



AI in situated action: a scoping review of ethnomethodological and conversation analytic studies

Jakub Mlynář¹ · Lynn de Rijk² · Andreas Liesenfeld² · Wyke Stommel² · Saul Albert³

Received: 26 June 2023 / Accepted: 5 March 2024
© The Author(s) 2024

Abstract

Despite its elusiveness as a concept, ‘artificial intelligence’ (AI) is becoming part of everyday life, and a range of empirical and methodological approaches to social studies of AI now span many disciplines. This article reviews the scope of ethnomethodological and conversation analytic (EM/CA) approaches that treat AI as a phenomenon emerging in and through the situated organization of social interaction. Although this approach has been very influential in the field of computational technology since the 1980s, AI has only recently emerged as such a pervasive part of daily life to warrant a sustained empirical focus in EM/CA. Reviewing over 50 peer-reviewed publications, we find that the studies focus on various social and group activities such as task-oriented situations, semi-experimental setups, play, and everyday interactions. They also involve a range of participant categories including children, older participants, and people with disabilities. Most of the reviewed studies apply CA’s conceptual apparatus, its approach to data analysis, and core topics such as turn-taking and repair. We find that across this corpus, studies center on three key themes: openings and closing the interaction, miscommunication, and non-verbal aspects of interaction. In the discussion, we reflect on EM studies that differ from those in our corpus by focusing on praxeological respecifications of AI-related phenomena. Concurrently, we offer a critical reflection on the work of literature reviewing, and explore the tortuous relationship between EM and CA in the area of research on AI.

Keywords Artificial intelligence · Conversation analysis · Embodiment · Ethnomethodology · Multimodality · Robotics · Social interaction · Video analysis

1 Introduction

The relationship between society and the range of “techniques and technologies that travel under the sign of AI” (Suchman 2023b) is permeated with paradox: AI seems to be concurrently a conceptual impossibility and a social reality (Jaton and Sormani 2023). On the one hand, from its outset, the notion of ‘artificial intelligence’ (AI) has been subject to a powerful conceptual critique of the distinctions and continuities between ‘artificial’ and ‘natural’, ‘human’ and ‘machinic’, and the centrality of ‘mind’ and ‘intelligence’ (e.g., Button et al. 1995; Coulter 1985; Dreyfus 1965). On

the other hand, AI has nonetheless become a social object: something that can be talked about (e.g., Mlynář et al. 2022; Petersson et al. 2022), for example seen as a promise or a threat (e.g., Kotásek 2015; Smith 2019), or attributed with societal agency (e.g., Bellon and Velkovska 2023; Collins 2018). These shifting discourses of AI and its social contexts have led to a diffuse range of empirical and methodological approaches to social studies of AI spanning many disciplines (Caluori 2023). From fields of research invested in advancing technology, to critical examinations of its effects, risks, and implications, Suchman (2023b: 2) points out that treating AI as a self-evident and unitary topic of study risks effacing the “work being done by the figure of AI in specific contexts”. The elusive concept of AI, coupled with its purported ubiquity and increasing encroachment into all aspects of everyday life (Elliott 2019; Pflanzner et al. 2023) has contributed to a ‘situational deficit’ (Marres and Sormani 2023) in social studies of AI that risks failing to “describe how AI features in the world as it is” (Brooker et al. 2019: 296).

✉ Jakub Mlynář
jakub.mlynar@hes-so.ch

¹ HES-SO Valais-Wallis University of Applied Sciences and Arts Western Switzerland, Sierre, Switzerland

² Radboud University, Nijmegen, The Netherlands

³ Loughborough University, Loughborough, UK

This article reviews the scope of ethnomethodological and conversation analytic (EM/CA) approaches to AI. In general terms, EM is a sociological program that examines and describes members' methods of producing mundanely recognizable social activities, while treating these everyday methods as topics of empirical study (Garfinkel 1967, 2002). CA applies these principles to empirically investigate the sequential organization of "talk-in-interaction" (Sacks 1992; Schegloff 2007), as well as the categorization work involved (see, e.g., Stokoe 2012). The shared focus and affinity of these two approaches rests in their "phenomenon-locating feature" (Wieder 1999: 168) through meticulous studies of the constitutive details of social order—although the specific ways of locating and accessing phenomena in EM and CA may differ, as we discuss below. These approaches bring about a conceptualization of AI as a phenomenon emerging in and through *situated action*, and amenable to detailed studies of human sociality and social interaction. As introduced by Suchman (1987, 2007), the term 'situated action' incorporates the principles of EM/CA and develops the notion of meaningful action as depending "in essential ways on its material and social circumstances" (2007: 70), inviting the study of "how people use their circumstances to achieve intelligent action" (ibid.).¹ Within social studies of AI, research informed by EM and CA draws focus on the forms of practical action and reasoning that constitute the detailed local organization of people's interaction with and among AI systems. The notion of situated action highlights how AI-based technologies may be used as a resource to produce actions in social situations, or constituted as social agents that engage in interactions rooted in distinct social contexts. Whereas human–computer interaction (HCI) tends to study retrospective accounts and perceptions of interactions with AI through, e.g., questionnaires or interviews, EM/CA studies "interactions themselves, as they unfold and are accomplished" (Tuncer et al. 2023: 2). This approach provides access to the constitutive detail of produced social orderliness that is the "normally thoughtless" (Garfinkel 2022b: 153), "unquestionable background of matters" (Garfinkel 1967: 173): tacit but observable aspects of the social life of AI.²

¹ In the first edition of the book (Suchman 1987), the formulation is slightly different in its first part: "... the term *situated action* ... underscores the fact that the course of action depends in essential ways upon the action's circumstances. Rather than attempting to abstract action from its circumstances and reconstruct it as a rational plan, the approach is to study how people use their circumstances to achieve intelligent action" (p. 35).

² Garfinkel's phrase "normally thoughtless" does not submit that people act without thinking. Rather, it underscores that EM/CA's phenomena consist of mundane and routine practices that are, for competent members of society, usually accomplished without preliminary planning or thorough deliberation. As such, they are also done without the need for explicit specification or explanation, and

Although EM/CA research also tends to prioritize the production of empirical research, here we take up Anderson and Sharrock's (2017) suggestion to review and reflect on collections of existing studies—in this case focusing on studies of AI in situated action. This emerging literature is scattered across disciplines, and has appeared under various methodological, topical, and field-specific banners including human–computer interaction (HCI), human–robot interaction (HRI), computer-supported cooperative work (CSCW), workplace studies, and interactional linguistics. Although originally developed in response to foundational issues in sociology, EM and CA are now embedded within various disciplinary domains reaching from linguistics to psychology and examining activities as diverse as coffee tasting, rock climbing, pediatric oncology, and court trials, among many others. Partly because of the resulting methodological differences between branches of EM and CA, and partly because of the vast range of specific phenomena and situations now glossed as 'AI', an exploratory scoping process is required to provide an overview of this body of work. In this article, we present and discuss the findings from our 'scoping review': a method used for mapping out a broad area of research that may turn out to include a heterogeneous collection of study designs, phenomena, and research objects (Arksey and O'Malley 2005). Bringing parts of this dispersed field together, we aim to trace trends and directions, consolidate significant findings, and showcase the distinctive contribution of EM/CA to the broader field of social studies of AI. We also reflect on the research procedure of the scoping review itself, and ask what we might infer from reflexively exploring the 'reviewability' of a prospective field of studies of AI in situated action.

2 Background: EM/CA and technology-in-action

In *Everyday Automation*, Pink et al. (2022: 1) describe how discussions of AI are "shrouded with narratives which highlight extreme and spectacular examples" rather than the mostly mundane experiences we have with automated technologies. Although anthropomorphic robots or self-driving vehicles might still carry a (temporary) sense of spectacle, EM/CA focuses specifically on how 'ordinariness' is produced and maintained (Sacks 1984b). Situated action, as an empirical and methodological focus, centers local methods of reasoning and social organization by asking what people manifestly *do* with technologies, and what kind

Footnote 2 (continued)

therefore characteristically "seen but unnoticed" (Garfinkel 1967: 37), "acknowledged but tacit and unexamined" by members (Garfinkel 2022c: 21). On the broader importance of considering "tacit knowledge" in the study of AI see, e.g., Gill (2023).

of everyday sense-making work is intertwined with these doings. Studying AI in this way enables researchers to ask if ‘smart devices’ and ‘intelligent machines’, as they are used and embedded in everyday life, present the much-vaunted profound transformations of the social world, and how, in practical terms, they might impact how we live and work. The EM/CA studies of AI we review here have contributed a systematic, reflective focus on how interactions unfold in ways that are demonstrably consequential for users (Reeves 2019b), by looking at how these technologies enable and constrain the practical organization of everyday social interaction. Before presenting our findings, we briefly introduce the relationship between EM and CA in the context of technology and computation.³

The field of EM/CA, broadly conceived, includes at least three distinctive but related strands of research: conceptual, conversational, and practical/self-instructive analysis (Sormani 2019). Historically, EM developed in the 1950s from the work of Harold Garfinkel (2019a [1959], 1967) and colleagues, drawing on Parsons’ systems theory (Garfinkel 2019c) and Schutz’s and Gurwitsch’s social phenomenology (Garfinkel 2021, 2022a).⁴ One of the central concerns of EM is the temporal and sequential achievement of ordinary activities (Coates 2022; Rawls 2005). The meaning of interactional conduct, here, is not established in advance, but is always situated in lived time, reflexively establishing the “witnessable order” (Livingston 2008) of activities through which sense is produced and recognized, discovered and abandoned, for all practical purposes. Harvey Sacks (1967, 1992) and colleagues later developed this aspect of EM as a ground-breaking approach to the study of language and social interaction.⁵ As a discipline, CA studies the orders of

talk-in-interaction (Schegloff 1988; Psathas 1995). Its unique “analytic mentality” (Schenkein 1978) is based on detailed scrutiny of audio-visual recordings of ‘naturally-occurring’ interactions, aiming to describe their local orders of organization such as turn-taking (Sacks et al. 1974), sequence organization (Schegloff 2007), categorization (Sacks 1972), and other features of ordinary interaction.

Although EM and CA share historical and philosophical origins, their divergence is a point of ongoing debate. Clayman et al. (2022) highlight Erving Goffman’s distinctive contribution to CA’s structural focus on the domain of social interaction.⁶ Button et al. (2022) argue that this focus has transformed CA and drawn its centre of gravity towards topics and concepts in linguistics. This has potentially sidelined more sociological aspects of ‘early’ CA such as membership categorization (Housley and Fitzgerald 2002: 59). Another point of divergence between EM and CA has been the development of applied CA (Antaki 2011) as a burgeoning social science research method engaged in developing interventions in e.g., communication training (Stokoe 2014), medical interaction (Robinson and Heritage 2014), or other settings rich in institutional talk. Similarly, Haddington et al. (2023) point out that both EM’s and CA’s engagements with new—often technologized—domains of social action have always provided opportunities for reconsideration of its methodological principles, issues, and research procedures. As we discuss below, combining EM and CA in the process of conducting a scoping review draws out the ‘heuristic tensions’ (Sormani and von Lehn 2023) between the ways different approaches and interpretations of this research legacy have evolved.

Whether considered together or separately, for the last four decades, EM and CA have offered rich insights into a range of technical fields spanning inception and conceptualization to design and evaluation including, e.g., the practical, interactional work of mathematicians (Greiffenhagen 2014; Livingston 1986), scientists (Garfinkel 2022c; Lynch 1993), and software developers (Suchman and Trigg 1993). Similarly, since EM/CA’s earliest studies of talk on the phone (e.g., Schegloff 1968), this approach has offered insightful perspectives on interactive technologies by revealing the intricate workings of interactional processes (Heath and Luff 2022; Mlynář et al. 2018). There is also a long tradition of EM/CA studies of computing, technology, and interaction with, through, and around machines. For example, Sudnow’s (1983) groundbreaking account of learning to play the video game *Breakout* combines phenomenology and EM,

³ For related discussions and overviews, see, e.g., Grudin (2009); Harper (2019); Moore (2012) Moore and Arar (2019); Moore et al. (2023); Randall et al. (2021).

⁴ In the late 1960s, Garfinkel had already recognized the importance of early chatbots such as ELIZA (Weizenbaum 1967) for studying how people make sense of the social world. In the 1980s and 1990s, he was also in close contact with researchers in the field of AI such as Yves Leclerc (1963) and Phil Agre (1997a). Although Garfinkel’s interests in computation and AI did not lead to publications during his lifetime, they can be reconstructed through archival materials (see Eisenmann et al. 2023a). For Garfinkel, at the time, ‘human–machine interactions’ were perspicuous settings in which the routine procedures of sensemaking in everyday life are made visible, observable, and investigable.

⁵ In addition to “the ethnomethodological foundations of conversation analysis” (Lynch 2000), it is also possible to consider “the conversation analytic foundations of ethnomethodology” (Lynch and Livingston 2017). Although EM historically precedes CA, and Garfinkel had profound impact on Sacks’ work, CA’s radical preoccupation with naturally organized phenomena and their formal structures, discovered by working through collections of empirical material, has also influenced EM’s developments (Garfinkel 2022b, c) such as the later ‘studies of work’ program (Garfinkel 1986).

⁶ Although Garfinkel was close with Goffman over many years (e.g., Rawls 2023), their approaches to the study of social order are radically different—while EM, CA and Goffmanian sociology can also be productively combined (see, e.g., Mondada and Peräkylä 2023; Sharrock 1999; Smith 2003).

reflexively detailing the process of achieving mastery. Following the EM principle of unique adequacy (Garfinkel and Wieder 1992) that urges researchers to obtain routine competences in the investigated activities, Sudnow “becomes the phenomenon” (Reeves et al. 2009: 209), studying the practical constitution and advancement of his own skillful playing (see also Sormani 2022). Suchman’s (1987) influential study of users’ work with the help system of a complex photocopier draws more on CA’s approach to audio-visual recordings of interaction⁷ to critique HCI models based on pre-established mental plans, showing that plans are resources that people use in situated actions. Suchman’s pioneering challenge to cognitivist conceptualizations of human action in AI points to EM’s fundamental reconceptualization of central topics in computation and technology—as taken up in Dourish and Button’s “technomethodology” (1998). Within these fields, however, EM/CA is more usually subsumed by the priorities of computer science through ‘user studies’ and providing ‘implications for design’ (see Dourish 2006). Despite its influence in research on human–machine interaction, EM/CA has not yet brought about a substantial transformation of the practical ways in which technologies are typically conceived, designed, developed, and tested (see Crabtree 2004).

Earlier ‘waves’ of AI research have also prompted substantial responses from EM/CA research (e.g., Gilbert and Heath 1985; Button et al. 1995), and have framed key empirical questions about how fundamental structures of talk-in-interaction might, as Schegloff (1980: 81) puts it, “enter into the participation of humans dealing with computers”. However, it is only relatively recently that various technologies commonly associated with AI have become such a routinized and pervasive part of everyday life (Hirsch-Kreinsen 2023; Pilling et al. 2022) that a sustained empirical focus on AI is starting to emerge within EM/CA more broadly. The findings of EM/CA research are both distinctive and complementary to the broader context of social studies of AI. They are distinctive in identifying previously neglected phenomena and describing them in detail. Yet they also offer a “praxeological respecification” (Button 1991; Garfinkel 1991; Hester 2009) of established themes in the social sciences such as cognition, emotions, knowledge, ethics, and trust. By focusing on practical action and reasoning in everyday and specialized settings, EM/CA explicates the taken-for-granted features of social scenes that are manifestly relevant for participants. While centering situated action, or

inter-action, rather than its individual participants (be they ‘humans’ or ‘machines’), these studies describe the methodical procedures for achieving concerted, orderly courses of action as well as dealing with troubles and misunderstandings. This scoping review aims to show how EM/CA studies of AI in situated social action help map and track the ways these technologies and discourses have interacted with everyday social life over the last four decades.

3 Conducting the scoping review of ‘AI’ in interaction

Drawing a boundary around ‘AI’ is already challenging enough (Caluori 2023), and even more so when developing a gloss that can circumscribe AI within the volatile field of ‘EM/CA’: itself an “increasingly incoherent bucket category” (Jenkins 2023: 5), with its own contested interpretations and definitions (Button et al. 2022). We, therefore, use a ‘scoping review’ method to “describe in more detail the findings and range of research in particular areas of study, thereby providing a mechanism for summarizing and disseminating research findings” (Arksey and O’Malley 2005: 21; cf. Munn et al. 2018). Whereas systematic reviews usually address well-established academic literature, a scoping review of EM/CA research (e.g., Mayor and Bietti 2017; Pilnick et al. 2018; Saalasti et al. 2023) lets us explore the breadth and scope of studies of AI in situated action. Since the scoping review process probes the feasibility of collecting and summarizing a body of work, it also foregrounds the methodological challenges of reviewing such an inherently diverse and particularized set of studies. Firstly, for EM/CA’s analytic descriptions “concreteness [should] not be handed over to generalities” (Garfinkel 1991: 15), so findings tend to resist straightforward summarization. Secondly, from an EM perspective, measurement and countability in bibliometrics and systematic literature searches are topics of study, not transparent analytic practices (Churchill 1971; Cicourel 1964; Bovet et al. 2011; Mair et al. 2022). Nonetheless, the scoping review presents an opportunity to synthesize a suppositional collection of studies, while considering the opportunities and limitations of this approach. In the present article, we begin with a review of 53 scientific communications that apply a range of EM/CA analytic principles and methods to study AI in interaction, taking stock of their specificity, contributions, and preoccupations. We draw our selection for review using a speculative gloss of ‘AI’ that, in this context, includes any studies that discursively frame a technological artifact as occupying a social role conventionally reserved for human interactants. These various technologies, including algorithms, robots, conversational interfaces, and self-driving vehicles, seem to be loosely related by intuitively evident, but

⁷ Although Suchman’s research strategy has shifted over the years from “communication” to include broader “sociotechnical configurations” (our thanks to an anonymous reviewer for pointing this out)—see, e.g., Suchman (2023a), but also the second edition of *Plans and Situated Action* (Suchman 2007)—the EM/CA underpinning of her most influential work is still manifest and relevant.

exhaustively unspecifiable “family resemblances” (Wittgenstein 1953: §65–71).⁸

Rather than focusing on specific types of AI-labeled technologies, here we follow Schwartz’s (1989: 199) concise characterization of AI systems as “social actors playing social roles” to explore how participants’ social actions incorporate discourses and practical interpretations of AI. This working definition of AI is intentionally ‘vernacular’ or even ‘naïve’ in the sense that it takes AI-labeled devices at face value without problematizing their ‘intelligence’. Examples would include technology that serves as a driver, a tutor, a student, a caller/answerer of the telephone, or a chess player.⁹ We avoid a technical definition of AI because many systems use computer science techniques that fall under the category of AI without this ever becoming apparent to ordinary users (e.g., text-to-speech or content-recommendation algorithms), and their AI-ness thus may not be demonstrably relevant from a members’ perspective. Depending on the specific application, AI techniques are also combined in heterogeneous ways. For example, a scripted social robot that uses Natural Language Processing (NLP) to deal with input and Natural Language Generation (NLG) to ‘speak’ its output technically uses AI to function, but the scripted ways it conducts itself within the interaction are not driven by AI. The implementation of AI in such a machine is incomparable to a system that uses AI to drive its central functions, such as the game system AlphaGo (Silver et al. 2016; see also Sormani 2023).

Our working notion of EM/CA was similarly ‘naïve’ in that we simply included any publications that the authors identified as contributions to EM and/or CA by explicitly claiming that affinity in the text. The studies reviewed here all focus on the local organization of practical action and reasoning around machines that are plausibly recognizable (to the people interacting with them) as a form of ‘AI’, through the detailed analysis of transcribed audio or video recordings of social interaction. In line with the approach of the scoping review, we avoid doing quality assessments of the reviewed studies (Arksey and O’Malley 2005: 22). These working definitions allowed us to begin the review process, while taking into account some key methodological implications and limitations, as we discuss below.

The review includes EM/CA work published in English, German, and French. We are aware that this excludes

relevant work published in Japanese and Chinese, amongst other languages, due to lack of our language competence. We started working on the review in 2021 and the last retrieval was on 21 December 2022, shortly after the onset of the current wave of public interest in AI based on the wide availability of large language models and ‘generative AI’. Our article thus offers a snapshot taken at the point in time when it was already clear that the topic would soon become even more prominent as further studies began to appear. In the discussion, we briefly reflect on the most recent directions in EM/CA research of AI in situated action.

Since our target studies fell between disparate fields and appeared in many different journals, conference proceedings, and (often less well indexed) edited collections, we used a range of specialist bibliographies and scholarly search engines. Following Mayor and Bietti (2017), we began collecting relevant texts using the *EMCA Wiki*,¹⁰ a specialist bibliography database that has been systematically archiving metadata of all publications in the field (primarily books and journal articles). The Wiki’s editorial policy considers a textual self-identification with or substantial relevance to EM and/or CA as the only criterion for inclusion. A search of the *EMCA Wiki* provided 76 studies that self-identify as related to or grounded in EM/CA and at the same time deal with various AI-related technologies. Of these, 30 texts presented findings that fell within our working definition of AI. To ensure that our collection was as complete as possible, we also used several academic search engines: *ACM*, *IEEE*, *LLBA*, *Springer*, and *Web of Science* (see Appendix 1 for the search strategy used). Only studies that focused on the local order of interacting with and around AI, and that employed the analytical orientations described above were included in our corpus. This secondary search yielded 18 further studies. Five more articles were found through ‘snowball’ sampling by examining references from the articles already collected in this way. We searched the text of these articles to ensure they either discussed EM/CA approaches or made explicit use of their conceptual and methodological apparatus.

This sequence of steps yielded our final corpus of 53 text units in total, published between 1994 and 2022 (13 were older than 10 years): 4 book chapters, 26 conference papers, and 23 journal articles.¹¹ While the conference papers were all published in venues linked to the fields of HRI (~42%), HCI (~50%), and HAI (~8%), most full-length articles were published in sociological journals (~58%). The reviewed studies appeared across a diverse range of disciplinary venues including linguistics, clinical medicine, philosophy,

⁸ One of the available labels for this ‘family’ is the acronym STAARA, meaning “Smart Technology, Artificial intelligence, Automation, Robotics, and Algorithms” (e.g., Zhang and Jin 2023). On the matter of defining ‘social’, ‘intelligent’ machines, see also Jones (2017).

⁹ Of course, many technologies can be said to perform work historically undertaken by humans. For example, ‘calculator’ was originally a human job title, but note that electronic calculators (as we understand the term today) are no longer also treated as human interactants.

¹⁰ Available on-line: https://emcawiki.net/EMCA_bibliography_database.

¹¹ For the sake of consistency and comparability, book-length monographs as well as conference contributions shorter than three pages were excluded from the collection.

psychology, engineering, and communication. We synthesized the studies with regard to four different aspects: the technology under examination (robot, voice assistant, etc.), how it operated (autonomous, Wizard of Oz, etc.),¹² how the experiment was set up or which settings/participants were studied, and which interactional phenomena were analyzed.

4 Findings: technologies, interactions, and praxeology of AI

We give a general overview of the results of our scoping review in Sect. 4.1, while Sect. 4.2 outlines the general trends we identified in the kinds of technologies and users involved in the EM/CA studies in our corpus. As a result of our inclusion criteria, most of these studies apply CA's conceptual apparatus, its approach to data analysis, and core CA phenomena such as turn-taking, repair, openings and closings.¹³ We summarize the empirical findings reported across this corpus centering on three key themes: opening and closing the interaction, miscommunication, and non-verbal aspects of interaction. A first, overall insight from our scoping review is the observation that by specifying a particular approach to empirical materials, our working definition excludes a significant body of research that adopts alternatively empirical EM/CA approaches to exploring the basis of AI as a social phenomenon.¹⁴ Since it may extend beyond the material of talk-in-interaction, this work

often engages more reflexively with the presuppositions of 'humanness' and 'artificiality' that underpin the construction of the interactional settings and roles featured in our scoping review corpus. We, therefore, discuss this important and complementary body of work in relation to the results of our scoping review in Sect. 5.1 of the discussion.

4.1 General trends

4.1.1 Technologies studied

The 53 studies in our corpus feature a wide range of technologies (see Fig. 1). Robots were studied most often ($n=27$) followed by Voice User Interfaces (VUI; $n=13$) and Virtual Agents (VA; $n=9$). One article investigated how technical agency is granted to an artifact by comparing interaction with a virtual agent (*Max*) to interaction with a walking aid (Krummheuer 2015a). Overall, there is a clear tendency in our corpus towards studies of technologies that involve the use of spoken language such as VUIs, VAs, and social robots.

The robots studied were mostly humanoid, although Muhle (2008) studied interaction with an *Aibo* robot dog and Pitsch and Koch (2010) presented a case study of a toddler interacting with an advanced toy robot dinosaur named *Pleo*. In both studies, the robot was programmed to act in a way that would resemble animal-like rather than human-like conduct. By contrast, Payr (2010, 2013) reports on a study of *Nabaztag*, a robot bunny and home companion that was programmed to perform the role of a health coach by greeting users, asking them about their day, and suggesting health related activities like exercise or weighing themselves. Most studies featured humanoid robots including one-off studies of *Lekbot*, *Robota*, *Robovie-R* and *BIRON* (all $n=1$), though some more widespread humanoid robots such as *Nao* ($n=6$), *Pepper* ($n=4$) and *Cozmo* ($n=3$) appeared in multiple studies. One study focused on the movements of industrial robot arms, to which the researchers added a screen that enabled it to present the user with different gaze patterns (Fischer et al. 2015).

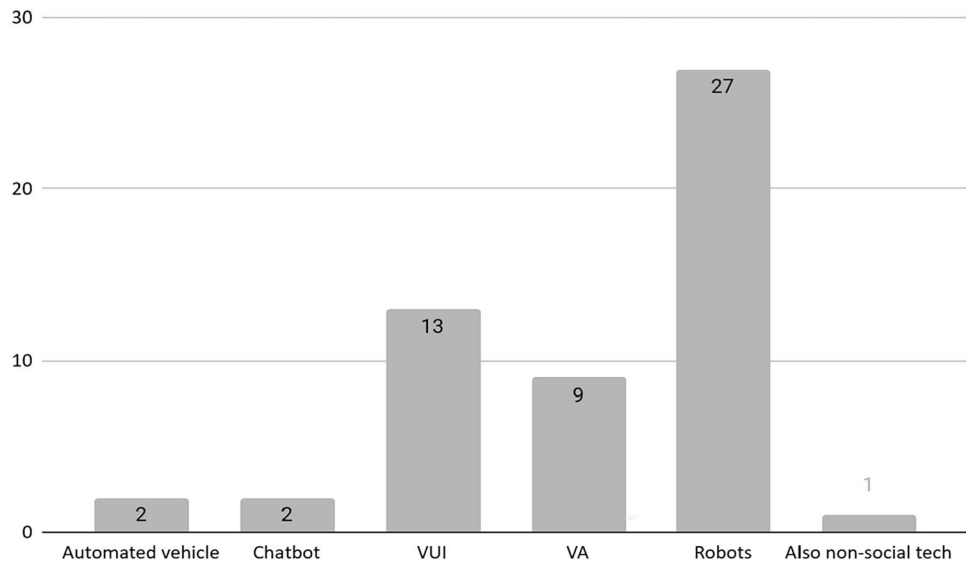
In the category of VUIs, two distinct types of technology are used. First, there are smart assistants such as *Alexa*, *Siri* and *Google Assistant* (Alač et al. 2020; Fischer et al. 2019; Porcheron et al. 2017, 2018; Velkovska et al. 2020). Second, there are telephone systems such as *Lenny* that simulate real callers (Sahin et al. 2017; Relieu et al. 2020), and telephone systems that act as operators of some sort (Aranguren 2014; Avgustis et al. 2021; Wallis 2008; Wooffitt 1994) or do automated interviews (Klowait 2017).

¹² "Wizard of Oz" is a setup where a human 'wizard' is, unbeknownst to the user, in control of the system's actions. It is often used to conduct proof-of-concept or user studies while the development of the technology is still ongoing. Conversely, an AI-based technological device is considered 'autonomous' when it can interact with a user without the on-going involvement of a human operator during the interaction.

¹³ This provides an immediate indication of what our scoping review selection criteria has ruled in—and out—of our corpus, and furnishes contextual cues for evaluating the other findings of the review. The centrality of recordings and transcripts in all these studies implies a specific "commitment to empirical phenomena" (Hilbert 1990) founded on ethnographic observations of the achievement of local orders in social interaction. For discussions of the relationship between ethnography and EM/CA see, e.g., the special issues introduced by Meier zu Verl et al. (2020), and Eisenmann et al. (2023b). The analyst, in this research context, steps back from their own involvement and specific competence in the scene's situated production, using "members' knowledge" (as competent interactants) as a resource for analysis, without necessitating further examination of their situated role. Such a perspective, implying an analyst's distance from an investigated 'subject', is incommensurable with the "radical empiricism" (Garfinkel and Rowan 1955: 8) of EM/CA studies that use necessarily reflexive methods such as self-instruction (e.g., Sormani 2020, 2023) or EM-informed phenomenology (e.g., Sudnow 1983), as well as hybrid studies of work (Lynch 2022), in which members' knowledge and practical involvement becomes a topic of inquiry (Pollner and Zimmerman 1970).

¹⁴ We thank our anonymous reviewer for emphasizing the importance of this issue in their constructive recommendations.

Fig. 1 Technology studied in the reviewed articles, including VUIs (Voice User Interfaces) and VAs (Virtual Assistants)



There was also a variety of VAs (sometimes also called Embodied Conversational Agents) studied in the corpus. For example, the agent *Max* consisted of a cartoon like 3D body projected onto a screen with which passersby in a shopping mall could interact using a keyboard while the agent gave verbal responses (Krummheuer 2008a, b, 2009, 2015a; b; Krummheuer et al. 2020). Two studies analyzed interactions with an agent that was somewhat similar to *Max*: a “Wizard of Oz” controlled agent named *Billie*, which also consisted of a cartoon-like body, visible on screen from the hips up, but rendered in a 2D style and able to interact with users entirely through speech (Cyra and Pitsch 2017; Opfermann et al. 2017). Lastly, two studies used more realistic-looking talking heads: a system controlled by a human wizard that was acting as a therapist for participants role-playing as patients (Torre et al. 2021), and an autonomous system asking a series of pre-recorded diagnostic questions in a memory clinic (Walker et al. 2020).

Lastly, some studies addressed automated vehicles (Brown and Laurier 2017; Pelikan 2021) or chatbots (Corti and Gillespie 2016; Jentzsch et al. 2019). While the relatively low number of automated vehicle studies could be explained by the technology only recently emerging for applications ‘in the wild’, it is noteworthy that there have been relatively few studies of chatbots despite these systems having existed for decades, and recently becoming commonplace and controversial in real-world contexts (cf. Eisenmann et al. 2023a).

4.1.2 Technological set-ups

In addition to the variety of technologies studied, the setup of the technologies varied. The vast majority of studies

addressed autonomous systems ($n=43$) (see Fig. 2). Within this category we counted all technology that was not manually controlled during the interaction. However, note that these autonomous systems had widely varying levels of interactional competence. Some could only speak pre-recorded lines (e.g., Walker et al. 2020), or perform a very basic script (e.g., Licoppe and Rollet 2020), or some mixture of both (e.g., Sahin et al. 2017; Relieu et al. 2020). Of these autonomous systems, 22 were robots, 12 were VUIs, 6 were VAs, two were automated vehicles, and one was a chatbot.

Aside from autonomous systems, a Wizard of Oz setup was also used in 11 studies (twice in Iwasaki et al. 2019). Although this type of setup does not technically involve AI-systems, we chose to include these texts on the basis that the absence of AI is not evident to the human participant. In addition, Wizard of Oz is a common technique used in the broader field of HCI to emulate human-like interactional roles and competences. In the following sections, we will point out some common tendencies in the Wizard of Oz-based studies to indicate how including these articles might have affected the overall trends we found. With regard to the kind of technologies used in this subset there were no clear trends (7 robots, 3 VAs, 1 VUI, 1 chatbot).

Some studies examined computer systems that were presented to the user as human (these were all autonomously functioning systems). Two studies address the *Lenny* system (Relieu et al. 2020; Sahin et al. 2017), a voice chatbot that can be used against unwanted callers such as telemarketers or scam calls. By playing pre-recorded lines when the caller is silent, this system is designed to create the impression that the caller is speaking to a human being. One study explored the impact of user expectations and mediation by having conditions ranging from ‘autonomous system’ to ‘disguised as human’ (Corti and

Fig. 2 Setup of the technology discussed in the reviewed articles. Note: All systems disguised as a human were autonomous and are included in that column as well

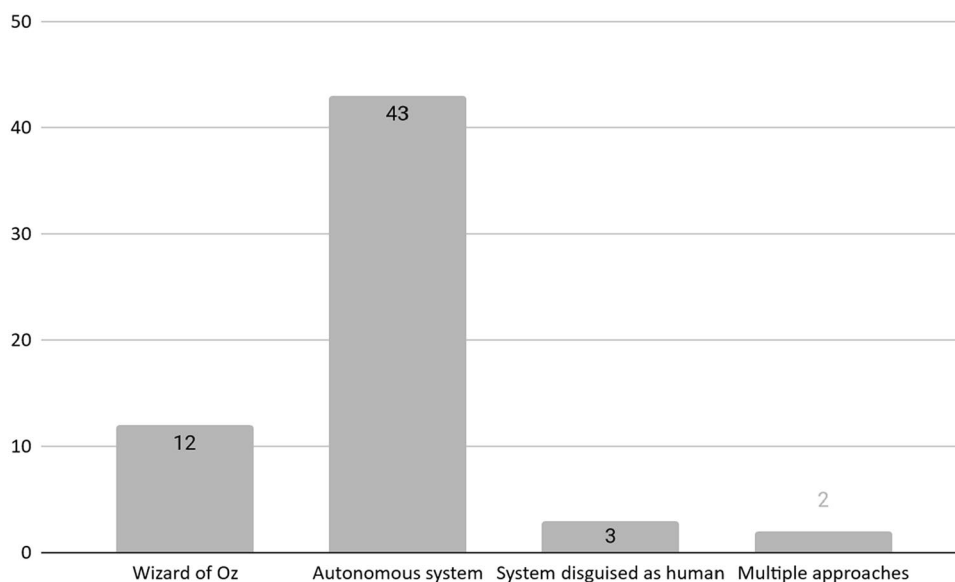


Table 1 Additional features of settings and participants in the reviewed papers

Feature	No. of papers
Public space (e.g., museum, university hallway, school, shopping mall)	15
Real-world telephone calls	5
Adult–child constellations (e.g., households, teacher–child, researcher–child)	10
Households (e.g., couples, families, dormitories)	8
Researchers are part of the interaction	8
People with (mild) cognitive impairments people (e.g., dementia, Acquired Brain Injury, Autism, Cerebral Palsy)	7
No specific additional aspect	19

Note: Categories in this list are not mutually exclusive, e.g., ‘Adult–child constellations’ overlaps six times with ‘Households’ and six times with ‘Researchers are part of the interaction’

Some studies used different datasets (e.g., Krummheuer 2015a; Pelikan et al. 2020). These datasets have been counted separately

Gillespie 2016). This study had participants interact with chatbots in four different conditions: chat vs. face-to-face (a human voicing chatbot responses), and a condition in which they were first informed that they would be interacting with a machine vs. an uninformed condition (Corti and Gillespie 2016).

Lastly, two studies used multiple approaches in their set up. One collected data with an autonomous robot and with Wizard-of-Oz-controlled robots (Alač et al. 2011). The other ‘multiple approaches’ study included a lab experiment with a Wizard of Oz set-up to discern desirable conduct for the robot, the findings of which were then used to program the robot and then to test the Wizard of Oz set-up in a field study (Iwasaki et al. 2019).

Our corpus showed the diversity of technologies studied, and the following trends within the reviewed body of research: robots are studied most frequently, and most studies focused on autonomously operating (rather than manually operated) systems. Next, we review the settings,

categories of human participants, and the activities involved in interaction with these technological systems.

4.1.3 Participants, activities, and settings

Aside from the technology used, there was also variation amongst the human participants studied with regard to age (e.g., studying a specific age group such as toddlers, students or older adults), languages spoken, and other factors (e.g., adult–child constellations, people with cognitive impairments, data collection in a public setting) (see Table 1). Participants were also engaged in a variety of activities with the technology including tutoring a robot to perform a simple task; playing a game with or through the technology; being coached by technology; encountering the technology in a daily activity (e.g., shopping mall); and routine use of already-owned technologies (e.g., querying *Alexa*) (see Appendix 2 for activities and a non-aggregated overview of all articles).

Table 2 Approach for data collection in the papers, including AVs (Automated Vehicles), VUIs (Voice User Interfaces) and VAs (Virtual Assistants)

Type of data collected	Papers (<i>n</i> = 53)	AV	Chatbot	VUI	VA	Robot
Naturalistic data (i.e., interactions with technology without researcher involvement, e.g., no participant recruitment)	11	2	0	7	0	2 ^a
Semi-elicited naturalistic data (i.e., interactions with technology through researcher involvement, but outside of lab setting)	23	0	0	5	5 ^b	13 ^c
Experimental data (i.e., participants recruited and invited to lab setting)	20	0	2	1	4	13 ^c

^aConcerns data of designer-instigated interactions between a robot and users, so while there was no researcher involvement, these are differently naturalistic from data with systems users might already acquired themselves

^bAll five papers used the same dataset

^cOne paper discussed both experimental and semi-elicited data (Iwasaki et al. 2019), hence 28 robots

The reviewed studies rarely addressed a specific age group (*n* = 37), with the category of ‘teenagers’ (12–18 years) being underrepresented; no studies focused on this group in particular. There were some trends in the corpus of studies regarding specific settings or participant groups (see Table 1). For one, some studies had participants with specific cognitive impairments, such as (mild) dementia, Acquired Brain Injury, Autism or Cerebral Palsy (*n* = 7). Furthermore, some studies specifically focused on interactions in which one or more children interacted with technology together with (an) adult(s). In one case, this adult was the researcher who was present to ensure the toddler would not break the robot but who also interacted with the child and the robot (Pitsch and Koch 2010), but generally the adult was a guardian or teacher. The adult–child category overlaps four times with studies looking at interaction between households (couples, dormitories, families) and technology (*n* = 6). Households also offer opportunities to study the technology in its ‘natural’ or designed-for habitat. By contrast, some studies had a researcher present during the interaction (*n* = 4), which is arguably less natural. In all these studies, the researcher’s conduct was unscripted and thus studied as part of the interaction. Lastly, many studies concerned data collected in a public space, such as a museum, university hallway, or shopping mall (*n* = 18), or used real-world telephone calls (*n* = 5).

In the Wizard of Oz studies in our corpus (*n* = 9), there were no clear trends in terms of participant categories (8 adult/non-specified, 2 older adult; 8 no additional features, 2 (mild) cognitive impairments).

While collecting data in everyday or institutional settings is in line with the approach of EM/CA, which generally takes ‘naturally-occurring’ and ‘naturally organized’ ordinary activities as its empirical material, many studies collected data in an experimental setting (*n* = 20) (see Table 2). Although, as Dourish and Button (1998: 406) note in their discussion of Suchman (1987), “laboratory studies are hardly the stuff of ethnomethodology”,

much of the research reviewed here has been done in labs. Other methods of data collection involved some researcher involvement, such as recruiting participants and/or putting the robot in its designed-for environment (*n* = 23). Relatively few studies used naturalistic data (*n* = 11), i.e., recordings of interactions that would have occurred without researcher involvement. This trend seems related to the technology’s occurrence in everyday life: automated vehicles and VUIs (including telephone systems) overwhelmingly used naturalistic data (automated vehicles [AVs] = 2 out of 2, VUIs = 7 out of 13), whereas interactions with robots and VAs were commonly collected through researchers’ involvement (VAs = 5 out of 9, robots = 13 out of 28) or experimental settings (VAs = 4 out of 9, robots = 13 out of 28).

4.2 Interactional phenomena

There were clear trends in the focal interactional phenomena explored by the studies in our corpus (see Table 3) with the three key topics being: (1) how interactions with AI devices are opened and closed; (2) miscommunication and how it is resolved (i.e., conversational repair); and (3) non-verbal communication and emotion displays.

4.2.1 Opening and closing interactions with AI in situated action

The studies in our corpus recurrently dealt with openings and closings in interactions with AI (6 out of 53, and a section in the analysis of 4 more papers). These included openings and closings with robots (*n* = 6) and telephone systems (*n* = 3). Most studies focused on openings while only two papers examined how interactions are closed (Licoppe and Rollet 2020; Payr 2010). In this section we outline how EM/CA studies of AI treat these foundational interactional phenomena (see e.g., Schegloff 1968; Schegloff and Sacks

Table 3 Overview of the key topics in the empirical articles

Key topic addressed	Main topic	Touches on
Opening and closing interactions with AI	6 (~11%)	4
Miscommunication and repair in interacting with AI	14 (~26%)	5
Non-verbal communication and emotion displays in interacting with AI	14 (~26%) (of which 2 on emotion)	5

Note: Key topics are not mutually exclusive

1973), as they appear to be reconfigured in encounters with AI.

Establishing mutual recognition and accessibility is core to opening an interaction, and is usually accomplished in human–human interaction through a multitude of verbal and non-verbal resources (e.g., see Kendon 1990; Pillet-Shore 2010; De Stefani and Mondada 2018). The studies in our corpus show that the same is true for human–AI interaction. In HRI, participants accomplish openings with gaze playing an important role, similar to openings in human interaction (Gehle et al. 2017; Pitsch et al. 2009). For example, a robot that restarts its sentence when it loses the addressee’s gaze is more successful in getting their attention and thus opening the interaction (Pitsch et al. 2009). Similarly, Iwasaki et al. (2019) found that a robot that returns a prospective users’ gaze during a greeting-and-opening sequence receives responses much more often than if it uses only verbalized greetings (e.g., “May I help you?”). They also suggest that people’s initial impressions and expectations of a robot’s perceptual capabilities significantly change their stance towards the robot and condition whether they will engage in a two-way interaction with it. Süssenbach et al. (2012) make a similar observation in their case study exploring pre-opening interactional activities such as how a robot is presented to a novice user by someone familiar with the system. They show that the user’s initial expectations are shaped by how the robot is first introduced. Both studies suggest that the initial framing of the robot and its abilities to display interactional gaze practices are an important resource in opening an interaction.

For telephone-based systems such as *Lenny* (Sahin et al. 2017), other kinds of paralinguistic resources such as hesitations, disfluencies, and other troubles of speaking are particularly important for creating a strong first impression during openings. *Lenny* is intended to ‘trap’ unsolicited spam, hoax, and telemarketer callers, all of whom are strongly incentivized to stay on the line, by engaging them in conversation with an automated agent. Despite using only pre-recorded turns, *Lenny* is remarkably successful at keeping this facade up as long as possible (average call times are just under 10 min). Apart from the caller’s tacit incentives to stay on the line, Sahin and colleagues (2017) suggest this remarkable success stems from *Lenny*’s openings displaying initial availability and willingness to talk before

complicating the interaction immediately by displaying troubles of speaking and hearing. While these troubles are unrelated to *Lenny*’s apparent willingness to continue, they still take time to resolve. In all these cases, the interactional goals and first impressions of the human interacting with the technology seem to strongly inform the success of the interactional opening in initiating (and then maintaining) ongoing interaction.

Two papers within our corpus address closing interactions with robots. Ending an interaction with a robot is accomplished in a variety of ways including leaving the interaction without doing a closing at all, i.e., walking away without any preparatory interactional work or even mutually acknowledging that the interaction has ended (Licoppe and Rollet 2020; Payr 2010). When closings are done by users, they involve multiple strategies such as the inclusion of pre-closing or closing-implicative moves (e.g., “okay”, see also Schegloff and Sacks 1973) and/or providing an account (e.g., “I have to go”, Licoppe and Rollet 2020). Humans also seem to make pre-closing moves without leaving room for the robot to respond (Licoppe and Rollet 2020). This suggests uncertainty in treating the robot as an ‘official’ interactant because a (pre-)closing sequence orients towards collaboratively closing the interaction, whereas denying the robot an opportunity to (dis)align with the closing does not (Licoppe and Rollet 2020). Reported closing conduct with robots also changes over time (as in human interaction, cf. Berger and Pekarek Doehler 2018). When mapping the closings of one participant over 10 days, Payr (2010) found the participant ended the interaction through both verbal and non-verbal closing moves, waiting for the robot to close, and leaving without closing. While still performing closings, leaving without closings became more frequent over time (Payr 2010).¹⁵ Payr (2010) also points out that the participant orients to social norms in her closings (e.g., providing justification for closing the interaction) and that instances in which the participant leaves without closing look more like turning off a machine (p. 480). So, how closings are done in human–robot encounters is tied to the system’s status in

¹⁵ Payr (2010) has a relatively small number of instances and thus treats this as an illustration of closing distribution, not as a general finding on the occurrence of closings over time in HRI.

the interaction, i.e., being treated (more) as an interactional partner or (more) as an object.

Notably, the papers on closings discuss how the technological system is, in many cases, disregarded as a social entity. Conversely, the papers concerning openings mostly address how a robot can get the user's attention in the first place, providing findings and suggestions as to what makes certain practices work (e.g., Gehle et al. 2017; Iwasaki et al. 2019; Pitsch et al. 2009; Sahin et al. 2017). On the one hand, this offers some key insights into common issues for HRI, e.g., that establishing mutual attention is not a given for these technological systems but requires specific perceptual and behavioral design. This is especially true for robots which, despite perhaps drawing attention or curiosity by virtue of their appearance *as* robots, are not easily able to communicate their availability for interaction (see Pelikan and Broth 2016). On the other hand, while closings are only studied in two of the papers in our corpus, their findings suggest that robots potentially struggle to sustain displays of sociality until the end of an interaction. This "problem of closings" (Schegloff and Sacks 1973: 292) may relate to some of the many issues of miscommunication in human–AI interaction documented in our corpus of EM/CA studies.

4.2.2 Miscommunication

Miscommunication in interacting with AI is another recurring topic in our corpus (14 out of 54, and a section in the analysis of 5 more papers).¹⁶ Of course, miscommunication is a pervasive concern for all participants in social interaction (see Jefferson 2018), which may partly explain why so many papers in our corpus take this as a focus of study. Since the reviewed studies take a fundamentally inductive approach that draws topics from their data (see Sacks 1984a), the prominence of this topic may also be due to the inability of many social-technological systems to sustain social interaction without frequent and unresolved miscommunication.

First, some papers focused on how to help a system identify moments of miscommunication. This is a significant practical issue because to be analyzed and resolved, moments of miscommunication first need to be identified. One paper focused on swearing to find moments of trouble in telephone interaction (Wallis 2008), another characterized a prototypical script and then identified deviations from the script as an indicator of miscommunication in interaction with a robot (Lohse et al. 2009). Krummheuer (2008a) focused on how displays of misunderstanding are done in interaction between humans and an Embodied Conversational Agent.

Second, studies focused on how humans adapt to interaction with an AI system over time by exploring moments of miscommunication. For example, a user may first orient to human social norms for timing their responses but, when this leads to trouble (e.g., the robot continuing its turn and thus overlapping with the user), users will adapt the turn-taking system by, among other things, leaving longer gaps before responding (Pelikan and Broth 2016). When using a self-driving vehicle, users were also found to learn the system's limitations and adjust their own conduct by monitoring the road during autopilot driving, then taking control in situations that, as they have learned from experience, the system tends to struggle with (Brown and Laurier 2017). Trouble may also escalate, with users interviewed by a robot first addressing the trouble by repeating or rephrasing their turn but, when this fails, using more extreme strategies such as resorting to scripted commands (e.g., 'skip') or changing their answer in a way that advances the robot's script (Stommel et al. 2022). Similarly, when facing complex interactional trouble, users of VUIs tend to prioritize restoring the progressivity of the interaction, rather than resolving the miscommunication (Fischer et al. 2019), which follows the broader preference for progressivity in many forms of human interaction (see Stivers and Robinson 2006; Heritage 2007). Lastly, in some cases humans do not appear to adapt to misbehaving technology even when they are experienced and well-informed about it. Pelikan (2021) described how an automated shuttle bus on public roads in Sweden was programmed to apply emergency brakes whenever it encountered a situation it could not handle, such as being overtaken by other road users. However, even after the bus had been on the road for 9 months with a sign on the back warning drivers to keep their distance to avoid triggering the emergency brakes, road users continued to maneuver around the bus, rendering it a static obstacle for other road users and leading to recurrent failures to coordinate shared road use smoothly (Pelikan 2021).

Miscommunication is also sometimes related to user expectations regarding system capabilities. For example, Corti and Gillespie (2016) found that people handle miscommunication differently when they are told that they will be communicating with a chatbot rather than a (presumed) human interactant, initiating other-repair significantly less frequently. Süssenbach et al. (2012) show that users assess the system's competencies step-by-step and that they differentiate between the robot's role as a social actor and the robot's role in that specific interaction (in their case, a fitness instructor). In order to learn more about the system when trouble arises, users also turn to system-external resources when available, such as a manual or a co-present expert such as, in the cases reviewed, the researcher or designer (see Alač et al. 2011; Arend et al. 2017; Muhle 2008). Muhle (2008) notes that this often entails the system being occasionally 'degraded' from being treated as a co-participant to

¹⁶ We use 'miscommunication' as a label that includes misunderstandings, among other troubles in talk, and include repair practices that work to resolve these troubles.

becoming a topic of conversation while users try to figure out how to continue interacting with the machine. With regard to the type of trouble occurring, there can be multiple issues. First, the machine can have trouble hearing (and/or transcribing) the user's voice input correctly or at all. Second, the machine may 'hear' but then fail to recognize and correctly interpret the input. Several articles found that when trouble occurs, users tend to treat this as a problem of 'hearing', despite the system not specifying the cause of the problem (e.g., Avgustis et al. 2021; Stommel et al. 2022). One suggestion for improving design for miscommunication is to provide the user with more relevant feedback on the nature of the problem (e.g., Porcheron et al. 2017; see also Button et al. 2015: 163–165, on run-time accountability, and more broadly also CA work on repair, e.g., Schegloff 1992; Drew 1997).

A key issue of miscommunication that many papers touch on is that the system often lacks access to the same information as the human and vice versa. These technical, perceptual, and design issues can range from sensors being unable to function in certain conditions that would yield no trouble for a human actor (e.g., sunlight preventing the autopilot from making the correct move, Brown and Laurier 2017) to sensors being (temporarily) shut off or not present at all (e.g., certain robots stop 'listening' when producing their turn so they don't get confused by their own audio, e.g., Pelikan and Broth 2016; Stommel et al. 2022).

4.2.3 Non-verbal conduct and emotive displays: human and machinic

Non-verbal conduct and displays of 'emotive involvement' (Selting 1994) were common topics in our corpus ($n = 14$, and formed an integral part of the analysis of 5 more papers).¹⁷ The studies of non-verbal conduct generally addressed systems that have a physical presence (AV = 2, robots = 13) though one study addressed a VA (Torre et al. 2021) and one addressed both face-to-face and text-based interaction (Corti and Gillespie 2016). The two papers on emotion looked at a robot's emotive displays (Pelikan et al. 2022) and patterns of emotive displays in customer calls with a telephone system (Aranguren 2014). Studies of non-verbal conduct described the use of interactional resources including gaze ($n = 7$), smiling ($n = 2$), and physical movement (Brown and Laurier 2017; Pelikan 2021). Sounds, gestures, and body posture/positioning were also addressed, on occasion, though always in the service of the wider analysis (in line with EM/CA findings that interactional resources are 'multimodally' intertwined, e.g., see Goodwin 2000; Mondada 2014).

Across our corpus, there is a key distinction between articles that focus on *human* non-verbal conduct or emotive displays versus those that focus on *machinic* non-verbal conduct or emotive displays. The former focuses on what humans do, either as something that could be used to improve robot design (e.g., showing that a robot sensitive to human gaze is more successful at securing human attention, Pitsch et al. 2009), or describing human non-verbal conduct during human–robot encounters (e.g., gaze and smile patterns between unacquainted children when interacting with a robot, Tuncer et al. 2022). Papers primarily exploring machinic non-verbal conduct or emotive displays focus on robot non-verbal conduct and humans' interactional responses (e.g., a robot applying a social gaze pattern helps users instruct the robot, Fischer et al. 2015). In this section, we discuss the papers in our corpus that deal with these interactional resources together, although we note here that these two approaches carry quite different theoretical, analytic, and design implications.

Most studies find that users tend to draw on their repertoire of practices from non-verbal human–human interaction when interacting with social technology. For example, the way gaze functions as a resource for managing availability for interaction in both human–human and human–AI openings holds true for many other interactional practices. Fischer et al. (2015) compared an industrial robot arm utilizing 'social gaze' (gazing at its human tutor when ready for instruction and otherwise gazing at the field of the task), with a robot that gazed only at the movements of its own arm. Using social gaze, the robot was able to solicit additional instructions from users more quickly than when it used simpler gaze patterns. A study by Pitsch et al. (2013) found that human tutors adjust the way they present instructions (e.g., pace of talk, pauses) depending on the robot's gaze, suggesting that optimizing gaze strategies for specific HRI instructional tasks could elicit more useful user input and more compliant robot conduct. The above, along with other studies of gaze in interactional openings (Pitsch et al. 2009), suggests that gaze and its timing are critical non-verbal interactional resources for managing mutual attention (see also Fischer et al. 2015).

The importance of the precise timing of embodied actions was an important finding for a range of non-verbal conduct. An experiment by Torre et al. (2021) used a virtual head with four different smiling conditions to show that humans do not, as some studies suggest, simply mimic the smiles and timing displayed by a VA. Instead, at smile-relevant moments in an interaction, human users smile in an affiliative way when the VA also produces a smile, and in a disaffiliative way if the VA fails to smile at the appropriate moment. A museum guide robot turning its head from the museum exhibit towards the addressed visitor when nearing turn completion was found to elicit more consistent and

¹⁷ We discuss these topics together because emotive displays often involve non-verbal cues.

nuanced non-verbal responses from the visitor than when the robot moved its head at less interactionally relevant points (e.g., in the middle of a turn constructional unit, Yamazaki et al. 2013). Similarly, Pelikan et al. (2020) found that ‘happy’ and ‘sad’ emotive displays by a Cozmo robot were treated as a response to the immediately preceding actions and that a ‘happy’ display had a different contingent effect on the ongoing interaction than a ‘sad’ display. They found that after a ‘happy’ display the interaction tends to proceed, while ‘sad’ displays function as a sort of repair initiation or “rewind button” where the user’s subsequent talk treats the display as an indication that something needs to be ‘fixed’ before the interaction can proceed (Pelikan et al. 2020). The importance of timing for the uptake of non-verbal cues also applies to automated vehicles where, for instance, the flashing and sound accompanying emergency braking was found to come too late to function as a warning both for the passengers inside to brace themselves, as well as for the cyclists outside (Pelikan 2021).

Gaze and smiling are also often discussed together. For example, Fischer et al. (2015) noted that users smiled more often in the interaction where the robot used social gaze (looking at the user when ready for instructions), and users generally smiled when gaze with the robot was re-established. One article also showed how robots can facilitate mutual gaze and smiling between (unacquainted) children (Tuncer et al. 2022), showing how the non-verbal conduct of users can be mediated by robot facilitation. These studies all point out that smiling and emotive displays by humans should be analyzed as performing a social function rather than interpreted as a reflection of emotional states as such.

Two studies in our corpus address mobile interaction, specifically automated vehicles in traffic. Road traffic is an interactional context in which communication is mostly non-verbal and where mutual understanding is critical. However, the two studies in our corpus show that understanding the conduct of other road users is still difficult for automated vehicles (Brown and Laurier 2017; Pelikan 2021). For instance, speeding up and slowing down are important indicators for the actions a traffic user is about to take (and, implicitly, for demonstrating their perception of the situation), which can lead to trouble when an automated vehicle does not use and/or is not sensitive to these kinds of social signals (Pelikan 2021). This can be an impediment to the smooth performance of even the most routine traffic maneuvers, such as overtaking (Pelikan 2021).

Moving to another modality, AI-based system sounds are also, on occasion, addressed in the corpus, although always as part of a larger analysis. For example, some social robots are designed with listening cues, eye lights, and bleeps designed to inform users when a robot stops and starts receiving input. However, users’ talk often overlaps with these bleeps (e.g., Pelikan and Broth 2016) and these sounds

regularly lead to confusion (Arend et al. 2017). These analyses suggest that non-verbal cues implemented to improve turn-taking in HRI do not necessarily facilitate turn-taking as intended. Conversely, when robot bleeps are done as part of a recognizable action sequence, users tend to interrupt their own speech and yield turn space to the robot (Pelikan et al. 2020). Potentially relevant to these contrasting findings is that Pelikan et al. (2020) studied Cozmo, a robot that only uses non-verbal sounds, whereas the other studies discussed a Nao that took verbal turns (Pelikan and Broth 2016; Arend et al. 2017). Users also sometimes mimicked the robot’s non-verbal sounds by, for example, producing a turn with a similar prosody to Cozmo’s after the robot made a ‘sad’ bleep (Pelikan et al. 2020) or mockingly imitating an Amazon Echo’s repetitive bleeps during interactional trouble (Fischer et al. 2019).

Some non-verbal interactional resources such as gestures, touch, body position, and bodily presence were discussed less often and as part of broader analyses rather than as the sole focus of any one study. Pelikan and Broth (2016) noted that gestures such as waves are sometimes mirrored by the user. Humans also sometimes use gestures to initiate closings, such as presenting a hand to initiate a handshake or waving (Licoppe and Rollet 2020; Alač 2016: 524). Humans also use touch when interacting with a robot, for example by petting the robot after a ‘happy’ or ‘sad’ display (Pelikan et al. 2020). The quality of touch can also indicate how a human orients towards the system, for example grabbing the neck of a robot dinosaur suggests that the robot is being accorded a more object-like status (Pitsch and Koch 2010). Several studies also analyzed how humans position their bodies in ways that indicate their position within a specific ‘participation framework’ (Goffman 1981)—e.g. Licoppe and Rollet (2020) or Alač (2016). Alač’s (2016) analysis of users’ touch and bodily positioning towards a robot also shows how they treat it both as a thing and as an agent. Lastly, with regard to bodily presence, Corti and Gillespie (2016) found that humans initiate other-repair more frequently when interacting with an embodied human co-participant rather than via text chat, even when subjects were told that the human in front of them was only echoing responses written by a chatbot.

Overall, the studies in our scoping review highlight the interactional contingencies of non-verbal communication and emotive displays. They extend existing findings from EM/CA research that show how emotion cannot be simplified into categories such as ‘smiling is happy’ (see also Peräkylä and Sorjonen 2012) or ‘one needs to gaze at someone else at all times’ (Rossano 2012). Across our corpus, the timing and action preceding these moves seems to be crucial to how the interaction unfolds. Gaze plays an especially important role in facilitating social interaction, from opening

and closing the interaction, to providing users with insight into the machine's functioning and how to interact with it.

5 Discussion: respecifying 'AI' as a worldly phenomenon

The previous section provided the results of our scoping review of EM/CA studies of technologized situated action. These studies all focused on the interaction patterns and sensemaking procedures involved in human interaction with and amongst 'intelligent machines'. As mentioned above, however, the selection criteria we developed for this scoping review led to a corpus that includes research mostly exploring the everyday interactional relevancies of AI *users*. The findings presented also reflect the predominance of CA within the broader contemporary field of EM/CA field. This has meant that so far, our review has excluded a wide range of EM/CA studies that examine and critique some of the presuppositions of conducting and grouping together these kinds of ethnographically observational studies, e.g., the notions of 'intelligence' and 'machines'. We now turn to discuss these findings in relation to a body of EM/CA work focused more on the professional relevancies of AI's creators and critics. In a sense, this ordering in our presentation follows the structure of classic EM works (e.g., Wieder 1974, see also Garfinkel 2022c) that first provide the results of an empirical study, and then investigate the constitutive features and conceptual presuppositions that make such an ethnography possible. We therefore begin this discussion with a narrative overview of some of the EM/CA studies excluded from our initial corpus, before discussing their intersections with and differentiations from the studies reviewed in Sect. 4.

5.1 The situated production of 'AI'

So far, our review has skirted the question of the 'artificiality' or 'autonomy' of AI technologies. What is it that makes AI-labeled devices and our interactions with them distinctively *what they are*, as socially situated worldly phenomena? Given the unstable definitions of AI (Caluori 2023; Sormani 2022), EM/CA's focus on the contingent, situated work of producing meaningful social objects as part of everyday and professional activities is ideally suited for asking such foundational questions. Indeed, from their outset in the 1980s, EM/CA-based studies of technology have offered a fundamental and critical respecification of established topics in engineering and computer science (Button et al. 1995; Coulter 2008), proposing that "AI's whole mentalist foundation is mistaken, and the organizing metaphors of the field should begin with routine interaction with a familiar world, not problem solving inside one's mind" (Agre 1997b: 149). However, through the scoping review process and evaluation

of its findings, our selection procedure excluded a body of EM/CA work that has methodologically engaged in forms of radical reflexivity and EM respecification (Pollner 1991, 2012) in favor of the predominant form of applied studies designed to address established discourses and practitioners in HCI/HRI research.¹⁸ Many such studies excluded from our initial corpus address a range of evidential materials and approaches that eschew or implicitly problematize the framing of 'user study' empiricism that many of the studies reviewed above share with HCI. As we outline below, Brooker et al. (2019) analyze chat transcripts and Python computer code; Sormani (2020) combines video analysis with reflexive self-instructive ethnography (building a 'do-it-yourself AI' kit by following the manual); or conducts instructive re-enactments of video demonstrations of an 'agent system' playing the computer game *Breakout* (Sormani 2022; cf. Sudnow 1983).

Radically reflexive and praxeological EM/CA studies offer a distinctive contribution to social studies of AI that couples the Garfinkelian (2002) 'hands-on' approach with the work of 'ordinary language philosophers' such as Ryle and Wittgenstein (Reeves 2017; Brooker et al. 2019; Sormani 2020, 2022; Mair et al. 2021).¹⁹ These studies follow Button et al.'s (1995) critique of central topics in cognitive science, psychology of mind, and linguistics that underpin the notion of 'thinking/talking machines'. They aim to problematize the conceptual foundations, assumptions, and presuppositions of the 'human-AI' interaction research discourses into which many of the EM/CA studies reviewed above were designed to fit. For example, Reeves (2017) points out that behind the ostensible engineering challenges of designing VUIs lie basic problems with the language and concepts we use for describing conversation itself, and methodological issues with applying CA findings derived from human-human interaction to 'human-machine' interaction.²⁰ Others highlight the lack of reflection and investigation into common

¹⁸ This exclusion may, in part, be due to the process of ethnomethodological respecification (i.e., reformulating the questions, materials, and philosophical starting points of a discipline as thoroughly praxeological problems) often being seen as tangential to the ostensible aims and topics of the field being respecified (Button 1991).

¹⁹ The relationship of ordinary language philosophy and EM/CA, not to mention ethnomethodological hybrid studies, is too complex to be covered here. Garfinkel, Sacks and their colleagues were familiar with and inspired by Wittgenstein's philosophy since the early 1960s (see, e.g., Garfinkel 2019a, b, c [1960]; Sacks 1992: 26), but the links remained mostly tacit. They have been made explicit by the 'Manchester school' of EM (Psathas 2008), and they are being developed further in the recent related work concerned with outlining a 'Wittgensteinian ethnomethodology' (e.g., Hutchinson 2022).

²⁰ Already in 1980, Schegloff has noted that "[i]n the design of computer interactants, and in the introduction of technological intermediaries in human-human interaction, the issue remains which type of person-person interaction is aimed for or achieved" (Schegloff 1980: 81)—i.e., that 'conversation' is but one of a number of many different

ways of speaking about ‘AI’ (e.g., Suchman 2023b) that ascribe psychological and agentic properties and contribute to ongoing conceptual/philosophical confusion on the nature of the phenomenon. For instance, an early study by Suchman and Trigg (1993) analyzes interaction between two AI researchers as they discuss technology and theory of mind. Their complex connections between the social world and its machinic representations recasts professional work in AI as series of interrelated re-representations. These start from the researchers’ experience of the world and extend through a textual scenario that stands as a proxy of the experience, to formalisms inscribing the scenario and its coded versions implemented in a machine, which is itself eventually reintroduced in the social world through interaction with human users.

Another radically praxeological approach involves ‘self-instructive practice’ through which, for example, Sormani (2020) engages in the activity of assembling a device advertised as ‘DIY AI’. In doing so, he encounters a series of unexpected problem–solution pairs that highlight problems of instructions and their enactment, as well as the tensions between marketing discourses and technical work. Similarly, Brooker and Mair (2022: 243) propose that social scientists engage in “hands-on ethnographic exploration of machine learning from within” by learning to code and doing “Programming-as-Social-Science” (Brooker 2019). Through these forms of radical praxeology that place AI in its practical contexts, we can study it as a social praxis involving configurations of humans, machines, and their interrelations²¹ rather than misattributing cognitive capacities to ‘ghosts in the machine’ (Brooker et al. 2019; Mair et al. 2021). Ziewitz (2017) adopts a similarly pragmatic EM approach to examining algorithms as instruction-delivering devices in an experimental study of walking where ‘decisions’ and ‘directions’ are grounded in an ad hoc algorithm rather than maps or conventional navigation systems. Algorithmic walks explore the conceptual and praxeological foundations of ‘AI’ and its social implications by showing how “any recourse to the figure of the algorithm is itself a practical accomplishment” (p. 12). These studies provide a foundation for a critical and deflationary approach to ‘AI’ rooted in the aim of technologists to build what Agre (1997b: 140) calls “suitably narratable systems” or, to use a more contemporary gloss, ‘explainable’ AI (see Albert et al. 2023a). Through conceptual inquiry, self-instructive practice, and other empirical engagements, these radically praxeological EM/CA studies

unpick the vernacular concepts of intentionality, agency, and accountability that underpin the constitutive metaphors of ‘AI’, and explicate how they are drawn upon in situated actions.

5.2 Heuristic tensions in EM/CA approaches to HCI

Having provided an overview of the studies missing from our scoping review, we see two distinct approaches emerging from a wider corpus of EM/CA studies of AI in situated action. As Dourish (2006: 544) argues, as well as providing findings that address the established frame of “implications for design” in HCI, EM/CA studies can defer and reflexively transform design-oriented analytic objectives into an “occasion for tacit theorizing”. On the one hand, more HCI-oriented studies in our corpus offer design recommendations to improve a specific technology (e.g., Wallis 2008; Opfermann et al. 2017; Pelikan and Broth 2016), often drawing on—and contributing to—theories, methods, and findings from human interaction research (e.g., Pelikan et al. 2020; Krummheuer 2015b; Gehle et al. 2017). To some extent, these studies take the anthropomorphic distinction between human and machine in HCI for granted, or at least side-step the issue to focus on interactional practices and contribute to established HCI discourses. On the other hand, studies that respecify HCI’s core topics and theories—often involving the same researchers—aim to deconstruct central issues of AI’s agency and artificiality (Pelikan et al. 2022; Krummheuer 2015a; Alač et al. 2011). Highlighting how AI systems are treated alternately as social agents or as material objects in interaction (Alač 2016; Gehle et al. 2017; Pelikan et al. 2022), these studies offer a fundamentally different approach to the way anthropomorphism is often seen as a ‘factor’ in HCI (Nass and Moon 2000; Heijseelaar 2023; Fischer 2021). This approach shifts focus from how well or badly machines might be designed to emulate human interaction to exploring the social uses of anthropomorphism in HCI, strictly resisting the conflation of “computational processes with human minds through a cognitivist/materialist/behaviourist lens” (Brooker et al. 2019: 273). These distinct approaches produce what Sormani and vom Lehn (2023), introducing a recent collection of studies developing Garfinkel’s legacy, call “heuristic tensions ... between analytic detachment and practical involvement” among EM/CA social studies of AI.

These tensions are present throughout our corpus in the distinction between ‘naturally occurring data’ and ‘naturally organized ordinary activities’ that characterizes EM and CA work (Lynch 2002). They are also methodologically embedded in EM/CA’s analytic reliance on meaning as interactionally and dynamically produced, moment by moment. Whether aiming to contribute to HCI or respecifying its premises, EM/CA provides a situated perspective on

Footnote 20 (continued)

speech-exchange systems for the organization of talk (see also Sacks et al. 1974).

²¹ See also more recent publications that explore the notion of socio-technical ‘assemblages’ in the context of EM/CA (Due 2023; Raudaskoski 2023).

AI design and prototyping (e.g., Suchman et al. 2002) that resists reductive reifications of meaning and technology-centric logics (Garfinkel and Sacks 1970). Technologists also face these tensions in implementations that take EM/CA findings into account. As Rollet and Clavel (2020) argue, a central design question for studies of AI in situated action remains: how, if at all, can technologists formalize the situated particulars of interaction sequences as ‘information’ that the machines can process? These considerations attest to the continuing relevance of Button et al.’s (1995: 196 ff.) powerful discussion of the “unformalizability of conversation” (see also Button 1990; Button and Sharrock 1995). Despite decades of innovation and technological advances, the studies in our corpus suggest that Suchman’s (1987) foundational questions about design for human–machine interaction remain fundamentally unresolved. We have also identified these heuristic tensions in our own reviewing process. Charting a ‘body of research’ within our own research domain requires us to adopt a position of analytic detachment and, as if it were possible to do so, to suspend reflexive inquiry into the practical involvements and premises of ‘doing scoping’. Nonetheless, in outlining the contribution of EM/CA studies of situated action and AI-based technology to the broader field of social studies of AI, the findings of our review suggest not only ‘implications for design’ of AI systems, but also implications for EM/CA research itself. In this sense, our review opens new trajectories for “navigating incommensurability” between EM, CA, and AI (Reeves 2022). Before returning to reflect on the scoping process, we outline some key points of intersection between the studies in our corpus and ask how they relate to the theoretical and methodological literature in EM/CA studies of AI and technology more broadly.

5.3 Implications for EM/CA research and for technology development

Most of the studies reviewed in Sect. 4 were written for an HCI and technology audience. Although the majority focused on humans interacting with autonomous robots, virtual assistants, and voice user interfaces, these studies could also contribute general findings back to EM/CA’s ‘core’ fields of human sociality, language, and interaction. As Schegloff (1987: 102) points out, even analysis of single episodes of interaction conducted in highly specialized circumstances can contribute to a systematic understanding of the “bedrock of social life”. While our review found that EM/CA studies of AI, mostly grounded in existing research in CA, tended to focus on beginnings and endings, miscommunication, and non-verbal and emotive displays, there were many more EM/CA phenomena mentioned in passing that could be expanded on, and some for which AI presents a particularly ‘perspicuous setting’ (Garfinkel 2002) for empirical analysis. For

example, studies of recipient design in talk to/with robots (Pelikan and Broth 2016; Avgustis et al. 2021; Tuncer et al. 2023) reveal users’ assumptions about the interactional competence of their (robotic) co-participants, and demonstrate the methods they use to make themselves understood given those assumptions. These findings, and the possibility of conducting them both ethically and systematically in an HRI context may have wider implications for applied EM/CA research in so-called ‘atypical’ interaction involving disabled people, whose competence and, as with AI, whose ‘intelligence’ and personhood are often called into question interactionally (Walton et al. 2020; Wilkinson 2019). If taken up more fully by EM/CA researchers, studies of AI in situated action could contribute valuable and ‘transferable’ understandings (Ziewitz 2017) of how displays of personhood, intelligence, agency and autonomy are avowed and ascribed in interaction (Antaki and Crompton 2015; Sidnell 2017; Pelikan et al. 2022).

Relatedly, our scoping review found that robots and VUIs receive much more attention than other AI-based devices and systems. This might be because these technologies are regarded as closer to face-to-face interaction, and therefore amenable to established EM/CA methods, theories, and conceptual frameworks. Indeed, much of the work reviewed in this paper comprises the application of concepts and findings from EM/CA studies of human–human interaction to the realm of interacting with ‘autonomous’ or ‘intelligent’ machines. On the other hand, we have also identified a set of studies that critically assess the very claims of ‘autonomousness’ and ‘intelligence’, and exploring the grounds of “the fantasy of the sociable machine” that has been a “touchstone for research in humanlike machines” (Suchman 2007: 235). These studies are closely related to what Sormani’s (2019) overview of ‘ethnomethodological analysis’ locates as conceptual analysis and practical/self-instructive analysis. They remind us that understanding ‘AI’ as a distinctive social phenomenon requires grasping it in its own terms—both as a professional technical domain (Suchman and Trigg 1993; Sormani 2020; Brooker and Mair 2022) and as an area of everyday action with its vernacular sense of ‘conversations’, ‘algorithms’, and ‘agency’ (Reeves 2019a; Pelikan et al. 2022; Ziewitz 2017; Housley et al. 2019; Velkovska and Relieu 2020). This brings us to the consideration of what EM/CA studies of AI-labeled technologies can contribute to AI development and evaluation.

The studies of existing AI-labeled technologies in our corpus most often took place in (semi-)experimental settings. EM/CA studies focus on exploring whether and how machines constitute proper interactional parties, or to what extent the human and non-human participants are treated differently in interaction (Arend et al. 2017; Licoppe and Rollet 2020; Reeves and Porcheron 2022). A situated approach to such ‘assessment’ of AI is especially useful in some contexts since the interactional requirements of specific contexts are

so variable. For example, in medical diagnosis Walker et al. (2020) show how a degree of ‘rigidity’ in the technological implementation of a survey-taking robot is useful, even if it may seem less human-like, since consistency in question design and performance might elicit more comparable and analytically useful answers to diagnostic questions. Similarly, Avgustis et al. (2021) propose that for some conversational agents used in service phone calls, having a more robot-like agent would reduce unmet user expectations and produce more fluent interactions. While Caluori (2023) points out that human-likeness is a definitional criterion of AI, these EM/CA findings suggest that it is only desirable to emulate human-like conduct when that outcome suits the practical requirements of the situation. In this regard, EM/CA studies could respecify the ‘uncanny valley’ (Mori 1970) as a thoroughly praxeological phenomenon as observable through interactional details.

EM/CA studies are also conducted at the level of technology implementation by mapping how participants may opportunistically and creatively (re)configure AI-labeled technologies for their own routine activities (see also Albert et al. 2023b). As technology becomes part of everyday life, research questions can move beyond the pre-defined experimental goals of a study to discover previously unimaginable phenomena in the data (Tuncer et al. 2022; Sacks 1984a). Pelikan (2021), for instance, points out that in the case of autonomous vehicles, coordination is often studied in restricted environments such as intersections. However, subtle coordination happens even in mundane activities such as overtaking, and here autonomous vehicles often struggle (see also Brown and Laurier 2017). Research in naturalistic settings also discovers new types of ‘user work’, such as coordinating multiple conversational agents in a household, and the asymmetry of their use within families may disrupt or reorganize established interactional practices (Velkovska et al. 2020; Albert et al. 2023b). One of the most fundamental recommendations of EM/CA is that ‘AI’, as a recognizable social phenomenon, is ‘enabled’ (Jaton and Sormani 2023) by various kinds of work on the part of AI’s ‘human users’. In his unpublished research on ELIZA and similar early ‘chatbots’ in the late 1960s, Garfinkel looked at “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” (Eisenmann et al. 2023a: 3). Since the early 1980s, EM/CA research has specified this form of accountability as a fundamental feature of human–machine interaction. Concurrently, the development of new technologies and their implementation in the social world is continually transforming the forms of social life being studied (see Mlynář and Arminen 2023), making novel topics available for detailed description and critical inquiry.

Having discussed the heuristic tensions between EM and CA studies of AI in our initial corpus and in the sub-set selected for our scoping review, and their implications across a range of fields, we return to a concluding reflection on the scoping review process.

6 ‘Doing scoping’: limitations and future directions

The work of conducting a ‘scoping’ literature review as an established method involves crafting representations of various empirical fields and research strategies, while glossing over their differences for the sake of a structured presentation of ‘results’. Nevertheless, as we noted above, the visualizable structures and describable trends that our review work uncovers in the reviewed domain of scientific literature seem deeply grounded in “the uneasy relationship between CA’s ethnomethodological origins and its development into an empirical social science” (Lynch 2002: 531). EM, and ethnomethodological CA, in many ways elude any easy ‘reviewability’ of their findings. One of the reasons is that the topics of inquiry and analyzed phenomena are never to be found in the textual items of EM/CA’s corpus of literature, and neither are they present in the accounts of how the texts came about. In their fullness, the phenomena are only to be encountered in the world, as part of the lived activities in which they originate and which they reflexively constitute. As we tried to show, the field of ‘AI’ can gain relevant insights from the EM/CA ‘approach’, but the crux of the work is to be done elsewhere, by working in the midst of the thing that is being ‘approached’. The EM imperative is to “see for yourself the infinite variety of everyday local methods of being in the world through collections of empirical demonstrations” (Brooker 2022: 5).

Developing Dourish and Button’s (1998) considerations of ‘technomethodology’, Crabtree (2004) notes that attempts to combine EM/CA with technology design “integrate a softer, more user-friendly version of ethnomethodological inquiry with other approaches to design”, thus placing EM in a “service-provider role having little or no strategic value or impact on design practice” (p. 196). Seeking a stronger position for EM/CA studies, Crabtree works out Garfinkel’s notion of ‘hybrid studies’, in which ethnomethodological analysis aims to contribute as much to the investigated domain (e.g., robotics, natural-language processing, machine learning) as it does to social science (see Eisenmann and Mitchell 2024; Garfinkel 2002, 2022a, b, c; Ikeya 2020). Indeed, some of the most recent developments in EM/CA studies of ‘AI’ have moved in precisely this direction (e.g., Ivarsson 2023; Saha et al. 2023), but further discussion of studies outside our reviewed corpus extends beyond the scope of this article. Other studies published after our

review ‘cut-off date’ also develop themes notably absent in our corpus of literature, while being profoundly relevant to interacting among ‘AI’ in various settings²²—such as the work of membership categorization (see, e.g., Sacks 1972; Fitzgerald and Housley 2015), which is connected to the assumptions and interactional procedures involved in (tacitly or explicitly) categorizing participants as either ‘human’ or ‘AI’ (Ivarsson and Lindwall 2023). Moreover, our review has noted a certain trend of EM/CA to focus on VUIs and robots, with chatbots being on the very margin. Considering the recent surge of societal interest and concern about large language models and their publicly available interfaces such as ChatGPT, we expect that more EM/CA studies will concentrate on this technology in professional and everyday activities in the near future.²³

The studies reviewed in this article represent an interaction-centered approach to empirical studies of AI technologies on the minute level of situated detail. This focus might invite criticisms previously leveled at EM/CA more broadly: that it is programmatically disinterested in generalization (in this case, e.g., across divergent technological systems, user groups, or usage scenarios), and/or unable to address contextual or social factors that occur outside of an instance of interaction (e.g., Billig 1999). But where these criticisms, many of which have been vigorously rebutted (e.g., Schegloff 1999), do accurately characterize EM/CA’s theoretical and analytic parsimony (e.g., Enfield and Sidnell 2017), this programmatic focus is often a useful intervention in more theory- and experiment-driven approaches within HCI and HRI. The principles and methodological procedures of EM/CA tend to lead away from theorizing, abstraction, and universally generalizable explanations, and instead to prioritize empirical inquiry. They also tend to prioritize ecological validity by studying interaction in situ and relying on evidence drawn from the participants’ own displays of understanding. For technology use, in-situ concerns are often identical with user concerns, which enables EM/CA studies to provide valuable insights for systems design (see Button 2012). An approach underpinned by an interactional, situated understanding of AI might ask which situations and which technologies are treated as ‘autonomous’, irrespective of their technical components or conformity with the norms and modalities of face-to-face interaction. This might facilitate a broader turn to ethnomethodological studies of

technologies that are less self-evidently amenable to interaction-analytic methods.

In sum, this scoping review has thrown up some challenges for the process of systematically reviewing EM/CA studies of AI. Parry and Land (2013), in their systematic review of CA healthcare research, note that “no pre-existing off the shelf approach [to literature reviewing] is adequate for handling conversation analytic evidence”. This challenge is partly due to the discontinuities in standards of evidence and conventions of reporting across the many areas (including HCI, sociology, linguistics, anthropology) from which we drew our corpus of studies. We are also aware of the incompleteness of our corpus of reviewed texts, as practitioners in EM/CA may not always explicitly claim affiliation to EM/CA at large. EM’s notion of ‘hybrid studies’, as well as ‘applied CA’, but also strategic decisions taken by authors for publication, may sometimes lead to the discursive disappearance of EM or CA in the published texts, which eventually makes them invisible to simple keyword-based search procedures, while perhaps still transparently relevant for EM/CA practitioners by other means.

7 Conclusion

This review has showcased the versatility of an ethnomethodological and conversation analytic approach to the study of interaction with ‘AI’. This approach has been applied to a wide range of technologies, user groups, and worksites. The findings and insights produced by these studies have highlighted and provided empirical backing for the importance of exploring locally established methods of reasoning through interaction with and around AI, rather than focusing on specific modalities, technologies, or design features. These studies highlight the interactional resources and methods people use for establishing and maintaining social order in their encounters with AI, and the constitutive particularity of diverse social settings (e.g., educational, medical, scientific, or other workplace-specific orders of activities). Generally, there seems to be a tendency in the field to study autonomous VUI and robot systems over other technologies, although there was a wide variety of ways in which systems were presented to user(s), the ages and constellations of user groups, the activities done with the system, and the complexity levels of the systems. Collection of naturalistic (non-experimental) data was relatively uncommon, which is noteworthy for the field, but seems related to the technology’s occurrence in everyday life.

With regard to reported findings, three interactional phenomena were recurrently addressed in the corpus. The first concerned opening and closing interactions with AI, showing that what happens before or at the potential start of interaction impacts whether and how the interaction

²² See the special issue of *Social interaction: Video-based studies of human sociality* (Ibnelkaïd and Avgustis 2023), which includes several EM/CA studies of “AI”; or Mlynář et al. (2024).

²³ For example, a special issue of *Discourse & Communication* on the topic of “Conversation Analysis and Conversational Technology” is currently under review with the prospect of being published in 2024.

unfolds. In addition, users close their interactions with a system in ways that orient differently to the agent-status of the machine. Miscommunication and repair was another recurrently studied phenomenon, with many studies showing that users quickly adapt to the system's perceived capabilities and, when trouble escalates, orient towards progressing the interaction above other interactional goals such as achieving what they were doing when trouble occurred (e.g., requesting information of the system or answering the system's question). A key issue of miscommunication that many papers touch on is that the system often lacks access to the same sensory information as the human and vice versa, with a recurrent suggestion to provide the user with more relevant feedback on the nature of the problem. Lastly, with regard to non-verbal communication and emotion displays, most studies find that users tend to draw on their repertoire of practices from non-verbal human-human interaction when interacting with social technologies, with gaze being an especially important resource for managing mutual attention. The importance of the precise timing of embodied actions was an important finding for a range of non-verbal conduct, including emotion displays, which extends existing findings from EM/CA research that show how non-verbal conduct and emotion displays cannot be simplified into categories such as 'smiling is happy' or 'one needs to gaze at someone else at all times'.

The main aim of our review has been to consolidate and provide an initial mapping of the burgeoning EM/CA literature on human-AI interaction, while identifying broad trends and gaps in its coverage thus far. While doing so, we have also attempted to provide a critical reflection on the work of reviewing, and have critically explored the relationship between EM and CA in the area of research on AI. Our focus on a relatively narrow subset of empirical literature sharing this general methodological approach allowed us to document and exemplify some trends that might be emblematic of the field as a whole. One is the prevalence of studies grounded in interactionist CA, and its applied variants, compared to much less frequent investigations aligned with the praxeological EM program (though both are often subsumed under the label 'EM/CA'). We found that these studies, as summarized above, are mostly examining a range of interactional phenomena already identified and described in previous CA studies into domains of social life other than interacting with AI-labelled technologies. The characteristic EM focus on the constitutive details of activities, i.e., laying out what exactly is distinctive about AI in situated action, seems to provide a complementary, affiliated, but in some cases incommensurable line of inquiry.

We have also highlighted some productive avenues for future research, and suggested how an EM/CA approach is well-placed to study the integration of AI technologies into ever more social settings, processes, and aspects of

our professional activities and everyday lives. AI-related technologies move from experimental 'sandboxes' and 'playgrounds' to routine activities embedded in the structures of everyday life, and they are recontextualized and reframed as people find ways to make them at home in their worlds. Over time, formerly exotic technological objects grow into unremarkable tools, while expertise for interacting with them becomes increasingly common. As our article has shown, EM/CA research allows us to specify—empirically, systematically, and in actual lived detail—how AI-labeled technology and social life mutually contribute to each other, in situ and in real time, explicating the mundane procedures by which a technology “is made at home in the world that has whatever organization it already has” (Sacks 1992: 549).

Appendix 1: Search strategy

AMC full-text collection

S1	Anywhere(“human–AI” OR “human–robot” OR “human–agent” OR “HRI” OR “virtual human” OR “social robot” OR “embodied conversational agent” OR “ECA” OR “artificial intelligence” OR “voice user interface” OR “chatbot”)	136,894
S2	abstract(“conversation analysis” OR “conversation analytical” OR “ethnomethodology” OR “ethnomethodological”) authorkeyword (“conversation analysis” OR “conversation analytical” OR “ethnomethodology” OR “ethnomethodological”)	44
S3	S1 AND S2	16
F1	Filter: Research Articles (Excl. = 2 posters)	14

IEEE

S1	allmetadata("human-AI" OR "human-robot" OR "human-agent" OR "HRI" OR "virtual human" OR "social robot" OR "embodied conversational agent" OR "ECA" OR "artificial intelligence" OR "voice user interface" OR "chatbot")	384,471
S2	("Abstract": "conversation analy*" OR "Abstract": "ethnomethodolog*") OR ("Author Keywords": "conversation analy*" OR "Author Keywords": "ethnomethodolog*")	86
S3	S1 AND S2	71

Springer

S1	anywhere("human-AI" OR "human-robot" OR "human-agent" OR "HRI" OR "virtual human" OR "social robot" OR "embodied conversational agent" OR "ECA" OR "artificial intelligence" OR "voice user interface" OR "chatbot")	870,865
S2	"conversation analysis" OR "conversation analytical" OR "ethnomethodology" OR "ethnomethodological"	6827
S3	S1 AND S2 (in search box: "human-AI" OR "human-robot" OR "human-agent" OR "HRI" OR "virtual human" OR "social robot" OR "embodied conversational agent" OR "ECA" OR "artificial intelligence" OR "voice user interface" OR "chatbot" AND "conversation analysis" OR "conversation analytical" OR "ethnomethodology" OR "ethnomethodological")	121
F1	Filter: Chapter, Article and Conference Paper/ Proceedings (Excl. = 52 Books, 6 Reference Works)	63

Web of science

S1	allmetadata("human-AI" OR "human-robot" OR "human-agent" OR "HRI" OR "virtual human" OR "social robot" OR "embodied conversational agent" OR "ECA" OR "artificial intelligence" OR "voice user interface" OR "chatbot")	413,345
S2	abstract("conversation analy*" OR "ethnomethodolog*") authorkeywords("conversation analy*" OR "ethnomethodolog*")	2288
S3	S1 AND S2	17

LLBA

S1	allmetadata("human-AI" OR "human-robot" OR "human-agent" OR "HRI" OR "virtual human" OR "social robot" OR "embodied conversational agent" OR "ECA" OR "artificial intelligence" OR "voice user interface" OR "chatbot")	3477
S2	abstract("conversation analy*" OR "ethnomethodolog*") identifier(keyword)("conversation analy*" OR "ethnomethodolog*")	3868
S3	S1 AND S2	16
F1	Limit to peer-reviewed	7

Appendix 2: Detailed information on each text unit

Article	Age group(s)	Additional feature	Language	Country	Interaction type
Alač et al. (2011)	12–36 months, adults	Adult–child constellation (toddler–teacher–researcher(designer)–researcher(ethnographer)), Researcher part of the interaction	English (native)	Unknown	No specific activity, just being with the robot in the same space (their day care)
Alač (2016)	12–36 months, adults	Adult–child constellation (toddler–teacher–researcher(designer)–researcher(ethnographer)), Researcher part of the interaction	English (native)	US	Preschool interactions, e.g., playing with the robot or the teacher encouraging children to do something with the robot, such as touch it
Alač et al. (2020)	Students	Households (e.g., couples, families, dormitories)	English (native)	US	Everyday use
Aranguren (2014)	N/A	Real-world telephone calls	French	Unknown	Real-world telephone calls
Arend et al. (2017)	N/A	Public space (museum), Researcher part of the interaction	English (non-native)	Luxembourg	Playing sports guessing game
Avgustis et al. (2021)	N/A	Real-world telephone calls	Russian		Call center phone calls
Brown and Laurier (2017)	N/A	Public space (traffic)	N/A (traffic)	Several countries	Traffic
Candello et al. (2020)	N/A	Public space (museum)	English (native)	UK presumably, that is where the Museum of Tomorrow is located	Learning something on a topic
Corti and Gillespie (2016)	N/A	No specific additional aspect	English (native)	UK	10 min talk (Ps spoke, their interlocutor either responded through text (chatbot or human via chat) or speech (humans own answers or echoing the chatbot)
Cyra and Pitsch (2017)	Seniors, persons with mild cognitive impairments (N/A), student control group	(Mildly) cognitive impaired people	English (unclear)	unknown	Assisted calendar schedule
Ferm et al. (2015)	Children (4–12), AND adults	Adult–child constellations, researcher present, (mild) cognitive impaired people (Cerebral palsy)	Swedish		Playing with the robot
Fischer et al. (2015)	N/A	No specific additional aspect	Unknown (non-verbal interaction focus)	Austrian	Tutoring the (WoOz controlled) robot in performing a simple task (moving/manipulating objects)
Fischer et al. (2019)	N/A	Households (couples and parents with children), adult–child constellations	English (native)	unknown	Everyday use

Article	Age group(s)	Additional feature	Language	Country	Interaction type
Gehle et al. (2017)	N/A	No specific additional aspect	Unknown (non-verbal interaction focus)	unknown	Walking through exhibit, robot explains
Iwasaki et al. (2019)	Fieldstudy: N/A, Experimental study: students	Fieldstudy: Public space (shop), Experimental study: no specific additional aspect	English (non-native)	Japan	Visiting a shop
Jentsch et al. (2019)	N/A	No specific additional aspect	English (non-native)	Luxembourg or Germany (many colleagues of researchers, but unclear which author)	Try to elicit basic information (e.g., What are your capabilities?)
Klowait (2017)	N/A	Real-world telephone calls	Russian		Telephone survey on voting behavior
Krummheuer (2008a)	N/A	Public space (shopping mall)	German		Voluntary interaction with Max in the shopping center (Max could present topic, make small talk or play a game)
Krummheuer (2008b)	N/A	Public space (shopping mall)	German		Voluntary interaction with Max in the shopping center (Max could present topic, make small talk or play a game)
Krummheuer (2009)	N/A	Public space (shopping mall)	German		Voluntary interaction with Max in the shopping center (Max could present topic, make small talk or play a game)
Krummheuer (2015a)	N/A	Dataset 1: Public space (shopping mall), Dataset 2: (Mild) cognitive impaired people, Public space (grocery walk (story + outside))	Shopping mall data: German, Walking help data: Danish		Voluntary interaction with Max in the shopping center (Max could present a topic, make small talk or play a game), walking help was used for walking (e.g., walk outside, grocery shopping)

Article	Age group(s)	Additional feature	Language	Country	Interaction type
Krummheuer (2015b)	N/A	Public space (shopping mall)	German		Voluntary interaction with Max in the shopping center (Max could present topic, make small talk or play a game)
Krummheuer et al. (2020)	N/A	(Mild) cognitive impaired people, Researchers are part of the interaction	Danish		Workshops during which programming/setting up the personalized device
Licoppe and Rollet (2020)	N/A	Public space (university hallway)	French		Interacting with Pepper voluntarily (Pepper has a very simple script, no activity)
Lohse et al. (2009)	N/A	No specific additional aspect	German		Teach the robot objects and locations in several rooms (guiding the robot through the room) NB general contact with robot was trained beforehand
Muhle (2008)	Students	No specific additional aspect	German		Interact with the robot dog, manual of the robot was available
Opfermann et al. (2017)	Elderly (70+), mildly cognitively impaired persons (N/A), student control group	(Mild) cognitive impaired people	English (native)	unknown	Schedule management setting
Payr (2010)	Older persons (50+)	No specific additional aspect	English (native)	UK	Receiving fitness instructions
Payr (2013)	Older persons (50+)	No specific additional aspect	English (native)	UK	Receiving fitness instructions
Pelikan et al. (2020)	Study 1: N/A, Study 2: Adults, children (4–12)	Study1: Households (couples, families), Researcher part of the interaction, Study2: Households (e.g., couples, families), Researcher part of the interaction, Adult–child constellations (households)	Study 1: German, Study 2: Swedish		Robot learning names and playing (e.g., giving a fist bump to the robot)
Pelikan (2021)	N/A	Public space (traffic)	N/A (traffic)		Traffic—automated vehicles
Pelikan et al. (2022)	Adults, Children (4–12)	Households (couples, families), Adult–child constellations	Swedish		Robot learning names and playing (e.g., giving a fist bump to the robot)
Pelikan and Broth (2016)	Students	No specific additional aspect	English (non-native)	Sweden	Game of charade

Article	Age group(s)	Additional feature	Language	Country	Interaction type
Pitsch and Koch (2010)	12–36 months, adults, Children (4–12)	Adult–child constellation (toddler–researcher), Researchers are part of the interaction	German		Play with the toy (e.g., pet, give ‘food’)
Pitsch et al. (2013)	N/A	No specific additional aspect	German		Tutoring a robot during a simple task (manipulating object)
Pitsch et al. (2017)	Adults, Children (4–12)	Households, Adult–child constellation (parent–child), Public space (museum)	German		Answering the robot’s questions about the exhibit
Pitsch (2020)	Adults, Children (4–12)	Households, Adult–child constellation (parent–child), Public space (museum)	German		Answering the robot’s questions about the exhibit
Pitsch et al. (2009)	N/A	Public space (museum)	Japanese		Robot opens interaction and explains paintings
Porcheron et al. (2018)	N/A	Households (couples and parents w. children)	English (native)	UK	Everyday use
Porcheron et al. (2017)	N/A	Public space (cafe)	English (native)	UK	Making a query during get together in café
Relieu et al. (2020)	N/A	Real-world telephone calls	English (native)	unknown (YouTube videos)	Making a phone call
Robins et al. (2004)	Adults, Children (4–12)	Adult–child constellation (experimenter–child), Researchers are part of the interaction, (Mild) cognitive impaired people (Autism)	English	UK	Minimal structure to the activity. The robot executed a preprogrammed sequence of movements (“dance”)
Rollet et al. (2017)	N/A	No specific additional aspect	French		Musical quiz
Sahin et al. (2017) (Lenny)	N/A	Real-world telephone calls	English (native)	unknown (YouTube videos)	Real-world telephone calls of telemarketers answered by LENNY
Stommel et al. (2022)	Seniors (70 +)	No specific additional aspect	Dutch		Health survey interview (recorded as part of experimental trial)
Süssenbach et al. (2012)	N/A	No specific additional aspect	German		Fitness instructions
Torre et al. (2021)	N/A	No specific additional aspect	English (native), English (non-native)	US	Therapy session (laboratory setting)
Tuncer et al. (2022)	Children (4–12)	No specific additional aspect	Multiple	Non-native Swedish, varying levels	Playing together
Velkovska et al. (2020)	N/A	Households (varying)	English (non-native), French	France	Everyday use (queries/requests etc.)
Walker et al. (2020)	N/A	(Mildly) cognitive impaired people (memory issues)	English (native)	UK	Doctor–patient interaction on memory issues
Wallis (2008)	N/A	No specific additional aspect	English (native)	US (probably)	Calling in to ask about flight scheduling and such

Article	Age group(s)	Additional feature	Language	Country	Interaction type
Wooffitt (1994)	students	No specific additional aspect	English (native)	UK	Experimental task wherein participants received a scenario based on which they called a system to acquire information
Yamazaki et al. (2013)	N/A	No specific additional aspect	Japanese		Museum exhibition in the lab

Acknowledgements In addition to all the researchers whose studies we reviewed, we acknowledge the contributions of many colleagues throughout the preparation of this text. Renata Topinková provided invaluable assistance with the analysis of bibliographic citation networks for an initial overview of the field. We received many inspiring comments and criticisms while presenting our work in October 2021 at the 6th Copenhagen Multimodality Day: Interacting with AI, in November 2022 at the Digital Meeting for Conversation Analysis (DMCA; <https://dmca.conversationanalysis.org/>), and in January 2023 at the meeting of the EMCA Artificial Intelligence Network (EMCAI; <https://emcai.conversationanalysis.org/>). We also extend our gratitude to two anonymous reviewers for their insightful and constructive critique and detailed suggestions. Any remaining shortcomings are our own.

Funding Open access funding provided by University of Applied Sciences and Arts Western Switzerland (HES-SO).

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 (1997a) Computation and human experience. Cambridge University Press, Cambridge
- Agre PE (1997b) Toward a critical technical practice: lessons learned in trying to reform AI. In: Bowker G, Gasser L, Star L, Turner B (eds) Bridging the great divide: social science, technical systems, and cooperative work. Erlbaum, Mahwah, NJ, pp 131–157
- Alaç M (2016) Social robots: things or agents? *AI Soc* 31(4):519–535. <https://doi.org/10.1007/s00146-015-0631-6>
- Alaç M, Movellan J, Tanaka F (2011) When a robot is social: Spatial arrangements and multimodal semiotic engagement in the practice of social robotics. *Soc Stud Sci* 41(6):893–926. <https://doi.org/10.1177/0306312711420565>
- Alaç M, Gluzman Y, Aflakoun 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 Aesthet* 9(1):3–53
- Anderson RJ, Sharrock W (2017) Has ethnomethodology run its course? Unpublished paper. <https://www.sharrockandanderson.co.uk/wp-content/uploads/2019/10/Run-its-Course-VII.pdf>. Accessed 13 Oct 2023
- Antaki C (ed) (2011) Applied conversation analysis: intervention and change in institutional talk. Palgrave Macmillan, Basingstoke
- Antaki C, Crompton RJ (2015) Conversational practices promoting a discourse of agency for adults with intellectual disabilities. *Discourse Soc* 26(6):645–661. <https://doi.org/10.1177/0957926515592774>
- Aranguren M (2014) Le travail émotionnel du client: La structure séquentielle des émotions dans les usages problématiques d'un serveur vocal. *Soc Sci Inf* 53(3):311–340. <https://doi.org/10.1177/0539018414523520>
- Arend B, Sunnen P, Caire P (2017) Investigating breakdowns in human robot interaction: a conversation analysis guided single case study of a human-NAO communication in a museum environment. *Int J Mech Aero Ind Mechatron Manuf Eng* 11(5):839–845
- Arksey H, O'Malley L (2005) Scoping studies: towards a methodological framework. *Int J Soc Res Methodol* 8(1):19–32. <https://doi.org/10.1080/1364557032000119616>
- Albert S, Buschmeier H, Cyra K, Even C, Hamann M, Licoppe C, Mlynář J, Pelikan H, Porcheron M, Reeves S, Rudaz D, Sormani P, Tuncer S (2023a) What 'counts' as explanation in social interaction? Six observations from an EM/CA approach. In: 2nd TRR 318 conference "Measuring Understanding", Paderborn University, 6–7 November 2023. <https://saulalbert.net/blog/what-counts-as-explanation-in-social-interaction/>
- Albert S, Hamann M, Stokoe E (2023) Conversational user interfaces in smart homecare interactions: a conversation analytic case study. In: Proceedings of the 2023 ACM conference on conversational user interfaces (CUI '23), July 19–21, 2023, Eindhoven, Netherlands. <https://doi.org/10.1145/3571884.3597140>
- Avgustis I, Shirokov A, Iivari N (2021) "Please connect me to a specialist": scrutinising 'recipient design' in interaction with an artificial conversational agent. In: Ardito C et al (eds) Human-computer interaction – INTERACT 2021. Lecture notes in computer science, vol 12935. Springer. https://doi.org/10.1007/978-3-030-85610-6_10
- Bellon A, Velkovska J (2023) L'intelligence artificielle dans l'espace public: du domaine scientifique au problème public: Enquête sur un processus de publicisation controversé. *Réseaux* 4(240):31–70. <https://doi.org/10.3917/res.240.0031>
- Berger E, Pekarek Doehler S (2018) Tracking change over time in storytelling practices: a longitudinal study of second language talk-in-interaction. In: Longitudinal studies on the organization of social interaction, Springer, pp 67–102
- Billig M (1999) Whose terms? Whose ordinariness? Rhetoric and ideology in conversation analysis. *Discourse Soc* 10(4):543–558. <https://doi.org/10.1177/0957926599010004005>

- Bovet A, Carlin A, Sormani P (2011) Discovery starts here? The ‘Pulsar Paper’, thirty years on—an ethnobiometric note. *Ethnogr Stud* 12:126–139. <https://doi.org/10.5449/idslu-001104720>
- Brooker P (2019) *Programming with Python for social scientists*. SAGE, London
- Brooker P (2022) Computational ethnography: a view from sociology. *Big Data Soc* 9(1):1–6. <https://doi.org/10.1177/20539517211069892>
- 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 Publications, London, pp 573–592
- Brooker P, Dutton W, Mair M (2019) The new ghosts in the machine: ‘Pragmatist’ AI and the conceptual perils of anthropomorphic description. *Ethnogr Stud* 16:272–298. <https://doi.org/10.5281/zenodo.3459327>
- Brown B, Laurier E (2017) The trouble with autopilots: assisted and autonomous driving on the social road. In: *Proceedings of the 2017 CHI conference on human factors in computing systems*, pp 416–429. <https://doi.org/10.1145/3025453.3025462>
- Button G (1990) Going up a blind alley: conflating conversation analysis and computational modelling. In: Luff P, Gilbert N, Frohlich D (eds) *Computers and conversation*. Academic Press, London, pp 67–90. <https://doi.org/10.1016/B978-0-08-050264-9.50009-9>
- Button G (1991) Introduction: ethnomethodology and the foundational respecification of the human sciences. In: Button G (ed) *Ethnomethodology and the human sciences*. Cambridge University Press, Cambridge, pp 1–9. <https://doi.org/10.1017/CBO9780511611827.002>
- Button G (2012) What does ‘work’ mean in ‘ethnomethodological studies of work?’: its ubiquitous relevance for systems design to support action and interaction. *Des Stud* 33(6):673–684. <https://doi.org/10.1016/j.destud.2012.06.003>
- 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, Oxford
- Button G, Crabtree A, Rouncefield M, Tolmie P (2015) *Deconstructing ethnography: towards a social methodology for ubiquitous computing and interactive systems design*. Springer, Cham
- Button G, Lynch M, Sharrock WW (2022) *Ethnomethodology, conversation analysis and constructive analysis: on formal structures of practical action*. Routledge, London
- Caluori L (2023) Hey Alexa, why are you called intelligent? An empirical investigation on definitions of AI. *AI Soc*. <https://doi.org/10.1007/s00146-023-01643-y>
- Candello H, Barth F, Carvalho E, Alves R, Cotia RAG (2020) Understanding how visitors interact with voice-based conversational systems. Design user experience and usability: design for contemporary interactive environments. In: 9th international conference DUXU 2020 held as part of the 22nd HCI international conference, HCII 2020, Copenhagen, Denmark, July 19–24 2020, *Proceedings Part II*. Springer International Publishing, pp 40–55
- Churchill L (1971) Ethnomethodology and measurement. *Soc Forces* 50(2):182–191. <https://doi.org/10.2307/2576936>
- Cicourel A (1964) *Method and measurement in sociology*. Free Press, New York
- Clayman SE, Heritage J, Maynard DW (2022) The ethnomethodological lineage of conversation analysis. In: Maynard DW, Heritage J (eds) *The ethnomethodology program: legacies and prospects*. Oxford University Press, New York, pp 252–286
- Coates L (2022) “The temporal ‘succession’ of here and now situations”: Schütz and Garfinkel on sequentiality in interaction. *Hum Stud* 45:469–491. <https://doi.org/10.1007/s10746-022-09632-8>
- Collins H (2018) *Artificial intelligence: against humanity’s surrender to computers*. Polity Press, Cambridge
- Corti K, Gillespie A (2016) Co-constructing intersubjectivity with artificial conversational agents: people are more likely to initiate repairs of misunderstandings with agents represented as human. *Comput Hum Behav* 58:431–442. <https://doi.org/10.1016/j.chb.2015.12.039>
- Coulter J (1985) On comprehension and ‘mental representation.’ In: Gilbert NG, Heath C (eds) *Social action and artificial intelligence*. Gower, Aldershot, pp 8–23
- Coulter J (2008) Twenty-five theses against cognitivism. *Theory Cult Soc* 25(2):19–32. <https://doi.org/10.1177/0263276407086789>
- Crabtree A (2004) Taking technomethodology seriously: hybrid change in the ethnomethodology-design relationship. *Eur J Inf Syst* 13(3):195–209. <https://doi.org/10.1057/palgrave.ejis.3000500>
- Cyra K, Pitsch K (2017) Dealing with long utterances: how to interrupt the user in a socially acceptable manner? In: HAI’17 proceedings of the 5th international conference on human agent interaction, pp 341–345. <https://doi.org/10.1145/3125739.3132586>
- De Stefani E, Mondada L (2018) Encounters in public space: how acquainted versus unacquainted persons establish social and spatial arrangements. *Res Lang Soc Interact* 51(3):248–270. <https://doi.org/10.1080/08351813.2018.1485230>
- Dourish P (2006) Implications for design. In: *Proceedings of the SIGCHI conference on human factors in computing systems (CHI’06)*. ACM. <https://doi.org/10.1145/1124772.1124855>
- Dourish P, Button G (1998) On technomethodology: foundational relationships between ethnomethodology and system design. *Hum Comput Interact* 13(4):395–432. https://doi.org/10.1207/s15327051hci1304_2
- Drew P (1997) ‘Open’ class repair initiators in response to sequential sources of troubles in conversation. *J Pragmat* 28(1):69–101. [https://doi.org/10.1016/S0378-2166\(97\)89759-7](https://doi.org/10.1016/S0378-2166(97)89759-7)
- Dreyfus H (1965) *Alchemy and artificial intelligence*. The RAND Corporation, Santa Monica, CA
- Due BL (2023) Situated socio-material assemblages: assemthodology in the making. *Hum Commun Res: hqad031*. <https://doi.org/10.1093/hcr/hqad031>
- Eisenmann C, Mitchell R (2024) Doing ethnomethodological ethnography. Moving between autoethnography and the phenomenon in “hybrid studies” of taiji ballet and yoga. *Qual Res* 24(1):32–59. <https://doi.org/10.1177/14687941221132956>
- Eisenmann C, Mlynář J, Turowetz J, Rawls AW (2023a) “Machine down”: making sense of human-computer interaction—Garfinkel’s research on ELIZA and LYRIC from 1967 to 1969 and its contemporary relevance. *AI Soc* [online first]. <https://doi.org/10.1007/s00146-023-01793-z>
- Eisenmann C, Meier zu Verl C, Kreplak Y, Dennis A (2023b) Reconsidering foundational relationships between ethnography and ethnomethodology and conversation analysis—an introduction. *Qual Res* [online first]. <https://doi.org/10.1177/14687941231210177>
- Elliott A (2019) *The culture of AI: everyday life and the digital revolution*, 1st edn. Routledge, London. <https://doi.org/10.4324/9781315387185>
- Enfield NJ, Sidnell J (2017) *The concept of action*. Cambridge University Press, Cambridge
- Ferm UM, Claesson BK, Ottesjö C, Ericsson S (2015) Participation and enjoyment in play with a robot between children with cerebral palsy who use AAC and their peers. *Augment Altern Commun* 31(2):108–123. <https://doi.org/10.3109/07434618.2015.1029141>

- Fischer K (2021) Tracking anthropomorphizing behavior in human-robot interaction. *ACM Trans Hum Robot Interact* 11(1):1–28. <https://doi.org/10.1145/3442677>
- Fischer K, Jensen LC, Kirstein F, Stabinger S, ErKent Ö, Shukla D, Piater J (2015) The effects of social gaze in human-robot collaborative assembly. In: Tapus A, André E, Martin J-C, Ferland F, Ammi M (eds) *Social robotics*. Springer, Cham, pp 204–213. https://doi.org/10.1007/978-3-319-25554-5_21
- Fischer JE, Reeves S, Porcheron M, Sikveland RO (2019) Progressivity for voice interface design. In: *Proceedings of the 1st international conference on conversational user interfaces*, pp 1–8. <https://doi.org/10.1145/3342775.3342788>
- Fitzgerald R, Housley W (eds) (2015) *Advances in membership categorisation analysis*. SAGE Publications, London
- Garfinkel H (1967) *Studies in ethnomethodology*. Prentice-Hall, Englewood Cliffs, NJ
- Garfinkel H (ed) (1986) *Ethnomethodological studies of work*. Routledge & Kegan Paul, London
- Garfinkel H (1991) Respecification: evidence for locally produced, naturally accountable phenomena of order*, logic, reason, meaning, method, etc. in and as of the essential haecceity of immortal ordinary society (I): an announcement of studies. In: Button G (ed) *Ethnomethodology and the human sciences*. Cambridge University Press, Cambridge, pp 10–19
- Garfinkel H (2002) *Ethnomethodology's program: working out Durkheim's aphorism* (edited by AW Rawls). Rowman & Littlefield, Oxford
- Garfinkel H (2019a [1959]) Common sense knowledge of social structures. A paper distributed at the session on the sociology of knowledge. In: *Fourth world congress of sociology*, Stresa, Italy, September 12, 1959. <https://doi.org/10.25969/mediarep/13805>
- Garfinkel H (2019b [1960]) Notes on language games as a source of methods for studying the formal properties of linguistic events. *Eur J Soc Theory* 22(2):148–174. <https://doi.org/10.1177/1368431018824>
- Garfinkel H (2019c) *Parsons' primer* (edited by AW Rawls). J.B. Metzler, Berlin
- Garfinkel H (2021) Ethnomethodological misreading of Aron Gurwitsch on the phenomenal field. *Hum Stud* 44(1):19–42
- Garfinkel H (2022a) A comparison of decisions made on four “pre-theoretical” problems by Talcott Parsons and Alfred Schütz. In: Maynard DW, Heritage J (eds) *The ethnomethodology program: legacies and prospects*. Oxford University Press, New York, pp 71–89
- Garfinkel H (2022b) Sources of issues and ways of working: an introduction to the study of naturally organized ordinary activities. In: Maynard DW, Heritage J (eds) *The ethnomethodology program: legacies and prospects*. Oxford University Press, New York, pp 141–161
- Garfinkel H (2022c) *Studies of work in the sciences* (edited by M Lynch). Routledge, London
- Garfinkel H, Rowan PK (1955) Letter to Dr. Leonard Broom, Editor of *American Sociological Review*, 5 July 1955. Available at the Harold Garfinkel Archive, Newburyport, Massachusetts, USA
- Garfinkel H, Sacks H (1970) On formal structures of practical action. In: McKinney JC, Tiryakian EA (eds) *Theoretical sociology: perspectives and developments*. Appleton-Century-Crofts, New York, pp 338–366
- Garfinkel H, Wieder DL (1992) Two incommensurable, asymmetrically alternate technologies of social analysis. In: Watson G, Seiler RM (eds) *Text in context: contributions to ethnomethodology*. Sage, New York, pp 175–206
- 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. <https://doi.org/10.1145/2909824.3020219>
- Gilbert NG, Heath C (eds) (1985) *Social action and artificial intelligence*. Gower, Aldershot
- Gill SP (2023) Why thinking about the tacit is key for shaping our AI futures. *AI Soc* 38:1805–1808. <https://doi.org/10.1007/s00146-023-01758-2>
- Goffman E (1981) *Forms of talk*. University of Pennsylvania Press, Philadelphia
- Goodwin C (2000) Action and embodiment within situated human interaction. *J Pragmat* 32:1489–1522. [https://doi.org/10.1016/S0378-2166\(99\)00096-X](https://doi.org/10.1016/S0378-2166(99)00096-X)
- Greiffenhagen C (2014) The materiality of mathematics: presenting mathematics at the blackboard. *Br J Sociol* 65(3):502–528. <https://doi.org/10.1111/1468-4446.12037>
- Grudin J (2009) AI and HCI: two fields divided by a common focus. *AI Mag* 30(4):48–57. <https://doi.org/10.1609/aimag.v30i4.2271>
- Haddington P, Eilittä T, Kamunen A, Kohonen-Aho L, Oittinen T, Rautiainen I, Vatanen A (eds) (2023) *Ethnomethodological conversation analysis in motion: emerging methods and new technologies*. Routledge, London
- Harper RHR (2019) The role of HCI in the age of AI. *Int J Hum Comput Interact* 35(15):1331–1344. <https://doi.org/10.1080/10447318.2019.1631527>
- Heath C, Luff P (2022) Technology in practice. In: Maynard DW, Heritage J (eds) *The ethnomethodology program*. Oxford University Press, New York, pp 398–419
- Heijselaar E (2023) The CASA theory no longer applies to desktop computers. *Sci Rep* 13:19693. <https://doi.org/10.1038/s41598-023-46527-9>
- Heritage J (2007) Intersubjectivity and progressivity in person (and place) reference. In: Stivers T, Enfield NJ (eds) *Person reference in interaction: linguistic, cultural and social perspectives*. Cambridge University Press, Cambridge
- Hester S (2009) Ethnomethodology: respecifying the problem of social order. In: Hviid Jacobsen M (ed) *Encountering the everyday: an introduction to the sociologies of the unnoticed*. Palgrave Macmillan, New York, pp 234–256
- Hilbert R (1990) Ethnomethodology and the micro-macro order. *Am Sociol Rev* 55(6):794–808. <https://doi.org/10.2307/2095746>
- Hirsch-Kreinsen H (2023) Artificial intelligence: a “promising technology.” *AI Soc*. <https://doi.org/10.1007/s00146-023-01629-w>
- Housley W, Fitzgerald R (2002) The reconsidered model of membership categorization analysis. *Qual Res* 2(1):59–83. <https://doi.org/10.1177/146879410200200104>
- Housley W, Albert S, Stokoe E (2019) Natural action processing. In: *HTTF 2019: Proceedings of the halfway to the future symposium 2019*, Article No. 34. <https://doi.org/10.1145/3363384.3363478>
- Hutchinson P (2022) Wittgensteinian ethnomethodology (1): Gurwitsch, Garfinkel, and Wittgenstein and the meaning of praxeological Gestalts. *Philos Sci* 26(3):61–93. <https://doi.org/10.4000/philosophiascientiae.3605>
- Ibnelkaid S, Avgustis I (2023) Situated agency in digitally artifacted social interactions: introduction to the special issue. *Soc Interact Video-Based Stud Hum Soc* 6(1). <https://doi.org/10.7146/si.v6i1.136855>
- Ikeya N (2020) Hybridity of hybrid studies of work: examination of informing practitioners in practice. *Ethnogr Stud* 17:22–40. <https://doi.org/10.5281/zenodo.4050533>
- Ivarsson J (2023) Dealing with daemons: trust in autonomous systems. In: Sorman P, vom Lehn D (eds) *The Anthem companion to Harold Garfinkel*. Anthem Press, London
- Ivarsson J, Lindwall O (2023) Suspicious minds: the problem of trust and conversational agents. *Computer Supported*

- Cooperative Work (CSCW) [online first]. <https://doi.org/10.1007/s10606-023-09465-8>
- Iwasaki M, Zhou J, Ikeda M, Koike Y, Onishi Y, Kawamura T, Nakaniishi H (2019) “That robot stared back at me!”: demonstrating perceptual ability is key to successful human-robot interactions. *Front n Robot AI* 6:85. <https://doi.org/10.3389/frobt.2019.00085>
- Jaton F, Sormani P (2023) Enabling ‘AI’? The situated production of commensurabilities. *Soc Stud Sci* 53(5):625–634. <https://doi.org/10.1177/03063127231194591>
- Jefferson G (2018) *Repairing the broken surface of talk: managing problems in speaking, hearing, and understanding in conversation*. Oxford University Press, Oxford
- Jenkins KN (2023) The neo-ethnomethodological program(s): on alignments with and departures from Garfinkel. *Symbolic Interaction* [online first]. <https://doi.org/10.1002/symb.676>
- Jentsch SF, Höhn S, Hochgeschwender N (2019) Conversational interfaces for explainable AI: a human-centred approach. In: Calvaresi D, Najjar A, Schumacher M, Främling K (eds) *Explainable, transparent autonomous agents and multi-agent systems (EXTRAAMAS 2019)*. Springer, Cham. https://doi.org/10.1007/978-3-030-30391-4_5
- Jones RA (2017) What makes a robot ‘social’? *Soc Stud Sci* 47(4):556–579. <https://doi.org/10.1177/0306312717704722>
- Kendon A (1990) *Conducting interaction: patterns of behavior in focussed encounters*. Cambridge University Press, Cambridge
- Klowait N (2017) A conceptual framework for researching emergent social orderings in encounters with automated computer-telephone interviewing agents. *Int J Commun Linguist Stud* 15(1):19–37
- Kotásek M (2015) Artificial intelligence in science fiction as a model of the posthuman situation of mankind. *World Lit Stud* 7(4):64–77
- Krummheuer AL (2008a) Zwischen den Welten: Verstehensicherung und Problembehandlung in künstlichen Interaktionen von menschlichen Akteuren und personifizierten virtuellen Agenten. In: Willems H (ed) *Weltweite Welten: Internet-Figurationen aus wissenssoziologischer Perspektive*. VS Verlag für Sozialwissenschaften, pp 269–294. https://doi.org/10.1007/978-3-531-91033-8_12
- Krummheuer AL (2008b) Die Herausforderung künstlicher Handlungsstrgerschaft. Frotzelattacken in hybriden Austauschprozessen von Menschen und virtuellen Agenten. In: Greif H, Mitrea O, Werner M (eds) *Information und Gesellschaft. Technologien einer sozialen Beziehung*. VS Research, pp 73–95
- Krummheuer AL (2009) Conversation analysis, video recordings, and human-computer interchanges. In: Kissmann U (ed) *Video interaction analysis. methods and methodology*. Peter Lang, Bern, pp 59–83
- Krummheuer AL (2015a) Technical agency in practice: the enactment of artefacts as conversation partners, actants and opponents. *PsychNology J* 13(2–3):179–202
- Krummheuer AL (2015b) Users, bystanders and agents: participation roles in human-agent interaction. In: Abascal J et al (eds) *Human-computer interaction—INTERACT 2015*. Springer, Cham, pp 240–247
- Krummheuer AL, Rehm M, Rodil K (2020) Triadic human-robot interaction: distributed agency and memory in robot assisted interactions. Companion of the 2020 ACM/IEEE international conference on human-robot interaction. ACM, New York, pp 317–319
- Lecerf Y (1963) Logique mathématique—machines de Turing réversibles. *C R Hebd Seances Acad Sci* 257:2597–2600
- Licoppe C, Rollet N (2020) “Je dois y aller”. *Analyses de séquences de clôtures entre humains et robot. Réseaux* 220–221(2–3):151–193. <https://doi.org/10.3917/res.220.0151>
- Livingston E (1986) *The ethnomethodological foundations of mathematics*. Routledge & Kegan Paul, Boston
- Livingston E (2008) *Ethnographies of reason*. Ashgate, Farnham
- Lohse M, Hanheide M, Pitsch K, Sagerer G, Rohlfing KJ (2009) Improving HRI design by applying systemic interaction analysis (SinA). *Interact Stud* 10(3):298–323. <https://doi.org/10.1075/is.10.3.03loh>
- Lynch M (1993) *Scientific practice and ordinary action: ethnomethodology and social studies of science*. Cambridge University Press, Cambridge
- Lynch M (2000) *The ethnomethodological foundations of conversation analysis*. *Text* 20(4):517–532
- Lynch M (2002) From naturally occurring data to naturally organized ordinary activities: comment on Speer. *Discourse Stud* 4(4):531–537
- Lynch M (2022) Garfinkel’s studies of work. In: Maynard DW, Heritage J (eds) *The ethnomethodology program: legacies and prospects*. Oxford University Press, New York, pp 114–137
- Lynch M, Livingston E (2017) *The conversation analytic foundations of ethnomethodology*. Talk at the 112th annual meeting of the American Sociological Association, Montreal, Quebec, Canada, 14 August 2017
- Mair M, Brooker P, Dutton W, Sormani P (2021) 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. <https://doi.org/10.1177/1468794120975988>
- Mair M, Sharrock WW, Greiffenhagen C (2022) Research with numbers. In: Maynard DW, Heritage J (eds) *The ethnomethodology program: legacies and prospects*. Oxford University Press, New York, pp 348–370
- Marres N, Sormani P (2023) Testing ‘AI’: do we have a situation?—a conversation. Working Paper Series—Collaborative Research Center 1187 Media of Cooperation. https://dspace.uni-siegen.de/bitstream/ubsi/2525/4/WPS_28_Marres_Sormani_Testing_AI.pdf
- Mayor E, Bietti L (2017) Ethnomethodological studies of nurse-patient and nurse-relative interactions: a scoping review. *Int J Nurs Stud* 70:46–57. <https://doi.org/10.1016/j.ijnurstu.2017.01.015>
- Meier zu Verl C, Kreplak Y, Eisenmann C, Dennis A (2020) Introduction. *Ethnogr Stud* 17:i–iv
- Mlynář J, Arminen I (2023) Respecifying social change: the obsolescence of practices and the transience of technology. *Front Sociol* 8. <https://doi.org/10.3389/fsoc.2023.1222734>
- Mlynář J, González-Martínez E, Lalanne D (2018) Situated organization of video-mediated interaction: a review of ethnomethodological and conversation analytic studies. *Interact Comput* 30(2):73–84. <https://doi.org/10.1093/iwc/iwx019>
- Mlynář J, Bahrami F, Ourednik A, Mutzner N, Verma H, Alavi H (2022) AI beyond deus ex machina – Reimagining intelligence in future cities with urban experts. In: CHI’22: conference on human factors in computing systems, New Orleans, USA, April 29–May 5, 2022. ACM. <https://doi.org/10.1145/3491102.3517502>
- Mlynář J, Depeursing A, Prior JO, Schaer R, Martroye de Joly A, Évéquoz F (2024) Making sense of radiomics: insights on human–AI collaboration in medical interaction from an observational user study. *Fron Commun*. <https://doi.org/10.3389/fcomm.2023.1234987>
- Mondada L (2014) The local constitution of multimodal resources for social interaction. *J Pragmat* 65:137–156. <https://doi.org/10.1016/j.pragma.2014.04.004>
- Mondada L, Peräkylä A (eds) (2023) *New perspectives on Goffman in language and interaction: body, participation and the self*. Routledge, New York
- Moore RJ (2012) *Ethnomethodology and conversation analysis: empirical approaches to the study of digital technology in action*. In:

- Price S, Jewitt C, Brown B (eds) *The SAGE handbook of digital technology research*. Sage, London, pp 217–235
- Moore RJ, Arar R (2019) *Conversational UX design: a practitioner's guide to the natural conversation framework*. ACM, New York
- Moore RJ, An S, Ren GJ (2023) The IBM natural conversation framework: a new paradigm for conversational UX design. *Hum Comput Interact* 38(3–4):168–193. <https://doi.org/10.1080/07370024.2022.2081571>
- Mori M (1970) The uncanny valley. *Energy* 7(4):33–35
- Muhle F (2008) 'Versteh ich grad nicht': Mensch-Maschine-Kommunikation als problem. *kommunikation @ gesellschaft* 9:21
- Munn Z, Peters MDJ, Stern C, Tufanaru C, McArthur A, Aromataris E (2018) Systematic review or scoping review? Guidance for authors when choosing between a systematic or scoping review approach. *BMC Med Res Methodol* 18:143. <https://doi.org/10.1186/s12874-018-0611-x>
- Nass C, Moon Y (2000) Machines and mindlessness: social responses to computers. *J Soc Issues* 56(1):81–103. <https://doi.org/10.1111/0022-4537.00153>
- Opfermann C, Pitsch K, Yaghouzadeh R, Kopp S (2017) The communicative activity of "making suggestions" as an interactional process: towards a dialog model for HAI. In: *Proceedings of the 5th international conference on human agent interaction (HAI '17)*. ACM, pp 161–170. <https://doi.org/10.1145/3125739.3125752>
- Parry RH, Land V (2013) Systematically reviewing and synthesizing evidence from conversation analytic and related discursive research to inform healthcare communication practice and policy: an illustrated guide. *BMC Med Res Methodol* 13:69. <https://doi.org/10.1186/1471-2288-13-69>
- Payr S (2010) Closing and closure in human-companion interactions: analyzing video data from a field study. In: *19th International symposium in robot and human interactive communication*, pp 476–481. <https://doi.org/10.1109/ROMAN.2010.5598625>
- Payr S (2013) Virtual butlers and real people: styles and practices in long-term use of a companion. In: *Trappl R (ed) Your virtual butler. Lecture notes in computer science, vol 7407*. Springer, Berlin
- Pelikan HRM (2021) Why autonomous driving is so hard: the social dimension of traffic. In: *Companion of the 2021 ACM/IEEE international conference on human-robot interaction*, pp 81–85. <https://doi.org/10.1145/3434074.3447133>
- Pelikan HRM, Broth M (2016) Why that Nao? How humans adapt to a conventional humanoid robot in taking turns-at-talk. In: *Proceedings of the 2016 CHI conference on human factors in computing systems*, pp 4921–4932. <https://doi.org/10.1145/2858036.2858478>
- Pelikan HRM, Broth M, Keevallik L (2020) 'Are you sad, Cozmo?': how humans make sense of a home robot's emotion displays. In: *Proceedings of the 2020 ACM/IEEE international conference on human-robot interaction*, pp 461–470. <https://doi.org/10.1145/3319502.3374814>
- Pelikan HRM, Broth M, Keevallik L (2022) When a robot comes to life: the interactional achievement of agency as a transient phenomenon. *Soc Interact Video-Based Stud Hum Soc* 5(3). <https://tidsskrift.dk/socialinteraction/article/view/129915>. Accessed 24 Aug 2023
- Peräkylä A, Sorjonen M-L (2012) *Emotion in interaction*. Oxford University Press, New York. <https://doi.org/10.1093/acprof:oso/9780199730735.001.0001>
- Petersson L, Larsson I, Nygren JM, Nilsen P, Neher M, Reed JE, Tyskbo D, Svedberg P (2022) Challenges to implementing artificial intelligence in healthcare: a qualitative interview study with healthcare leaders in Sweden. *BMC Health Serv Res* 22:850. <https://doi.org/10.1186/s12913-022-08215-8>
- Pflanzer M, Dubljević V, Bauer WA, Orcutt D, List G, Singh MP (2023) Embedding AI in society: ethics, policy, governance, and impacts. *AI Soc* 38:1267–1271. <https://doi.org/10.1007/s00146-023-01704-2>
- Pillet-Shore D (2010) Making way and making sense: including newcomers in interaction. *Soc Psychol Q* 73(2):152–175. <https://doi.org/10.1177/0190272510369668>
- Pilling M, Coulton P, Lodge T, Crabtree A, Chamberlain A (2022) Experiencing mundane AI futures. In: *Lockton D, Lenzi S, Hekert P, Oak A, Sádaba J, Lloyd P (eds) DRS2022: Bilbao, 25 June–3 July, Bilbao, Spain*. <https://doi.org/10.21606/drs.2022.283>
- Pilnick A, Trusson D, Beeke S, O'Brien R, Goldberg S, Harwood RH (2018) Using conversation analysis to inform role play and simulated interaction in communications skills training for healthcare professionals: identifying avenues for further development through a scoping review. *BMC Med Educ* 18(1):267. <https://doi.org/10.1186/s12909-018-1381-1>
- Pink S, Berg M, Lupton D, Ruckenstein M (eds) (2022) *Everyday automation: experiencing and anticipating emerging technologies*. Routledge, Oxon/New York
- Pitsch K (2020) Répondre aux questions d'un robot: Dynamique de participation des groupes adultes-enfants dans les rencontres avec un robot guide de musée. *Réseaux N° 220–221(2)*:113–150. <https://doi.org/10.3917/res.220.0113>
- Pitsch K, Koch B (2010) How infants perceive the toy robot Pleo. An exploratory case study on infant-robot-interaction. In: *Second international symposium on new frontiers in human-robot-interaction (AISB). The Society for the Study of Artificial Intelligence and the Simulation of Behaviour*
- Pitsch K, Kuzuoka H, Suzuki Y, Sussenbach L, Luff P, Heath C (2009) "The first five seconds": contingent stepwise entry into an interaction as a means to secure sustained engagement in HRI. In: *RO-MAN 2009—the 18th IEEE international symposium on robot and human interactive communication*, pp 985–991. <https://doi.org/10.1109/ROMAN.2009.5326167>
- Pitsch K, Vollmer A-L, Mühlhig M (2013) Robot feedback shapes the tutor's presentation: how a robot's online gaze strategies lead to micro-adaptation of the human's conduct. *Interact Stud Soc Behav Commun Biol Artif Syst* 14(2):268–296. <https://doi.org/10.1075/is.14.2.06pit>
- Pitsch K, Gehle R, Dankert T, Wrede S (2017) Interactional dynamics in user groups: answering a robot's question in adult-child constellations. In: *Proceedings of the 5th international conference on human agent interaction (HAI '17)*. Association for Computing Machinery, pp 393–397. <https://doi.org/10.1145/3125739.3132604>
- Pollner M (1991) Left of ethnomethodology: the rise and decline of radical reflexivity. *Am Sociol Rev* 56(3):370–380. <https://doi.org/10.2307/2096110>
- Pollner M (2012) The end(s) of ethnomethodology. *Am Sociol* 43(1):7–20. <https://doi.org/10.1007/s12108-011-9144-z>
- Pollner M, Zimmerman DH (1970) The everyday world as a phenomenon. In: *Douglas JD (ed) Understanding everyday life: towards a reconstruction of sociological knowledge*. Aldine Publishing, Chicago, pp 80–103
- Porcheron M, Fischer JE, Sharples S (2017) 'Do animals have accents?': talking with agents in multi-party conversation. In: *Proceedings of the 2017 ACM conference on computer supported cooperative work and social computing*, pp 207–219. <https://doi.org/10.1145/2998181.2998298>
- Porcheron M, Fischer JE, Reeves S, Sharples S (2018) Voice interfaces in everyday life. In: *Proceedings of the 2018 CHI conference on human factors in computing systems*, pp 1–12. <https://doi.org/10.1145/3173574.3174214>
- Psathas G (1995) *Conversation analysis: the study of talk-in-interaction*. Sage Publications, Thousand Oaks

- Psathas G (2008) Reflections on the history of ethnomethodology: the Boston and Manchester ‘schools’. *Am Sociol* 39(1):38–67. <https://doi.org/10.1007/s12108-008-9032-3>
- Randall D, Rouncefield M, Tolmie P (2021) Ethnography, CSCW and ethnomethodology. *Comput Support Coop Work* 30(2):189–214. <https://doi.org/10.1007/s10606-020-09388-8>
- Raudaskoski PL (2023) Ethnomethodological conversation analysis (EMCA) and the study of assemblages. *Front Sociol* 8. <https://doi.org/10.3389/fsoc.2023.1206512>
- Rawls AW (2005) Garfinkel’s conception of time. *Time Soc* 14(2–3):163–190. <https://doi.org/10.1177/0961463X05055132>
- Rawls AW (2023) The Goffman-Garfinkel correspondence: planning “On Passing”. *Etnogr Ric Qual* 1:175–218
- Reeves S (2017) Some conversational challenges of talking with machines. In: Talking with conversational agents in collaborative action, workshop at the 20th ACM conference on computer supported cooperative work and social computing
- Reeves S (2019a) Conversation considered harmful? In: CUI ’19: proceedings of the 1st international conference on conversational user interfaces. Article No. 10. <https://doi.org/10.1145/3342775.3342796>
- Reeves S (2019b) How UX practitioners produce findings in usability testing. *ACM Trans Comput Hum Interact* 26(1):1–38. <https://doi.org/10.1145/3299096>
- Reeves S (2022) Navigating incommensurability between ethnomethodology, conversation analysis, and artificial intelligence. *arXiv.org*. <https://doi.org/10.48550/arXiv.2206.11899>
- 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 Publications, London, pp 573–592
- Reeves S, Brown B, Laurier E (2009) Experts at play: understanding skilled expertise. *Games Cult* 4(3):205–227. <https://doi.org/10.1177/1555412009339730>
- Relieu M, Sahin M, Francillon A (2020) Une approche configurationnelle des leurres conversationnels. *Réseaux* 220–221(2):81–111. <https://doi.org/10.3917/res.220.0081>
- Robins B, Dickerson P, Stribling P, Dautenhahn K (2004) Robot-mediated joint attention in children with autism: a case study in robot-human interaction. *Interact Stud* 5(2):161–198. <https://doi.org/10.1075/is.5.2.02rob>
- Robinson JD, Heritage J (2014) Intervening with conversation analysis: the case of medicine. *Res Lang Soc Interact* 47(3):201–218. <https://doi.org/10.1080/08351813.2014.925658>
- Rollet N, Clavel C (2020) “Talk to you later”: doing social robotics with conversation analysis. Towards the development of an automatic system for the prediction of disengagement. *Interact Stud* 21(2):268–292. <https://doi.org/10.1075/is.19001.roll>
- Rollet N, Jain V, Licoppe C, Devillers L, Gamberini L, Spagnolli A, Jacucci G, Blankertz B, Freeman J (2017) Towards interactional symbiosis: epistemic balance and co-presence in a quantified self experiment. In: *Symbiotic interaction: 5th international workshop*, Sep 2016, Padua, Italy, pp 143–154. https://doi.org/10.1007/978-3-319-57753-1_13
- Rossano F (2012) Gaze in conversation. In: Sidnell J, Stivers T (eds) *The handbook of conversation analysis*. Wiley-Blackwell, Chichester, pp 308–329
- Saalasti S, Pajo K, Fox B, Pekkala S, Laakso M (2023) Embodied-visual practices during conversational repair: scoping review. *Res Lang Soc Interact* 56(4):311–329. <https://doi.org/10.1080/08351813.2023.2272528>
- Sacks H (1967) The search for help: no one to turn to. In: Shneidman ES (ed) *Essays in self-destruction*. Science House, New York, pp 203–223
- Sacks H (1972) An initial investigation of the usability of conversation data for doing sociology. In: Sudnow D (ed) *Studies in social interaction*. The Free Press, New York, pp 31–74
- Sacks H (1984a) Notes on methodology. In: Heritage J, Maxwell Atkinson J (eds) *Structures of social action: studies in conversation analysis*. Cambridge University Press, Cambridge, pp 2–27
- Sacks H (1984b) On doing “being ordinary.” In: Heritage J, Maxwell Atkinson J (eds) *Structures of social action: studies in conversation analysis*. Cambridge University Press, Cambridge, pp 413–429
- Sacks H (1992) *Lectures on conversation (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. <https://doi.org/10.2307/412243>
- 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*. <https://doi.org/10.23987/sts.122892>
- Sahin M, Relieu M, Francillon A (2017) Using chatbots against voice spam: analyzing Lenny’s effectiveness. In: *Thirteenth symposium on usable privacy and security (SOUPS 2017)*, pp 319–337
- Schegloff EA (1968) Sequencing in conversational openings. *Am Anthropol* 70(6):1075–1095. <https://doi.org/10.1525/aa.1968.70.6.02a00030>
- Schegloff EA (1980) What type of interaction is it to be? In: *ACL’80: proceedings of the 18th annual meeting on Association for computational linguistics*. Association for Computational Linguistics, Stroudsburg, PA, pp 81–82
- Schegloff EA (1987) Analyzing single episodes of interaction: an exercise in conversation analysis. *Soc Psychol Q* 50(2):101–114. <https://doi.org/10.2307/2786745>
- Schegloff EA (1988) Description in the social sciences I: talk-in-interaction. *IPrA Papers Pragmat* 2(1–2):1–24. <https://doi.org/10.1075/iprapip.2.1-2.01sch>
- Schegloff EA (1992) Repair after next turn: the last structurally provided defense of intersubjectivity in conversation. *Am J Sociol* 97(5):1295–1345. <https://doi.org/10.1086/229903>
- Schegloff EA (1999) “Schegloff’s texts” as “Billig’s data”: a critical reply. *Discourse Soc* 10(4):558–572. <https://doi.org/10.1177/0957926599010004006>
- Schegloff EA (2007) *Sequence organization in interaction: a primer in conversation analysis*. Cambridge University Press, Cambridge
- Schegloff EA, Sacks H (1973) Opening up closings. *Semiotica* 8(4):289–327. <https://doi.org/10.1515/semi.1973.8.4.289>
- Schenkein J (1978) Sketch of an analytic mentality for the study of conversational interaction. In: Schenkein J (ed) *Studies in the organization of conversational interaction*. Academic Press, New York, pp 1–6
- Schwartz RD (1989) Artificial intelligence as a sociological phenomenon. *Can J Sociol* 14(2):179–202. <https://doi.org/10.2307/3341290>
- Selting M (1994) Emphatic speech style: with special focus on the prosodic signalling of heightened emotive involvement in conversation. *J Pragmat* 3(4):375–408
- Sharrock W (1999) The omnipotence of the actor: Erving Goffman on “the definition of the situation”. In: Smith G (ed) *Goffman and social organization: studies in a sociological legacy*. Routledge, London, pp 119–137
- Sidnell J (2017) Distributed agency and action under the radar of accountability. In: Enfield NJ, Kockelman P (eds) *Distributed agency*. Oxford University Press, Oxford, pp 87–96. <https://doi.org/10.1093/acprof:oso/9780190457204.003.0010>
- Silver D, Huang A, Maddison CJ, Guez A, Sifre L, van den Driessche G, Schrittwieser J, Antonoglou I, Panneershelvam V, Lanctot M, Dieleman S, Grewe D, Nham J, Kalchbrenner N, Sutskever I, Lillicrap T, Leach M, Kavukcuoglu K, Graepel T, Hassabis D (2016) Mastering the game of Go with deep neural networks and tree search. *Nature* 529:7587. <https://doi.org/10.1038/nature16961>

- Smith G (2003) Ethnomethodological readings of Goffman. In: Javier Treviño A (ed) *Goffman's legacy*. Rowman & Littlefield, Lanham, pp 254–283
- Smith BC (2019) *The promise of artificial intelligence: reckoning and judgment*. MIT, Cambridge, MA
- Sormani P (2019) Ethnomethodological analysis. In: Atkinson P, Delamont S, Cernat A, Sakshaug JW, Williams RA (eds) *SAGE research methods foundations*. <https://methods.sagepub.com/foundations/ethnomethodological-analysis>. Accessed 1 Oct 2022
- Sormani P (2020) 'DIY AI'? Practising kit assembly, locating critical inquiry. *Ethnogr Stud* 17:60–80. <https://doi.org/10.5281/zenodo.4050539>
- Sormani P (2022) Remaking intelligence? Of machines, media, and montage. *TecnoScienza* 13(2):57–85. <https://doi.org/10.6092/issn.2038-3460/17579>
- 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>
- Sormani P, vom Lehn D (eds) (2023) *The Anthem companion to Harold Garfinkel*. Anthem Press, London
- Stivers T, Robinson JD (2006) A preference for progressivity in interaction. *Lang Soc* 35(3):367–392. <https://doi.org/10.1017/S0047404506060179>
- Stokoe E (2012) Moving forward with membership categorization analysis: methods for systematic analysis. *Discourse Stud* 14(3):277–303
- Stokoe E (2014) The conversation analytic role-play method (CARM): a method for training communication skills as an alternative to simulated role-play. *Res Lang Soc Interact* 47(3):255–265. <https://doi.org/10.1080/08351813.2014.925663>
- Stommel W, de Rijk L, Boumans R (2022) 'Pepper, what do you mean?' Miscommunication and repair in robot-led survey interaction. In: Proceedings of the 31st IEEE international conference on robot and human interactive communication (RO-MAN), pp 385–392. <https://doi.org/10.1109/RO-MAN53752.2022.9900528>
- Suchman L (1987) *Plans and situated actions: the problem of human-machine communication*. Cambridge University Press, Cambridge
- Suchman L (2007) *Human-machine reconfigurations: plans and situated actions*, 2nd edn. Cambridge University Press, Cambridge
- Suchman L (2023a) Imaginaries of omniscience: automating intelligence in the US Department of Defense. *Soc Stud Sci* 53(5):761–786. <https://doi.org/10.1177/03063127221104938>
- Suchman L (2023b) The uncontroversial 'thingness' of AI. *Big Data Soc* 10(2):20539517231206790. <https://doi.org/10.1177/20539517231206790>
- Suchman L, Trigg RH (1993) Artificial intelligence as craftwork. In: Chalklin S, Lave J (eds) *Understanding practice: perspectives on activities and context*. Cambridge University Press, Cambridge, pp 144–178
- Suchman L, Trigg R, Blomberg J (2002) Working artefacts: ethnomethods of the prototype. *Br J Sociol* 53(2):163–179. <https://doi.org/10.1080/00071310220133287>
- Sudnow D (1983) *Pilgrim in the microworld*. Heinemann, London
- Süssenbach L, Pitsch K, Berger I, Riether N, Kummert F (2012) "Can you answer questions, Flobi?": interactionally defining a robot's competence as a fitness instructor. In: Proceedings of the 21st IEEE international symposium on robot and human interactive communication (RO-MAN), pp 1121–1128. <https://doi.org/10.1109/ROMAN.2012.6343899>
- Torre I, Tuncer S, McDuff D, Czerwinski M (2021) Exploring the effects of virtual agents' smiles on human-agent interaction: a mixed-methods study. In: 2021 9th international conference on affective computing and intelligent interaction (ACII), pp 1–8. <https://doi.org/10.1109/ACII52823.2021.9597445>
- Tuncer S, Gillet S, Leite I (2022) Robot-mediated inclusive processes in groups of children: from gaze aversion to mutual smiling gaze. *Front Robot AI* 9. <https://doi.org/10.3389/frobt.2022.729146>
- Tuncer S, Licoppe C, Luff P, Heath C (2023) Recipient design in human-robot interaction: the emergent assessment of a robot's competence. *AI Soc* (online first). <https://doi.org/10.1007/s00146-022-01608-7>
- Velkovska J, Relieu M (2020) Pourquoi ethnographier les interactions avec les agents conversationnels? *Réseaux* 2–3(220–221):9–20. <https://doi.org/10.3917/res.220.0009>
- Velkovska J, Zouinar M, Veyrier C-A (2020) Les relations aux machines conversationnelles: vivre avec les assistants vocaux à la maison. *Réseaux* 220–221(2):47–79. <https://doi.org/10.3917/res.220.0047>
- Walker T, Christensen H, Mirheidari B et al (2020) Developing an intelligent virtual agent to stratify people with cognitive complaints: a comparison of human-patient and intelligent virtual agent-patient interaction. *Dementia* 19(4):1173–1188. <https://doi.org/10.1177/1471301218795238>
- Wallis P (2008) Revisiting the DARPA communicator data using conversation analysis. *Interact Stud* 9(3):434–457. <https://doi.org/10.1075/is.9.3.05wal>
- Walton C, Antaki C, Finlay WML (2020) Difficulties facing people with intellectual disability in conversation: initiation, co-ordination, and the problem of asymmetric competence. In: Wilkinson R, Rae JP, Rasmussen G (eds) *Atypical interaction: the impact of communicative impairments within everyday talk*. Springer, Cham, pp 93–127. https://doi.org/10.1007/978-3-030-28799-3_4
- Weizenbaum J (1967) Contextual understanding by computers. *Commun ACM* 10(8):474–480. <https://doi.org/10.1145/363534.363545>
- Wieder DL (1974) *Language and social reality: the case of telling the convict code*. Mouton, The Hague
- Wieder DL (1999) Ethnomethodology, conversation analysis, microanalysis, and the ethnography of speaking (EM-CA-MA-ES): resonances and basic issues. *Res Lang Soc Interact* 32(1–2):163–171
- Wilkinson R (2019) Atypical interaction: conversation analysis and communicative impairments. *Res Lang Soc Interact* 52(3):281–299. <https://doi.org/10.1080/08351813.2019.1631045>
- Wittgenstein L (1953) *Philosophical investigations*. Blackwell, Oxford
- Wooffitt R (1994) Applying sociology: conversation analysis in the study of human-(simulated) computer interaction. *Bull Méthodol Sociol* 43(1):7–33
- Yamazaki A, Yamazaki K, Ikeda K, Burdelski M, Fukushima M, Suzuki T, Kurihara M, Kuno Y, Kobayashi Y (2013) Interactions between a quiz robot and multiple participants: focusing on speech, gaze and bodily conduct in Japanese and English speakers. *Interact Stud* 14(3):366–389. <https://doi.org/10.1075/is.14.3.04yam>
- Zhang X, Jin H (2023) How does smart technology, artificial intelligence, automation, robotics, and algorithms (STAARA) awareness affect hotel employees' career perceptions? A disruptive innovation theory perspective. *J Hosp Market Manag* 32(2):264–283. <https://doi.org/10.1080/19368623.2023.2166186>
- 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>

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