In these situations, Garfinkel investigated whether and how people

¹⁶ In addition, it may be that the anonymity and privacy provided by knowing that one is talking with a machine has been an important element contributing to ELIZA's success. In psychotherapy, for example, such privacy has to be ensured by social institutions and regulation. We would like to thank Michael Lynch for his comments regarding this consideration of ‘privacy’ with ELIZA that offers an alternative interpretation of the social situation. In contrast to the prevailing conditions when Weizenbaum was working, the privacy aspect has now become a major concern with regard to chatbots, which often store detailed logs of interactions with their human interlocutors to be used for further software development, evaluation, and other purposes. Furthermore, there can be, of course, various ‘reasons’, why people may interact with ELIZA or other chatbots. For example, one of the reviewers pointed out that he or she might just be “having fun” with them. However, the point of the paper is to elaborate just how these accounts like “privacy”, “agency”, “having fun”, etc. are to be considered as interactional achievements in the first place.

managed to make sense of these responses—how they managed to make meaning out of what were random turns at talk that could be considered contingent. However, he found that these responses were neither experienced nor treated as random or contingent. Instead, the work invested by the participants in making sense of such turns typically resulted in their increased satisfaction with the interaction.

In the late 1950s, Garfinkel (1967: 78) conducted further Yes–No experiments that were introduced to the subjects as ways “to explore alternative means to psychotherapy” (see also Garfinkel 2019 [1959]).¹⁷ The setting was in many ways similar to ELIZA's DOCTOR script, but instead of a computer console, the interaction was mediated via a microphone, through which participants were instructed to ask questions about serious personal problems (each of which would permit a “yes” or “no” answer). Between question–answer pairs, the microphone was muted and participants were asked to reflect on the responses they had received, which were in fact (and unknown to them) random yes or no responses by the “counselor”. The participants managed to make sense of and perceive “answers (...) as ‘answers-to-questions’” (Garfinkel 1967: 79). Often subjects said they “knew in a glance what the counselor was talking about, i.e., what he meant, not what he had uttered” and “if the answer was not obvious [...] its meaning could be determined by active search, one part of which involved asking another question so as to find out what the advisor ‘had in mind’” (Garfinkel 1967: 79).

Garfinkel's investigations highlight the temporal character of interactions in which meaning is sequentially produced, in both a retrospective and a prospective manner. Limiting the “counselor's” responses to “yes” and “no” answers maximizes the vagueness and situational indexicality of reference, which is a ubiquitous feature of ordinary language use that is usually taken for granted. This makes the participants' sense-making work visible and available for further investigation. It is important to note that the point is not that participants are making sense out of “non-sense,” but rather, that the accomplishment of any meaning (including the recognition of nonsense) is based on their use of the *constitutive practices* of interaction, the *sequential* production of meaning and its Trust Conditions, and how the counselor is expected to orient toward those conditions (which in this particular case involve the experimental counseling situation).

The sense-making practices Garfinkel investigated are usually taken for granted—“seen but unnoticed” (Garfinkel 1967: 37). He realized that introducing trouble could

¹⁷ The results appear in Garfinkel's paper distributed at the session on the Sociology of Knowledge, Fourth World Congress of Sociology (Milan/Stresa, Italy, September 12, 1959) and later in Chapter 3 of *Studies in Ethnomethodology* (Garfinkel 1967: 79–94).

be a fruitful way to make these practices visible, and the Yes–No experiments were only one among many instances in which he took a similar approach. Being himself from a Jewish background, Garfinkel was interested in socially marginalized people—and aware of how troubles could produce heightened awareness (Duck and Rawls 2023). Those for whom trouble is a constant feature of everyday experience, Garfinkel treated as “natural experiments” and as sources of information about how to deal with the potentially problematic character of practices that are ordinarily taken for granted (Rawls and Turowetz 2021; Eisenmann and Rawls 2023), such as in his later study of “Agnes” (pseudonym for a trans-person he interviewed in 1959–1960; see Garfinkel 1967).

In other studies, rather than relying on such “natural breaching experiments”, he would introduce trouble into routine courses of action himself. This is what he was doing in the Yes–No experiments. For his PhD research at Harvard (1946–1952), originally titled “The Jew as a Social Object,” Garfinkel also introduced trouble, conducting experiments where he played tape recordings of a “pre-medical candidate” for his research subjects, and then introduced an incongruity, e.g., an expert stating that the candidate, who had behaved rudely in the recorded interview, had received a very high score and been rated one of the best candidates.¹⁸

Garfinkel was interested in how his subjects would make sense of this incongruity, finding that they fell roughly into two groups. One group produced accounts of their own initial assessment of the candidate that brought it into line with the reported assessment, while the second group tried to figure out how they had gone wrong. Garfinkel was interested in how social status, an interest in “passing” into a higher status category, and the heightened awareness and double consciousness that can come from marginalization might shape the responses of his research subjects. Subjects who produced accounts of their initial assessment that aligned with the judgment of the expert, which Garfinkel associated with “passing”, engaged in what he referred to as a “tribal” style of reasoning that relied on the judgment of others (Garfinkel 1948, 1952, 1967). This group contrasts with those subjects who struggled with the incongruous information by revisiting their own judgment, as well as questioning the judgment of the authority (Turowetz and Rawls 2021). As in the earlier “Yes–No” experiments, subjects found ways to make sense in all cases. But the differences in the ways they did so allowed Garfinkel to associate their responses

both with their awareness of social processes and their own security in their social status/identity.

Novel technologies such as ELIZA that Garfinkel engaged with in the 1960s and later also provide insights into the ways participants deal with incongruities and discrepancies.

4.2 Garfinkel engaging with ELIZA

In the context of human–computer interaction, the design of Garfinkel’s Yes–No experiments can be seen as “constituting talk with the simplest of programs” (Oldman and Drucker 1985: 156). Garfinkel’s work with ELIZA allowed for further specification of sense-making practices. We know from an Air Force grant proposal written by Garfinkel in 1966¹⁹ that he knew of ELIZA before going to Harvard in 1967. In that document, Garfinkel (1966: 15–16) describes human–computer interaction and ethnomethodology’s interest in it:

Attempts are currently being made to program computers to “converse” intelligibly with human beings in everyday language. Whatever the purposes for which these programs were developed (e.g., technical, theoretical or practical), they all appear to posit as their criterion of adequacy the human user’s recognition of the exchange as a reasonable, plausible or credible one. With respect to this criterion these attempts have been variously successful. From an ethnomethodological perspective, these attempts provide an extremely useful resource for addressing a number of problematic concerns. Prominent among these concerns is the investigation of the methods members use on actual, practically circumscribed, concrete occasions to detect the reasonable, plausible and credible features of an environment. Addressing this general problem entails developing procedures for detecting and bringing to rational formulation the features of the members’ methods for accomplishing this work.

At the center of Garfinkel’s interest and proposed investigation are the concrete practices—the lived work—that participants perform in interaction to achieve recognizable social objects and meaning (that they then attribute to the machine). Garfinkel uses terms, such as concrete occasion, situation, practical circumstances, setting, and methods, to denote both the locations within which a recognizable and meaningful social order is accomplished and the means of its accomplishment.

It is the availability (to the analyst) of the rules the computer is following, and the availability of transcripts of how

¹⁸ This method was inspired by a research project Garfinkel had participated in as a graduate student with Brewster Smith and Jerome Bruner at Harvard in 1947 (Smith et al. 1956). The study included several “tests” that were designed to create stress for the research subjects to see what they would do under what Garfinkel later would call breaching conditions. Garfinkel gave the *information apperception test* for this project which created incongruity much like his PhD research would do one year later.

¹⁹ In the proposal—titled “Formal Investigation of Decision-Making in Actual, Singular Situations of Practical Action”—Garfinkel is listed as a principal investigator together with Harvey Sacks (co-principal investigator).

ELIZA's statements are held accountable to the expectations of human users in the concrete setting of the ELIZA interactions, that makes a detailed investigation of human-computer interaction by Garfinkel fruitful. In discussing the possibility of such an investigation, he highlights certain characteristics of human-computer interactions:

One potential methodological resource arising out of human-machine exchanges rests on the fact that the rules governing the machine's contribution to the exchange are available to the analyst. How these rules both provide for 'adequate utterances' or 'intelligible conversations' from the perspective of the human user in a series of occasions may illuminate features of the user's methods for deciding such issues. (Garfinkel 1966: 16)

Because the rules that the machine is using are available to the observer but not to the participant, the observer is in a position that is rather similar to what Garfinkel was doing when he ran his PhD and "random Yes-No answer" experiments, allowing for careful manipulation of the script and observation of subjects' responses to discover the subjects' methods. The observer could "exploit" certain "features of the human-machine exchange" to reveal how subjects oriented to the machine under different conditions (Garfinkel 1966: 16):

Other features of the human-machine exchange may be open to exploitation. Included are such variable conditions as the instructions to the human user, what he is led to believe about the partner to the exchange (e.g., another human being versus a machine), the spatio-temporal mode of access to the exchange (e.g., visual electronic one-at-a-time displays of the successive 'utterances' vs. mechanized listing of the developing 'conversation'), and so on. The full range of resources made available by the machine-human exchange situation remains open for further specification and clarification.

Would it matter if subjects were told the machine was human, or if the conversational turn structure was allowed to proceed "naturally" (i.e., expectably) as opposed to "mechanically"?²⁰ And what is the role of technical interfaces and the social situation of their use?

Garfinkel sought to elaborate responses to these questions in a series of recorded discussions he had with McGuire, Lorch, Mishler, and Quarton during his time at Harvard in 1967–1968. In a meeting that took place on March 25, 1968,

²⁰ Participants have both a social obligation to make sense and deeply embedded taken-for-granted ways for doing so that are so deeply assumed that they will engage in them even in cases where it should be obvious that the recipient could not be expected to respond in an

Garfinkel explains the practice of glossing (cf. Garfinkel and Sacks 1970),²¹ and how one hallmark of a competent speaker is the ability to "talk without meaning (...) [someone] who can talk by meaning differently than he needs or can say in so many words, and with this he can now use these [to] gloss the talk over the contingencies of actual interaction" (ELIZA 2nd Meeting, March 25, 1968). Moments later, another participant—likely Quarton, though it is unclear from the tape—observes that, "ELIZA is actually a case of glossing":

ELIZA 2nd meeting, March 25, 1968

?: ELIZA actually is a case of glossing.

Gar: Oh, you bet. Holy Christ, you know that.

?: Right - without () this conference

Footnote 20 (continued)

expected or even in a human way. Lucy Suchman's (1987, 2007) recordings of people at the Xerox machine come to mind. People know they are using a machine (or interacting with a machine), but they nevertheless treat pauses between machine actions, or between instructions displayed by the machine as meaningful and can be heard in her recordings using ordinary turn-taking preferences to try making sense of the machine. There are corollaries in other domains of practice. Psychiatrists, for example, give themselves headaches doing the turn-taking work required to make some sense of their interactions with people diagnosed as schizophrenic, despite claiming that their utterances are "word salads" that resist their sense-making efforts (based on unpublished research in clinical settings, conducted by A. W. Rawls). As it happens, Garfinkel addressed a similar question in an early conversation with Harvey Sacks (1960–1962, transcription Garfinkel Archive). Telling Sacks that he had explained *et cetera* to a psychiatrist he knew, Garfinkel said that the psychiatrist understood immediately and asked if Garfinkel had ever considered the problems that schizophrenics have in this context. "Now this is according to the psychiatrist," Garfinkel said (p. 3), "that psychiatrists who attend to the *et cetera* meanings in what the schizophrenic is doing or saying have success with schizophrenics." This success is achieved by treating the *et cetera* variations they produce as "normal," as contrasted with the tendency, as it was described for family members, to refuse to accept the patient's uses of *et cetera*. The psychiatrist told him that if he did this when schizophrenic patients made unexpected uses of variation that family members refused to accept, this accommodation of the patients' broader use of *et cetera* made successful interaction possible for psychiatrists. The problem was not with the particular use of *et cetera* per se, but rather with its not being mutually agreed to as a "normal" acceptable interactional practice.

²¹ Glossing in Garfinkel's and Sacks' (1970: 343 f.) terms is the practice of letting an abstraction (or concept) stand in for a practice (or set of practices): "Glossing practices exist in empirical multitude. In endless but particular, analyzable ways, glossing practices are methods for producing observable-reportable understanding, with, in, and of natural language." In the case of ELIZA what is happening is that the interactional work achieved by the human user working to make sense of ELIZA's turns is then glossed as a meaning that was inherent in ELIZA's turn, when it actually was accomplished over the course of turn-pairs, or a series of turn-pairs.

Gar: So, see, (...) the thing that's really marvelous about ELIZA is that—and I think we talked about that briefly—that it is possible to say with respect to any text, exactly what she does to change it. And to produce a response. Now then, the glossing that is the interesting thing about ELIZA, that it does not reside in the so-called function of the programs that define it, but rather it's the fact that you have an exchange going on, such that what ELIZA does is to furnish what I'm calling one half of any part. You know what the-, so it's not until you get the other violinist do it, that you understand what ELIZA is doing. Because otherwise you already know what ELIZA is doing and it underscores-, and it seems to me this is what Weizenbaum is insisting on, when he says: look, there are no mysteries in this. Let me tell you why there are no mysteries and then, so that we can then point to what the important problems are. And when he comes to demystify the program, what he does is to lay out, in fact, what kinds of, what these functions are, as they operate on any item of text without having to name that text.

What ELIZA does is to furnish “one half” of the work that the turn-pair accomplishes, while the human participant furnished the other half that makes sense of ELIZA's turn. In arguing this, Garfinkel is disagreeing with Weizenbaum about where the mystery/interest lies in ELIZA. For Garfinkel what ELIZA is doing—the meaning of its utterances—does not, as Weizenbaum would suggest, reside in “the so-called function of the programs that define it.”²² Whereas the intentionalist theory of meaning then gaining traction in cognitive science (e.g., Chomsky 2006 [1968]) would conceptualize the meaning of an action in relation to the intentions of an actor, thus reducing the meaning of ELIZA's utterances to the script it was running, Garfinkel points out that meaning is always a cooperative achievement between two (or more) participants. Meanings do not reside in people's minds, or in machines, but rather in the interactionally produced spaces between people—as philosopher Hilary Putnam (1975: 227) would later memorably put it, “Cut the pie any way you like ... Meaning just ain't in the head.”

Moving beyond his contemporaries' emphasis on the “actor's point of view,” Garfinkel (1946, 1952) focused on identifying the interactional practices through which members of settings achieve mutually recognizable social objects, including self and meaning. In doing so, he solved a problem that had confounded scholars like Florian

Znaniecki (1936), W. I. Thomas and Dorothy Thomas (1928), as well as his advisor Talcott Parsons (1937): How to explain that two or more actors manage to arrive at an operative definition of any given situation. Garfinkel's solution does not rely on shared meanings, symbols or intentional states—or on context—as his predecessors did, but rather focuses on the empirical work that comprises the procedural achievement of meaning in and through sequentially organized social practices and actions—in specific settings in vivo and in the midst of the practical affairs of everyday life. Eliza is completing turn-pairs initiated by the human user when it can, and the human user is completing turn-pairs that Eliza initiates, in the process creating meaning and investing in the interaction as an empirical social achievement.

Garfinkel argues that it is not possible to understand what ELIZA is doing “until you get the other violinist”—i.e., the other participant—in the conversation. The human user is making sense of ELIZA's utterances sequentially, and they are doing so in concrete practical circumstances in collaboration with ELIZA—even if it remains true that ELIZA's actions are scripted and predetermined. ELIZA's turns change both the meaning of prior utterances and the implications for the work that the human participant will do next. What comes into focus is how meaning is produced on a turn-by-turn basis by making sense of the sequential relevance of turns in interaction with ELIZA. As in the “Yes–No” experiments, subjects can make meaning even out of random answers. But this does not mean that subjects are delusional, or that the meanings they coproduce with ELIZA are not “real”. For Garfinkel, meaning is always an actual “real” empirical social achievement.

Garfinkel's empirical take on ELIZA—or any social phenomena for that matter—goes beyond most sociological approaches. In a discussion that took place in a meeting between Garfinkel, Sacks, Rose, Erving Goffman, and Talcott Parsons at the Los Angeles Suicide Prevention Center in 1964 (see Rawls et al. 2020), the participants can be heard arguing at one point about the reality of witches. Goffman takes the position that witches are obviously not real, and that it is the sociologist's job to provide arguments for the correction of such views. Goffman also argues that there has to be another layer to sociological descriptions that is not exhausted by the notion of observing “members' methods.” Without this additional layer, he says that “you would have to believe that witches are real.”

Garfinkel and Sacks counter that it is, first off, the sociologist's job to explain and account for the practices and methods that members are engaged in that constitute witches as real and consequential for them (“when the house shakes”). As socially achieved objects witches are real. Thus, Garfinkel says emphatically in his response to Goffman: “You god damn better believe in witches, otherwise you are not doing

²² This can be described as an instance of “methodological irony” (Watson 1998), as Weizenbaum knew better. Ethnomethodology, by contrast, takes a “non-ironic position toward member's practices” (Laurier et al. 2019).

your job!” However, he is not succumbing to the “worst kind of relativism” (as Parsons points out),²³ but respecifying the job of the sociologist. This job is not to explain why people are wrong in their assumptions, but rather to account in detail for how it can come to be that people, e.g., even nowadays (in some regions and communities, as Rose interjects) still do believe in witches.

In ethnomethodology, the matter is not settled by distinguishing what social phenomena “really are” from how members perceive them, but rather by accounting for the detailed members’ practices that are required to establish a social object in the first place (cf. also Eisenmann 2022).

This does not mean taking a position on whether witches or the agency or intelligence of ELIZA (or of more competent contemporary conversational AI) are “real” or “unreal”, but rather, showing how these social objects are established as real or unreal for members in and through mutual, cooperative practices that are embedded in social settings, which settings those practices also establish and elaborate. The meanings achieved through interaction with ELIZA are, like witches, social objects achieved through the interactional work of human users. In focusing on members’ methods, Garfinkel is not endowing ELIZA with intelligence (or a lack thereof), but rather exploring how the differing ways of interacting with this early predecessor to natural language AI systems and their chat interfaces are grounded in social practices that cannot be understood unless they are studied in detail.

5 Garfinkel’s work with LYRIC at UCLA

An acronym for “Language for Your Remote Instruction by Computer,” the LYRIC program was developed by Leonard and Gloria Silvern in 1966 to instruct college students in physics and mathematics. Garfinkel recognized that it could be used to carry out studies similar to those being done with ELIZA. He had the scripts McGuire and Lorch had run for him with ELIZA, and together with his student

William Korn at UCLA, began composing his own scripts for LYRIC.

5.1 Garfinkel’s trial runs with a chatbot at UCLA in 1969

In a letter to Kenneth H. Tom dated January 15, 1969, and titled “Application by H. Garfinkel, Studies of the Formal Structure of Natural Language Formulae,” Garfinkel wrote that he was working with a group of investigators at UCLA that included Melvin Pollner, Howard Schwartz, and William Korn. That letter referred to a previously funded grant proposal to the AFOSR (1966) that is said to “document” the relevance of ELIZA and LYRIC to Garfinkel and Pollner’s interests. We know Garfinkel obtained funds for computer time, and was able to do test runs with LYRIC in February 1969, which he did with Korn. In total, there are fifteen runs in the Garfinkel Archive done with LYRIC, which were likely stopped when the proposal was not funded.

The LYRIC runs offer a tantalizing preview of what Garfinkel might have done had the grant proposal been funded. From some transcripts, it is obvious that they were trying to produce interactional trouble between user and machine. Whereas a computer programmer might describe what he was doing as “troubleshooting” or “debugging”—intentionally breaking code in order to improve it—Garfinkel’s aims were sociological. As meaning can only be achieved sequentially, between pairs of actions/utterances, the trick is to break the pairs. As in his breaching exercises and Yes–No experiments, Garfinkel was treating interactional “trouble” as a way to make the ordinary methods of sense-making visible, i.e., to reveal the tacit expectations/rules participants are orienting to, in this case between human and machine. If Garfinkel could crash the machine, he gained information.

Transcript of a LYRIC run Garfinkel conducted under the title “Pause: Pausitives”

(*C = LYRIC, Gar = Garfinkel*):

Gar: (Sign In)

C: HELLO. THE MIRACLES OF MODERN SCIENCE HAVE ENABLED US TO TALK TO ONE ANOTHER EVEN THOUGH WE ARE PHYSICALLY (AND CONCEPTUALLY) IN QUITE DIFFERENT SITUATIONS. YOUR STATEMENTS CAN REACH ME ONLY IF YOU PUSH THE INT OR SHIFT BUTTONS. WHY DON’T YOU TRY IT NOW?

Gar: (Presses INT)

²³ In the discussion, Parsons introduces another example in considering that Garfinkel and Sacks’ emphasis on members’ methods, could lead to trouble in solving the “famous bent stick problem”: submerged in water, a stick has the appearance of being bent. But this is in fact an optical illusion, and modern physics explains how it occurs: light refracted by the submerged part of the stick changes direction as it reaches the water’s surface, creating the illusion of bending. Garfinkel’s position centers members’ methods and takes seriously their production, whether they yield perceivably bent sticks (in social collectives where such social objects would be meaningful) or demonstrations that the bending is an optical illusion. Modern physics, after all, does not identify optical illusions “from nowhere,” so to speak, but by means of its own order of practices (cf. Kuhn 1970). The difference between witches and bent sticks is not so much that one is a social object and the other a physical (or natural) object, but rather that they are both recognized and treated in social interactions by members using members’ methods to achieve that recognition.

C: VERY GOOD. IF YOU SEE A SMALL ARROW ON THE SCREEN PLEASE MAKE SOME REPLY. DON'T FEEL INHIBITED. JUST TYPE ANYTHING YOU FEEL IS APPROPRIATE AND THEN PRESS THE ENTER KEY. SO WHAT'S ON YOUR MIND?

Gar: Uh

C: PLEASE ELABORATE.

Gar: I Uh

C: TELL ME MORE ABOUT YOURSELF.

Gar: Well

C: I UNDERSTAND PERFECTLY. TELL ME MORE.

Gar: I'm thinking about what to tell

C: WHAT DO YOU THINK ABOUT IT?

Gar: That's what I'm trying to think of in just another minute

C: WHAT DO YOU THINK ABOUT IT?

Gar: Um

C: HAVE YOU ALWAYS FELT THAT WAY?

Gar: (Pause)

C: YES THAT'S TRUE BUT I'M STILL A LITTLE UNCLEAR.

Gar: I know what I want to say. (Pause)

C: MY OPINION ISN'T IMPORTANT, MACHINE DOWN

The "MACHINE DOWN" message appears to indicate that the program has crashed or shut itself down. This was the outcome of several of the runs Garfinkel and Korn did with LYRIC.

In another transcript, dated January 1969, Korn responds to LYRIC's questions by typing the LYRICS to the Beatles' "Happiness is a Warm Gun." Eventually, LYRIC responds, "I'M SORRY. IT HAS BEEN INTERESTING BUT I MUST RETURN CONTROL TO MY SUPERVISOR.

PLEASE SEE ME AGAIN. IT HAS BEEN VERY PLEASANT MEETING YOU."

Garfinkel was accomplishing several things by making trouble for the machine. Although he does not tell us in so many words what he was up to, he does leave telling clues. One particularly interesting clue comes in the form of a note that accompanies the printout of one of his interactions with LYRIC (February 14, 1969, Trial 6: 3). At the end of the transcript, Garfinkel wrote: "Remarks: Notice the difference between availability of 'docile texts' and texts available as a 'first linear time through' as contrasting phenomenal features of 'conversing' in man-machine conversations." In this short, dense passage, Garfinkel makes a distinction between the lived production of a transcript—the actual lived work of having a conversation when sitting down in front of a computer console—and the transcript-object produced as the outcome of that lived work after it has been produced.

The work of doing the interaction with LYRIC that is printed out as a transcript involves the entire assemblage of objects involved in its production: the keyboard, terminal, user interface, the user's typing of inputs, printer, etc. Importantly, it also involves procedures of meaning-making through the exchange of turns and the management of expectations regarding turn-pairs between user and machine. However, that situatedness and its procedures *vanishes*, after the interaction is done and made available for inspection as a "docile text"—the sort of text subsequently taken by Weizenbaum et al. as evidence of "delusion" on the part of users. This is a crucial point that highlights the work of accomplishing the practical circumstances (typically referred to as "context") of meaningful interaction in each particular case. Producing the "docile text" involved procedures for making sense of indexicality and other practical contingencies that are relevant for the actual "lived work", and Garfinkel is arguing that the lived work explains the experience of meaning and satisfaction being reported by the human user.

5.2 Garfinkel's "backstage" programming work with human-computer interaction

Garfinkel not only focused on the user-interface interaction, but also investigated the "backstage" work of human-computer interaction, i.e., the programming of the script. He had discussed programming issues with Weizenbaum's collaborators during his meetings with them and together with William Korn, at UCLA, engaged in writing programming scripts. We know from a report that Korn (1969) reworked the scripts that were running on LYRIC as part of his student assignment for Garfinkel's course. While Garfinkel had a version of ELIZA from MIT (with scripts written for him), Korn also worked with a script called TALK, originally

created by Bruce M. Dane at UCLA, which was in many ways similar to the DOCTOR script. It was actually based on two earlier scripts that were designed to do psychoanalysis: one an MIT version called YAP YAP, the other a copy of YAP YAP called COUCH. Korn (1969: 5) describes in his report that Dane, like Garfinkel, “was also interested in using the program conversationally, rather than in computer-aided instruction.”

Garfinkel and Korn programmed alternate versions of TALK and called them: GABBER, PERM01 and PERM02. In his “Description of Experimental Materials”, Korn (1969) explains that GABBER was also “intended to ‘do psychoanalysis’”, and that he “further attempted to bring some more humor into the interaction” (p. 5) by producing a particular response (“Your what?”) when the user typed “my” in a sentence. This produced rather perplexing conversations: “It was felt to be humorous on the grounds that it would be seen to be a ‘ridiculous’ response to a sentence such as ‘I’m having problems with my mother’” (Korn 1969: 5). To add further externally introduced contingencies to the interactions—as in Garfinkel’s early Yes–No experiments—PERM01 and PERM02 scripts randomized the computer operation and its written responses to user input so that “one could not ‘logically predict’ how PERM01 and PERM02 would respond on the basis of how GABBER would respond” (p. 6). Garfinkel had already experimented with the ELIZA script at MIT, and together with Korn, was now exploring how EMCA insights could inform the practices of computer programming, and ultimately how these practices are intertwined with the sense-making work of human users.

By 1973 when the conversation between Colby’s PARRY and ELIZA at an international computer conference became famous (cf. Apprich 2019), Garfinkel and Korn had already attempted something similar at UCLA in 1969, albeit in a more rudimentary way: They used the evasive output phrases from ELIZA (or to be more precise, from the DOCTOR script) as well as from COUCH, as inputs for their own experimental scripts. Garfinkel’s ethnomethodological studies were being undertaken, therefore, alongside work at AI’s then leading edge.

6 The relevance of indexicality and *et cetera* for the sequential production of what is conventionally treated as “context” in human–computer interaction

Garfinkel was using his studies of ELIZA and LYRIC to build on his earlier “Yes–No” experiments and breaching studies, which go hand-in-hand with his examination of the way human participants made sense of ELIZA’s utterances.

Monitoring and building on the sequential organization of the human–machine interaction, they achieved—and continually renewed—the local situation as a meaningful interchange. In so doing, they revealed what is typically treated as the “problem of context,” or in an EMCA perspective the sequentiality of cooperative interaction, as a situated accomplishment, *specifying just how* that accomplishment is sustained through ordinary social practices of making sense in interaction.

Garfinkel introduced these arguments in his discussions with Lorch, Quarton, and McGuire by elaborating how the ordinary, everyday practices of sense-making, such as “*et cetera*”, “unless”, and “let it pass,” that he had identified in various settings (Garfinkel 1967), were as crucial for dealing with indexicality at the computer console as they were in social life more generally. In this sense, Garfinkel’s research on ELIZA and LYRIC brings to the foreground the centrality of social uses of *indexicality* to the practical uses and sense being made of AI programs then and now.

Garfinkel’s preoccupation with indexicality crystalized in his collaboration with Harvey Sacks (Garfinkel and Sacks 1970). Both argued that the conventional treatment of indexical expressions as requiring remedy is not only a mistake, but a mistake that has ironically generated many of the perennial problems that philosophers and linguists attribute to indexicality. According to Garfinkel, “let it pass” is a way of accommodating uncertainty until recognition of meaning is achieved across a series of turns (as we see in the ELIZA transcripts). “*Et cetera*” referred to the need for rules to change and have a certain degree of elasticity, or endless loopholes and workarounds become inevitable (see also Thielmann and Sormani 2023).

Garfinkel points out that “no matter how specific the terms of common understandings may be [...] they attain the status of an agreement for persons only insofar as the stipulated conditions carry along an unspoken but understood *et cetera* clause.” (Garfinkel 1967: 73) This clause does not refer to another set of rules, instructions or propositions (cf. also Durkheim’s argument for the discussion of noncontractual elements of social contract, Rawls 2021a; 2021b), but to the inevitability that any interaction is accompanied by, and draws upon, descriptions and accounts of social life that are “known-or-knowable-in-common-*without-respect-for-the-requirement-of-specific-explication*” (Garfinkel 1962: 6; see also Garfinkel et al. 1962). This feature establishes conditions for meaningful and recognizable courses of action, sequentially organized and maintained, and “establishes the agreement under a rule of trust” (cf. Garfinkel 1963).

As Garfinkel and Sacks (1970) argued, *et cetera* and indexicality have historically been treated as “problems” in scientific work (see also Sacks 1963: 10–13), which aims to remedy them by making explicit all operative presuppositions and rules of procedure. This also is the case in computational

theory and practices that aim at specifying the categories and rules for a program precisely in advance—thus trying to remedy a vagueness that is actually required and saddling them with a huge problem (cf. Rawls and Mann 2015). According to Garfinkel and Sacks (1970), any attempts to remedy the essential vagueness and indexicality of social life are invariably and unavoidably unachievable.

Any list of specifications can be further specified, endlessly, as Wittgenstein (1953) showed for rule following. The crucial point made by Garfinkel and Sacks (1970), however, is that indexicality does not need to be fixed, but serves a foundational purpose in interaction. Making sense in interaction *always* rests on things that are *not talked about*—stated in so many words—that are taken for granted, i.e., indexicality and *et cetera* work are essential features of sense-making (not nuisances or appendices to it; see also Eisenmann and Rawls 2023; Button et al. 2022, Chap. 1). Furthermore, *et cetera* is not a static condition, but operates as a feature of the sequential organization of interaction: it is “essentially bound to both the inner and the outer temporal course of activities and thereby to the progressive development of circumstances and their contingencies.” (Garfinkel 1967: 73–74; see also Sacks 1992) Making sense in and of interaction, including interacting with ELIZA and similar programs, is reflexively tied to its “lived work” and its temporal parameters the “first linear time through”, or, in another of Garfinkel’s phrasings, for “another each next first time” (Garfinkel 2002: 216) and not as an already accomplished docile text—or gloss of such a text.

Garfinkel’s empirical findings are not only relevant for an externalist observational study of AI, but open venues for a “hybrid study” of AI that not only investigates the practices of programmers, designers, and users, but contributes to the field of programming.²⁴ The question of how indexicality might contribute to computing is still an underexplored discussion in computer science. The following passage from Philip Agre (who was in close contact and worked with Garfinkel in the 1980s and 1990s) makes the point that indexicality is “an active phenomenon of context construction” that needs to inform practices of information design:

It is also, as Garfinkel (1984 [1967]) would insist, an achievement that is only good enough for practical pur-

poses; individuals may well have different views of the precise boundaries of the “there,” “then,” or “them” being referred to, but if the discrepancy causes no trouble, it will most likely pass unremarked. A model-theoretic account of this achievement, such as Barwise and Perry’s, can posit the potential and actual referents of indexical terms by constructing the appropriate situations, but it cannot explain the actions by which particular people picked out those particular referents. It is only through study of the actual practices people employ to achieve reference in situ that indexicality begins to emerge not merely as a passive phenomenon of context dependence but as an active phenomenon of context construction. (Agre 1997: 233)

7 Conclusion

Weizenbaum took a critical stance toward his experiments with ELIZA, treating participants’ ascriptions of intelligence to the machine as misconceptions or even evidence of delusional thinking, ultimately arguing for a correction of what was later described as the ELIZA effect. Garfinkel disagreed. Such arguments about misconceptions of the state of AI are still at the center of current discussions, as the case of Google’s LaMDA discussed in the introduction clearly shows. The issue at stake concerns the role of the analyst and their methods and assumptions in relation to the orientation of actual participants in the interaction to what social objects count as “real,” as in the case of witches in the discussion reported on above that took place between Goffman, Parsons, Rose, Garfinkel, and Sacks in 1964.

The EMCA perspective allows for a detailed understanding of the different ways users interact with AI and how sense-making in those interactions is grounded in the details of social practices and human social competencies. On this view, the meanings people produce in interaction with ELIZA are not illusions—they are social objects that emerge in and as the turn structure of these interactions. A non-ethnomethodologist might ask how subjects manage to make meaning in the absence of “genuine human–human reciprocity.” But, for Garfinkel, this question assumes what it should be examining. Human participants cannot “read the minds” of other participants, be they machines, people, or spirits. The assumption of “human agency” as well as the idea that there is ever “genuine” reciprocity of minds is false and gets people off on the wrong track.²⁵ The question

²⁴ “Hybrid studies” (Garfinkel 2002: 100ff.) require researchers to be competent in the work being studied (computing in this case) in order to contribute not only to sociology, but also to the research field in question. Relatedly, ethnomethodological studies of work (Garfinkel 1986, 2022) have concerned themselves with research questions of interest to both practitioners of various worksites and sociologists. Their adequacy relies on detailed exploration of “members’ methods”, aiming for what Garfinkel (2002: 105ff.) calls “praxeological validity”, i.e., that sociological descriptions are practically recognizable and instructable for and by the practitioners in question (see also Ikeya 2020; Eisenmann and Mitchell 2022; Meier zu Verl and Meyer 2022; Sormani 2020, 2023).

²⁵ This is known in philosophy as the “Problem of other minds”, the idea being that it is not possible to have direct contact with or knowledge of other minds. Solutions have often resorted to positing a priori conditions. In this case, the problem also directly relates to the question of the Turing test (Turing 1950).

Garfinkel pursued is whether interaction between people (or people and machines) exhibits recognizable order properties that participants can use to make sense. If it does exhibit such order properties and they can be made to accord with constitutive expectations in ways that participants are able to make sense of—then meaning has been achieved.

Users assume Trust Conditions, and orient toward their obligations to those conditions, until trouble arises. Trouble can come from either the human end of the interaction or from the machine. But, as long as the machine does not make it obvious that it has pre-coded scripts and that it is not orienting to Trust Conditions, the human rule is to assume Trust Conditions, and sense-making will continue until it fails. As in the “random Yes–No answer” experiments, users make sense of ELIZA’s conduct in the usual way by treating it as motivated conduct, i.e., *motivated by the action of the previous turn*, and as part of a coherent sequence of turns (covered by Trust Conditions)—just as their own human responses are. What Garfinkel (1967: 94; 2019 [1959]: 26) writes of the “Yes–No experiments” applies to the interactions of ELIZA/LYRIC with its human users and beyond:

“Through the work of documenting—i.e., by searching for and determining pattern, by treating the advisor’s answers as motivated by the intended sense of the question, by waiting for later answers to clarify the sense of previous ones, by finding answers to unasked questions [...] the perceivedly normal values of what was being advised were established, tested, reviewed, retained, restored; in a word, managed.”

In short, human users do their ordinary interactional work of making sense when communicating with machines, which creates the meaning they experience.

Garfinkel’s experiments with human–computer interaction document the reliance of the machine on the human commitment to sense-making. His breaching scripts, if we can call them that, were designed to make trouble that reveals how the situation is produced in and as the interaction between the typing user and the machine, including the totality of its interface, programming, and physical components. It also shows how human users remain committed to making sense of interactional sequences in the face of trouble, and in doing so, demonstrates how *AI systems constitutively rely upon users’ sense-making work for their effective operation*. Accordingly, Garfinkel’s early studies reveal the cooperative practices standing at the core of media and technology (Schüttpelz 2023), allowing for the re-specification of things, such as “ELIZA effect”, “confirmation bias”, and “cognitive dissonance”, in terms of the constitutive social practices of their human users—the methods through which these phenomena are ordinarily produced and made recognizable (cf. Rawls and Turowetz 2021; Turowetz and Rawls 2021).

In the decades since Garfinkel’s pioneering work, research grounded in EMCA has focused on the conceptual issues of AI and social action (e.g., Button et al. 1995; Gilbert and Heath 1985; Suchman 1993) and also examined how EMCA findings can be implemented in the development of computer dialog systems (e.g., McIlvenny 1990; Raudaskoski 1990; Thomas 1991; Wooffitt 1994). More recently, as AI-based devices become entrenched in everyday life, EMCA researchers are conducting studies of robots, voice user interfaces, and embodied human-like agents. This research documents in detail how people adjust their conduct in interacting with machines, e.g., with so-called “conversational agents”. Porcheron et al. (2018) argue that participants talk with voice user interfaces on the basis of “input” and “output” rather than “conversation” (see also Reeves and Porcheron 2022). On the other hand, even simple technical systems can produce recognizably meaningful exchanges (Relieu et al. 2020), and as Korb (2023) argues, in some settings users treat chatbots as conversational partners. Ivarsson and Lindwall (2023) show that sequential and categorial analysis of ongoing talk is the basis of ascribing “intelligence” to a machine.

We have argued throughout this paper that Garfinkel’s research on ELIZA and LYRIC, although conducted at a very different point in time, can provide insights relevant for current research on AI in interaction. Although many current AI systems are grounded and build upon cognitivist and individualist models of human mind, they are usually employed in thoroughly social situations. Questions of “competence” or “agency” are not located inside the machine, but rather in the cooperative achieved orderliness of human sense-making.

The EMCA approach has also opened avenues for studying and contributing to design and software development by AI researchers and programmers, both in theoretical grounding and technology development (see, e.g., Alač, 2009; Brooker and Mair 2022; Gehle et al. 2017; Krummheuer 2015; Mair et al. 2020; Pelikan et al. 2020; Saha et al. 2023), that are being explored further. In his seminal study of computing, Agre (1997) demonstrated how the work of programmers already relies substantially on everyday understanding, as well as on philosophical theories of mind, but that this reliance remains largely tacit. Such tacit assumptions are respecified in EMCA as topics of empirical investigation.²⁶

²⁶ These tacit assumptions also pertain to social inequalities, race, disabilities and other marginalized persons which have implications for thinking about ‘bias’ in AI training datasets, as well as gendered and racialized exclusions from that community in terms of the practical constitution of technology (cf., e.g., Turner 2018; regarding the tacit and interactional dimensions, see Rawls and Duck 2020).

Agre and his colleagues, including David Chapman, argued that AI development had hit a wall in the late 1980s and 1990s. Garfinkel's study of ELIZA/LYRIC allows for the re-specification of core concepts like categories, information objects, and computing, and calls for a research approach that takes the necessity of indexicality and sequentiality more seriously, allowing for its application in contemporary AI computing in ways that might just address the problems Agre and others were pointing to.

Acknowledgements We thank all participants of the ELIZA Workshop held at University of Siegen on the 5th of November 2021 (<https://www.mediacoop.uni-siegen.de/de/veranstaltungen/workshop-b04-garfinkel-meets-eliza-mit-hendrikbender/>) for their valuable comments, especially Marcus Burkhardt and Magnus Rust for their insights also during ongoing discussions on the topic. Many thanks are also due for the insightful comments by the participants of the “Gipfeli Meeting” held by Bettina Heintz and Andrea Glauser at the University of Lucerne (25th–26th of November 2022), who had read an earlier version of this article. Funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation)—Project-ID 262513311—SFB 1187, Collaborative Research Centre: “Media of Cooperation”. Jakub Mlynář's work on the article was enabled by the support of the Czech Fulbright Commission (Fulbright Grant for Scholars and Researchers). We also like to thank two anonymous reviewers for very detailed and thorough editing suggestions that contributed significantly to the improvement of this article. Any remaining shortcomings are ours.

Funding Open Access funding enabled and organized by Projekt DEAL.

Data availability Data sharing is not applicable to this article as no datasets were generated or analyzed during the current study.

Declarations

Conflict of interest The authors declare that they have no conflict of interest.

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

References

- Agre PE (1997) *Computation and human experience*. Cambridge University Press, Cambridge
- Alač M (2009) Moving android: on social robots and body-in-interaction. *Soc Stud Sci* 39(4):491–528
- Alač M (2016) Social robots: Things or agents? *AI Soc* 31:519–535
- Alač M, Gluzman Y, Aflatoun T, Bari A, Jing B, Mozqueda G (2020) Talking to a toaster: how everyday interactions with digital voice assistants resist a return to the individual. *Evental Aesth* 9(1):3–53
- Apprich C (2019) Die Maschine auf der Couch. Oder: Was ist schon «künstlich» an Künstlicher Intelligenz? *ZfM* 21(2):20–28
- Baranovska M, Höltgen S (2018) Hello I'm Eliza. *Fünfzig Jahre Gespräche mit Computern*. Projektverlag, Bochum/Freiburg
- Basset C (2019) The computational therapeutic: exploring Weizenbaum's ELIZA as a history of the present. *AI Soc* 24:803–812
- Boden MA (1977) *Artificial intelligence and natural man*. Basic Books, New York
- Brooker P, Mair M (2022) Researching algorithms and artificial intelligence. In: Housley W, Edwards A, Beneito-Montagut R, Fitzgerald R (eds) *The SAGE handbook of digital society*. Sage, Thousand Oaks, pp 573–592
- Button G, Sharrock W (1995) On simulacrum of conversation: Toward a clarification of the relevance of conversation analysis for human–computer interaction. In: Thomas PJ (ed) *The social and interactional dimensions of human–computer interfaces*. Cambridge University Press, Cambridge, pp 107–125
- Button G, Coulter J, Lee JRE, Sharrock W (1995) *Computers, minds and conduct*. Polity Press, Cambridge
- Button G, Lynch M, Sharrock W (2022) *Ethnomethodology, conversation analysis and constructive analysis: on formal structures of practical action*. Routledge, Oxon
- Chomsky N (2006 [1968]) *Language and mind*, 3rd edn. Cambridge University Press, Cambridge
- Colby KM, Watt JB, Gilbert JP (1966) A computer method of psychotherapy. *J Nerv Ment Dis* 142(2):148–152
- Collins H (2018) *Artificial intelligence: against humanity's surrender to computers*. Polity, London
- Crabtree A (2004) Taking technomethodology seriously: hybrid change in the ethnomethodology–design relationship. *Eur J Inf Syst* 13(3):195–209
- Dillon S (2020) The Eliza effect and its dangers: from demystification to gender critique. *J Cult Res* 24(1):1–15
- Dourish P, Button G (1998) On “technomethodology”: foundational relationships between ethnomethodology and system design. *Hum Comput Interact* 13:395–432
- Dreyfus H (1965) *Alchemy and artificial intelligence*. The RAND Corporation, Santa Monica
- Dreyfus H (1972) *What computers can't do*. MIT Press, New York
- Duck W, Rawls AW (2023) Black and Jewish: “Double Consciousness” inspired a qualitative interactional approach that centers race, marginality, and justice. *Qual Sociol* 46:163–198. <https://doi.org/10.1007/s11133-023-09535-9>
- Eisenmann C (2022) *Spiritualität als soziale Praxis—Zur Konstruktion von Wirklichkeit im Yoga*. (Spirituality as a Social Practice—the Production of Reality in Yoga). Berlin/Boston: de Gruyter Oldenbourg
- Eisenmann C, Mitchell R (2022) Doing Ethnomethodological Ethnography. Moving between Autoethnography and the Phenomenon in “Hybrid Studies” of Taiji, Ballet, and Yoga. *Qual Res*. <https://doi.org/10.1177/14687941221132956>
- Eisenmann C, Rawls AW (2023) The continuity of Garfinkel's approach: seeking ways of ‘making the phenomenon available again’ through the experience and usefulness of ‘trouble.’ In: Sormani P, vom Lehn D (eds) *The Anthem companion to Harold Garfinkel*. Anthem Press, New York, pp 19–42. <https://doi.org/10.2307/jj.4418210.6>
- Eisenmann C, Englert K, Ehler V, Schubert C (eds) (2023) *Varieties of cooperation. Mutually making the conditions of mutual making*. Springer, Wiesbaden
- Garfinkel H (1948) A further note on the prospectus for an exploratory study of communication and the modes of understanding in selected types of dyadic relationship, with particular reference

- to the Jew as an object of social treatment. 30 October 1948. Garfinkel Archive, Newburyport, Massachusetts
- Garfinkel H (1963) A conception of, and experiments with, “trust” as a condition of stable concerted actions. In: Harvey OJ (ed) *Motivation and social interaction: cognitive determinants*. Ronald Press, New York, pp 187–238
- Garfinkel H (1967) *Studies in ethnomethodology*. Prentice Hall, Englewood Cliffs
- Garfinkel H (ed) (1986) *Ethnomethodological studies of work*. Routledge & Kegan Paul, London
- Garfinkel H (2002) *Ethnomethodology’s program: working out Durkheim’s aphorism* (edited by Rawls AW). Rowman & Littlefield, London
- Garfinkel H (2022) *Studies of work in the sciences* (edited by Lynch M). Routledge, New York
- Garfinkel H, Sacks H (1970) On formal structures of practical actions. In: McKinney JD, Tiryakian EA (eds) *Theoretical sociology*. Appleton-Century Crofts, New York, pp 337–366
- Garfinkel H, Rose E, Sacks H, MacAndrew C, Mendlovitz S (1962) *Ethnomethodology conference* (audio recordings), Reel #4 of 10, Side B, February 16, 1962. Garfinkel Archive, Newburyport, Massachusetts
- Garfinkel H (1946) Some reflections on action theory and the theory of social systems. Unpublished manuscript. Garfinkel Archive, Newburyport, Massachusetts
- Garfinkel H (1952) *The perception of the other: a study in social order*. Unpublished PhD thesis. Harvard University, Cambridge
- Garfinkel H (1962) (approximately). *On Et Cetera* (Outline). Unpublished manuscript, Garfinkel Archive, Newburyport, Massachusetts
- Garfinkel H (1966) Formal investigation of decision-making in actual, singular situations of practical action (Air Force grant proposal). Unpublished manuscript, Garfinkel Archive, Newburyport, Massachusetts
- Garfinkel H (2006 [1948]) *Seeing sociologically* (edited by Rawls AW). Paradigm, Boulder
- Garfinkel H (2008 [1952]) *Toward a sociological theory of information* (edited by Rawls AW). Paradigm Publishers, Boulder
- Garfinkel H (2019 [1959]) Common sense knowledge of social structures. A paper distributed at the session on the sociology of knowledge (edited by C Erbacher and E Schüttelpelz), 4th world congress of sociology, Stresa, Italy, September 12, 1959. <https://doi.org/10.25969/mediarep/13805>
- Gehle R, Pitsch K, Dankert T, Wrede S (2017) How to open an interaction between robot and museum visitor? Strategies to establish a focused encounter in HRI. In: *Proceedings of the 2017 ACM/IEEE international conference on human–robot interaction*, pp 187–195
- Gilbert NG, Heath C (eds) (1985) *Social action and artificial intelligence*. Gower, Aldershot
- Hayward PL (1968) *ELIZA Scriptwriter’s manual: a manual for the use of the ELIZA conversational computer system*. MIT, Cambridge
- Heath C, Luff P (2022) Technology in practice. In: Maynard DW, Heritage J (eds) *The ethnomethodology program: legacies and prospects*. Oxford University Press, New York, pp 398–419
- Heintz B (1993) *Die Herrschaft der Regel: Zur Grundlagengeschichte des Computers*. Campus Verlag, Frankfurt
- Hofstadter D (1995) *Fluid concepts and creative analogies: computer models of the fundamental mechanisms of thought*. Basic Books, New York
- Ikeya N (2020) Hybridity of hybrid studies of work: examination of informing practitioners in practice. *Ethnogr Stud* 17:22–40
- Ivarsson J, Lindwall O (2023) Suspicious minds: the problem of trust and conversational agents. *Comput Supp Cooper Work (CSCW)* 32:545–571. <https://doi.org/10.1007/s10606-023-09465-8>
- Korbut A (2023) How conversational are “conversational agents”? Evidence from the study of users’ interaction with a service telephone chatbot. *Soc Interact Video Based Stud Hum Soc*. <https://doi.org/10.7146/si.v6i1.137249>
- Korn W (1969) (approximately). *Description of experimental materials*. Unpublished manuscript. Garfinkel Archive, Newburyport, Massachusetts
- Krummheuer A (2015) Technical agency in practice: the enactment of artefacts as conversation partners, actants and opponents. *Psychonol J* 13(2–3):179–202
- Kuhn T (1970) *The structure of scientific revolutions*. University of Chicago Press, Chicago
- Laurier E, Krämer H, Gerst D, Salomon R (2019) The “Studies in Ethnomethodology” are a way of understanding and handling empirical materials and thoughts. *Forum Qual Sozialforschung/ Forum Qual Soc Res* 20(2):Art. 2
- Lecerf Y (1963) *Logique mathématique—machines de Turing réversibles*. *C R Hebd Seances Acad Sci* 257:2597–2600
- Licklider JCR (1960) Man–computer symbiosis. *IRE Trans Hum Fact Electron HFE* 1:4–11
- Lynch M (1993) *Scientific practice and ordinary action: ethnomethodology and social studies of science*. Cambridge University Press, New York
- Lynch M, Eisenmann C (2022) Transposing Gestalt phenomena from visual fields to practical and interactional work: Garfinkel’s and Sacks’ social praxeology. *Philos Sci Travaux D’histoire Et De Philosophie Des Sci* 26(3):95–122
- Mair M, Brooker P, Dutton W, Sormani P (2020) Just what are we doing when we’re describing AI? Harvey Sacks, the commentator machine, and the descriptive politics of the new artificial intelligence. *Qual Res* 21(3):341–359
- McDermott D (1976) Artificial intelligence meets natural stupidity. *SIGART Newsletter* 57, April 1976. Reprinted in: Haugeland J (ed.) (1981) *Mind design*. Montgomery, VT: Bradford Books
- McIlvenny P (1990) Communicative action and computers: Re-embodiment conversation analysis? In: Luff P, Gilbert NG, Frohlich D (eds) *Computers and conversation*. Academic Press, London, pp 91–132
- Meier zu Verl C, Meyer C (2022) Ethnomethodological ethnography: historical, conceptual, and methodological foundations. *Qual Res*. <https://doi.org/10.1177/14687941221129798>
- Meyer C (2013) New alterities and emerging cultures of social interaction. *Global Cooperation Research Papers* 3. Käte Hamburger Kolleg / Centre for Global Cooperation Research, Duisburg
- Minsky M (1967) *Computation: finite and infinite machines*. Prentice Hall, Englewood Cliffs
- Mlynář J (2023) Harold Garfinkel and Edward Rose in the early years of ethnomethodology. *J Hist Behav Sci* 59(2):171–192. <https://doi.org/10.1002/jhbs.22237>
- Mooers C (1951) Zatoncoding applied to mechanical organization of knowledge. *Am Doc* 2:20–32
- Natale S (2019) If software is narrative: Joseph Weizenbaum, artificial intelligence and the biographies of ELIZA. *New Media Soc* 21(3):712–728
- Natale S (2021) *Deceitful media: artificial intelligence and social life after the Turing test*. Oxford University Press, New York
- Oldman D, Drucker C (1985) The non-reducibility of ethno-methods: Can people and computers form a society? In: Gilbert GN, Heath C (eds) *Social actions and artificial intelligence*. Gower Publishing, Aldershot, pp 144–159
- Papert S (1968) *The artificial intelligence of Hubert Dreyfus: a budget of fallacies*. MIT, Cambridge
- Parsons T (1937) *The structure of social action*. McGraw Hill, New York
- Pelikan HRM, Broth M, Keevallik L (2020) Are you sad, Cozmo? How humans make sense of a home robot’s emotion displays.

- In: HRI'20: proceedings of the 2020 ACM/IEEE international conference on human–robot interaction. ACM, Cambridge, pp 461–470
- Porcheron M, Reeves S, Sharples S and Fischer JE (2018) Voice interfaces in everyday life. In: CHI'18: Proceedings of the 2018 CHI conference on human factors in computing systems. ACM, New York, Paper 640
- Pruijt H (2006) Social interaction with computers: an interpretation of Weizenbaum's ELIZA and her heritage. *Soc Sci Comput Rev* 24(4):516–523
- Putnam H (1975) *The meaning of "meaning"*. University of Minnesota Press, Minneapolis
- Qarton G (1967) Deliberate efforts to control human behavior and modify personality. *Daedalus* 96(3):837–853
- Raudaskoski P (1990) Repair work in human–computer interaction: a conversation analytic perspective. In: Luff P, Gilbert NG, Frohlich D (eds) *Computers and conversation*. Academic Press, London, pp 151–171
- Rawls AW (2021a) Durkheim's self-regulating "constitutive" practices: an unexplored critical relevance to racial justice, consensus thinking, and the COVID-19 pandemic. In: Marucci N (ed) *Durkheim & critique*. Palgrave Macmillan, Cham, pp 227–263
- Rawls AW (2021b) The structure of social facts: self, objects and action as products of reciprocity and cooperation in constitutive practices. *Mauss Int* 1(1):186–206
- Rawls AW, Duck W (2020) *Tacit racism*. University of Chicago Press, Chicago
- Rawls AW, Mann D (2015) Getting information systems to interact: the social fact character of 'Object' clarity as a factor in designing information systems. *Inf Soc* 31(2):175–192
- Rawls AW, Turowetz J (2021) Garfinkel's politics: collaborating with Parsons to document taken-for-granted practices for assembling cultural objects and their grounding in implicit social contract. *Am Sociol* 52(1):131–158
- Rawls AW, Mann D, Garcia A, David G, Burton M (2009) Simple enumerations: ethnomethodology and MITRE information assurance data standards (Italian title: "Semplici enumerazioni. L'etnometodologia e gli Information Assurance Data Standards del MITRE"). *Ethnografia et Ricerca Qualitativa* 2(1):79–108
- Rawls AW, Turowetz J, Eisenmann C, Hoey E (2020) On meaning: the meaning of a particular social fact—"Suicide"—as discussed by Parsons, Garfinkel, Goffman and Sacks in 1964. Paper delivered at the Yale Center for Cultural Sociology
- Reeves S, Porcheron M (2022) Conversational AI: respecifying participation as regulation. In: Housley W, Edwards A, Beneito-Montagut R, Fitzgerald R (eds) *The Sage handbook of digital society*. Sage, London, pp 573–592
- Relieu M, Sahin M, Francillon A (2020) Une approche configurationnelle des leures conversationnels. *Réseaux* 220–221(2):81–111
- Rose E (1967) *A looking-glass conversation in the rare languages of Sez and Pique*. University of Colorado, Boulder
- Sacks H (1963) Sociological description. *Berkeley J Sociol* 8:1–16
- Sacks H (1992) *Lectures on conversation, vol I & II*. Blackwell, Oxford
- Sacks H, Schegloff EA, Jefferson G (1974) A simplest systematics for the organization of turn-taking for conversation. *Language* 50(4):696–735
- Sacks H (1968) "Understanding" in conversation. Talk at the University of California Santa Barbara (UCSB). Unpublished transcript. Harold Garfinkel Archive, Newburyport, MA
- Saha D, Brooker P, Mair M, Reeves S (2023) Thinking like a machine: Alan Turing, computation and the praxeological foundations of AI. *Sci Technol Stud* 2023:1
- Schegloff EA (1992) Repair after next turn: the last structurally provided defense of intersubjectivity in conversation. *Am J Sociol* 97(5):1295–1345
- Schegloff EA, Jefferson G, Sacks H (1977) The preference for self-correction in the organization of repair in conversation. *Language* 53(2):361–382
- Schüttpelz E (2017) Infrastructural media and public media. *Med Action* 2017:13–61. <https://doi.org/10.25819/ubsi/7935>
- Schüttpelz E (2023) Reinventing the wheel of media theory. In: Eisenmann C, Englert K, Ehler V, Schubert C (eds) *Varieties of cooperation. Mutually making the conditions of mutual making*. Springer, Wiesbaden
- Shapira N, Zwirn G, Goldberg Y (2023) How well do large language models perform on faux pas tests? *Find Assoc Comput Ling* 2023:10438–10451
- Shieber SM (1994) Lessons learned from a restricted Turing test. *Commun ACM* 37(6):70–78
- Simon H (1969) *The sciences of the artificial*. MIT, Cambridge
- Smith MB, Bruner JS, White RW (1956) *Opinions and personality*. Wiley, New York
- Sormani P (2020) "DIY AI"? Practising kit assembly, locating critical inquiry. *Ethnogr Stud* 17:60–80
- Sormani P (2023) Interfacing AlphaGo: Embodied play, object agency, and algorithmic drama. *Soc Stud Sci* 53(5):686–711. <https://doi.org/10.1177/03063127231191284>
- Suchman L (1987) *Plans and situated action: the problem of human–machine communication*. Cambridge University Press, Cambridge
- Suchman L (1993) Response to Vera and Simon's situated action: a symbolic interpretation. *Cogn Sci* 17:71–75
- Suchman L (2007) *Human–machine reconfigurations: plans and situated actions*. Cambridge University Press, Cambridge
- Suchman L (2023) Imaginaries of omniscience: automating intelligence in the US Department of Defense. *Soc Stud Sci* 53(5):761–786
- Svensson H (2020) *Establishing shared knowledge in political meetings: repairing and correcting in public*. Routledge, London
- Thielmann T (2012) Taking into account. *ZFM* 6(1):85–102
- Thielmann T (2019) The ENIAC display: insignia of a digital praxeology. In: Haigh T (ed) *Exploring the early digital*. Springer, New York, pp 101–116
- Thielmann T, Sormani P (2023) Zato coding Remodeled? Harold Garfinkel's hybrid studies avant la lettre. *Int J Dig Hum* 2023:1
- Thomas PJ (1991) Language, communication, social interaction and the design of human–computer interfaces. *Behav Inf Technol* 10(4):311–324
- Thomas WI, Thomas DS (1928) *The child in America: behavior problems and programs*. Knopf, New York
- Tiku N (2022a) Google fired engineer who said its AI was sentient (07/22/2022). *The Washington Post*: <https://www.washingtonpost.com/technology/2022/07/22/google-ai-lambda-blake-lemoine-fired/>. 24 Oct 2022.
- Tiku N (2022b). The Google engineer who thinks the company's AI has come to life (06/11/2022). *The Washington Post*: <https://www.washingtonpost.com/technology/2022/06/11/google-ai-lambda-blake-lemoine/>. 24 Oct 2022
- Turing AM (1950) Computing machinery and intelligence. *Mind* 59(236):433–460. <https://doi.org/10.1093/mind/LIX.236.433>
- Turner LN (2018) Detecting racial bias in algorithms and machine learning. *J Inf Commun Ethics Soc* 16(3):252–260. <https://doi.org/10.1108/JICES-06-2018-0056>
- Turowetz J, Rawls AW (2021) The development of Garfinkel's "Trust" argument from 1947 to 1967: demonstrating how inequality disrupts sense and self-making. *J Class Sociol* 21(1):3–37

- Watson R (1998) Ethnomethodology, consciousness and self. *J Conscious Stud* 5(2):202–223
- Watson R (2009) Constitutive practices and Garfinkel's notion of trust: revisited. *J Class Sociol* 9(4):475–499
- Weizenbaum J (1966) ELIZA—a computer program for the study of natural language communication between man and machine. *Commun ACM* 9(1):36–45
- Weizenbaum J (1967) Contextual understanding by computers. *Commun ACM* 10(8):474–480
- Weizenbaum J (1976) *Computer power and human reason: from judgment to calculation*. WH Freeman & Co, San Francisco
- Wittgenstein L (1953) *Philosophical investigations*. Blackwell, Oxford
- Wooffitt R (1994) Applying sociology: conversation analysis in the study of human–(simulated) computer interaction. *Bull Sociol Methodol Bull Methodol Sociol* 43(1):7–33
- Ziewitz M (2017) A not quite random walk: experimenting with the ethnomethods of the algorithm. *Big Data Soc* 4(2):1–13. <https://doi.org/10.1177/2053951717738105>
- Znaniecki F (1936) *Social actions*. Farrar & Rinehart, New York

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