3 Methodology: Encountering Human Computation Ethnographically

> The anthropological tradition of extended field research stems in part from the recognition that in human affairs, the relation between beliefs and practice is invariably complex. In any given situation, what people believe that they should do, what they report that they do, and what they can be seen to do by an outside observer may all differ somewhat. However, in the absence of systematic participant observation, such disparities are difficult to detect. If we base our study of science solely on scientists' self-reports, we may fail to realize what the reported actions or tools actually consist of; if we look only at observed practice, we may miss what particular objects or actions mean to the scientists involved; and if we limit ourselves to introspection about a particular problem addressed by scientists, we may learn little or nothing about how the scientists themselves approach the problem. (Forsythe [1993] 2001h, 36)

In my investigation of how HC-based CS are formed in the interplay of different human and nonhuman actors and into the intraversions of their human–technology relations, I followed the tradition of cultural and digital anthropological inductive ethnographic fieldwork based on (constructivist) grounded theory (Glaser and Strauss [1967] 1971; Charmaz 2000; 2014), inspired by Forsythe's (2001f.) pioneering research on medical informatics and AI in the 1980s and 90s and her reflections on ethnographers in these fields. Owing to her work and that of other ethnographers discussed in the related work section, my research builds on their achievements in STS research, ethnographic reflection, and new collaborative approaches.[1] I conducted field research over the course of nearly three years in Germany and the US. My ethnographic approach, which draws on

1 However, since Forsythe's research, a lot has also changed, such as the field of AI research itself. Human computation had not yet been born as a sub-research field of AI. Another difference in my field work experience lies in the openness of my research partners toward other epistemologies.

praxiography,[2] combines multiple perspectives and methods and considers assemblages and their human–technology relations across different scales and dimensions.[3]

I started from the perspective of designers and developers of HC-based CS but also included the perspectives of participants and researchers. Since HC-based CS assemblages are always sociomaterial and formed by human–technology relations, I also included nonhuman actors in my ethnographic analysis, including code, infrastructures, data, and how it circulates. This required navigating between different perspectives, diverse practices, and processes through which the relations between these elements became observable, as well as the application of experimental methods to include the relevant materials (e.g., source code and large digital textual chat data) in my analysis to supplement the more well-established methods, such as participant observation and qualitative interviews, I applied.

In this chapter, I first discuss methodological considerations that include co-laborative ethnography inspired by praxiography, and the methodology of (constructivist) grounded theory. This is followed by its operationalization in my research process, which includes reflections on my role in the field. Finally, I individually introduce each of the main methods I worked with.

Praxiographically Inspired Co-Laborative Ethnography

I followed a co-laborative (Niewöhner 2014) ethnographic approach to analyze HC-based CS assemblages in their complexity and their formation in the interplay of different human and nonhuman actors. I, first, discuss ethnography inspired by praxiography before turning to co-laboration.

Ethnography is particularly well-suited for investigating my research questions regarding sociotechnical phenomena as its intensive participation and observation enables the tracing of situations, interactions, and conflicts in their dynamics (Knecht 2013, 86). It is characterized by the collection of various complementary empirical material allowing for the inclusion of multiple perspectives (Amann and Hirschauer 1997, 16; cited in Knecht 2013, 86). Furthermore, due to its processuality and openness, the ethnographic approach can adapt dynamically in the course of research to shifting research questions and developments in the field (Pink et al. 2016, 11). Finally, ethnography allows for the encounter of aspects that had not been initially anticipated in the research design due to its participating, immersive, and open approach to the phenomena studied (Knecht 2013, 86). For these reasons, ethnography as an approach is always multi-methodical and

2 Following ethnologist Michi Knecht, I conceive praxiography as a specific position within ethnography (Knecht 2012, 249) I discuss this relation between praxiography and ethnography in more detail when discussing participant observation below.

3 My approach could, therefore, also be considered "assemblage ethnography," as described by anthropologist and sociologist Ayo Wahlberg, which studies configurations forming assemblages "across scales, sites, and practices" (2022).

-perspectival. I combined virtual digital ethnography with on-site in-person ethnography, also referred as "hybrid ethnography," (Przybylski 2021) in my research.[4]

My ethnographic approach is inspired by praxiography, which emerged as a new methodical approach in STS research and ANT, and aimed at moving beyond binary ontologies (such as the distinction between nature and culture) (Knecht 2012, 249; *cf.* Niewöhner 2019a), which fail to grasp how humans and nonhumans together shape sociotechnical phenomena. Praxiography describes the "*observation and description of actors and objects in action and interaction*" (Mol 2002b, 53–55; Hirschauer 2004, 73; Law 2004, 59; cited in Knecht 2013, 98, emphasis i.o.). As such, similar to ethnography, it focuses on processual and praxeological perspectives (Knecht 2013, 96), and is, therefore, particularly valuable for the analysis of sociotechnical assemblages and relations between humans and technology. In this sense, praxiography and ethnography have much in common (Knecht 2012), with both forms of knowledge production leading to a research practice that continuously reflects and critically reassembles its concepts and practices in the encounter with the field (Knecht 2012, 262).[5] However, unlike ethnography, praxiography places less emphasis on reflection the social dimensions of researchers' participation in the field. This approach can result in overlooking social dimensions (Knecht 2013, 97),[6] as well as aspects of meaning-making and perception. Therefore, while building on central perspectives and approaches from praxiography, with participant observation as one important method, this work does not pursue a radical praxiographical approach. An additional pragmatic reason for this is that HC-based CS projects come into being through the practices of various actors and materialities distributed across the world, in different time zones, which made it impossible to directly observe all practices through participant observation.

The understanding of praxis in this work follows Beck, Niewöhner, and Sørensen (2012), who define praxis as an analytical perspective regarding "human coexistence as manifoldly situated" (2012, 33). According to this understanding, ethnographic practice focuses on what may be experienced directly but simultaneously does not neglect their material, historical, and cultural conditions (Beck, Niewöhner, and Sørensen 2012, 33). "Praxis thus includes different time horizons, different spaces and the material-objective contexts in the analysis of the concrete *how* of human coexistence" (Beck, Niewöhner, and Sørensen 2012, 33, emphasis i.o.). Building on Beck's complex situational analysis of the use of technology, as discussed in the last chapter, I focused in my processual ethnographic approach on the dynamics of HC-based CS and analyzed both the sociotechnical systems and their surrounding practices (Beck 1997; *cf.* Bareither a.o. 2013, 32).

---

4 On digital ethnographic approaches and the combination of virtual and on-site ethnography, see, e.g., Boellstorff et al. (2012); Miller and Horst (2012); Pink et al. (2016); Fleischhack (2019); Przybylski (2021).

5 Different perspectives and understandings of praxiography exist (e.g., Mol 2002b; Niewöhner 2017). I aim in this section to describe how praxiography influences my ethnographic approach in a pragmatic way without going into details about different praxiographic approaches in general. For a historical overview of the emergence and different developments of practice theory see, e.g., Beck (2019) and Sørensen and Schank (2017).

6 Praxiography in its radical understanding focuses only on the practices, their constellations and structures, and not on actors or social collectives (Knecht 2012, 249).

In order to conduct research *with* rather than *on* my research partners (e.g., Ingold 2020, 210), I followed a collaborative approach with the Human Computation Institute. My approach in this was inspired by Forsythe's research on AI and Suchman's fieldwork at Xerox's Palo Alto Research Center and her engagement with some projects there, which she described as an "involved but also constructive engagement" (Suchman 2021, 70–71). I aimed at constructively but also critically contributing to the development of HC-based CS projects.

It has become common knowledge in cultural anthropology that we, as researchers, construct the phenomena we study. By asking certain questions, approaching the phenomena in specific ways, and defining boundaries—setting cuts, as Barad has described it (2007a)[7]—we not only describe but always also create our research fields and realities (Law 2004, 5). Design and futures anthropologist Sarah Pink wrote that "[w]hat was conventionally called 'the ethnographic field' is ongoingly made and remade through our active participation as ethnographers in collaboration with research participants, other stakeholders in research and future readers and viewers" (Pink 2018, 201). Ethnographic knowledge must, therefore, also always be considered both partial and situated (Haraway 1988). However, this understanding of the ethnographer's construction of the research field also means we (ethnographers) cannot sneak away once our research has been published. It has "real material consequences" (Barad 1996, 183) and, therefore, to stay with Barad, "[w]e are responsible for the world of which we are a part, not because it is an arbitrary construction of our choosing but because reality is sedimented out of particular practices that we have a role in shaping and through which we are shaped" (Barad 2007b, 390).

Working together with the Human Computation Institute, therefore, played a major role in my research. Even though I take a multi-perspectival approach, including both the participants' and researchers' perspectives, as well as the other two case studies of Foldit and ARTigo, the institute is a core focus and starting point of this work. My collaborative approach was inspired by that of co-laboration developed by Niewöhner (2014), which builds a complement to the ethnographic approach with the goal of strengthening and expanding anthropological reflexivity and the production of new knowledge (Niewöhner 2019b, 26–27; Bieler, Bister, and Schmid 2021, 88). Niewöhner defines co-laboration as "temporary, non-teleological, joint epistemic work aimed at producing disciplinary reflexivities not interdisciplinary shared outcomes" (Niewöhner 2016, 2; 2019b, 27). In this way, it differs from forms of collaboration that follow shared work objectives. Instead, it builds on shared epistemic work and on the "exchange between different epistemic cultures" (Klausner and Niewöhner 2020, 162; *cf.* Niewöhner 2016).[8] Niewöhner specifies that "practices of co-laborating help to diversify existing notions of reflexivity and critique, thereby broadening the analytical spectrum and adding interpretative degrees of freedom" (2016, 2). It is particularly useful in research fields characterized by high degrees of reflection (Niewöhner 2016, 41).

I engaged in a field wherein scientific methods already guided the development of HC systems, reflexive analysis, and publications, even if not ethnographically. The ques-

7 "Different agential cuts produce different phenomena" (Barad 2007a, 175).

8 "People do different things through the same process" (Niewöhner 2019b, 32).

tions that brought me to the field, regarding, for example, the subjectivities in HC-based CS, how actors involved strived to contribute to something greater, and the societal implications overlapped with questions already guiding the institute's work. This presented opportunities to think about them together from different perspectives and disciplinary viewpoints.

Finally, regarding co-laboration, I would like to highlight the aspect of moving between roles in the field. Through co-laboration, Niewöhner writes, clear roles, power relations, and relationships continuously change or break down between the parties involved from which "[a]mbiguous and searching moments emerge" (2019b, 39–40). Despite the advantages resulting from my collaborative engagement, such as gaining a deeper understanding of the phenomenon, this ambiguity and the uncertainties also presented challenges during my research process. Questions of how I could meaningfully and constructively but, at the same time, critically contribute to the development of new hybrid human–AI systems accompanied my research process, and, at times, proved to be more difficult than expected. These questions are not new to anthropological STS research in the field of computer science and AI, having already been discussed, for example, by Forsythe when "studying up" was still uncommon in ethnographic research (2001f). Forsythe mentions how this (at that time) new fieldwork context of studying up does not only "create new kinds of vulnerability, but the risks to both anthropologist and informants may extend far beyond the fieldwork itself" ([1999] 2001d, 125). It also creates new "dilemmas for the anthropologist as conflicting loyalties pull her in opposite directions and the collapsed roles of participant, observer, critic, employee, and colleague collide with one another" (Forsythe [1999] 2001d, 125).[9] Such dilemmas sometimes further complicated my own role within the field beyond the common challenge of closeness and critical distance to the field in ethnographic research. I sometimes, for instance, felt like I missed the chance to contribute my perspective based on my ethnographic observations simply because I was too slow in deconstructing and reflecting on current developments. Here, open discussion and reflection of my role both with team members of the Human Computation Institute and with colleagues in STS and cultural anthropology proved to be important to learn from these dilemmas experienced and reflect on the divergent different temporalities of HC development and cultural anthropological knowledge production.

Constructivist Grounded Theory

My co-laborative ethnographic approach builds on the methodology of (constructivist) grounded theory. How ethnographers approach the phenomenon they study, what they "bring" with them, and their position—as I have discussed above—in some ways, constructs the research object in the first place. Hultin aptly describes how her "position as a researcher [must be] understood as a genealogical line (rather than a dot) entangled with and conditioning what 'data' can become" (Hultin 2019, 100). This entanglement frames

9 Similar concerns and challenges have been described by European ethnologist and sociologist Rolf Lindner for ethnographic research in general (1981).

what ethnographers learn, even if it does not necessarily determine the knowledge to be gained from the field (Charmaz 2009, 48). According to Kathy Charmaz, this holds equivalently for theorization: "how we theorize reflects our interactions before we begin and those occurring within and beyond the field" (Charmaz 2009, 48). She, therefore, concludes that "[t]heorizing arises through analytic thinking about our field experiences, not merely recording and synthesizing them" (Charmaz 2009, 48). The experiences we (ethnographers) make in the research process, the methods we apply, and our knowledge and emotions also feed into the analysis, not only the (in my case) transcribed, digitized empirical data that I gathered and analyzed. Not all information contributing to theorization can be digitized in the first place, instead, it unfolds in the interaction with the field site and its actors. Grounded theory as a systematic but flexible method (Charmaz 2014, 1) allows not only for incorporating these diverse materials, experiences, and reflections but particularly encourages circular movements between data collection and analysis. It is, therefore, particularly well-suited for studying processes.

Grounded theory goes back to sociologists Barney G. Glaser and Anselm L. Strauss, who first developed it in their publication *The Discovery of Grounded Theory: Strategies for Qualitative Research* ([1967] 1971). Since then, grounded theory has been developed further by not only Glaser and Strauss themselves but by other researchers who, at times, also distanced themselves from the positivism that was underlying Glaser and Strauss' understandings of the method (Charmaz 2014, 12). In my research, I follow a *constructivist* grounded theory approach, first mentioned in Charmaz (2000). While embracing important aspects of Glaser and Strauss' approach, such as inductivity, comparison, emergence, open-endedness, and iterativity (Charmaz 2014, 12–13), it—in contrast to Glaser and Strauss' method (or "conventional" [Charmaz 2009, 52] grounded theory)—"highlights the flexibility of the method and resists mechanical applications of it" (Charmaz 2014, 13). It aligns with the understanding of situated and partial knowledge and construction of the field described above (Charmaz 2014, 13).

Starting from inductive research, grounded theory is a "comparative, iterative, and *interactive* method. The emphasis in grounded theory is on analysis of data; however, early data analysis informs data collection" (Charmaz 2012, 2, emphasis i.o.). It jointly considers data collection and analysis and, therefore, allows for the construction of middle-range theories that are "'grounded' in their data" (Charmaz 2012, 2).

Data analysis was part of my research process from the early collection of empirical data. Even though, as stated above, analysis cannot be reduced to coding digital data, such as interview transcripts, fieldnotes, media articles, and chats, I gathered most of the data, as well as my fieldnotes, in the qualitative data analysis software MAXQDA (VERBI – Software. Consult. Sozialforschung. GmbH, n.d.). My analysis of this data followed the two main phases described by Charmaz as the "initial phase," which included line-by-line coding of practices, processes, meanings, perceptions, and imaginations, and a "focused, selective phase that uses the most significant or frequent initial codes to sort, synthesize, integrate, and organize large amounts of data" (Charmaz 2014, 113). This approach is also well established in ethnographic coding (Breidenstein et al. 2020). Working with codes, which Star described as "transitional objects" ([2007] 2015, 130) using psychoanalyst Donald Woods Winnicott's (1965) term, helps one to understand the phenomenon studied better, while, at the same time, abstracting from it and creating something new

(Star [2007] 2015, 130). Notably, but not only, in the second phase, comparing codes with each other or with other data allowed me to further abstract and connect my codes to existing research (Charmaz 2012, 4). By then going back into the field and collecting more data with the already emerging categories in mind, these codes densified further. The second phase of my field research in Ithaca, NY, in 2022, was particularly constructive for my engagement with the data. This process of condensing and further abstracting codes is referred to as theoretical sampling in grounded theory terminology.

> Theoretical sampling stretches the codes, forcing other sorts of knowledge of the object. The theory that develops repeats the attachment-separation cycle [that already characterizes codes], but in this sense taking a code and moving it through the data. In so doing, it fractures both code and data. (Star [2007] 2015, 130)

This process of oscillating between analysis and data collection was complemented by writing notes, or memos, in addition to fieldnotes, in which, for example, I captured connections between data or helpful theoretical concepts and how they related to my data.

Following the discussion of methodological considerations guiding my research, the next section describes its operationalization in the research process.

Doing Research On, With, and Among Researchers and Developers

I first learned about HC by way of a computer science class by Bry at LMU Munich. I came across its description in early 2019 while searching for classes to attend in the following term: "Human Computation is a novel and interesting branch of Computer Science that intends to incorporate human intelligence to solve problems computers alone typically have problems to deal with." (Bogner n.d.) I chose to attend this class intrigued by this brief description of HC and the topics to be covered, including GWAPs, "Social Behavior Analysis," and "Participation and Ethics," (Bogner n.d.) which piqued my interest regarding questions about *how* humans and computers are combined in HC, what the image of humans and their intelligence is, as well as what the role of games was in this. From this first encounter with HC, my initial research interest focused on the human–technology relations and the role of the human in the sociotechnical systems that emerge in the "interferences" (Dippel and Fizek 2017a; 2019) of play and science.

I started exploring HC-based CS games by participating in various examples, including Stall Catchers, Foldit, and ARTigo, which I then selected as examples to study due to their visibility in online CS, and their differences in the scientific research area and modes of engaging participants. I gained my first ethnographic experiences with the ARTigo project, developed at LMU Munich by Bry and his colleagues in cooperation with the Institute of Art History at LMU Munich. Around the end of 2019 and beginning of 2020, I interviewed team members of the project from both the department of computer science and the art history institute. While getting into contact with the ARTigo team was straightforward, gaining access to ARTigo participants proved difficult. When I began my research, the project had already been ongoing for over a decade. Although the initial years were prosperous, the team's activities were reduced to platform maintenance after

the funding period ended. The software had become partially outdated by that time, and there was no longer much active research being conducted on it, as team member Emilia explained in our interview (Nov. 8, 2019). ARTigo could be played without registration,[10] and there was no communication feature for contacting other participants. I describe my attempts to acquire interview participants from ARTigo's player base below.

At around the same time, in the fall of 2019, Bry introduced me to Michelucci, the director of the Human Computation Institute. As the only institute solely focused on the development of HC systems (to the best of my knowledge, and at the time), the institute presented a good starting point for learning how GWAPs or HC-based CS systems were being developed and maintained and how human–AI relations were envisioned by scientists and developers. After an initial call and the exchange of several emails to identify possible paths of working together, I joined the institute at the end of 2019 as an intern—specifically with the title of "ethics intern," a title that was suggested by the institute. In one of my first email exchanges with Michelucci, in which we tried to determine how we could best collaborate, I had described my research interest, which, at that time, was still broad and not yet clearly defined, to lie (among other aspects) in questions of digital ethics in the sense of what dignified contributions to HC systems were and could be and how participants could gain recognition through their participation. Despite the effect that my title would at times predefine or bias the questions and topics I would be actively included in, my research partners allowed me to participate in the institute's different working contexts and to get to know all the different perspectives and tasks related to developing and maintaining HC-based CS projects.

As an intern at the Human Computation Institute, I was listed on the institute's website and my intern role was, thus, visible to participants and the general public. Later, when I became involved in community management for Stall Catchers, my initial username, which I had chosen at the beginning of my engagement with HC-based CS, was changed to my name plus the suffix "+HCI" to identify me to participants as an institute contact. After a brief introduction in the chat at the beginning of my fieldwork, a more thorough introduction, supported by the institute, followed in August 2021 in the form of a blog post describing my work with the institute and ethnographic research in general (Vepřek 2021a). The institute also sporadically reported on my research via blog posts and tweets (e.g., Santander 2022).

My contributions to the institute's projects included a variety of activities. Initially, my primary role was to assume the position of community manager for the new soon-to-be-relaunched Dream Catchers[11] project (Ramanauskaite 2020). However, my responsibilities soon evolved into a diverse set of tasks ranging from assisting Ethical Review Board (or Institutional Review Board) applications, community management of Stall Catchers, for example, by helping organize Catchathons, to contributing to load and performance testing and optimization efforts of the platform's software and infrastructure. In addition, members of the biomedical engineering laboratory identified me, at least in part, with the institute during my fieldwork, as became clear during laboratory

10 Unfortunately, no public data on user and platform statistics exist.

11 At the time of writing this work, Dream Catchers has yet to be relaunched for various reasons, including prioritization of other projects and insufficient resources.

meetings, in which open questions for the institute would sometimes be directed toward me.[12]

My work with the institute also included collaborative work on the question of how to improve or adapt ethical review to online (HC-based) CS projects. We organized, for example, a workshop at the ECSA conference 2020 that brought together researchers, practitioners, and participants to discuss this topic. Furthermore, we designed a new ethical review workflow together with Institutional Review Board operations and administrator Patricia Seymour (Vepřek, Seymour, and Michelucci 2020). I then further developed and designed this workflow in my computer science master's thesis as a *Collaborative and Adaptive Ethical Review* platform (Vepřek 2022b). Even though this collaborative research also informed my understanding of the field of HC-based CS and is an important example of how I contributed to shaping the field through my engagement, this work only plays a marginal role in the ethnographic research to which I return in the following.

Engaging in forms of experimental "worldings" (Tsing 2011; Niewöhner et al. 2016) helped me reflect on my epistemological assumptions and thinking constraints brought to the field, and opened up new insights into and possibilities for joint reflection of the entanglements of relations within HC systems. This joint reflection with members of the Human Computation Institute and the researchers happened throughout the research process. However, it also built the explicit focus of my second research visit to Ithaca in 2022, while the first visit in 2021 focused on participant observation and collecting data. In the following, I describe the two research visits and their process in detail and include my approach to the Foldit and ARTigo examples.

From the beginning of my engagement with the institute, two research visits of about three months each to Ithaca were planned. Due to the COVID-19 pandemic, which began shortly after the start of my fieldwork, plans to join the institute in Ithaca had to be canceled and remade several times. Therefore, I initially began the collaboration and fieldwork remotely, conducting most of my ethnographic research online.[13] Because of the very digital and web-based nature of large parts of the institute and the projects being studied, this did not (only) have disadvantages, as for most members of the institute, this form of engagement was also similar to their everyday experience with the Human Computation Institute. Except for annual in-person meetups, the team had been working remotely, communicating exclusively through web-based instant messaging services, email, and virtual meetings. Over the course of my research, the institute consisted of a core team of around eight members including the roles of developers, CS coordinators and strategic advisors, and a larger group of affiliated researchers. Since

12 My experience sometimes also resembled social and cultural anthropologist and migration scholar Maria Schiller's, who did research as a "research trainee" in municipal organizations of three European cities (2018, 67).

13 On remote ethnography, see, e.g., Postill (2017), and for a recent account on remote ethnography taking into account ethnographic fieldwork in times of the COVID-19 pandemic, see Podjed and Muršič (2021). Anthropologist Dan Podjed and cultural anthropologist Rajko Muršič point to the advantages remote ethnography can have, for example, by allowing interview participants to remain in their own environment, creating a more relaxed interview situation via digital communication technologies (2021, 45).

the institute is a nonprofit research organization, it is strongly dependent on funding. Accordingly, the team's size varied throughout my engagement.

During this first remote part of my fieldwork, I also conducted qualitative interviews with Stall Catchers participants and analyzed readily available data, such as the in-game chat and Stall Catchers' source code. These methods will be explained in more detail below. In addition, as the second comparative HC-based CS example, I conducted interviews with the US–based Foldit team in early 2020. For the latter, I used video conference services and followed Foldit's developments by playing the game occasionally and reading its newsletter and regular scientific and game updates shared on the website.

While remote engagement and digital methods helped me to gain valuable insights for my research, this approach also made some aspects of Stall Catchers and the Human Computation Institute's working practices invisible. A considerable part of the Stall Catchers project remained in the dark, most notably the biomedical research conducted at the Schaffer–Nishimura Lab. The Alzheimer's disease research underlying Stall Catchers, including the scientists' practices, the wet lab, experiments with mice, microscopes, lasers, dyes, medication, as well as the connected work on Stall Catchers' data pipeline, remained invisible from my screen in Germany. Moreover, and especially before the COVID-19 pandemic times, the Human Computation Institute organized various in-person events where people would, for example, come together to participate in Stall Catchers during the special Catchathon events, another aspect of the project I viewed as important for my research that required physical presence.

Therefore, as soon as was possible given the COVID-19 pandemic restrictions, my remote digital ethnographic research was complemented by two research visits to Ithaca. The first three-month field research period took place from August to October 2021, its foci being the collaboration with the institute and ethnographic data collection through participant observation and interviews.

In some ways, my collaboration with the Human Computation Institute during these three months continued as before with weekly Zoom meetings with team members but with the added benefit of in-person meetings and conversations with Michelucci several times a week, which often included or evolved into hours of informal conversations. I also had the opportunity to accompany Michelucci to meet researchers interested in collaborating to develop a new HC-based CS game to support their research, which proved insightful to better understand how HC-based CS projects are started and imagined before they are actually built.

During this three-month period, I participated in meetings and events of the biomedical engineering laboratory at Cornell University and observed the working practices of Alzheimer's disease research in the laboratory. The focus was on practices related to Stall Catchers, such as work on the data pipeline, including testing new software tools for preparing data or the manual analysis of research images. Researchers also walked me through the individual steps and related code structures. In addition to practices related to the data pipeline, I also observed and participated in onboarding meetings for new laboratory members, and observed scientific practices in Alzheimer's disease

research, such as craniotomies,[14] drug and treatment injections at night for behavioral experiments with the mice the next day, imaging sessions, and, most challenging for me personally, mouse euthanasia and tissue collection. Even though my analysis focuses on human–technology relations and the assemblage starting from the imaging sessions and excludes the preceding research practices in the laboratory, insights from observing practices beyond this focus, such as the animal research practices, were important to situate these relations and the assemblage in their broader context of biomedical Alzheimer's disease research. These participant observations were complemented by qualitative interviews with laboratory members working with Stall Catchers. Taken together, the time spent in person at the Human Computation Institute and Schaffer–Nishimura Lab were invaluable to my research, especially thanks to the patience my research collaborators showed me.

Following this first research visit, my focus shifted to transcribing and analyzing the empirical material collected. During this time, I reduced my active involvement with the institute and the frequency of conducting participant observations both at the institute and on the Stall Catchers platform, although Michelucci and I continued a biweekly meeting schedule, and I occasionally attended larger team meetings and events. One exception to this reduced engagement was during the *Danish Science Festival* and the *Engaging Citizen Science Conference* in April 2022 which took place in Aarhus, Denmark. Eleven team members of the institute came together there to participate in the festival and conference and connect (some meeting during this event in person for the first time).

My second research visit to Ithaca, NY, lasted from October to November 2022. This time, I concentrated on discussing my observations and preliminary findings with my research partners. To avoid overly formal interview situations and, instead, concentrate on exchanging ideas, I decided to omit audio recording sessions and relied on taking field notes, which I later copied into my diary. The second field phase also allowed me to learn about further developments in the laboratory on the data pipeline since my last visit in 2021. During the twice-weekly meetings with Michelucci, in addition to discussing a variety of topics related to HC, Stall Catchers, the institute, and the insights I had gained over the course of my fieldwork, we also discussed early iterations of ideas, themes, and theses that are now included in this work.[15] I also met individually with all researchers from the laboratory who had participated in my research during my first field phase to discuss my observations. These meetings were of great value for refining my insights and were essential in giving them the opportunity to share their perspectives on my preliminary findings. I draw on these conversations and reflections extensively in later chapters of this work. Toward the end of my visit, I finally gave a presentation on my work at the laboratory's weekly meeting, which was also attended by Human Computation Institute team members. The presentation of my work and subsequent discussion presented a

14 Craniotomies are surgeries in which a part of the skull is removed from the head of a mouse and windows are installed to allow subsequent imaging of blood flow in the brain.

15 In one session, for example, we discussed play/science entanglements in HC-based CS, in another, my ideas on intraverting human–technology relations, and in yet another, we reflected on our collaboration and what ethnographers could bring to the field of HC.

helpful learning process in my ethnographic work (Klausner 2015, 49) and provided valuable insights into how my research partners perceived it. It also included joint reflection on how projects such as Stall Catchers could or should be run. During these conversations and meetings with research partners, the roles shifted between me as the ethnographer, learning from the institute's members and the researchers, and my research partners approaching me with questions about how to improve Stall Catchers and the collaboration between all different actors.

In the following section, I discuss the main individual ethnographic methods I applied in the course of this study.

A Toolkit of Methods for Emerging Hybrid Systems

Participant Observation

Participant observation[16] is one of the core research methods in ethnographic fieldwork and formed the basis of my analysis to examine everyday practices[17]—both use and development practices—of HC-based CS assemblages and their human–technology relations. Participant observation made it possible to analyze the practices' tacit and embodied knowledge as well as the processes, situations, and conflicts related to them. In this way, the innovation potentials and "tactics" (Certeau [1980] 2013) of routinized practices also became analyzable (Beck 1997, 346).

Participant observation at the Human Computation Institute was crucial for gaining insights into what it means to build "sustainable participatory systems" (Human Computation Institute, n.d.). Building such systems involves a wide range of tasks and practices beyond engineering, such as conducting meetings, team communication, writing papers, giving talks, software maintenance, management (of people, processes, and systems), and fundraising. These activities are similar to the tasks performed by members of the expert system laboratories studied by Forsythe in the 1980s and 1990s ([1993] 2001h, 23–24) and are generally representative of the work conducted in (university) research laboratories. Additionally, work at the Human Computation Institute also included community outreach and translating biomedical knowledge into content for the broader public in the form of blog posts and other formats, in other words, science communication. Due to my collaboration with the institute, I was able to closely follow developments, such as the first experiments with AI bots in Stall Catchers, and to do so from the different perspectives of both institute's team members and the project's participants. I followed participants and researchers in their engagement with and the

16 Participant observation can be traced back to anthropologist Bronislaw Malinowski ([1922] 2013) and has since been further developed into one of the core research methods in ethnographic fieldwork (*cf.* DeWalt and DeWalt 2011).

17 Bareither described this focus on practices as one of the strengths of cultural anthropology of technology because of its "sensitivity for everyday, routinized, culturally encoded and in social negotiated processes integrated *practice* of actors with and in relation to technology" (2013, 31, emphasis i.o.).

relations they enter into within the HC-based CS system, as well as the data moving through and along these relations and infrastructures.

While my engagement involved participation in the institute's everyday practices, focusing on the development and maintenance practices at the institute and the different game platforms, my involvement in the laboratory was typically limited to observation, as the focus of my research was not on the wet laboratory practices but rather on human–*technology* relations directly connected to the HC-based CS project. In studying the CS platforms of the studied examples from the users' perspective, I actively participated in all projects as a participant. I recorded my experiences in a field diary to better understand the participants' experiences and relations with the platforms. Here, participating included playing the games, engaging in the project's communication channels, and reading blog and forum posts as well as other updates provided by team members and participants on the project's websites. By being "digitally co-present" (Hamm 2011), I could move within the same digital space as and together with the projects' participants.

While participating in ARTigo and Stall Catchers was relatively easy, I experienced a difficulty that many Foldit beginners face, namely the steep learning curve of the game. My participation in Foldit can be described as experimental, since after completing the tutorials, I mostly tested different approaches to folding proteins by following tips other participants had shared with me in our conversations, experimenting with different recipes, and using other algorithmic tools provided by Foldit.

Because Stall Catchers formed my primary research example, I here also contributed to all Stall Catchers events, such as Catchathons. By engaging not only as a Human Computation Institute team member and observing ethnographer, but also as a participant, I had the chance to experience the team spirit that emerges during such events, especially in the final hours of a challenge when the competition between different game teams heats up.

Participant observation was also my primary approach to including materialities, technologies, nonhuman actors, and entities, such as data, in my analysis, and specifically to focus on the relations between them and human actors. I here want to briefly point to some selected instances highlighting the importance of this approach for my investigation. Researcher–technology relations, for example, only became fully comprehensible after observing how researchers use and work on the infrastructures, software, and microscopes, how imaging data were manually analyzed, or how data were transformed and prepared via the data pipeline before being sent to Stall Catchers. Similarly, understanding how data is analyzed on the Stall Catchers platform required following the flow of data between human input, the interface and servers, code and its algorithms, as well as the database and developer's interventions.[18]

Taken together, participant observation allowed for the generation of situated knowledge, since "to observe is not to objectify; it is to attend to persons and things, to learn from them, and to follow in precept and practice. Indeed there can be no observation without participation—that is, without an intimate coupling, in perception and action,

18 The approach of following actors, commodities, or ideas is especially prominent in multi-sited ethnography (Marcus 2009).

of observer and observed" (Ingold 2000, 108; cited in 2014, 387–388).[19] Through participant observation, I gained insights into the everyday practices of the different actors, of which it was specifically valuable to analyze the interplay and different sociotechnical relations they enter. It also revealed tacit, implicit, and embodied knowledge often difficult to articulate, for example, in interviews (Beck, Niewöhner, and Sørensen 2012, 19).

However, the method of participant observation also has limitations when it comes to phenomena that include globally distributed actors, making specific participant observations impossible, or when it comes to including discursive and historical contexts that also shape such phenomena. My investigation, therefore, draws on qualitative interviews, media analysis, and exploratory methods, such as source code and in-game chat analysis, to cover these shortfalls.

Qualitative Interviews

To gain access to the ideas, perceptions, and values behind practices in line with a "practice of understanding" (Jeggle 1995, 56) typical of a cultural anthropological approach, I conducted qualitative semi-standardized interviews across all three projects. All interviews were guided by a questionnaire, which served as a flexible guide to structure the interview but did not determine the conversational flow (Schmidt-Lauber 2001). I followed the topics my interview partners cared about or brought up on their own accord, including questions in the conversation according to the specific situational flow (Schmidt-Lauber 2001, 176) rather than actively steering the conversation back to a fixed agenda. Interviews, therefore, differed in their form, length, and depth. I talked to participants, researchers, developers, project leads, community liaisons, and other team members involved in HC-based CS projects. The interviews were conducted in English, German, and Dutch.

Qualitative interviews opened the door to certain dimensions of knowledge and perceptions, which, especially in the case of CS participants, would not otherwise have been possible to include in my research. It was through the interviews, for example, that I gained a sense of the importance of a personal connection to Alzheimer's disease for many Stall Catchers participants and how deeply entangled these personal experiences were with their contribution to HC-based CS. The global distribution of the phenomena studied made it effectively impossible to observe participants' playing practices at their physical location, therefore, remote interviews were vital to covering this gap. In-person interviews were invaluable in other settings: I only understood the complexity of the Stall Catchers project at the laboratory by being there and talking with biomedical researchers. Sitting down with them, I learned how much they cared about their research, sometimes because of the same connections to Alzheimer's that the participants described, and the problems they faced. At the same time, I gained access to different layers of designers and developers' imaginations that went beyond the inscribed and programmed ones and were not accessible for me by reading code. Sociotechnical imaginar-

19 My approach could also be considered to carry autoethnographic traits insofar as I include my personal experiences, positions in the field and their reflections in my analysis; on autoethnography, see, e.g., Ploder and Stadlbauer (2013); Caivano and Naumes (2021).

ies of HC drive the design and development of such systems but, at the same time, often break down or do not hold in everyday life. Interviews help access these different layers and narrations.

It should be noted that these narrations are self-descriptions that are contextualized in the research partner's sociocultural, embodied, and material or technological environment. At the same time, they are polished external presentations adapted to the interview situation (Froschauer and Lueger 2020, 236). In my conversations with Stall Catchers participants, for example, interview partners sometimes addressed me not only as an ethnographer but also as a team member of the Human Computation Institute. At times, I may have been seen as something of a mediator who would pass on both praise and criticism. Because information is filtered and interview partners apply discursive modes of representation that do not necessarily correspond to the practical mode of being and conscience (Beck 1997, 346), qualitative interviews require a critical approach to the empirical material. The relationship between representation and observable action is complex (Geertz 1973; 1983; Forsythe 2001g, 139). Beck talks about "translation errors" that are to be expected and points to the problem that these can only be partially corrected by comparison with observations (1997, 346). Keeping these limitations in mind, qualitative interviews, particularly in combination with the other methods applied in this research, form a practical approach to HC-based CS projects.

I conducted a total of 64 interviews, of which 49 were oral interviews and 15 written interviews and questionnaires with follow-up questions. Most of these were within the Stall Catchers case study: 28 interviews were conducted with Stall Catchers participants, five with researchers of the laboratory, and five with the Human Computation Institute team members. Thirteen interviews were conducted with Foldit participants[20] and six with Foldit team members and researchers. Finally, two ARTigo participants and five ARTigo team members were interviewed, as well as three representatives of funding institutions in the US that fund or have funded some of the projects. Given the scope of this study, the insights from the latter serve as contextual background information.

Although I had initially planned to conduct a comparable number of interviews with participants for each case study, it proved difficult to find and get in contact with ARTigo participants. Even with the support of the ARTigo team, who allowed me to post calls for participation via ARTigo's Twitter and Facebook accounts, it was difficult to acquire participants. The call was answered by one ARTigo participant. Another participant interview was acquired via the snowball principle. An additional problem was that it was not possible to work directly with the ARTigo platform during these interviews in 2020 because the server was down at the time, which limited the interactivity in some ways. Furthermore, the platform does not include any in-game communication features (such as a forum or chat), making it challenging to analyze participants' experiences and practices by other means. A small written survey in which participants were asked to play ARTigo and share their experiences was conducted in January 2021 with fellow students

20 While 28 Stall Catchers participants immediately answered an open call for participation in my research via email, recruitment of interview partners from Foldit participants was slow. In addition to posting a call in the game's forum, I actively sent more than 50 interview requests to individual participants via the game's platform and Discord server.

as part of an explorative study for a computer science master's seminar at LMU Munich. We created a small qualitative questionnaire which we shared in our networks to compare the user experience of different games and because there was no data on the participants' perspective on ARTigo.[21] Due to these circumstances and the resulting limited empirical material, ARTigo was not included as a comparative study to the same level as Stall Catchers or Foldit in my final analysis. However, a new iteration of the ARTigo platform was launched during the later course of my research introducing a new stage of the project and an interesting turn regarding my analysis of how HC-based CS are formed and change over time. I discuss this development in Chapter 6, less empirically than analytically, applying my theoretical concept of intraversions to this development and exploring its explanatory potential using this example.

All but one interview with participants were conducted via video conference services, such as Zoom or Skype, or via a landline due to the geographic distance and the COVID-19 pandemic. Although in-person interviews allow for a richer communication and interaction context, the digital and phone interviews were particularly suitable for this research since they allowed CS participants from all over the world to contribute and even increased the possibilities of participation (Markham 2005, 801; *cf.* Hengartner [2001] 2007, 201).

The majority of participants interviewed were based in the US, followed by five interview participants based in Germany, and three in Belgium. In addition, one interview partner contributed from each of the following countries: Australia, Brazil, the Netherlands, Nigeria, and Singapore. The location of one participant was unknown.[22] The average length of interviews was about one hour, with the shortest being 21 minutes and the longest being two hours and 16 minutes. It is worth noting that interviews with participants of the case studies analyzed are necessarily limited to those who actively volunteered to contribute to my research by responding to my call for interview participants. Hence, the viewpoints represented are those of participants that were indeed willing to share their experiences with Stall Catchers, Foldit, or ARTigo and the motivations that drive them.

The interviews with team members, researchers, and funding institution representatives were mostly conducted via video conferencing services, except for conversations with ARTigo team members and those with the researchers at the Schaffer–Nishimura Lab.[23] Even though I spent three months with the Human Computation Institute in

21 A total of ten participants in the age range between 20 and 64 contributed to our written questionnaire. Even though we did not collect personal information in addition to age, it is very likely that most participants had a computer science background and were based in Germany. Five participants had not heard of ARTigo previously, three had heard of it but never played it, and two had played it before our small study.

22 It should be noted that I left it open to my interview partners to indicate their age and gender, as these categories did not form a focus of my research. The age range generally included individuals between around 20 and 75, with a slight majority of participants being over 40. Slightly more participants identified as women across the three different case studies.

23 As in the case of interviews with participants, gender and age distributions are not considered in my analysis. However, I would estimate that the distribution of different genders was roughly representational of the general population among researchers and team members of Stall Catchers

Ithaca, NY, which allowed many in-person conversations, most interviews with its team members were also conducted virtually, as the institute's team is distributed around the world.

All interviews were audio recorded and transcribed. The quotes used in this work were minimally smoothed to improve readability, on the one hand, and protect the privacy of my interview partners, on the other hand. This means that expressions such as *um*, *uh*, *you know*, and *like* as well as word repetitions, if not purposefully repeated, were omitted from the quotes. These expressions can transport important meanings to which cultural anthropology is sensitive, similar to hesitations or pauses while speaking. However, since not all interview partners were native speakers, the usage of such expressions and terms varied across all interviews, potentially resulting in identifiable speech patterns. This was especially noticeable as I translated the non-English interview quotes used in this work.

Another strategy I employed to protect the privacy of research participants is the use of pseudonyms and random gender changes. Pseudonyms were chosen arbitrarily, with the best effort to maintain the overall representation of origin and diversity of real names across all data. Furthermore, due to some roles at the Human Computation Institute and the laboratory being unique, I partly generalized roles—such as "researcher" for different positions including experimentalists, PhD students, postdocs, and developers—and duplicated positions that are unique to an individual to multiple representations of the same research partner in order to make deanonymization more difficult. These obfuscations do not affect the overall results of my research, and I specify the position in cases where the role is relevant. For these reasons, only a few of my research partners' names, whom I refer to by their last names, were not pseudonymized.[24] These exceptions were necessary due to their unique field position and public appearances. Anonymization could, therefore, not be guaranteed. Consent for this was granted by all individuals affected.

Finally, I jointly refer to users, players, and citizen scientists as "participants" since some participants reject the term "game" and do not identify as "players." I also prefer "participants" over "users" and "citizen scientists" to emphasize their active role in shaping the HC-based CS systems.

In addition to the qualitative interviews, informal conversations with Human Computation Institute team members, researchers from the laboratory, and a few participants were of great value to my research. In these cases, I took field notes from which I quote in this work when conversation partners agreed.

While qualitative interviews, informal conversations, and participant observation formed the core methods of my research, it also involved the analysis of existing data, such as external perspectives on the projects studied (e.g., in media articles), self-presentations of the project's teams, and, most importantly, the Stall Catchers source code and

---

contributing to my research. For the research partners of Foldit and ARTigo, the majority seemed predominantly to align with male.

24 The decision to refer to research partners whose names I have not pseudonymized by their last names while using first names as pseudonyms for others is purely pragmatic and not meant to imply any hierarchical distinction between them.

data pipeline as socio-technological foundation and constituent entities of the project. The methods used in connection with these are described in the following. As will become clear, human–technology relations and the sociomaterial assemblage only became understandable in the combination of the different methods which revealed the entanglements of the different human and nonhuman actors.

Experimental Approaches to Infrastructures, Code, and Digital Chat Data

In the course of this research, I analyzed a variety of textual material. This included media articles about the different HC-based CS examples, blog posts, and the case studies' own websites and forums. The websites' content and project descriptions were valuable sources, revealing how the teams wanted to represent themselves to the world. In the course of my collaboration with the Human Computation Institute and due the openness of its members toward my ethnographic approach, I also received access to the institute's workspaces and digital infrastructures, such as its primary communication space on Slack (Slack Technologies, LLC n.d.), the Github repository,[25] databases, Stall Catchers' admin spaces, and, where required for certain collaborative work, computational infrastructure. The analysis of these text-based and infrastructural sources followed an exploratory approach with the aim of including them as supplementary material to support or contrast with other observations and empirical materials.

Particularly crucial was the access to computer code and digital chat data from the institute. I focused on analyzing the code of Stall Catchers, which is a key component in HC-based CS assemblages (Mackenzie 2006, 2) to understand how user–technology relations unfold. Through the analysis of code, the intra-actions of participants and technology can be traced and the underlying and (consciously or unconsciously) inscribed design logics of developers and designers revealed (Koch 2017b, 117).

The analysis of computer code is a relatively new area in social sciences and humanities, with emerging subdisciplines like software studies, digital STS, critical data studies, and CCS (e.g., Fuller 2008; Vertesi and Ribes 2019; Marino 2020; Hepp, Jarke, and Kramp 2022).[26]

Code is multidimensional and, as such, requires a multiperspectival analysis to fully appreciate its sociotechnical embeddedness, its becoming and implications, as well as its

25 *Github* (GitHub, Inc. n.d.) is a hosting service for version control using the distributed version control system *Git* and software development, which allows the distributed development of software in teams.

26 Today, ethnographic studies including computer code often refer to these fields, which provide a rich repository of useful and important methodical approaches. However, depending on their scientific situatedness, they follow specific research interests. The CCS, for example, are strongly influenced by and emerged from the field of literary studies. By comparison and despite the expressed need for it, an ethnographic approach that focuses on the practices and meaning-making processes, for example, has not yet been established (e.g., Carlson et al. 2021; Vepřek et al. 2023). In order to change this, the *Code Ethnography Collective*, a group of researchers from mainly cultural anthropology and STS has met regularly since 2021 to discuss ethnographic approaches to computer code (Code Ethnography Collective n.d.).

inscriptions and the various practices associated with it (Vepřek et al. 2023). In my ethnographic research, code analysis played an important role in qualitative interviews and walkthrough sessions (Light, Burgess, and Duguay 2017).[27] Additionally, it was also crucial in participant observations during special events and laboratory sessions. This also extended to my collaborative engagement tasks, such as performance and load testing of the platform. The different approaches, thereby, supported each other because insights from participant observation revealed interesting starting points for focused code analysis. My focused code analysis approach (first discussed in Carlson et al. 2021; Vepřek et al. 2023) builds on the CCS method (Marino 2016; 2020) of reading code in a critical way. That is, as Marino writes,

> to explore the significance of the specific symbolic structures of the code and their effects over time if and when they are executed (after being compiled, if necessary), within the cultural moment of their development and deployment. To read code in this way, one must establish its context and its functioning and then examine its symbols, structures, and processes, particularly the changes in state over the time of its execution. (Marino 2020, 23)

Unlike CCS, however, my research used code analysis as an additional perspective, particularly focusing on the *flow of code in practice*, which is why I call it *focused* code analysis. I selectively analyzed code sections related to algorithms and I/O (input/output) operations. Accordingly, I traced selected I/O operations or sequences of function calls, which I had previously identified as interesting starting points. This way, I followed the flow of the code during its run-time operation, entangled with and in relation to other elements of the sociotechnical assemblage. I included specific code sections in MAXQDA and annotated them similarly to other empirical material, such as interview transcripts, to analyze selected code blocks.

By integrating this analysis with other ethnographic methods, I traced the flow of actions of human–technology relations within and beyond the text. This approach follows an understanding of software "structured as a distribution of agency" (Mackenzie 2006, 19).

Due to the dynamic nature of code, it was only through the combination of the different approaches and methods described in this chapter that I was ultimately able to gain an in-depth understanding of the participant–technology relations as they unfolded in the sociotechnical assemblage. This combination also revealed the different development, maintenance, and use practices and how values, norms, and imaginations of future hybrid human–AI systems were inscribed into the Stall Catchers project and guided its implementation. Since Stall Catchers' source code is proprietary, I do not include actual code samples in this work, except for one instance which the Human Computation Institute kindly granted permission to include. However, whenever necessary to make an argument or provide an example, I describe insights from analyzing the code. An example of this can be found in Chapter 7, which focuses on trust. In this

27 I also included the code underlying the data pipeline and ML model at the Schaffer–Nishimura Lab in the form of walkthroughs with researchers and readings on the laboratory's Github repository.

context, integrating code into my analysis enabled me to comprehend how algorithmic mechanisms contribute to preventing cheating and ensuring data quality.

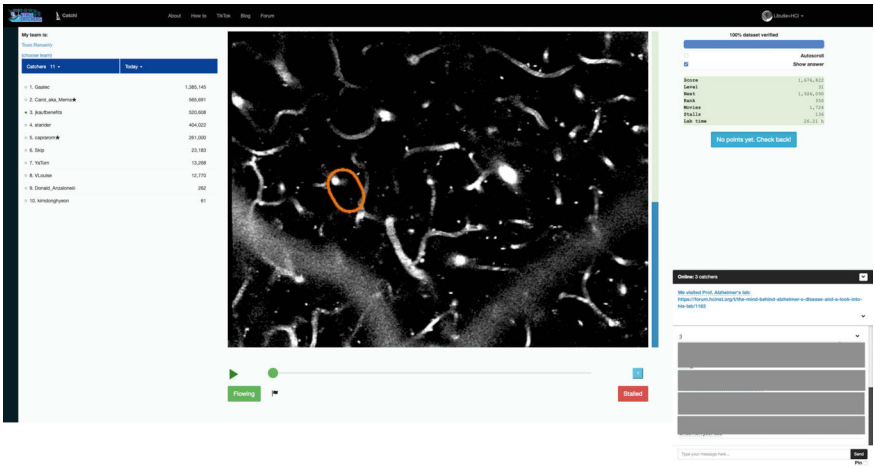
Finally, as a second exploratory method used in this work, the analysis of Stall Catchers' in-game chat data during my fieldwork in 2020, particularly in the midst of the COVID-19 pandemic, provided valuable insights into the project and participants' perspectives. Yet, it required a more exploratory approach because, despite advances in digital text data analysis in the digital humanities and computational social sciences (Lemke and Wiedemann 2016; Franken 2023), qualitative analysis of large datasets like chat records remains challenging for qualitative and inductive approaches. While methodological considerations exist for qualitative approaches to the analysis of written online conversations (e.g., Schirmer, Sander, and Wenninger 2015; Nam 2019), the question of how to make large amounts of data manageable for qualitative analysis remains generally little discussed. Conducting a manual analysis of the entire chat record following the grounded theory approach proved to be arduous, if not infeasible, due to the large amount of data involved, particularly considering its supplementary role in the overall scope of my research. The chat record in 2020 contained around 17,000 messages, spanning a period of around one and a half years. To manage this, I explored a new approach using relational databases[28] for (re)organizing and analyzing the extensive chat data, making it more accessible for qualitative study.

The Stall Catchers' chat, located in the lower right corner of the game's interface (see Figure 1), forms a fundamental aspect of Stall Catchers in that it is the primary method of direct communication between participants. Conversations ranged from play practices and Alzheimer's disease research to questions on game functionalities that experienced participants answered. At the same time, the chat provided a direct communication channel for the participants to the Stall Catchers team and vice versa. Even though not all participants actively engaged in the chat, it was important for the motivation and

28 Digital databases are collections of electronically stored information that can be maintained, accessed, and updated via a database management system, which functions as an interface between users or programs and the database. However, databases are more than mere information repositories; they are a consistently organized set of data whose informational patterns allow us to ask various questions—so-called queries—to the data (Quamen and Bath 2016). Among the different types of databases, relational databases are currently the most common form. They built upon a relational model (Schubert 2007, 35). This means that data are stored as so-called relations, which are typically represented as tables with rows (describing an object, such as a chat message) and columns (describing attributes or characteristics of a chat message, such as the sender's name). In these tables, each row describes an object, such as a message in the case of the chat analysis. Each column describes an attribute of the object, such as the sender id or the content of a message. In relational databases, object attributes can reference other objects by referencing their identifiers (typically the so-called primary key) in one of their columns. It is important to mention that databases are not neutral storage media but provide a specific perspective on data, as has been shown by sociologist Christine Hine (2006). Storing data in a structured database keeps it more manageable than simply using spreadsheets or other unstructured documents. The main advantage becomes apparent when it comes to analyzing and transforming data: with a relational database, one can interact with the data in complex and highly specific ways, which allows the answering of certain questions much more easily than a manual or spreadsheet-based approach could. For relational databases, this is commonly done via SQL.

contribution of some, as participants used it for mutual encouragement. Through chat analysis, I uncovered themes that were not mentioned in interviews, such as the playful reinterpretations of research data as artworks (see Chapter 5). Moreover, the analysis of chat messages provided valuable indirect and exploratory access to the field.

Figure 1: Stall Catchers' main UI



Source: Screenshot taken by LHV on Mar. 24, 2024 (https://stallcatchers.com/virtualMicroscope)

In another work (Vepřek 2023a), I provide a comprehensive discussion of my methodology, including a discussion of the chat format's particularities and addressing the ethical considerations and challenges involved in analyzing digital chat data. A brief summary is provided in the following for the purpose of this section. The process involved the four steps of data acquisition, cleaning, structuring into a relational database, and restructuring for analysis with SQL queries. My approach aimed to conducting a content-based analysis that focuses on the narration of play practices and meanings of Stall Catchers. I extracted sets of conversational contexts involving a specific player in order to gain a deeper understanding of the perspective of some of these individual Stall Catchers players via the chat. The extracts included entire conversations between participants with only minimal interspersed sections where other users engaged in the chat without actively contributing to the conversation of interest. I maintained the conversation dynamics by including a window of messages before and after each message of the user in focus. This way, I changed the superficially linear structure of the data into separate contextually linear excerpts. Occasionally, I revisited the data to isolate topical fragments of discussion or perform other small supporting analyses.[29] In the main analysis, these

29 The approach could also be extended to specifically isolate (or exclude) contexts where specific words or phrases were mentioned, making it easier to drill down on specific topics of conversation (in general, or once again involving a particular user). Since the results of a query are presented in virtual tables that can be exported and the queries do not operate on or change the dataset itself, it is possible to return to the dataset at any time and query it with different foci. Such flexibility

extracted conversations were exported into a format suitable for loading into MAXQDA for manual analysis, where they presented an access point to participants' perspectives during gameplay.

Chat analysis, despite its potential for ethnographic research, also has some limitations and challenges that are important to consider. These include the lack (or subtlety) of nonverbal elements in written conversational data, the question of how representative chat data are for the overall participant base, and significant questions of privacy (for a more detailed discussion see Vepřek 2023a).[30] I chose in this work to anonymize the chat data and combine several player identities into one. I also refrained from directly linking any personal interview data to the chat data or quoting directly from the chat in this work. Despite these limitations, if contextualized and included in a multi-perspectival and -methodological ethnographic approach, I expect this approach of restructuring and querying digital textual data with relational databases to have potential for qualitative ethnographic research beyond the specific analysis presented here. It is conceivable to apply this method to other textual data, such as tweets or comment threads, or to design queries that focus not on the textual content but on metadata, such as timestamps, to learn more about the temporal flows of chat communication, or other aspects and phenomena.

After having discussed the methodological foundation and its operationalization of my research, I now turn to the first empirical chapter, in which I analyze the imaginaries behind HC that guide and shape the development of HC-based CS.

is particularly helpful for ethnographic and inductive research, where fieldwork phases alternate with analysis and reflection and do not follow a linear scheme.

30 The Stall Catchers chat is semi-public; it is accessible to all registered Stall Catchers participants. Because registration only takes the provision of an email address and accepting the terms and conditions, the chat can be considered almost public. Participants are informed about this in the terms and conditions.