8 Conclusions

> Hybridization […] is not to be rejected, neither is it the greatest danger, but it does deserve the greatest care. (Dorrestijn 2012b, 240)

In 2021, for the first time in the history of Stall Catchers, an AI bot named GAIA was included alongside human participants in the HC-based CS game to analyze Alzheimer's disease research data. This first encounter between humans and AI bots in Stall Catchers presented at the beginning of this book initially appeared as a snapshot, independent of Stall Catchers' historical becoming and its future directions. Yet, the encounter merely represented one stage in the evolution of the Stall Catchers' assemblage and its continuously intraverting participant–technology relations. This human–AI bot encounter was first shaped by the Human Computation Institute's imaginaries of HC, the design of Stall Catchers as it was built in part upon the Stardust@home project, and the laboratory's practices for manually analyzing Alzheimer's disease research data. Importantly, the encounter was also formed by preexisting participant–technology relations in Stall Catchers, including the data annotation performed by human participants, while other algorithms, in turn, review and adjust for individual participants' skill levels. Human participants' annotations were then used to train ML algorithms, forming the basis for the AI bots' inclusion as artificial participants on the platform. At the same time, the HC system's intrinsic play/science entanglements created the encounter as a fun, yet serious situation, introducing competition among human and nonhuman participants, albeit mitigated through the common pursuit of advancing Alzheimer's disease research. However, the encounter was not solely defined by historical developments, path dependencies, and previous relations. Instead, with AI bots on the leaderboard, human participants actively related to the new configurations using tactics such as redeeming points and forming human alliances *against* the bots. Simultaneously, participants also shaped and created the possibility for new human–technology relations by seizing specific algorithmic affordances not actually intended by the system's design. Through their motivations and experiences that brought them to Stall Catchers, participants ascribed new meanings to the project while rejecting specific inscriptions, such as that of referring to Stall Catchers as a game. With all of this in mind, that initial encounter presents only one

key moment in Stall Catchers' long chain of intraversions, one which specifically marks the emergence of a new dynamic in participant–technology relations. This new dynamic resulted in direct productive cooperation between humans and AI in pursuit of the CS project's overarching goal of bringing about a cure for Alzheimer's disease.

Although Stall Catchers and the Human Computation Institute served as the primary focus of my research, specific aspects regarding how human–technology relations in HC-based CS intravert only become clear when considering the additional case studies of Foldit and ARTigo.

More precisely, the analysis of Foldit showed particularly well how human–technology relations in HC-based CS continue to intravert over long time periods—over almost 20 years in this case. Beginning with the distributed computing project Rosetta@Home, Foldit underwent significant cycles of intraversions from its inception. Born from an idea regarding how to improve an automated system through the introduction of human creativity, the project was subsequently reconfigured to allow humans to create and deploy customized, automated tools. Foldit navigated through these major changes without ever becoming obsolete, even, and perhaps most notably, in the face of stark advances in AI capabilities, such as AlphaFold's solution to the problem of protein structure prediction. Instead of rendering it obsolete, AlphaFold's capabilities were soon subsumed into Foldit in the form of an additional tool, with Foldit's own purpose shifting from advancing protein structure prediction to, instead, furthering progress on protein design, as well as other unsolved problems. Notably, the most impactful intraversions in the case of Foldit often emerged out of the engagement of the participants themselves, rather than being introduced only by the system's developers. This provided a contrast between Foldit and Stall Catchers (where more, but not all, of the innovations were introduced by the Human Computation Institute team).

The analysis of ARTigo specifically highlighted how the overall purpose of HC-based CS assemblages transformed with its intraverting human–technology relations and by what means these changes unfold. By concentrating on the larger scale of the overall system's life cycle, I demonstrated how the platform required reimplementation and how participant–technology relations needed reconfiguration to remain both engaging for participants and at the edge of technological development. ARTigo persisted not only through the intraversions in its relations and alongside advancements in computer vision toward new goals. To an extent, it could be said that ARTigo *had to* undergo changes due to secondary factors, such as continued active participant engagement and motivation along with gradual personnel changes to the team developing and maintaining the system.

Various processes of territorialization and deterritorialization shape the complex interplay of different human and nonhuman actors in HC assemblages. As an example, I discussed trust as an ongoing territorialization process. Instead of understanding trust as a cognitive category, it must be considered a sociomaterial practice, since trust is negotiated in the intra-actions of humans and technologies. Given the changing nature of HC-based systems, new questions of trust emerge while disrupting existing trust-establishing practices. Ultimately, possibilities for new intraversions emerge at the horizon of human–technology relations in HC. Towards the conclusion of this study, for instance, the next intraversion of humans–AI relations in Stall Catchers remained open. However,

it seems clear that even if AI eventually achieves the accuracy required to fully fill in for human participants in the analysis of research data, human participation in Stall Catchers is unlikely to become obsolete. New ideas regarding how to include human participants in other steps in Alzheimer's disease research already exist at the Human Computation Institute and in the biomedical laboratory.

As Gray and Suri (2019) argue, newly identified problems will continue to arise with the advancement of AI and the automation of increasing numbers of tasks: "Thus, there is an ever-moving frontier between what machines can and can't solve. We call this the paradox of automation's last mile: as machines progress, the opportunity to automate something else appears on the horizon" (Gray and Suri 2019, 175). The paradox lies in the continuous repetition of this process—taking over tasks performed by humans and simultaneously creating new ones humans can solve. "In other worlds, as machines solve more and more problems, we continue to identify needs for augmenting rather than replacing human effort" (Gray and Suri 2019, 175–176). It is clear for the Human Computation Institute that as ML models become more accurate in analyzing Stall Catchers data, the future involvement of humans must necessarily look differently than that of AI bots. If the intraversions of Foldit prove instructive here, this may result in human engagements moving to a higher level of abstraction and creativity, for example, by creatively combining different AI capabilities or solving a new, higher-order problem. As the frontier of AI shifts along with technological advancements, the system might even shift its purpose altogether as new projects become possible, addressing problems which cannot yet be solved in fully automated ways. In this way, HC systems can effectively remain at the edge of AI in pursuit of ever-evolving goals and capabilities.

This study aimed to investigate the formation of HC-based CS assemblages in the interplay between human and nonhuman actors. To do so, I analyzed how HC advocates and designers imagine them as new forms of HI, building a counter-imaginary to AGI and drawing their own understanding of how human–technology relations *ought* to evolve to create desirable futures. At the same time, however, these imaginaries are, nevertheless, rooted in shared AI paradigms. Imaginaries materialize in infrastructuring practices, which, in turn, shape HC designers' visions of the future. However, once other actors (including participants and materialities that bring with them the potential to fail) join and engage within HC, these systems are negotiated in practice through the interferences of play and science, the motivating role of the system's overarching purpose, and what human–technology relations in these systems *should* look like, just as they are shaped by different territorialization and deterritorialization processes. These HC-based CS assemblages, as I have shown, are temporally consistent but, nevertheless, fluctuating compositions of heterogeneous entangled human–technology relations. In turn, these relations are reconfigured within this complex interplay as they continuously transform through future-thinking, everyday adaptations, failures, and intentional redesigns. I introduced the concept of intraversions to analyze how these relations unfold and transform, and how subject/object positions, tasks, responsibilities, and, thus, power are distributed across these relations and along these developments. The concept of intraversions emerged from my co-laborative ethnographic observations and experiences in the field and is particularly suited for the investigation of HC systems developed as transitory, always open, and, in fact, intended, to change again. The transformations of hu-

man–technology relations in HC(-based CS) assemblages—that is, intraversions—can be described as forward circular movements that unfold along temporal developments, which are both instantaneous and gradual. As an analytical tool, intraversions allow one to account for both of these dimensions and combine them in the analysis. The concept helps to trace *how* the frontier of automation described by Gray and Suri (2019) moves by focusing on HC's continuously changing human–technology relations along their whole "life span" and beyond, as new relations and even entire sociotechnical systems emerge from existing ones, resembling assemblages that, as they extend, develop into new assemblages (Deleuze and Guattari 2013, 7).

Engaging at the Edge of Artificial Intelligence

I aimed in this work to not only contribute to the body of knowledge of digital anthropology and STS. It is also my hope that the results of this study and my engagement in the field, collaborative from the beginning, may serve as an example of how ethnographers can play active roles in informing and shaping the design, implementation, and maintenance of emerging technologies through constructively accompanying their everyday practices. To that end, in what follows, I highlight three examples of how this research can contribute to the development of HC systems, followed by more general arguments for why and how digital anthropologists and STS researchers should actively engage in the development of such technologies.

First, the analysis of HC imaginaries demonstrated that the *human* in the loop does not strictly refer to an embodied, and socially and culturally embedded human being, but rather to an information-processing unit understood primarily as part of a crowd, given motivation via the context of a game. My investigation of the becoming of assemblages demonstrated how cognition—just like agency, power, and responsibility—is distributed across human–technology relations and cannot be attributed to human actors alone. Furthermore, human actors cannot be reduced to their "intelligence" or "thinking," given that human experience relies on *being in the world*, which is always embodied and affective. Additionally, I discussed the importance of extending "the loop" when referring to the humans in the loop to include developers, researchers, and, depending on the specific example, other actors involved. If the goal of HC for human–AI hybridization is to form the future of AI, this analysis presents an attempt and further invitation to reflect on the understandings and imaginations of the human in the loop and the interplay of the humans and the technologies behind HC systems. As Dorrestijn describes, hybridization should be given the greatest care (2012b, 240).

Second, while the concept of intraversions primarily serves as analytical tool, it can, nevertheless, also function as a reflective tool aiding in the development of HC by providing an understanding of how humans and technology are configured and how these relations evolve over time. Such configurations and relations depend on existing relations, path dependencies, different actors' intentions, and various other forces such as serendipity acting upon, through, and with them. The concept of intraversions helps us to see how sociotechnical systems and their human–technology relations become what they are, suggesting that, while the specific potential paths along which they develop may

vary significantly, they are, indeed, likely to change again in very particular ways in the future, possibly offering instructive insight for their design. It may, thus, also help the creators of these systems reflect upon and reconsider whether the path currently followed is desirable for the long-term evolution of the system.

Third, assemblage thinking can contribute to an understanding along the lines of what Bateson described in her foreword to the *Handbook of Human Computation* (2013). She argues that HC, with its crowd-based approach integrating a diversity of human and nonhuman actors, could contribute to a shift in "attitudes away from the fetish of individual autonomy and [teach] us, by implication, to recognize that we are connected parts of a larger whole, this is a goal to be pursued" (Bateson 2013, vii). Looking at HC systems as assemblages brings forth the various human and nonhuman elements and relations involved in the formation of these as well as the continuous forces and processes acting upon them, bringing them together and tearing them apart. In this context, intraversions, as a reflective tool, make the complex interplay and intra-actions between humans and technology in HC assemblages visible. Via this perspective, humans are not independent actors autonomously interacting with other humans and technology but a part of and situated within larger entanglements.

Cultural anthropological research, with its focus on everyday life and historical becoming, can connect past, present, and future. As Suchman (2007b) argues, plans do not unfold as linear sequences of action but, in practice, unfold rather unpredictably, given that they are always situated in sociocultural contexts. Ethnographic analysis can help us understand where AI systems come from, bringing together the different perspectives involved, and, thereby, pointing to potential futures. Cultural and digital anthropology can contribute by critically accompanying and contextualizing the developments of emerging technologies to steer them in directions preferable to all of the actors involved.

To do so, ethnographic analysis must be embedded in the development practice of emerging technologies (such as HC) itself. Collaborative approaches are, of course, well-established practice in ethnographic research. However, with a few (yet notable) exceptions, such as works by Suchman (2007b; 2021), Forsythe (2001f), and Pink (e.g., 2023),[1] participatory ethnographic research into the development of computational systems remains lacking, especially in the domain of AI research. Given how AI is becoming a greater presence in our everyday lives, and not only in the public discourse, I think it is crucial for digital anthropologists and STS researchers to not only accompany software development as advisors on cultural or ethical questions but to also actively engage in the very processes of creating socio-computational systems. Digital anthropologists and STS researchers must, thus, contribute to shaping the future. As Pink argues, "We must become players in the same futures-focused space as other stakeholders in the future

1 Additionally, two ongoing research projects should be mentioned here: Kinder-Kurlanda and colleague's ongoing work within NoBIAS: Artificial Intelligence without Bias (2020–2024), a Marie Skłodowska-Curie Innovative Training Network (ITN) (NoBIAS n.d.) and cultural anthropologist Sarah Thanner's collaborative research within the VIGITIA project (Physical-Digital Affordances Group University of Regensburg n.d.).

of emerging technologies, create new collaborations and bring different, diverse and everyday stories to the centre" (Pink 2023, 11).

The extent of domain-specific knowledge or training required for such work depends on the specific context and research field, but I argue that it does not always necessarily involve learning how to code, as often suggested. Instead, it can be much more important to understand a broader range of underlying conceptual levels more generally. In fields adjacent to AI, these may include concepts such as computation, the architecture and functioning of computers, the basic and emergent behavior of algorithms, or the cognitive ideas underlying the implementation of AI models. In other words, ethnographic research on computer science and AI cannot and should not treat the technical content of these fields as primarily about understanding and producing source code. Beck (2012) argues in his article on social neuroscience that anthropology must recognize that cognitive science eludes simple characterizations; instead, it combines various scientific fields, theoretical approaches, and methods. As a result, anthropology should join the various endeavors "that all somewhat fumble in the dark to better understand the distributed, self-organizing 'systematics' of cognitive phenomena" (Beck 2012, 114) and contribute to an ever-refining discourse:

> The unwanted alternative would be that an undifferentiated critique might deepen the split between the natural and social sciences and, even more important, might prevent possibly productive challenges that the new findings and modes of explanation in the neurosciences might have for established modes of thought and research practices cultivated by the social sciences and the humanities (Beck 2012, 114).

By adopting Beck's viewpoint and applying it to the field of AI, cultural anthropology ought not to confine AI to a singular discipline, theoretical concept, or method, and, consequently, should refrain from evaluating and criticizing it based on such limited perspectives. Instead, it should examine specific fields of research and applications and attempt to develop an understanding of both general and case-specific underlying concepts and assumptions. This will allow the discipline to engage in critical yet constructive ways (Suchman 2021, 70–71) in these fields to explore and contribute to shaping how we want to live and organize our everyday lives with AI.

Beyond the Edge

In this work, I analyzed HC as a phenomenon *at the edge* of AI, focusing on the specific and emerging branch of HC, more specifically considering its applications in the field of CS. I chose such a concrete focus on a specific area of AI because I believe grasping it as a whole is unfeasible and perhaps even counterproductive, given the diverse understandings, missions, and persistent lack of a cohesive definition for "intelligence," be it human or artificial. I think it is important for researchers in STS (and other areas) to broaden our focus beyond the imaginaries of and efforts to create "strong AI" or AGI but to also examine other existing and emerging approaches in the field, which increasingly inform and contribute to both developments and discourse in and around AI. The analysis of con-

crete fields of practice allows one to gain an in-depth understanding of the paradigms, visions, and intentions of the various actors, as well as the meanings and failings shaping the fields. Such an understanding is necessary to actively intervene, an undertaking which may be gaining in importance today given the vertigo-inducing (Dippel 2021) pace with which developments in AI are increasingly impacting our everyday lives. My research represents "a study of a culture in the making" (Turkle 2005b, 23). Much like in Turkle's investigation of computer culture in the late 1970s and 1980s, my research subject formed a moving (perhaps even accelerating) target. Today, the speed of developments continues at unprecedented levels, pushing the edge of AI further.

To conclude I turn *to the edge* of AI itself and situate my work in the broader context of AI research and the field's advancements. Taking a broader view, the edge of AI in recent years has been marked by significant advances driven largely by the emergence of highly powerful "foundation models" (Bommasani et al. 2022). This is perhaps most notable in OpenAI's advances in the domain of large language models with GPT-3 and, subsequently, GPT-4 (as well as ChatGPT, OpenAI's GPT-powered AI chatbot that popularized these developments in late 2022). These models significantly exhibit advanced capabilities in text comprehension, reasoning, problem-solving, and a host of other tasks, prompting some researchers to describe them as featuring first "Sparks of Artificial General Intelligence" (Bubeck et al. 2023), a characterization that invokes new imaginaries as well as reviving old ones. These advancements have reignited and accelerated debates about the automation and obsolescence of human labor, previously centered around menial tasks but now broadening their scope to include advanced domains of knowledge work, extending into medicine, finance, and even software engineering itself. While it is too soon to determine how these developments will unfold and what impact they might have, thus far, the deployment of these new AI systems has been marked primarily by what must be described as HI: "AI assistants" and "co-pilots" have been developed for myriad domains, including customer support, marketing, insurance, medicine, and software engineering, all with the promise of helping professionals to do *more*, rather than to help them do *nothing*.[2] In addition to such purpose-built AI assistants, another theme pointing in the direction of increasing human–AI hybridization can be found in the development of software frameworks that actively make affordances for building tools combining AI and human capabilities. One example of this is found in LangChain, an early but now highly popular Python framework for "building applications with LLMs through composability" (Chase 2022). One of the types of "agents" available in this framework is in fact *humans*—described as useful in helping other AI agents: "[h]uman[s] are AGI so they can certainly be used as a tool to help out [an] AI agent when it is confused" (Chase 2023), states the framework's documentation, introducing even newer and further configurations of human–AI relations in which humans and AI both take diverse subjective positions, sometimes simultaneously.

2 For specific examples, it is instructive to review the US-based technology startup accelerator Y Combinator's list of startup companies in its winter 2023 cohort (n.d.). A cursory glance suggests that the single biggest unifying theme is AI, typically represented as either domain-specific AI-powered software, functionally specific AI assistants or tools for the development of or with AI.

Today, it seems clear that some—perhaps many—specific currently existing jobs and tasks humans perform across a multitude of domains can indeed be automated using AI. While this book offered a deep dive into only a few specific domains, it is also clear that the primary theme has remained one of *enhancement*, not *replacement*, and my analysis suggests that humans will continue to play a role in these systems, turning to ever-newer tasks and goals to pursue. More precisely, the question of the "level of involvement" of humans or AI may be insignificant, if instead we think of these systems as truly hybrid human–AI assemblages. To that end, I offer intraversions as an analytical tool to trace and analyze how the relations between humans and technology unfold and transform alongside continued developments and progress in AI, and to engage constructively and with the greatest care with human–AI systems in the making.