Producing Humans: An Anthropology of Social and Cognitive Robots

Louise Veling M.A. M.Sc.

A thesis presented in partial fulfilment of the requirements for the degree of Doctor of Philosophy in Anthropology

Maynooth University

Department of Anthropology

April 2022

Head of Department: Professor Hana Cervinkova

Supervisors: Dr. Mark Maguire, Dr. Steve Coleman

Table of Contents

List of Figures

Summary

In this thesis, I ask how the human is produced in robotics research, focussing specifically on the work that is done to create humanoid robots that exhibit social and intelligent behaviour. Robots, like other technologies, are often presented as the result of the systematic application of progressive scientific knowledge over time, and thus emerging as inevitable, ahistorical, and a-territorial entities. However, as we shall see, the robot's existence as a recognisable whole, as well as the various ways in which researchers attempt to shape, animate and imbue it 'human-like' qualities, is in fact the result of specific events, in specific geographical and cultural locations. Through an ethnographic investigation of the sites in which robotics research takes place, I describe and analyse how, in robotics research, robotics researchers are reflecting, reproducing, producing, and sometimes challenging, core assumptions about what it means to be human.

The dissertation draws on three and a half years of ethnographic research across a number of robotics research laboratories and field sites in Ireland, the United Kingdom, and the United States between April 2016 and December 2019. It also includes an investigation of the sites where robotics knowledge is disseminated and evaluated, such as conferences and field test sites. Through a combination of participant and non-participant observation, interviews, and textual analysis, I explore how the robot reveals assumptions about the human, revealing both individual, localised engineering cultures, as well as wider Euro-American imaginaries.

In this dissertation, I build on existing ethnographies of laboratory work and technological production, which investigate scientific laboratories as cultural sites. I also contribute to contemporary debates in anthropology and posthumanist theory, which question the foundational assumptions of humanism. While contemporary scholarship has attempted to move beyond the nature/culture binary by articulating a multitude of reconfigurations and boundary negotiations, I argue that this is done by neglecting the body.

In order to address this gap, I bring together two complementary conceptual devices. First, I employ the embodiment philosophy of Maurice Merleau-Ponty (2012; 1968) particularly his emphasis on the body as a site

of knowing the world. Second, I use the core anthropological concept of the 'fetish' as elaborated by William Pietz (1985). By interrogating the robot as 'fetish', I elaborate how the robot is simultaneously a territorialised, historicised, personalised, and reified object. This facilitates an exploration of the disparate, and often contradictory nature, of the relations between people and objects.

In my thesis, I find many boundary reconfigurations and dissolutions between the human and the robot. However, deviating from the relational ontology dominant in the anthropology of technology, I discover an enduring asymmetry between the human and the robot, with the living body emerging as a durable category that cannot be reasoned away. Thus, my thesis questions how the existing literature might obscure important questions about the category of the human by focusing disproportionately on the blurring and/or blurred nature of human/non-human boundaries. Ultimately, I argue for a collaborative and emergent configuration of the human, and its relationship with the world, that is at once both relational and embodied.

This dissertation is structured as follows. An initial introductory chapter is followed by a chapter documenting the literature review and conceptual framework. This is followed by four chapters that correspond to the four aspects of the fetish in Pietz's model: Historicisation, Territorialisation, Reification and Personalisation. These chapters alternate between scholarly sources and ethnographic data. In Historicisation, using existing scholarship, I trace the history of the robot object, including the continuities and discontinuities that led to its creation, as well as the futures that are implicated in its identity. This is followed by the Territorialisation chapter, in which ethnographic data is used to interrogate the robot's materiality, as well as the spaces in which it is built, modified, and tested. The next chapter, Reification, considers the robot as a valuable object according to institutions and the productive and ideological systems of Euro-American imaginaries. This chapter integrates ethnographic detail with existing scholarship to focus on contrasts between the dominant image of imminent super-human intelligence and the human interventions and social relationships necessary to produce the illusion of robot autonomy. Finally, the chapter Personalisation brings ethnographic attention to the intensely personal way that the robot-as-fetish is experienced in an encounter with an embodied person, understood through the lens of Merleau-Ponty's embodiment philosophy. In the final chapter, I draw together the various strands to articulate how understanding the robot as a fetish, underscored by Merleau-Ponty's embodiment phenomenology, can provide useful resources for developing an alternative understanding of the human in anthropology without dissolving it all together.

Acknowledgements

My PhD research was very much dependent on the guidance, assistance, and support of a broad range of people and groups since it began in February 2015. I am particularly indebted to my supervisor Dr. Mark Maguire for taking on this atypical project, for his invaluable advice, and for providing me with the drive to keep going. I would also like to thank my second supervisor, Dr. Steve Coleman for his unique perspective and knowledge of the subject. I am grateful to the Department of Anthropology at Maynooth University for their support. I would also like to thank the Wenner-Gren foundation for funding the US leg of my fieldwork.

I owe a huge debt of gratitude to the robotics community, and their astonishing openness and kindness. In particular, I would like to thank Dr. Conor McGinn, Dr. Frank Broz, Professor Alan Winfield and Professor Illah Nourbakhsh for sharing their knowledge, and for their advice and support. I would also like to thank the many people who participated in my fieldwork and interviews, who cannot be named but to whom I am eternally grateful.

I would like to thank my family and friends, especially Mandy, Anna, Sinéad, Laoise, Louise W and my parents, Mary and Leo, for their help, support, and encouragement. Love and thanks to my husband, Anthony, without whose unending knowledge and support this project would not have been possible. Finally, a big thank you to my daughters, Anú and Étain, whose incredible energy, and creative and unique insights helped to shape my project.

Chapter One: Introduction

Topic and context

In the past decade, there has been a surge of interest in, and funding dedicated to, projects related to Artificial Intelligence (AI) and robotics. Each day sees a proliferation of articles in media about the latest exciting advances on the path to human-like intelligence, as well as warnings about the possible risks, from a loss of jobs to existential concerns for the future of humanity. It would seem that the hitherto elusive dream of building intelligent machines seems finally to have arrived, or at least to be imminent. The implications of such a development would stretch far beyond the impact of technology itself and constitute a radical shift in how the human is known, and in our relationship with the world.

Although my background is not overtly technical, I had spent the decade before starting this research project working alongside technical teams both in industry and academia. In industry, I had been part of the management teams overseeing the integration of new technical systems and saw first-hand the experience of those who had to use the new system. I was also responsible for teams delivering digital content for large online publishing companies. In more recent researcher roles, I had worked alongside computational linguists and machine learning researchers to create software programmes for commercial application. In all of these roles, I had noticed a huge discrepancy between how the human was conceptualised by the technical and management teams (as an information processor, inert until activated, resistant to change) and the realities on the ground (people as creative, social, resourceful). This discrepancy was more than theoretical; it affected the strategies, the processes and the practices people were expected to adopt, as well as how they were perceived by management and customers. This often led to profound feelings of frustration and alienation, compounded and amplified by the durability of the technical systems in which they were inscribed. Furthermore, existing dominant theories of economics and information provided no useful tools to understanding how people interacted with computers. The single path to understanding how a person experienced a system, and to successfully design a new one, was to be present with those using the system in real time, to watch their actions, their frustrations, and to listen to their ideas. As a researcher, this insight led me to two disciplinary areas that had previously been unknown to me: the method of observation as it is practiced in anthropology, as well as the phenomenological philosophy of Merleau-Ponty, which emphasises the primacy of embodied experience.

Like many people of my generation, I grew up immersed in science fiction stories in which human-like androids were envisioned as part of our future social worlds. From the delightful yet indispensable C-3PO and R2-D2 in *Star Wars*, to the dystopian AI HAL in *2001: A Space Odyssey* and the eponymous anti-hero of the *Terminator* films, robots had become part of our wider cultural imaginary, a communal thought experiment in which humans could explore the best and the worst of what we consider to be our humanity. Despite representing the ultimate futuristic technology, they also provided a comforting and nostalgic continuation of an ancient narrative of humanity's ongoing and inevitable progress towards an end-state, in which we increasingly gain control over nature and ultimately come to dominate all worlds. Technologies, it seemed, shaped both our everyday worlds and intimate daily experiences, as well as fuelling our collective imaginations and assumptions. And yet, it proved almost impossible for me to fully grasp the concept. Any attempt I made to define, or bound, it invariably led to an ever-expanding concept that eventually encompassed whole worlds.

AI and robotics research offers an extreme example through which to investigate our entanglements with technology. On the one hand, robots have much in common with other technologies. They have the potential to extend our reach into the world, and to increase our comfort and convenience. Like other technologies, robots are deeply interwoven with economics and politics, and may be used to concentrate power and to shift it (Winner 1980). They are shown to have a strong 'bio-power' effect, affecting bodies and immersing them in social relations of power, inclusion, and exclusion (Foucault 2010). Like other technologies, concerns have been raised about the specifics of their technological capability and design. These range from immediate concerns, such as issues of deception and attachment with potentially vulnerable robot users (Turkle 2011), the threat to jobs (Acemoglu and Restrepo 2017), surveillance (Eubanks 2019; O'Neil 2017), the implications of semi-autonomous weapons (Suchman 2016), to longer-term existential concerns (Bostrom 2002). Thus, as is true for other technologies, the ongoing proliferation of AI and robotics technologies in our lives may represent both an opportunity for the enhancement of life, as well as a tool to negatively affect people's lives, including increasing inequality and suffering in society.

On the other hand, humanoid robots occupy a unique position in the contemporary. First, they are a powerful symbol of advanced technology and human ingenuity. In literature and popular culture, they provide complex mirrors of our culture(s), collective identities, and existential fears. The potency of their symbolic power is also connected to their physical nature, and the strong reactions that are evoked in the concrete encounter by their shape and movement in space. Further, no other technology so overtly seeks to replicate, and thus understand, ourselves as humans in the world. In a robotics laboratory, core ontological and epistemological assumptions are enacted. As Lucy Suchman (2007) observes, robots act as a 'doubling or mimicry in the machine' (226), a revelatory site that both discloses and produces assumptions about the human, including what is considered essential and what is considered 'Other', as well as assumptions about gender, class and race. By building and modifying robots, robotics researchers are thus reflecting, reproducing, producing, and sometimes challenging, core assumptions about what it means to be human. AI and robotics research is thus at once a project of engineering, and of philosophy. The optimism surrounding current developments in AI and robotics make it a timely focus to bring ethnographic and anthropological attention to this, often opaque and esoteric, subject matter and makes the robotics laboratory a strategic research site in which to explore these themes.

Focus and scope

In this dissertation I ask: ‘How is the human produced in robotics research?’ It focuses specifically on the work that is done to create humanoid robots that exhibit social and cognitive behaviours. In order to investigate the subject matter, I conducted a three-and-a-half-year ethnographic investigation across multiple field sites in Ireland, the United Kingdom, and the United States where robotics research is conceived, carried out, disseminated, and tested. I also attended and presented at a social robotics conference in Tsukuba, Japan. My research comprised of both participant and non-participant observation, as well as interviews. I carried out informal conversation and interviews with over 100 people in the field, carrying out formal, semi-structured interviews with 16 roboticists. My informants comprised of robotics researchers with a background in either engineering or computer science, or both. Most of my informants are senior roboticists with significant experience and standing in the community. However, I also gathered perspectives from a number of younger roboticists at undergraduate and graduate level. Although, throughout my fieldwork, I visited many labs, interacted with many researchers, and encountered many robots, my primary research site was a robotics laboratory in Trinity College Dublin, Ireland. This team were building ‘Stevie’, a social and care robot.

This research builds on existing ethnographies of laboratory life (Latour and Woolgar 1986; Knorr-Cetina 2014; Lynch 1985; Traweek 1992), as well as studies focused specifically on AI and/or Robotics research (Suchman 2007; Adam 1998; Helmreich 2000; Kember 2003; Richardson 2015) which investigate scientific laboratories as cultural sites. It also contributes to contemporary debates in posthumanism theory, which question the foundational assumptions of humanism (Haraway 1991; Hayles 1999). Both of these research trajectories have questioned the foundational concept of the human as it is known in anthropology and seek instead to explicate ways in which the boundary between humans and non-humans are unstable and subject to dissolution. For theorists in these fields, advances in AI and robotics research are proposed as evidence for the increasing obsolescence of the human as a core category of ontological significance, e.g. (Haraway 1991).

However, I depart from this line of reasoning in this study. Through a detailed ethnographic investigation of the work of robotics, this project finds an enduring asymmetry between the human and the robot, with the living body emerging as both a core category and enduring material reality that cannot be reasoned away. This project thus questions how the existing literature might obscure important questions about the category of the human by focusing disproportionately on the blurring and blurred nature of human/non-human boundaries. This study thus not only to traces how boundaries are dissolved and reconfigured in robotics research, but also ways in which they resist dissolution. In this way, it contributes to the discussion of how they might be usefully reconfigured and restored as a central anthropological concept.

In order to do this, this study uses two core concepts to structure the analysis and develop the argument. The first concept is that of the 'body-subject', which is drawn from the phenomenological philosophy of Merleau-Ponty (2012: 1968), particularly his emphasis on the body as a site of knowing the world. I also draw on Dreyfus (1992) and his reading of phenomenology in the context of AI technologies. This concept forms the foundational ontology for this dissertation, as well as an alternative to the dominant image (both in anthropology and in robotics) of human-object relations. The second concept is the core anthropological concept of the 'fetish', specifically the preliminary model elaborated by William Pietz (1985).

Pietz is a historian and independent scholar whose detailed scholarship on the origins and history of the fetish concept launched a revival of the concept. In a series of papers, Pietz (1985; 1987; 1988) developed a historical account of the fetish based on an analysis of the use of the word in Euro-American scholarship since it emerged in the cross-cultural spaces on the West African coast in the 16th century, highlighting the spatial and historic specificity of the term. This aspect of Pietz's work is widely cited, e.g. (Graeber 2005; Latour 2010; Taussig 1993a), inspiring a small sub-field and literature on the fetish (Apter and Pietz 1993; Spyer 1998). However, outside of this narrow area, Pietz is almost unknown, even something of a mystery. As well as the detailed historical account of the fetish, in the original article Pietz also proposed a ‘preliminary theoretical model of the fetish’ (Pietz 1985, 7), developed from the recurrent themes in the fetish discourse, which he tentatively proposed as revealing a ‘truth’ about all historical objects. However, this part of Pietz’s work has received little attention. It is likely that Pietz’s historical and linguistic analysis found a more fertile ground in the poststructuralist era 1980s and 90s in than any attempt to articulate any ‘theory’ of a human universal. More recently, a number of scholars have highlighted this aspect of Pietz work, e.g. (Sansi Roca 2015; Braune 2020), however, little work has been done to develop it in earnest.

In my study, Pietz’s fetish concept both serves to incorporate a Merleau-Pontian conception of the human, as well as drawing attention to the multifaceted, multi-layered, and often contradictory ways in which the robot is simultaneously a territorial, historical, reified and personal object. Ultimately, this study argues for a reconfiguration of the figure of the human that is simultaneously relational *and* embodied.

Meet Stevie

Stevie the robot is just two years old, yet in his short life as a social and care robot he has already become an international superstar. Stevie is one of a handful of robots including R2D2 and the medical robot Da Vinci, to have appeared on the cover of *Time* magazine, see (Purtill 2019). He is also listed as one of the top 100 inventions of 2019. His appearance is humanoid but not at all human. He looks like a child robot, or a robot that a child would draw. His voice, on the other hand, is that of an adult male. His accent is English.

Figure 1: Stevie the Robot on the cover of *Time* magazine (TIME 2019)

Stevie was created by a team from the Robotics Lab in Trinity College in Dublin, Ireland. I first met Daniel[1], Stevie's creator and head of the team, at the International Conference on Social Robotics (ICSR) in Tsukuba Japan in November 2017, where I was giving a paper. Stevie had been entered in to, and won, one of the prizes in a robot design competition. After returning to Ireland, I started to visit the lab at Trinity regularly over the next two and a half years and it became my primary field site. As well as 'hanging out' in the lab and with the team, I also collaborated with them on human-robot interaction (HRI) studies in Ireland and travelled with them to a retirement community in the US to carry out an on-site field study. I also collaborated with them on a journal article, see (Veling and McGinn 2021).

However, Stevie and the team at Trinity were not my first introduction to robotics. I had started my research by travelling to the Artificial Intelligence and Simulated Behaviour (AISB) conference in Sheffield in April 2016 hoping to find a project where I could investigate

[1] Throughout the thesis, I have given synonyms to interlocutors who I introduce by first name only.

my topic of interest further. I was initially interested in computational creativity, a field that attempted to recreate creativity using computational and machine learning techniques. However, I was quickly side-tracked once I met roboticists working on cognitive robotics, whose focus was on building robots that were both intelligent and social. This would remain my core focus for the next four years. Some of the roboticists whom I met at the conference would become key interlocutors, opening up field sites for me in the UK and the US.

Methods

This dissertation brings ethnographic attention to the work of robotics research to consider the problematic of the human-object relations, and how the human is known. The ethnographic material in this study is grounded in fieldwork conducted between April 2016 and October 2019. While my primary field site was the Robotics Laboratory in Trinity College in Dublin, during this time I also pursued a number of inquiries simultaneously, travelling between field sites in Ireland, the United Kingdom and the United States. My fieldwork led me to robotics labs, conferences, industry boardrooms, care centres, and retirement communities. Throughout the duration of my fieldwork, I also worked as a research fellow in a technology research centre co-founded by Maynooth University and Intel, researching issues related to organisations and technology change.

This project takes a multi-sited ethnographic approach (Marcus 1995), integrating a phenomenologically-grounded theoretical framework with ethnographic research in order to develop a multi-factorial account of the robot and robotics research. Rather than 'conventional single-site location, contextualized by macro-constructions of a larger social order' (1995, 95), multi-sited ethnography incorporates multiple sites of observation. Additionally, multi-sited ethnography eschews macro-constructions of larger social order in favour of more integrative concepts 'that cross-cut dichotomies such as the "local" and the "global," the "lifeworld" and the "system"' (1995, 95). As objects, robots embody these contradictions and multiplicities, just as the work of robotics spans multiple territories. Robots are at once physical and local entities, but also part of a wider, global, hyperreality of connectivity and networks. They are experienced both at a personal and visceral level, as well as forming part of wider social and cultural imaginaries. Robotics researchers are spread across many field sites and engaged in many different types of work, including designing and installing the physical components, programming behaviours, holding meetings, engaging in demos and networking, writing papers, attending conferences and testing the robots in community settings. As I found during my research, the sites where AI and robotics research take place are at once subject to disparate histories, geographies, institutional arrangements, and individual interests, dispositions, and motivations, while also bound together by common histories, discourses, assumptions and norms.

Thus, rather than fieldwork situated in a single, bounded field site, my study follows the cultural production of robots across a number of localities and temporalities. It pays particular ethnographic attention to discourses and material practices, tracing associations and connections and paying attention to the continuities and discontinuities between and across sites. The 'culture' that I am investigating in this project is thus also the culture that I am constructing with this work, emerging from the connections that I followed and the spaces that were opened up to me (Marcus 1995, 96). Thus, I find communities of roboticists that are geographically and culturally specific, while also sharing a common identity with the wider robotics community.

My field sites were chosen through a combination of approaches, including a combination of exploratory, emergent and opportunity sampling (Patton 2015). When I set out, I hoped to find an exemplar case. Ideally, I was looking for a group that was building a complete, humanoid robot from scratch. I also hoped to include a group from one of the 'big four' US robotics centres, MIT, Carnegie Mellon, Stanford University, and Stanford Research Institute (SRI). My strategy involved attempting to insert myself into robotics communities in public settings (primarily at conferences) and then following leads, connections, and associations (Marcus 1995). In order to negotiate and maintain access, I looked for opportunities to make myself useful to robotics communities through opportunities for collaboration on HRI studies and publications, or by offering to give seminars. During my fieldwork, many roboticists, particularly those working in HRI, were becoming interested in qualitative research and ethical issues and I was seen as being useful to them in these areas. In this way, I managed to give seminars, co-author papers, and support teams in running interaction studies, which were important factors in gaining and maintaining access, as well as building trust and rapport.

Although I had some technical experience and had worked in a technological institute, when starting out, I had little experience of the work of AI and robotics and no contacts in the field. I initially approached colleagues in the computer science and engineering departments of the University in which I was based, however, rather than representing a group or a lab, the work being done on this topic was being carried out by individuals and was not their primary focus. However, it was one of these colleagues who suggested that I travel to the conference on Artificial Intelligence and Simulated Behaviour (AISB) in Sheffield in April 2016 to get an insight into the variety and scope of work currently being done in the area. At AISB, I was taken aback by the openness and approachability of many in the (various) fields that were connected to the topic. It was at this conference that I met two prominent British cognitive roboticists, who introduced me to the field and with whom I have maintained contact throughout my project. One of these was to become a key interlocutor, opening up a field site at Bristol Robotics Lab (BRL) in the South of England, as well as introducing me to contacts at the Robotics Institute at Carnegie Mellon University in Pittsburgh.

In November 2017, one of my papers was accepted at the International Conference on Social Robotics (ICSR), in Tsukuba in Japan, see (Veling 2017). This conference had a much narrower focus than AISB, featuring HRI studies, with groups from psychology and design, as well as robotics. This time, I was able to move among attendees as a peer, rather than a somewhat suspicious outsider who was studying them. It was here that I met two more key interlocutors: Daniel, whose team at Trinity was to become my primary field study, and Gerry, an assistant professor at a UK university. Gerry had previously hosted philosophers at his lab, leading me to assume, rightly as it turned out, that he would be open to hosting me too.

Between December 2017 and December 2019, I was able to maintain a regular presence at the Robotics Laboratory at Trinity, as well as carrying out a number of visits to field sites in the UK and the US. I also conducted a series of interviews at three of the field sites: Trinity, Heriot-Watt, and Carnegie Mellon. The Robotics Lab at Trinity was one of the rare labs that worked on building a complete robot. While there, I mostly observed the work and chatted with researchers, I also participated in the lab studies (both as researcher and participant), attended meetings, and collaborated on papers. As we will see, the roboticists rarely work directly on the robot as a whole. Instead, their work involves programming, electrics, finding components, building parts, fixing, modifying, testing and experimenting, as well as writing papers, demonstrating the robot, attending meetings and media communications. In October 2018, I assisted the team in conducting a series of pilot focus groups with a charity for older people living alone in Dublin. In August 2019, I travelled for a week with the team to a retirement community on the East Coast of the United States to conduct an extended evaluation of the robot. The team in Dublin were not part of a larger institutional or national robotics drive and had to source funding piecemeal for individual projects. They were continually engaged with media and funders, to expand their reach and the interest of investors, and they identified more with innovators and entrepreneurs than with other academics. Because of this, in a way that was not typical for robotics teams, they had also included potential users in the design of the robot since the earliest prototypes.

Throughout 2018 and 2019, I travelled between four different robotics labs: one in Ireland, two in the UK and one in the US. In May 2018, I spent several days at research laboratories in the UK, Heriot-Watt in Edinburgh, Scotland, and Bristol Robotics Lab (BRL) in Bristol, England. The Robotics Lab at Heriot-Watt is part of the School of Mathematical and

Computer Sciences. The team's focus here is on different types of robotics, including swarm robotics, and drones, but their big focus, and the place that they hope to distinguish themselves, is on the inter-disciplinary work of HRI, for which they collaborate closely with cognitive scientists and psychology researchers. The teams work mostly with off-the-shelf robots that are standard across European universities, robots such as Nao, iCub, and FLASH.

At Heriot-Watt, I was given a desk among PhD students and, in what was to become a trend across my field sites, this group of informants were the most suspicious of my presence and least likely to include me in their activities. I was also a PhD student, but older than they were, an anthropologist placed in their midst by more senior faculty, possibly to 'study' them. Luckily, Gerry was willing to allow me to accompany him throughout my days there and so I got a good impression of the daily life of the roboticists, which, between meetings, demoing, supervising students, writing papers and social events, had a lot less to do with actual, physical robots than I had previously imagined. During my time there, I was asked to deliver a seminar on using qualitative research for robotics work, which proved extremely effective in building trust and getting to know people.

Like the team in Dublin, the teams at Heriot-Watt spent a disproportionate amount of their days demonstrating the robots for media, schools, and to potential funders. The team were very open to collaboration with other disciplines, and most of the HRI project also included researchers from psychology. The roboticists with whom I interacted were extremely thoughtful and philosophical about their work, distancing themselves from the fantastical discourses of imminent human-like intelligence, and emphasising the practical nature of their work. They were, however, constrained by the norms of the HRI field, predominantly restricting their studies to statistical and physiological metrics, and generally eschewing critical reflection in their publications. I returned to Heriot-Watt again in June 2019 to continue my fieldwork and conduct a series of interviews. By the time I returned to Heriot-Watt in 2019, I was greeted as an old friend.

On my way from Edinburgh to the South of England, I stopped off in London to visit a roboticist contact who I had met at AISB. Previously a prominent academic, he had moved from a British university to DeepMind in London. DeepMind is a British company specialising in AI research, which is now wholly owned by Alphabet Inc., Google's parent company. Before arriving, I was sent a non-disclosure agreement to sign. On the day of my visit, I waited for a long time in the plush reception area, watching looped advertisements about the company and its achievements. Eventually, I was ushered into a meeting room and met with my (previously open but now very guarded) contact. We exchanged a few pleasantries, but I succeeded in extracting little about the company. I asked whether I could have a look around, whereupon I was brought to the rooftop to see the view, and quickly ushered out.

I continued on my journey south to Bristol Robotics Lab (BRL). BRL was established in 2006 and is a collaboration between the University of Bristol and the University of the West of England and is the biggest robotics laboratory in the UK. The lab is a dedicated, state-of-the-art space, situated in an old Hewlett-Packard building and covers an area of over 4,600 sq. metres. It has a much wider focus than at either Trinity or Heriot-Watt, each cube hosts a different research area: swarm robotics, social robots, 'soft robotics' and tactile sensing robots, each featuring different scientific approaches to robotics. In the centre of the space is a large aerodrome for drones and aerial robots. Researchers at BRL are engaged in some radically innovative projects, including 'self-sustaining' robots, which 'feed' on slugs, flies, and urine. The technologies derived from this research have been used to power electric-lit toilets for use by women in India. There is an emphasis on pro-social research, with researchers working on collaborative robots and attempting to embed simulated ethical reflection into robots. I was also given the opportunity to do a seminar to which all of the roboticists in the lab were strongly encouraged to go, not just those involved in HRI. This was because, as one of the directors of the institute told them firmly, all of their work was essentially 'social'.

In October of the same year, I visited Carnegie Mellon for two weeks. Carnegie Mellon's Robotics Institute (CMRI) is part of the school of computer science and, founded in 1979, one of the oldest robotics departments in the world. It is huge, with over 50 full time faculty members, spanning seven floors and several buildings. I had been given a desk with the team working at the Community Robotics, Education, and Technology Empowerment, or CREATE, lab. This lab had a very different research agenda than the rest of the Robotics Institute, having been established with the goal of refusing military funding. From my base at the lab, I was able to explore the wider Robotics Institute and hold interviews with faculty, as well as understand the work being done at CREATE.

At Carnegie Mellon, I had perhaps grown too comfortable with myself, and did not prepare a seminar. At first, I found it difficult at first to get people to talk to me. After a few days of uncomfortable and suspicious looks, I asked whether I could introduce myself formally at the start of one of their meetings. I made sure to talk about some of the issues that I had heard over the preceding days, which at the CREATE lab included a focus on empowerment through technology. The effect was immediate. For the rest of my time there, I was included in meetings, invited to observe project work, as well as invited along for coffee and lunches.

As we will trace in the following chapter, the culture and history of CMU was unique compared to the other labs that I visited. There was both a dedicated Robotics Institute and a separate, dedicated HRI lab, which was part of the wider human-computer interaction (HCI) effort. Ethical, or 'pro-social', robotics also comprised a distinct research group. The main research institute is a highly prestigious institution, attracting top talent from across the US, and indeed the world. During my visit, some of the world's most powerful technology companies were in the process of setting up permanent offices on its periphery, including Apple, Disney, and Facebook. The work is primarily military-funded, but increasingly also funded by industry. As we will see, the roboticists that I interview here see themselves as the direct inheritors of the founders of the project of AI, such as CMU's Alan Newell and Herbert Simon and MIT's Marvin Minsky, echoing their early aphorisms, such as Minsky's alleged description of humans as 'meat machines'. Additionally, roboticists here were more likely to express disdain for what they considered the marginal work of robotics, including ethics, philosophy, and HRI. However, as we will trace in this study, despite the geographical and cultural specificities of the individual robotics labs, I also found a robotics community united by common historical narratives, identities, and mythologies.

Throughout my fieldwork, I captured observations in field notes daily. When starting out, I mostly took handwritten notes, which I later transcribed. However, I later found that it was often possible to write field notes directly into a digital journal on my laptop, as many people in the lab would have their laptops open as they worked, and expected me to have the same. Thus, field notes were gathered using a combination of these methods. Once captured, they were archived for subsequent analysis. The fieldwork was supplemented by more than 100 informal, open-ended, unstructured interviews, as well as more formal, semi-structured interviews. The interview protocol was based on insights gathered during the previous two years of fieldwork and analysed according to the theoretical framework that I had chosen, as is described in the next chapter. Interviews lasted approximately an hour and were audio-recorded. After the interviews, I transcribed the recordings. Along with the field notes, all interview transcripts were anonymised. This involved removing identifying material and replacing it with a code. Another document provides a link between the participants and the code, which is stored separately, and password protected. Along with the archived field data, all interview data gathered was collated, and iteratively coded using QSR NVivo 11 to identify themes and patterns.

The research underwent ethical review and was granted approval through the university's ethics process. Before each of the interviews, I gave interviewees a consent form and an information sheet, which explained the purpose of the research, the anonymisation process, and asked their permission to record. It also explained that they could withdraw from the interview at any time.

Outline

This dissertation is structured as follows. After this initial introductory chapter, Chapter Two consists of a literature review and theoretical discussion, setting out the conceptual framework that provides the structure for the rest of the thesis. The literature review section is divided into four sections. The first section looks at early anthropological theory, in which technology was seen as a driver of evolution and culture. The second section is focused on scholarly attempts to reconcile the social and the material. Next, in the third section, I focus on more recent attempts to conceptualise human-object relations, including works drawing on phenomenology and traditional anthropological concepts, such as animism and fetishism. Finally, the fourth section gives an overview of Merleau-Ponty's (2012) concept of embodiment and Pietz's (1985) concept of the fetish, which together provide an overarching theoretical framework for this project. The following four chapters (Chapter Three, Four, Five and Six) make up the main body of this work, drawing together ethnographic description with scholarly work to focus on four key aspects of the robot, corresponding to the four aspects of the fetish in Pietz's model: Historicisation, Territorialisation, Reification and Personalisation. In the final chapter, I draw together the various strands to articulate the complex, pluralistic, and often contradictory ways in which objects, of which the robot is an exemplar, are imbricated in the cultural world. This ultimately leads me to posit a conceptualisation of the human that is both relational and embodied. This final chapter draws out some of the implications of this way of understanding the category of the human for anthropology.

Chapter Two: Literature and Theoretical Approach

This study is an anthropological investigation of the work of robotics research. It follows similar studies in anthropology and related disciplines, particularly Science and Technology Studies, or STS, which aim to treat scientific and technological laboratories as traditional culturally meaningful settings. These include pioneering studies of laboratory life (Latour and Woolgar 1986; Knorr-Cetina 2014; Lynch 1985; Traweek 1992), as well as more recent works focused on AI, Artificial Life (AL) and robotics (Suchman 2007; Adam 1998; Helmreich 2000; Kember 2003; Richardson 2015). In particular, I build on the ethnographic works by Suchman (2007) and Kathleen Richardson (2015), both of whom carried out fieldwork at MIT Robotics Lab in the early 2000s, focusing on humanoid robots and the day-to-day practices of robotics researchers. Both authors also develop their own, albeit conflicting, theories about human-machine relations. My study, by contrast, is based on fieldwork in a number of different robotics field sites and posits an alternative conceptualisation of both the human and human-machine relations.

This alternative conceptualisation originated in my intuition, stemming from my professional work at the human-machine interface, that there is a type of knowing that has eluded theorists across academic fields, which I subsequently came to recognise in the embodiment philosophy of Merleau-Ponty. This intuition was developed and refined as my research progressed. Further, over the course of my fieldwork, the concept of the fetish, particularly as articulated by Pietz (1985), gradually grew in prominence. However, it was not until I was at the robotics conference in Japan in 2015 that I found myself explaining my research in its terms. As we will see, Pietz's fetish concept accommodates an analysis of the contradictory nature of our relationship with technology, which is simultaneously historical and embodied, material and reified. It therefore allows for an articulation of the human-machine boundary where previous discussions have faltered. However, before picking up this thread, I will first situate the framework in the wider anthropological literature by tracing how, from its inception, anthropology has dealt with technology, and the object in general.

Technology in Evolutionary and Cognitive Anthropology

The concept of 'technology' in anthropology has not been a stable one. In early anthropological theory, technology was considered an expression of a particular culture or epoch, rather than something reducible to specific objects or artefacts. Early anthropological theory, which aspired to create a scientific account of culture comparable to the physical sciences, featured evolutionary accounts of societal development and progress, with technology as a key driver. In these accounts, separate, bounded cultures were divided into increasingly progressive groups dependent on their relationship to it (Tylor 1871; Morgan 2003). The artefacts of pre-modern cultures, on the other hand, were the focus of a separate field of museum anthropology (e.g Stocking 1985), see (Hicks 2010).

In *Primitive Culture*, English anthropologist Edward Burnett Tylor (1871) developed the theory that culture evolved from the simple to the complex, passing through a series of progressive stages from an initial 'savage' state, through 'barbarism', and finally to 'civilisation'. In the US, Lewis Henry Morgan (2003) developed a comparable seven-stage evolutionary scheme, in which culture was seen as progressing from 'lower savagery', through to 'modern civilisation', largely based on its technological achievements. Thus, for Morgan, modern civilisation was distinguished by such technological developments as the telegraph, steam engine, and spinning-jenny. For early anthropological theory, cultural evolution was thus perceived as linear and natural, and determined both by increasing complexity and increasing systematisation. Inherent in these theories was the assumption that European and American cultures were at an advanced stage of evolutionary development.

As Miller (1987) has traced, with the emergence of fieldwork in the early 20th century, the focus in anthropology changed. Instead of collecting artefacts and displaying them in museums, anthropologists were encouraged to immerse themselves in the living cultures of those they wished to understand. As he points out, Malinowski (2014) maintained that studying technology alone was to fetishise it: ‘The canoe is made for a certain use, and with a definitive purpose; it is means to an end, and we, who study native life, must not reverse this relationship, and make a fetish of the object itself’ (116). Thus, although in mainstream anthropological literature technology featured in the detail of everyday life, such as tool use (Boas 1955), irrigation (Leach 1959; Geertz 1972), art (Boas 1955; Leach 2001; Levi-Strauss 1988), and objects of value (Mauss 2002), technology itself was not considered intrinsically of interest. Instead, the key focus was on its role in social institutions and social facts, and later on, on structure, language, and discourse.

Despite this predominant focus, a number of anthropological theorists continued to pursue the ambition of aligning anthropology closer to the physical sciences. Anthropology was one of six core fields that contributed to founding the discipline of ‘cybernetics’; an interdisciplinary field that closely aligned human cognitive and communication processes with those of machines. Early cyberneticists included mathematicians Norbert Wiener and Alan Türing, physiologists Warren McCulloch and Arturo Rosenblueth, as well as anthropologists Gregory Bateson and Margaret Mead.

As we will trace in the next chapter, cybernetics and its associated ‘behaviourist’ paradigm would be superseded by the ‘cognitivist’ programme of AI. According to a cognitivist view of mind, people come to understand the world by translating data received by the senses into symbolic representations using processes similar to that of a computer program (Johnson-Laird 1988).

These developments led to the establishment of ‘cognitive anthropology’, whose proponents sought to bring scientific cognitivist theories and tools to bear on longstanding anthropological questions, see (Bloch 1998; Boyer 1994; Sperber 1996; Atran 2004). More recently, alternatives to the traditional cognitive model have been posited, including connectionism and embodiment. Connectionism is based on the computing systems of ‘artificial neural networks’, or ANNs, inspired by the biological neural networks of animal brains. Connectionists propose that, rather than a centralised processor, the mind is instead made up of many small, simple processors operating in parallel, which greatly accelerates the process (Bloch 1991, 191). A physiological and neurological reading of the concept of ‘embodiment’ in this literature is also seen has having the potential to understand culture as grounded in the biological body (Downey 2007, 13). This alignment between theories of the human and technical systems is a major theme in my work and will be discussed in detail in the following chapters.

A separate, but related, approach was taken by Leslie A. White (2007), who combined the cognitivist view of human culture with the deterministic and evolutionist theory of technology of Tylor and Morgan. White hoped that this combination would serve to further the scientific rigour of the anthropological approach, and its contribution to general scientific theory. White’s account of cultural evolution is linear, sequential, and both biologically and technologically determined: first man developed neurologically, then he developed the ability ‘to symbol’, and this brought culture into existence (2007, 6). Culture, in turn, progressed in relation to its ability to use technology to harness energy (2007, 338). Thus, for White, [t]ool use in humans is ‘a cumulative and progressive process’ and ‘[s]ocial evolution comes as a consequence of technological development’ (1943, 7, 353). Technology, in this view, takes on a moral dimension, as it then becomes a society’s duty either to foster the effective operation of the technology, or to restrain and thwart it (1943, 347). As we shall see in the following chapters, this view of technology in society persists today in futuristic, techno-utopianism.

These views were also espoused by younger anthropologists, including early Marshall Sahlins, a student of White’s. In *Evolution and Culture*, Sahlins (1994) distinguishes between ‘specific’ and ‘general’ evolution. Sahlins defines the distinction as, on the one hand, the ‘phylogenetic, rambling, historic passage of culture along its many lines’, while on the other, the ‘passage from less to greater energy transformation, lower to higher levels of integration, and less to greater all-round adaptability’ (38). In this way, Sahlins attempts to reconcile both evolutionary and diffusionist explanations of technological change. For White and early Sahlins, this technological determinism served as a Marxist counter to what they saw as the prevailing ideology of genetic determinism, which sought to explain culture and human activity as the result of a universal ‘human nature’, a product of economic theory of rational choice and competitive self-interest (‘Interview with Marshall Sahlins’ 2008).

Both of these lineages, the cognitivist and the evolutionist, are reflected in the view of technology as a driver of culture in the linguistic and media theory of Walter Ong (2012) and Jack Goody (1973; 1977). For both of these scholars, writing is a technology that has transformed both human consciousness and society. According to Ong, the technology of writing includes the ‘use of tools and other equipment: styli or brushes or pens, carefully prepared surfaces, such as paper, animal skins, strips of wood, as well as inks or paints, and much more (2012, 80-81). Writing is ‘completely artificial’ being ‘governed by consciously contrived, articulable rules’ and ‘the most momentous of all human technological inventions’ (2012, 81, 84). Similarly, Jack Goody (1973; 1977) identifies writing as ‘technologies of the intellect’, which, he maintains, have radically transformed both cognition and culture. For Goody, technology is the key factor in the distinction between ‘modern’ and ‘primitive’ thought (Goody 1973, 11). Thus, technical processes, like language, are considered evidence of the evolution of a specifically human form of intelligence. Although Ong is careful not to tie his work too closely to the explicitly evolutionist projects of Tylor and Morgan, he nonetheless refers to the ‘evolution of consciousness’ (2012, 172, 174, 176). Goody is less circumspect, explicitly tying his project to evolution, which he described as ‘simply long-term change’ (Goody 1973, 1). As well as evolution, Ong and Goody’s work also reflects the influence of a cognitivism, in which human thought is considered continuous with computer processing.

Thus, we can see two lineages emerging in anthropological approaches to technology, veering between idealism and scientism. On the one hand, in the mainstream approach, a focus on technology is eschewed in favour of a focus on structure, language, and discourse. On the other hand, technology is seen as a bridge to developing anthropology as a positivist science, and the search for abstract and universal laws. Despite these differences, as Tim Ingold (2000) has pointed out, the human/computer analogy has also found its way into interpretive and cultural anthropology. Thus, he shows how Geertz (1973) equates cultural control mechanisms, such as plans, recipes, rules, and instructions, with computer programmes (44). As we will explore in the next section, this assumption is also a key part of a symmetrical, posthuman anthropology.

Reconciling the Social and the Material

A radically different view of technology emerged in the 1970s and 80s when a number of scholars started to bring ethnographic attention to the West, in particular focusing on the scientific laboratory. Scientific laboratories were treated as a traditional ethnographic setting, drawing attention to the social and cultural aspects of experimental laboratory work, e.g. (Latour and Woolgar 1986; Knorr-Cetina 2014; Lynch 1985; Traweek 1992). This was what Bruno Latour called a 'symmetrical anthropology', in which people who are traditionally considered 'non-modern' and those who see themselves as 'modern' are both subject to the same investigative methods (Latour 1993). These studies revealed how activities previously assumed to be linear and stepwise, instead involved 'tinkering' (Knorr-Cetina 2013) and 'bricolage' (Levi-Strauss 1966; Latour and Woolgar 1986). Technology, in this context, was seen to be an instrument, not for discovering knowledge, but rather for producing it. Scientific reasoning and decision-making were declared 'indexical', context-dependent and post-hoc, rather than rational and predictable (Knorr-Cetina 2013). These insights even led some to conclude that nothing 'scientific' happens in scientific laboratories (Knorr-Cetina 2013; Latour 1983). Not surprisingly, much of this intellectual development was met with pushback from others in the scientific community. In response, Latour and other scholars distanced themselves from their early overtly social emphasis. Latour and Woolgar, for example, dropped the adjective 'social' from the subtitle of 1979 edition of *Laboratory Life* (1979), publishing it simply as the 'construction' of scientific knowledge in the second edition (1986).

As an alternative to either a technologically or socially determined emphases, Latour, along with Callon and Law, developed Actor Network Theory, or ANT, in a series of papers in the late 1980s (Callon 1986; Latour 1987; Law 1987). The 'network' represented a novel techno-social arrangement, in which, at least nominally, neither the social, nor the technical elements are privileged. In a network, emphasis shifts from a focus on individual entities to the relations between them. Action and agency are thus no longer viewed as a property of individual humans or objects, rather they emerge from their interactions (Callon and Latour 1992). Gradually, symmetrical anthropology gave way to a wider principle of symmetry, a 'flat' or 'symmetrical ontology', in which it is not just modern and pre-modern humans subject to the same investigative methods, but also humans and 'non-humans'.

Other accounts of laboratory studies took a more critical position, reflecting on issues such as class, gender and race (Rapp 1988; Martin 1991; Downey, Dumit, and Williams 1995). Central to this literature is Donna Haraway's (1991) figure of the 'Cyborg'. The Cyborg is 'a rhetorical strategy and a political method' (149) offered as a counter to critical dystopian readings of technology. The Cyborg is a 'hybrid' of imagination and material reality, part-human and part-machine, reorienting feminist accounts to the positive potential of science and technology in order to blur and transform (oppressive) boundaries and identities and instead offer a utopian vision of a post-gender, post-essentialist world.

These developments in laboratory studies, along with the publication in English of Marcel Mauss' (2002) *The Gift*, contributed to a renewed interest in, and reconceptualisation of, technology and materiality in anthropology (Hicks 2010). Publications such as Arjun Appadurai's (1986) *Social Life of Things,* Daniel Miller's (1987) *Material Culture and Mass*

Consumption and Marilyn Strathern's *The Gender of the Gift (2001)* reimagined objects as social, and situated within wider networks of exchange. Instead of embracing a fully symmetrical view, Appadurai (1986) proposes a 'methodological fetishism', in which objects are temporarily 'subjectified' and imagined as subjects with social lives of their own:

> [W]e have to follow the things themselves, for their meanings are inscribed in their forms, their uses, their trajectories ... even though from a *theoretical* point of view human actors encode things with significance, from a *methodological* point of view it is the things-in-motion that illuminate their human and social context. (1986, 5)

Thus, as well as an emphasis on an object's meaning, a novel concern for the object's physicality, or the 'concrete, historical circulation of things' (1986, 5) also emerged. This, for Appadurai, 'is in part a corrective to the tendency to excessively socialize transactions in things, a tendency we owe to Mauss' (1986, 5).

Similarly, in *Material Culture and Mass Consumption*, Miller (1987) makes a case for a new field of 'Material Culture' which will take as its focus the 'problem of artefact as a single example of cultural form' (15), distinct from the dominant linguistic tradition. Like Appadurai, he also emphasises the importance of the object's physicality:

> An analysis of an artefact must begin with its most obvious characteristic, which is that it exists as a physically concrete form independent of any individual's mental image of it. This factor may provide the key to understanding its power and significance in cultural construction.

The objects of Miller's focus are mundane industrial and domestic consumer goods, and the spaces in which they are encountered, such as the home and the supermarket. These are goods that, he claims, 'mediate social relations silently, in a kind of "ordering of the unconscious world"' (1987, 99). Here, Miller's focus is on the consumption, rather than the production, of objects. His focus is motivated by what he claims are the political implications of an abstract view of objects, which have led to 'nihilistic and global critiques of "modern" life' (1987, 4), rather than a deeper analysis 'at the micro-level of the actual relationship between people and goods' (1987, 4). For Miller, such a deeper analysis will reveal 'a perspicacity and subtlety in mass behaviour which is a far cry from passivity, illusion and denigration implied in many self-proclaimed radical perspectives' (1987, 5). Ultimately, this will facilitate a move away from the lack of positive suggestions to how we might actually benefit from the proliferation of products of industrial society, and thus avoids 'providing only conservative and nihilistic assessments of the ubiquity and thus the inevitability of oppression' (1987, 6).

Miller presents the consumption of objects as integral to the creation of social identities, in which, through a process of identification and disavowal with the material object, a person becomes aware of their own being. Building on Hegel's concept of 'objectification', he describes a series of processes including externalisation, or self-alienation, meaning an expansion of the self into the world, and subsequent 'sublation' or reincorporation during consumption of the expanded and alienated part of the self (1987, 12). Thus, for Miller, it is at the point of consumption that meaning is conferred onto the object.

The early social emphasis of the Material-Cultural turn was accused both of an excessive humanism and of perpetuating normative conceptions of human identity (Hicks 2010). This might account for Miller's subsequent position in opposition to what he calls the 'tyranny of the subject', or any privileging of the human over the non-human (Miller 2005, online). Instead, he advocates for 'burying the corpse of our imperial majesty: society' and by building, in its place, 'a dialectical republic in which persons and things exist in mutual self-construction and respect for their mutual origin and mutual dependency' (2005).

A similar concern with the shifting boundary between subjects and objects is evident in Strathern's (2001) ethnographic study in Highland New Guinea, published around the same time. Strathern argues that, for the Melanesian people, traditional Western concepts, such as 'society' and 'individual' simply do not apply. Building on Mauss' concept of personhood, see (Mauss 1997), she argues that Melanesian persons are conceived as 'dividuals', being part of a larger whole. Persons are thus simultaneously 'multiple' and 'partible' (2001, 185). Specifically, it is in the exchange of objects in the context of the Melanesian gift society that the boundaries between subjects and objects are disturbed: personhood is distributed, and humans become dividuals, with objects gaining subjectivity. Strathern's redescription of the person has been so influential on subsequent ethnographic accounts that Sahlins has noted, 'the Strathernian "dividual" is threatening to become a universal form of pre-modern subjectivity' (Sahlins 2013, 25).

Strathern's (1990; 1992; 1996) subsequent publications focus more directly on the problem of studying artefacts and the possibilities and conceptual transformations that might be brought about by new technologies, such as new reproductive technologies. Following Haraway, Strathern (1992) maintains that these technologies have the potential to challenge and transform the foundational categories of nature, culture, personhood and kinship, as well as having the power to lay bare Euro-American imaginaries. Strathern also finds common cause with Latour's symmetrical and networked approach, which, for Strathern, is isomorphic with the concept of the hybrid, 'if we take certain kinds of networks as socially expanded hybrids then we can take hybrids as condensed networks' (1996, 523).

Other theorists approached materiality by taking an interest in the body and the concept of 'embodiment'. Just like objects, anthropological theory traditionally treated the body as symbolic, and a medium for societal relations, e.g. (Douglas 2003; Turner 2008). Following Michel Foucault (2012), other theorists emphasised the historicity of the biological body, e.g. (Butler 2011). New feminist accounts drew attention to the absence of the body in traditional Western philosophies (Adam 1998; Kember 2003). As Alison Adam (1998) points out, much of women's lives and experiences are to do with bodies (134). And yet, the problem of reconciling a new interest in the body and the radical anti-essentialism of STS and anthropological theory remained an open one. Adam worries that 'in order to eschew essentialism we may end up disembodying the body' (Adam 1998, 135). Indeed, as Katherine Hayles (1999) has articulated, this posthuman view 'configures human being so that it can be seamlessly articulated with intelligent machines' (3). According to Hayles, the cybernetic construction of the posthuman, more than any other critique, has had the most impact in downplaying and erasing the role of the body and embodiment as an essential element of the human (1999, 3). This is, she maintains, a feature that cybernetic posthumanism has in common with the liberal humanist subject, whose universal human essence transcends the body, the body instead being its possession (1999, 4).

Posthuman, anti-essentialist perspectives, with their 'symmetrical' or 'flat' ontology, have remained a persistent theme in Anthropological theory characterised by problematising previously taken-for-granted distinctions between people and things, and a proliferations of non-essentialist constructs, such as networks, hybrids, and cyborgs. Instead of the human, the 'social' is conceived as a dialectical relation between human and non-humans, understood symmetrically. Thus, as Jan Heiss and Albert Piette (2015) observe, anthropology, ostensibly the study of '*anthropos*', is left with no theory of the human as a point of departure (10).

Methodologically, this theoretical orientation has manifested in an ethnographic commitment not to study individual entities or essences, but to follow how boundaries are produced and reproduced in practice, e.g. (Suchman 1985; 2007; Downey 1998; Helmreich 2000). However, despite claiming to focus on how those categories are produced or dissolved in practice, it is difficult not to suspect that the categories are being both produced and unmade in the analysis. Furthermore, despite claiming to dissolve dualistic boundaries, underlying these posthuman works is often an implicit commitment to a disembodied, cognitivist view of mind, in which human perception and understanding is the result of symbolic representations. This can often contribute to assumptions that human-like machine intelligence is real, or at least imminent. In a familiar passage, Haraway (1991) maintains:

> Late twentieth-century machines have made thoroughly ambiguous the difference between natural and artificial, mind and body, self-developing and externally designed, and many other distinctions that used to apply to organisms and machines. Our machines are disturbingly lively, and we ourselves frighteningly inert. (152)

However, early confidence that, within a generation, machines would be capable of human-like intelligence has not materialised. As we will explore in the following chapters, there is broad agreement amongst the roboticists in my study that, even if it is ever possible to replicate human-like intelligence, it will not be within their lifetime, see also (Ford 2018). Exceptions to this are techno-futurists like Nick Bostrom and Ray Kurzweil, whose pronouncements and predictions we will examine in the following chapters.

Further, in symmetrical posthumanism, as human exceptionalism, intentionality and agency are thrown into doubt, objects are elevated to 'integral and active threads of an intersecting scientific, social and discursive world' (Conty 2018, 5). As archaeologist Severin Fowles (2016) has pointed out, despite purporting to be 'symmetrical', there is a 'methodological focus on the non-human side of the dialectic' (20). For Fowles, the shift in the anthropological gaze downwards, away from humans and towards things is a response to the crisis of representation. Things, unlike humans, are comparatively uncontroversial and persistently 'Other'. The discourse of the human subaltern is therefore redeployed in defence of the non-human object (2016, 12). In this way, the symmetrical approach is sometimes extended to encapsulate a view of objects as having a moral dimension (Latour 2002).

As we have seen, within Anthropological theory there is a persistent fear of being charged with dystopian and Luddite views of technology. In order to take technologies and objects seriously, it has become necessary to eliminate any distinction between them and living things, or else be accused of anthropocentrism, essentialism, or excessive anti-materialism. This can also lead to an unfounded optimism about the possibilities for transformation and revolution through technology, which have not materialised, e.g. (Haraway 1991; Strathern 1992; Turkle 2005). As Suchman (2007) articulates, symmetrical ontology arose in the social sciences and humanities, which had previously excluded 'facts of nature' and 'technology' from its analysis (269). Outside of these fields, however, such as in ‘the context of technoscience and engineering’, ‘the situation is in important respects reversed’ (2007, 269). Indeed:

> Far from being excluded, ‘the technical’ in regimes of research and development are centered, whereas ‘the social’ is separated out and relegated to the margins. It is the privileged machine in this context that creates its marginalized human others. (2007, 269-270)

Furthermore, as well as the objectification of humans, the subjectification of objects also leads to ignoring their material facticity: their physical form and durability, their historical and temporal specificity, and their ‘complex sensuality’ (Henare 2003, 57).

Despite theoretical and methodological differences, the radically anti-essentialist and anti-anthropocentric fields of laboratory studies and cultural anthropology, as well as the scientific perspective of cognitive anthropology, have found a common ground/alliance (albeit an uneasy one) in a symmetrical and relational ontology, and in the figure of the posthuman. As we shall trace throughout this dissertation, far from being neutral, dissolving boundaries between humans and technology has far-reaching implications for both.

Thinking Anew: Outside the Network

Even among its most ardent proponents, the relational and symmetrical ontology has received some pushback. Latour (2005) has proposed the concept of ‘plasma’ to explain what necessarily sits in the ‘background’, outside of the network. Plasma is ‘*in between*’ and ‘*unknown*’ and it is ‘that which is not yet formatted, not yet measured, not yet socialized, not yet engaged in metrological chains, and not yet covered, surveyed, mobilized, or subjectified’ (2005, 224). Similarly, Suchman (2007) finds ‘enduring asymmetries of person and machine’ which people ‘inevitably rediscover … in practice’ (13), even a ‘durable dissymmetry’ (270). Thus, original posthumanist scholars have called for a return to realism (Latour 2004) and a necessary reconfiguration of boundaries, despite their instability (Barad 2003; Suchman 2007).

The critical power of the symmetrical view is also coming under scrutiny. Latour (2004) has acknowledged that simply obliterating existing boundaries has rendered criticism ineffective when dealing with pressing societal and ecological issues, such as the denial of scientific consensus concerned with climate change. Suchman (2019), reflecting on her work with the United Nations Convention on Certain Conventional Weapons (CCW), observes that 'calling central human actors to account for the effects of those systems rests on the possibility of articulating relevant normative and legal frameworks' (55). This means that conceptions of 'agency' need to 'reinstate human deliberation at the heart of matters of life, social justice, and death' (2019, 55).

However, it is not just for holding the powerful to account that a category of human is needed. According to Sherry Ortner (2006), 'a post-structuralist "anti-humanism" poses real problems for attempts of subalterns (in the Gramscian sense) to attain the privilege of becoming subjects in the first place' (109). As Africana Philosopher Lewis Gordon (1998) points out, the dominant group can afford to give up their humanity because theirs is presumed (39). Similarly, Ruha Benjamin's (2019b), in her exploration of the intersection of race and technology writes: '*…posthumanist visions assume that we have all had a chance to be human*.' [original emphasis] (32). Furthermore, other critical scholars have pointed out that the concept of the human at which posthumanism takes aim is a narrow one: the post-Enlightenment subject, or 'Enlightenment Man' (Atanasoski and Vora 2015, 5). Other scholars have pointed out how alternative conceptualisations of the human in paralleled genealogies of thought that have been ignored (Weheliye 2008, 321; Jackson 2013, 670).

A growing body of literature is engaged in trying to reconfigure boundaries, rather than to simply erase them. As we shall trace, theorists are drawing on various sources such as phenomenology, Charles Sanders Peirce's 'pragmaticism', the alternative metaphysics of Gilles Deleuze, and traditional anthropological concepts, such as animism, fetishism, and mimesis, as an alternative to theorising the complexity of human-object entanglements with the world.

Michael Taussig (2010; 1993b) uses the concepts of the 'fetish', and later 'mimesis' and 'alterity' to theorise the complexity of human/object relations, including technology. In *The Devil and Commodity Fetishism*, Taussig (2010) explores the social significance of the devil image in the folklore of plantations workers in Western Columbia. This is theorised by building on the 'commodity fetishism' of Karl Marx (2008) whereby:

> …the social character of men's labour appears to them as an objective character stamped upon the products of that labour… in order, therefore, to find an analogy, we must have recourse to the mist-enveloped regions of the religious world. In that world the productions of the human brain appear as independent beings endowed with life, and entering into relation both with one another and the human race. So it is in the world of commodities with the products of men's hands. This I call the Fetishism which attaches itself to the products of labour… (43)

For Taussig, the devil, as fetish, is used to mediate 'the conflict between capitalist and pre-capitalist modes of objectifying the human condition' (2010, xvi). The devil contract is thus a recognition of the magical and metaphysical aspects of Western capitalism and market economies, and their role in structuring everyday life, which had been misunderstood (or denied) in the West (2010, 113). Taussig's aim is to illuminate aspects of Western culture through the prism of the Indigenous population for whom these categories have not yet become reified. Thus, Taussig reveals how Western capitalist cultures remain blind to 'the social basis of essential categories', which include 'time, space, matter, cause, relation, human nature, and society itself' (2010, 4). Borrowing from György Lukács, Taussig maintains that within capitalism, fetishism is a 'phantom objectivity', in which objects appear animated and humans, in turn, are objectified' (2010, 31). By contrast, the fetishism of pre-capitalist societies 'arises from a sense of organic unity between persons and their products' (2010, 37). Taussig's overarching commitment, however, remains to a relational configuration. For Taussig, it is the domination of individuated 'things' over the 'relational gestalt' (2010, 35), which is a gross deception, and ignores that they are the 'embodiments and concretizations of relationships that bind them to a larger whole' (2010, 36).

Taussig's project to apply concepts traditionally associated with 'primitive' cultures to understand 'modern' worlds continues in his book, *Mimesis and Alterity* (Taussig 1993b). In this work, Taussig uses the example of the mimetic figurines of the Cuna people of the San Blas Islands off the coast of Panama, through which they mimic the colonial Other. For Taussig, mimesis is a universal human faculty and mode of perception. However, as I will argue later, in this book his focus has shifted to include the centrality of the body in perception.

Alfred Gell (1992; 1998) draws on the semiotics of Charles Sanders Peirce, Strathern's distributed personhood, and the Tylorian concept of animism to develops a theory of art as technology, with both the agency and the power to enchant. Art is made through an enchantment of technology, extending sociality into material forms and hiding the reality of its production (Gell 1992). Instead of an aesthetic response, for Gell (1998), art elicits a deeply personal response, triggering terror, awe or fascination (6). Objects do not have their own agency, rather in order to have agency, Gell argues, they must be entangled within social relations (1998, 7). Objects thus distribute and extend human agency, acting like as 'secondary agents' (1998, 20), or, in Peircean terms, 'indexes' (1998, 13), of human agency. Like Taussig, Gell uses traditional anthropological concepts that separated the modern from the non-modern, arguing that 'the Tylorian concept of animism can be made into a more serviceable analytical tool if it is abstracted from the essentially pejorative contact of Victorian positivistic thought' (1998, 121). Ultimately, once again, for Gell, the relational view subsumes both the person and the object, 'it does not matter, in ascribing "social agent" status, what a thing (or a person) "is" in itself; what matters is where it stands in a network of social relations' (1998, 123).

In a series of publications, Tim Ingold (1997; 2000; 2011) sets out to develop an anthropological account of technology, which, he claims, has been 'one of the most undeveloped aspects of the discipline' (Ingold 1997, 106). He takes aim at previous evolutionary, relativist, and more recent symmetrical accounts, with their idea of object agency. Instead, Ingold proposes that ‘technology’ has emerged within a specifically Western ‘machine-theoretical cosmology’, in which:

> …society is considered to be the mode of association of rational beings, nature is the external world of things as it appears to the reasoning subject, and technology is the means by which a rational understanding of that external world is turned to account for the benefit of society. (1997, 130)

He thus argues that the concept of ‘technology’ belongs solely to the contemporary West, being not an artefact or tool, but rather ‘the epistemological conditions for society’s control over nature by maximising the distance between them’ (Ingold 2000, 314). In order to develop an alternative schema, Ingold draws on phenomenology, ecology theories, as well as the concept of animism.

Ingold is critical of the dominant view of culture, Geertz’s ‘webs of significance’, as understanding culture as belonging to a ‘realm of discourse, meaning and value’ where ‘culture is conceived to hover over the material world but not to permeate it’ (2000, 340). In this view, Ingold argues, sensory data from nature are only rendered meaningful by being passed through the medium of cultural representations. For the cognitivists, Ingold says, this mediating framework is held in the mind, for the ‘relativists’, these cultural control mechanisms are social and exterior to the person (2000, 159). As Ingold points out, in the conventional account, there are two kinds of nature, neither of which we can access directly: ‘“really natural” nature (the object of study for natural scientists) and “culturally perceived” nature (the object of study for social and cultural anthropologists)’ (2000, 41). In practice, this means that, according to this view, nature can only be accessed through formal analytical tools, rather than directly as lived experience. Instead, Ingold (2002) argues that people know their environments directly (40). As we shall see, this is a key aspect of the embodiment philosophy of Merleau-Ponty that I will draw on in this study.

Ingold draws on both Heidegger and Merleau-Ponty, adopting their phenomenological concept of ‘being-in-the-world’, which takes the ‘animal-in-its-environment rather than the self-contained individual’ as its point of departure' (2000, 76, 186). This situated perspective is characterised by Ingold as 'dwelling', a concept borrowed from Heidegger (2000, 154). His is, as Viveiros de Castro has called it, 'a practical, immanent phenomenology' (2012, 92), in which a person is 'a singular locus of creative growth within a continually unfolding field of relationships' (Ingold 2002, 4-5). Rather than being preceded by cognitive activities of planning and design, forms 'arise within the current of their involved activity, in the specific relational contexts of their practical engagement with their surroundings' (Ingold 2000, 186). Ingold is also critical of the treatment of the body in anthropology, accusing anthropologists of 'a tendency to treat body praxis as a mere vehicle for the outward expression of meanings emanating from a higher source in culture or society' (2000, 169). Instead, he follows Merleau-Ponty in advocating a view of the body as 'the *subject* of perception' (2000, 169).

Ultimately, Ingold (2011) proposes the concept of 'meshworks', borrowed from Henri Lefebvre, as an alternative to Latour's 'network'. He criticises Latour's network with its lines of connections and points, and instead takes Deleuze and Guattari's (1987) focus on 'lines of becoming' (Ingold 2011, 83). At the point of coalescence, the lines become interwoven, forming bundles of 'haecceities', which possess an individuated character, but actually constitute 'events' (Deleuze and Guattari 1987, 262-264). According to Ingold, Latour's 'network' approach perpetuates a residual essentialism, treating objects as concrete entities existing prior to the relations, allowing for no 'mutuality without prior separation' (2011, 70). For Ingold, there is no entity apart from the relation. Objects occur. The meshwork is thus the world we inhabit, emerging from the entanglement of interwoven lines, not in the connecting points' (2011, 63). For Ingold, this view is comparable to an animist lifeworld, in which 'beings do not propel themselves across a ready-made world but rather issue forth through a world-in-formation, along the lines of their relationships' (2011, 63).

Ingold's (2013) recent work is no longer concerned with either technology or the human, rather he continues to follow Deleuze and Guattari to 'think from materials', 'to find "the consciousness or thought of the matter-flow"' (94). Ingold advocates for 'an anthropology that has been liberated from ethnography' (2013, 6), in which, as Fowles (2016) has noted, 'exploring the affordances of sand and sticks is considered a form of participant observation' (20). As Fowles has articulated, Ingold has 'replaced anthropology with material philosophy, effectively evacuating matters of culture and politics and obviating the crisis of representation altogether' (2016, 20). Ultimately then, despite alighting momentarily on the centrality of the artefact and the body, Ingold's work extends the symmetrical posthumanist concept of object agency, or 'the progressive subjectification of the object world' (2016, 20).

Although a radically anti-essentialist, relational ontology dominates anthropological theory, there are rare exceptions. Anthropologist Signe Howell (2016), who carried out her ethnographic research with the Chewong people of the Malaysian rainforest in the 1970s, uses evidence from her research to argue against the posthuman dissolution of '*anthropos*'. For Howell, recent trends to assign symmetry and equivalence between human and non-humans is itself a form of anthropomorphism (2016, 45). Howell shows how Chewong metaphysics, while offering an alternative ontological orientation to the sharp boundary between humans and nature, nonetheless maintain certain boundaries and separations. Indeed, she identifies 'the principle of separation' as a 'dominant principle' among the Chewong (2016, 48). For the Chewong, every perceptible object is a potential subject. Thus, there are potential boundaries, as well as potential commensurability, between conscious subjects (*ruwai*) and non-conscious objects, as well between humans and other living beings.

Howell draws on the work of embodiment philosopher Mark Johnson to further analyse both the continuity, and separation of, conscious beings (2016, 53). Consciousness, according to Chewong metaphysics, is characterised by rationality, intentionality, and emotionality. It is also 'constituted in and through the body', and therefore 'it is not a matter of indifference which *ruwai* inhabits which body' (2016, 52). For the Chewong, in some context non-humans and humans are continuous, in other contexts, or from other perspectives, they are separate. Thus, we can learn from the relations in the Chewong human/non-human sociality which is 'predicated upon these two principles of connectedness and differentiation, or recognition and separation' which further inform 'a semantics of equality' rather than a hierarchical ordering of elements (2016, 51). For Howell, an alternative conception of the boundaries between humans and the world does not necessarily entail a complete dissolution of them. She thus argues for retaining the category of the human in anthropology:

> I cannot see how we can drop the '*anthropos*' from our discipline without throwing the baby out with the bathwater. We are both immanent in nature and, through our reflections upon it, transcendent to it. (2016, 59)

Theoretical Framework

In order to carry out this investigation, I take a similar approach to other theorists who have attempted to reconfigure the boundaries by combining phenomenological insights with traditional anthropological concepts, specifically, that of the 'fetish'. However, I diverge from previous theorists in two key ways. First, I maintain a commitment to the concrete specificity of both the body and the object. In order to do this, I have grounded this study in the embodiment philosophy of Merleau-Ponty (2012; 1968) particularly his emphasis on the body as a site of knowing the world. As I will argue, while many theorists have drawn on the concept of the body and the concept of embodiment, the deep insights of Merleau-Ponty's alternative ontology are often overlooked.

Second, I draw on a novel framework for theorising the fetish: the 'preliminary theoretical model' articulated by Pietz (1985). Although a focus on the level of experience is necessary in order to provide a basis from which other insights may be developed, it alone is not sufficient to capture a complete picture of the human-object relation and the object's social significance. As well as the experience of the concrete encounter, Pietz's fetish concept also opens up a space for a consideration of the historical, territorial, and social status of the robot. This model provides a pluralistic analytical scheme that acknowledges for the multifarious, and often incommensurate, ways in which the relations between humans and the object-as-fetish may be traced. Although, as I will describe, a number of theorists have drawn on Pietz's historical analysis of the fetish (Taussig 1993a; Hornborg 2001; Graeber 2005; Latour 2010), none has yet applied this theoretical model to understand its utility. Thus, in this project, I explore how this novel framework, underscored by the embodiment philosophy of Merleau-Ponty, might sensitise us to alternative understanding of human agency, intelligence, and boundary separations/resistance in ethnographic settings.

Merleau-Ponty's embodiment philosophy

In this study, I take an explicitly Merleau-Pontian interpretation of the 'body' and 'embodiment', which I describe in detail in this section. In particular, I emphasise how it contrasts with the image of the human, both in robotics and in anthropology, which has traditionally been one in which experience of the world is mediated through symbolic representations. The dominant image manifests variously, including an adherence to the view of the human as information processor and rational actor in robotics, to the dominance of structuralist and linguistic approaches in anthropological theory. As I argue in this study, by articulating a radically alternative conception of how the human relates to the world, Merleau-Ponty's (2012; 1968) embodiment philosophy, and his related concepts of the 'flesh' and the 'chiasm', offer an alternative ontology. As articulated by philosopher of science Andrew Feenberg (2014), phenomenology allows us to distinguish lived experience from a 'naturalistic ontology', positioning the naturalistic ontology of the West as an abstraction from a richer experience (279-280).

Merleau-Ponty's 'body' is an active and living body and the fundamental site of knowing for the person (Merleau-Ponty 2012). The body, as understood by Merleau-Ponty, is not the object of physical science; rather it is a 'body-subject' or unity of body and mind, and the basis for our experience. It is a holistic unity that is never exhaustively available to knowledge, instead it precedes and surpasses it (2012, 231), The body 'is not a collection of particles, each one remaining in itself, nor yet a network of processes defined once and for all' (2012, 229). It is not something that is owned by the person, a container for the person, or a medium through which the person communicates. Instead, it *is* the person:

> Therefore the body is not an object … Whether it is a question of another's body or my own, I have no means of knowing the human body other than that of living it … Thus experience of one's own body runs counter to the reflective procedure which detaches subject and object from each other, and which gives us only the thought about the body, or the body as an idea, and not the experience of the body or the body in reality. (2012, 231).

Just as the mind and body are not separate, the living body and the world also form a phenomenal unity, which is always 'already there'. The body is the site of this unity. The body-subject is thus both embodied and situated; it is both personal and integral, immersed in the world and continuous with it. The body-subject is not separate or prior to the world, instead each body is an 'incomplete individual', who *has* the world (2012, 408).

The body is 'our general medium for having a world' (2012, 169). It is through the needs of our body and its relation to the world that we come to know and understand other people and things (2012, 216). By interacting with the world through the body, we both perceive it, and actively constitute it. The body is in the world 'as the heart is in the organism', and it is generative, 'it breathes life into it and sustains it inwardly, and with it forms a system' (2012, 235).

Therefore, to know the body (and thus also ourselves, others and the world), we must return to our experience of it in the moment (2012, 109). For Merleau-Ponty, 'the first philosophical act' is:

> …to return to the world of actual experience which is prior to the objective world, since it is in it that we shall be able to grasp the theoretical basis no less than the limits of that objective world, restore to things their concrete physiognomy, to organisms their individual ways of dealing with the world, and to subjectivity its inherence in history. (2012, 66)

By returning to our experience of it, we find that we experience the body simultaneously as subject and object (2012, 109). Similarly, as we will examine in detail in this thesis, we find that we do not perceive others as objects, instead, 'the sentient subject … enters into a sympathetic relation to them' (2012, 248). Thus, by being in a particular 'body and situation', we find ourselves to be in 'an intersubjective field', which through our spatial and temporal bodily specificity, we are simultaneously 'all the rest' (2012, 525). There is thus kinship and continuity between the embodied person and the world.

For Merleau-Ponty, the social and cultural world does not mediate our experience, rather experience is already social and cultural: 'It is impossible to superimpose on man a lower layer of behaviour which one chooses to call 'natural', followed by a manufactured cultural or spiritual world. Everything is both manufactured and natural in man' (2012, 220). Instead:

> …the social is already there when we come to know or judge it. An individualistic or sociological philosophy is a certain perception of co-existence systematized and made explicit. Prior to the process of becoming aware, the social exists obscurely and as a summons. (2012, 422)

This pre-objective unity (of our bodies with the world and with others) is experienced prior to our reflection on the world, or our conceptions of it, and is known to us intuitively. It is the basis for our existence, our perception, and our action in the world and is this foundation that makes all other types of knowledge possible, including rational and reflective thought. This embodied knowing does not negate our reflective capacity; rather it negates it as the starting point of experience. In our pre-objective experience of the world, we do not focus on individual elements or concepts, rather we are faintly aware of the world on the 'perceptual horizon', or 'background', which remains more or less indeterminate.

According to Dreyfus (1992), this aspect of Merleau-Ponty's phenomenology had big implications for the project of symbolic AI. According to the conventional view, we passively receive meaningless information from a stable world through our senses, which we then process as an input to our planning and decision-making. For Dreyfus, this confuses 'this human world with some sort of physical universe. There is no reason to suppose that the human world can be analysed into independent elements' (1992, 232). Instead, embodied beings encounter a world that is already ordered in a way that makes sense to them, endowed with significance, relevance and concerns (1992, 261). We each perceive the world that we have constructed (socially, culturally, but also bodily, and in interaction with the world) in a way that breathes life, sense and meaning into it: [m]y body is geared into the world when my perception presents me with a spectacle as varied and as clearly articulated as possible' (Merleau-Ponty 2012, 292). It is our bodies that confer meaning on this indeterminate background, organising and unifying our experience of the world, enabling us to bypass cognitive, formal analysis (Dreyfus 1992, 235-249). We are thus 'master players' in our own perceptual worlds, engaged in an ongoing process of 'creative discovery' in which 'the world reveals a new order of significance which is neither simply discovered not arbitrarily chosen' (1992, 274, 277). Instead of individual senses processing a single stimulus, the body-subject integrates and synthesises the experience:

> …any object presented to one sense calls upon itself the concordant operation of all the others. I see a surface colour because I have a visual field, and because the arrangement of the field leads my gaze to that surface—I perceive a thing because I have a field of existence and because each phenomenon, on its appearance, attracts towards that field the whole of my body as a system of perceptual powers. (Merleau-Ponty 2012, 370-371)

The body-subject thus perceives the world directly, and directly experiences what is relevant. This experience is prior to conventional ideas about knowledge and is not mediated symbolically either internally (in the mind) or externally (culturally, socially or through the 'extended' mind). It is, as Thomas Csordas (1990) observes, not pre-cultural but pre-abstract. For Dreyfus (1992), '[t]here is indeed a world to which we have no immediate access… atoms and electromagnetic waves. But the world of cars and books is just the world we *do* immediately experience' (268-269).

As we move around this world, we continually refine it, aiming to get a better handle on it, to bring any discord between our bodies and the world into harmony, to find an equilibrium in our orientation towards the world (Merleau-Ponty 2012). Merleau-Ponty gives an example of how the body and the situation move us around a painting in a gallery:

> For each object, as for each picture in an art gallery, there is an optimum distance from which it requires to be seen, a direction viewed from which it vouchsafes most of itself: at a shorter or greater distance we have merely a perception blurred through excess or deficiency… this is not in virtue of any law or in terms of any formula, but to the extent that I have a body, and that through that body I am at grips with the world. And just as perceptual attitudes are not known to me singly, but implicitly given as stages in the act which leads to the optimum attitude… (Merleau-Ponty 2012, 352-353)

Dreyfus (2001) has summarised this as an 'optimal grip', allowing us to engage in, after Heidegger, 'skilful coping' with everyday situations (259, 252):

> …in our skilled activity we are drawn to move so as to achieve a better and better grip on our situation ... acting is experienced as a steady flow of skilful activity in response to one's sense of the situation. When one's situation deviates from some optimal body-environment gestalt, one's activity takes one closer to that optimum and thereby relieves the 'tension' of the deviation. One does not need to know what that optimum is in order to move towards it. One's body is simply solicited by the situation to lower the tension. (255)

Thus, instead of the fully intentional agent acting on the environments, it is the world that draws the body in through its 'solicitations'.

Just as the body is not the body as theorised by biology or psychology, similarly its unity with the world cannot be fully captured by a concept, method, or measurement. Instead, we 'are involved in the world, and with others, in an inextricable tangle' (Merleau-Ponty 2012, 528). Merleau-Ponty's embodiment philosophy is one that defies formal description. Aspects of it are necessarily holistic, vague, and indeterminate. This insistence on an inherent ambiguity and vagueness highlights the limitations of a science and a philosophy that aims at complete knowledge.

Merleau-Ponty's embodiment philosophy instead offers a re-imagination of the human that is simultaneously integral and of the world, in which the body is central and cannot be reasoned away. It is the core subject-object through which the world is experienced, through which the social is enacted, and the world is modified. It is through the body that all else gains significance and meaning for us (2012, 273). Thus, for Merleau-

Ponty, by ‘remaking contact with the body and with the world’, we will also ‘rediscover’ ourselves (2012, 239).

Merleau-Ponty’s philosophy shares many of the characteristics and features of posthumanist theory. Both attempt to move beyond the inherited dualism and assumptions of liberal humanism and attempt to redefine the relationship between humans and nature in a less hierarchical and more harmonious way. However, Merleau-Ponty’s account of perception does not dissolve into a network, in which boundaries disappear and the entities within it are inert and interchangeable. The ‘body-subject’ remains imbued with life and meaning, and it remains deeply personal and carnal. In the end, there is always the body, ‘a flesh that suffers when it is wounded’ (Merleau-Ponty 1968, 137).

While in *Phenomenology of Perception*, Merleau-Ponty (2012) focuses on perception, in *The Visible and the Invisible*, he further elaborates his theory into a ‘chiasmic ontology’ ‘as an alternative to humanism, naturalism or theology’ (1968, 274). The ‘chiasm’ represents a novel metaphysical expression of our worldly entanglement, or ‘intertwining’, in which we are involved with the world in a relation of ‘pre-established harmony’ (1968, 262). The body is now expressed as the ‘flesh’, a term that incorporates the ontological continuity, or kinship, between the sensing and the sensed, subject and object, self and the world:

> Flesh lines and even envelops all the visible and tangible things with which nevertheless it is surrounded, the world and I are within one another, and there is no anteriority of the *percipere* to the *percipi*, there is simultaneity. (1968, 123)

Merleau-Ponty’s embodiment philosophy thus represents a radical departure from the characterisation of the human liberal subject, indeed from the entire metaphysical humanist project based in its privileging of human reason and a naturalistic ontology. Although Merleau-Ponty does focus on the human, his philosophy does not end at the boundary of an individualistic view of the human, nor does he make assumptions of a hierarchical and dominant form of human exceptionalism. Indeed, his concept of bodily perception, experience, interiority, reflexivity, and culture are extended beyond the human, leading scholars to designate his philosophy as ‘protoecological’ (Westling 2010). Yet, the living body always remains central to his philosophy, resisting abstraction, and generalisation.

Merleau-Ponty’s phenomenology has received mixed responses in anthropological theory. The phenomenological concepts of ‘embodiment’ and ‘situatedness’ have been used to argue for subjectivity and agency beyond the human (Suchman 2007; Hayles 1999; Ingold 2000; Bennett 2010). Others have used Merleau-Ponty’s concept of embodiment to articulate alternative relations without erasing the category of the human altogether. Anthropologist Michael Jackson (1983) has called for the use of Merleau-Ponty’s concept of the ‘lived body’, along with Bourdieu’s concept of ‘habitus’ to ‘move away from the unduly abstract semiotic models which have dominated anthropological research’ and a ‘tendency to interpret embodied experience in terms of cognitive and linguistic models of meaning’ (327). Instead, Jackson focuses on the ‘embodied character of lived experience, such as movement, interactions and bodily praxis, a view that is consonant with indigenous and pre-literate understandings’ (1983, 339). As Jackson observes, many of his own ethnographic insights are derived from practical, physical engagements with his informants:

> Many of my most valued insights into Kuranko social life have followed from comparable cultivation and imitation of practical skills: hoeing on a farm, dancing (as one body), lighting a kerosene lantern properly, weaving a mat, consulting a diviner. (1983, 340)

This ‘practical mimesis’, for Jackson, holds the potential to ground a common understanding between people: ‘While words and concepts distinguish and divide, bodiliness unites and forms the grounds of an empathic, even a universal, understanding’ (1983, 341).

More recently, Jackson (2002) has focused more on the phenomenological concept of ‘intersubjectivity’. Through this lens, he explores how humans extend subjectivity into the extra-human world, in contrast with what he has labelled the ‘cognitive schemata and communicative “rationality”’ that he identifies with the anthropology of human-machine interaction, particularly Suchman’s early fieldwork (2002, 333). Jackson (2005), and later with Albert Piette (2015), proposes a new subfield of ‘existential anthropology’, which recognises that life is ‘irreducible to the terms with which we seek to grasp it’ (2005, 9). It therefore ‘does not reduce the human to a specific assemblage of social, cultural, psychological, historical, and biological characteristics’ (Jackson and Piette 2015, 25). Ultimately, then, existential phenomenology is:

> …a method for exploring the tension and dialectic between immediate and mediated experience, reducing reality neither to some purely sensible mode of being nor to the theoretical language with which we render existence comprehensible. (2015, 11)

Similarly, Csordas (1990) combines Merleau-Ponty’s embodiment philosophy with Bourdieu’s practice theory to argue for a ‘paradigm of embodiment’ in anthropological theory, in which ‘the body is not an *object* to be studied in relation to culture, but is to be considered as the *subject* of culture, or in other words as the existential ground of culture’ (2).

Although Taussig (1993b) does not draw on Merleau-Ponty, at least explicitly, we can see parallels between his articulation of the second meaning of ‘mimesis’ with Merleau-Ponty’s concept of embodied perception: it is ‘a palpable, sensuous, connection between the very body of the perceiver and the perceived’ (21). Taussig asks why vision is privileged, while ‘other sensory modalities are, in Euro-American cultures at least, so linguistically impoverished yet actually so crucial to human being and social life’ (1993b, 26). These include ‘tactility and tactile knowing, and what I take to be the great underground of knowledges locked therein’, as well as that which is ‘conveyed in the mysterious jargon of “proprioception”, but also … the virtual wordlessness of pain’ (1993b, 26). In this ‘corporeal understanding’, he argues, ‘you don’t see so much as be hit’ (1993b, 30). For Taussig, as for Merleau-Ponty, in perception ‘the senses cross over and translate into each other’ (1993b, 57).

However, more often, both Merleau-Ponty’s work and the broader phenomenological project has been the subject of criticism in anthropological theory, especially in two key areas. One of these is the argument that, methodologically, it focuses attention exclusively on subjectivities and consciousness, or on the empirical ‘body’, to the detriment of social, discursive and structural factors (Desjarlais and Throop

2011). Nonetheless, as we have seen, although Merleau-Ponty's work does centre on the perception, and thus the experience of the individual person, it is not the disembodied subjectivity, nor the 'body', of empirical science, that many of these critiques suppose. Merleau-Ponty's concepts of the body, and subsequently also the 'flesh' and the 'chiasm', are as relational as they are individual. Nonetheless, it is true that a focus on individual experience may lead to ignoring those factors that are not apparent in the experience. This critique thus highlights the need to understand the experience in the context of wider historical and structural forces.

The second key criticism comes primarily from fellow phenomenologist Emmanuel Levinas (1994) and is phenomenology's perceived neglect of the problem of 'alterity'. As we have seen, Merleau-Ponty's embodiment philosophy, as well as his subsequent chiasmic ontology, presumes a foundation of common understanding. In contrast, Levinas (1999) insists on a 'radical alterity', or absolute otherness, 'such that I cannot presume anything about the other from my experience of myself; indeed, I cannot know the other at all' (111). For Levinas (2002), it is the face of the Other that is prior even to experience, and remains an irreducible alterity (515). It cannot be understood as a thing, as something that can be contained, comprehended and encompassed, or incorporated into the self, rather it 'commands a gathering – or a proximity… more ancient and aware than knowledge or experience' (2002, 515, 534). For Levinas, the face is the ground of ethics, a fundamental responsibility to the other person that precedes ontology (2002, 517-520). Although Merleau-Ponty does reject the idea of an Other that is forever inaccessible and incomprehensible, as a number of scholars have pointed out, his concept of embodiment is not closed to alterity (Lueck 2012; Reynolds 2002). Nevertheless, Levinas' critical interjection reminds us that despite a common embodied existence, we cannot presume that the Other is fully knowable to us, or that they can be absorbed into our self-identity.

Pietz's Fetish Theory

As we have seen, Pietz (1985; 1987; 1988) develops an account of the fetish concept that has been responsible for a revival of the concept both within anthropology and beyond. In this study, I draw on a less-well known aspect of Pietz's work, his 'preliminary theoretical model of the fetish' (Pietz 1985, 7), developed through an historical and linguistic analysis of the fetish concept as it has been used in the literature since the 16th century. Pietz's model is developed from recurrent themes in the fetish discourse, and has four aspects: territorialisation, historicisation, reification, and personalisation. In this study, I use Pietz's model of the fetish, underscored by the embodiment phenomenology of Merleau-Ponty, to integrate a historical and structural reading with one that takes both materiality and embodiment seriously.

Unlike previous uses of the concept, Pietz understands the fetish, not as a fixed object with a prior model or truth, but instead as a 'radically historical object that is nothing other than the totalized series of its particular usages' (1985, 7). Pietz finds that the fetish concept emerged from a novel social formation in the cross-cultural spaces on the West African coast during the 16th and 17th centuries. In these sites, Christian feudal, African lineage, and merchant capitalist, each representing radically different social orders and value systems, encountered one another and interacted over the course of several centuries (1985, 5-6). Pietz identifies fetish discourse as consisting of three stages. The first stage culminated in a general theory of primitive religion at the start of the eighteenth century. The second stage was elaborated as a general Enlightenment theory and adopted by philosophers in the late eighteenth century. The third stage was the fetish as it was used and developed in twentieth-century popular and social scientific discourses (1985, 5).

Like others have done, Pietz acknowledges that the fetish is overwhelmingly used as a pejorative term: 'fetish discourse is a critical one about the false objective values of a culture from which the speaker is personally distanced' (1985, 14). Despite this, he argues, the emergence of the idea of the fetish in these specific historical and geographical conditions 'marks the breakdown of the adequacy of the earlier discourse' (1985, 6). This, he argues, 'represents the emerging articulation of a theoretical materialism quite incompatible and in conflict with the philosophical tradition' (1985, 6). Specifically, this new problematic concerns 'the capacity of the material objects to embody – simultaneously and sequentially – religious, commercial, aesthetic and sexual values' (1985, 6). Pietz's fetish is thus not an 'indigenous' concept, rather it is a concept developed by the West to explain relations that they felt but could not describe.

Pietz' fetish model is drawn from this analysis, characterised by four recurrent themes in the discourse essential to the concept. These are: 'historicisation', 'territorialisation', 'reification', and 'personalisation', see Figure 2 below.

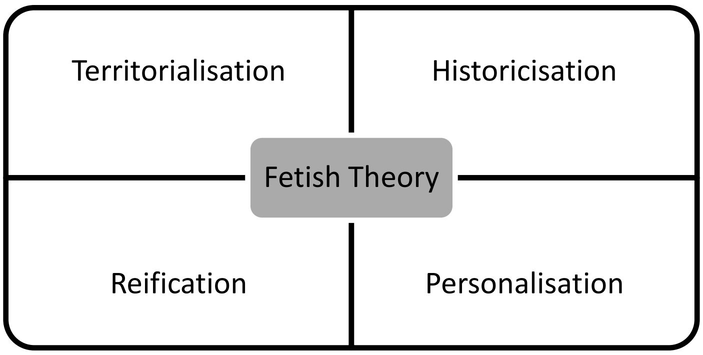


Figure 2: Characterisation of William Pietz's 'Model of the Fetish' 1985

First, the fetish is 'historical', it is the fixation of a unique and unrepeatable originating event bringing together previously heterogeneous elements into a novel identity (1985, 7). Second, the object is 'territorialised', it is an irreducible material object in geographical space unifying otherwise unconnected multiplicities (1985, 12). Third, it is 'reified', it is 'recognizable as a discrete thing (a *res*) because of its status as a significant object within the value codes proper to the productive and ideological systems of a given society' (1985, 12). Finally, and significantly for the phenomenological orientation of this study, the fetish is 'personalised': it 'evokes an intensely personal response from individuals' (1985, 12), with an emphasis on their 'embodied status' (1985, 10).

For Pietz, previous attempts to classify the fetish and dismiss it 'as proper object with its own singular significance' are flawed (1985, 5-6). He is critical of both ‘particularist’ or ‘universalists’ accounts. Particularists, he argues, dismiss the fetish as a ‘corrupt genus’ obscuring true meaning. This appears often to be the case in anthropology, in which theorists from Tylor to Mauss, opt instead for more palatable concepts such as ‘animism’ in Tylor’s case or, for Mauss (2002), the concepts of ‘*mana*’ and ‘*hau*’. Universalists, on the other hand, attempt to subsume the fetish into ‘alleged human universal tendencies’ such as privileging phallic symbolism or errors of logical type (Pietz 1985, 6). Further, Pietz is critical of Marxist and structuralist interpretations of the fetish, which ignore the relation of the fetish to the embodied person, of which, he argues, labour theory of value is only one example (1985, 10). Instead, taken as a whole, the fetish is a ‘locus of a sort of primary and carnal rhetoric of identification and disavowal’ (1985, 14). Specifically, ‘each fetish is a singular articulated identification … unifying events, places, things, and people, and then returning them to their separate spheres’ (1985, 13). The ‘structured relationships that are thereby established’, constitute, for Pietz ‘the phenomenological fabric … of immediate prereflective experience’, specifically, ‘the “flesh” in Merleau-Ponty’s sense in *The Visible and the Invisible*’ (1985, 13).

More recently, in part in response to Pietz’s scholarship, the fetish as a concept has experienced a resurgence in anthropological theory. Although Taussig does not explicitly build on Pietz’s model, he does refer to his genealogy (Taussig 1993a). As we have seen, Taussig (2010; 1993b) adopts a primarily Marxian interpretation of the fetish in which the fetish is a reification, concealing exploitative relations, which he contrasts with a ‘pre-capitalist’ fetishism. David Graeber (2005) builds on Pietz’s articulation of the fetish as emerging from an inter-cultural confrontation to show how the fetish may be understood as ‘social creativity’ or ‘the creation of new social forms and institutional arrangements’ (407). Graeber argues that creativity is ‘not an aspect of the objects at all’, rather it is ‘a dimension of action’, thus the fetish object is ultimately ‘only the medium’ (2005, 425) and its materiality is of little consequence. Latour (2010) also explicitly builds on Pietz’s work, maintaining that it supports his symmetrical anthropology, particularly his insistence on the continuity between ‘primitive’ artefacts and the practices of science and technology. Latour is highly critical of what he calls the 'anti-fetish' tradition, or those who see the fetish simply as reification, revealing a hidden truth, or fact, behind the fetish. Instead, he argues for an approach that emphasises objects' historicity, paying attention to how events generate new agencies, rather than pre-supposing any inherent agencies (Sansi Roca 2015).

These diverse theorists recognise the potential of the fetish to inquire into alternative formulations of how humans relate to the world. As Taussig (2010) articulates, by using the concept of 'fetishism' to investigate our own 'modern' culture, we might 'become sensitive to the suppositions and ideological character of our own culture's central myths and categories, categories that grant meaning as much to our intellectual products as to our everyday life' (6). However, despite embracing the fetish concept, none of these theorists have responded to Pietz's insistence on both the territoriality of the fetish or the primacy of the body in the encounter as I do in this study.

Embodiment and the Fetish

By combining a Merleau-Pontian embodiment perspective of the human, with a conceptualisation of the robot as fetish-object as articulated by Pietz, I have constructed a novel framework for understanding human-object relations through the prism of robotics research. Pietz's model offers a mechanism to ensure that the need to rebalance and reintegrate the body and the object does not blind us to that which is not experienced in the face-to-face encounter, in particular structure and alterity. The model recognises the complexity of human-object relations and offers a solution beyond the dialectic and circular arguments of structure versus agency, nature versus culture, and between materiality and semiotics. It recognises the historical and constructed nature of the fetish object, as well as its irreducible materiality. It also acknowledges the social and symbolic value of the reified object. Uniquely, and crucially for my research, it emphasises the centrality, indeed the primacy, of an embodied connection between the self and the fetish object, quite distinct from the role of the fetish as a social significant object. As we have seen, Pietz explicitly references Merleau-

Ponty's concept of the 'flesh' to articulate the relationships that are established by the 'singular articulated identification' between the embodied person and the fetish object (Pietz 1985, 13). Furthermore, the embodied experience in the 'singularly fixating encounter' is:

> …'stripped of all symbolic value' and, paradoxically because of this degradation from any recognizable value code, becomes a crisis moment of infinite value, expressing the sheer incommensurable togetherness of the living existence of the personal self and the living otherness of the material world. (1985, 12)

In this study, I consider whether contemporary humanoid robots can be considered a fetish according to Pietz's model, and whether, as Pietz claims, the fetish might represent, not just humanoid robots, but technology in general, or even the 'truth' of the 'total collective material object' (1985, 14). In so doing, I will also examine the utility of Pietz's fetish concept as a novel theoretical materialism, divergent from the philosophical tradition. The four chapters that make up the main body of this dissertation accord with Pietz's four aspects of the fetish.

First, the chapter Historicisation considers the historical aspect of the robot-as-fetish. In this chapter, I also draw on the work of Deleuze (1994) and Foucault (1977; 1980; 2002; 1990) to write a 'history of the present', tracing the continuities and discontinuities of the robot-as-fetish. I also examine the futures that are implicated in its identity. The next chapter, Territorialisation, considers the material fact of the robot. Specifically, it focuses on the robot's 'untranscended materiality': investigating it as an irreducible material object in geographical space unifying otherwise unconnected multiplicities and subject to temporal and spatial constraints. The robot is a concrete reference-point, and a site in which specific cultural imaginaries are materialised. As well as the robot object, this chapter also investigates the spaces in which robotics research takes place using the concept of 'space' as articulated by Lefebvre (1991).

In the next chapter, Reification, I consider the robot as a valuable object according to the institutions and the productive and ideological systems specific to Euro-American imaginaries. Specifically, in this chapter I focus on the image of machine intelligence as analogous to human intelligence and its role in developing and maintaining cultural identities. First, drawing on the concept of the fetish as articulated by Marx (2008) and Taussig (2010), in this chapter I look at how this image is used to elide the human work of animating machines. Second, drawing on Taussig's (1993b) elucidation of the concept of 'mimesis', I focus on how the symbol of the robot and the performance of animacy may also be used to play, to enchant, to learn, to educate, and to balance social dynamics.

In the Personalisation chapter, I explore the personal way that the robot-as-fetish is experienced by the embodied person. The 'personal' experience of the fetish object is prior, in Pietz's words 'even more basic' and 'first of all', to the other three themes, and the conditions for their existence (Pietz 1985, 10, 11). In this chapter, I investigate the experiences of those who encounter the robot, as well as other entanglements between the embodied person and the robot. I argue that the lens of 'personalisation', and its characterisation of the experience of the fetish encounter, reveals an under-theorised and little recognised, yet vital, aspect for understanding human-object relations, to which Merleau-Ponty's embodiment phenomenology is key. In order to further elaborate and develop this theme, Taussig's (1993b) reading of Walter Benjamin's concepts of 'mimesis' and 'alterity' is used. Taussig's concept proves a particularly useful lens through which to view encounters with robots, while also allowing us to consider the identifications, as well as the disavowals, between the person and the object. Finally, in this chapter, I contrast the ways in which the encounter is typically theorised, as anthropomorphic projections of mental models and using Theory of Mind (ToM), with Merleau-Ponty's embodiment concept.

In Chapter 7, I draw together the four strands to develop a full account of the robot as fetish, exploring how the four aspects are brought together to offer a distinct insight into the human-object relation. In this chapter, I also articulate some of the implications of this novel framework for understanding the category of the human for anthropology. Ultimately, I argue for reconfiguring the human as simultaneously relational *and* embodied.

Chapter Three: Historicisation

Stevie, the robot, is an historical object. He did not spring fully formed from the imaginations of his creators, nor is he the product of a systematic and methodological application of science. The events that led to Stevie's creation were neither inevitable, nor predictable. Instead, the facts of Stevie's existence: his appearance, the way he is constructed, and the technologies that are used to implement his capabilities, are the result of specific historical events and social conditions. By excavating Stevie's material structure, we can uncover layers of history, revealing the materials, technologies, histories, political, social and economic influences, theories, ideologies and assumptions that are embodied in his form.

The robot-as-fetish, a 'composite fabrication', it is 'the fixation or inscription of a unique originating event that has brought together previously heterogeneous elements into a novel identity' of 'articulated relations' (Pietz 1985, 7, 8). This 'event' is the encounter, which 'brings together and fixes into a singularly resonant unified intensity … a particular object or arrangement of objects, and a localised space' (1985, 12). Pietz's theme of 'historicisation' is informed by the fetish concept in the work of both Deleuze (1994) and Foucault (2007), as well as their articulation of history as a dimension of the present (Deleuze 1994; Foucault 1977). These histories are not linear, progressive, and inevitable, instead they are discontinuous, emergent, and unique. Foucault's archaeological (2002), and subsequent Nietzschean-inspired genealogical approach to history (1977; 1980; 1990; 2020), focus on present-day phenomena and practices to uncover what he terms a '*dispositif'*, or apparatus of power-knowledge. In such an arrangement, relations are established between 'a thoroughly heterogeneous ensemble consisting of discourses, institutions, architectural forms, regulatory decisions, laws, administrative measures, scientific statements, philosophical, moral and philanthropic propositions' (Foucault 1980, 194). For Foucault, examples of *'dispositifs'* are the historical construct of sexuality (Foucault 1990), the reformative prison system, and the American death penalty (Foucault 2020). Similarly, as Pietz shows, the ‘heterogeneous components appropriated into an identity by a fetish are not only material elements; desires and beliefs and narrative structures establishing a practice are also fixed (or fixated) by the fetish (Pietz 1985, 7).

This chapter will thus focus on the ‘historicisation’ aspect of the robot-as-fetish, in which I trace contingent processes and power arrangements, chance events, design decisions and material circumstances that led to the identity as robot, and as fetish. I will also consider the futures that are projected by the robot object, as well as by the wider community of interested parties, including techno-utopians, futurologists, science fiction writers, transhumanists and posthuman scholars. Recognising the robot as an historical object also focuses attention on the transformed temporalities, geographies, materialities, practices and social arrangements that are set in motion by the novel identify of the robot. Some of these are further developed in the subsequent chapters of Territorialisation, Reification, and Personalisation.

The Age of the Automaton

Robots have a lineage stretching back far into pre-history of human attempts to represent and reproduce life and human likeness; for aesthetic purposes, for entertainment, to inspire religious practice, and in the pursuit of craft, knowledge and understanding. However, human-likeness as ‘technology’ is a more recent development. The view of technology as a symbol of advancement, and the progress of civilisation, dates back to Classical Greece. In her study of luck and ethics in Greek tragedy and philosophy, the philosopher Martha Nussbaum (2001) shows how themes of reducing risk or ‘exposure to luck’ dominated Greek philosophy. Plato’s *Protagoras* paints a picture of humans in nature as isolated, exposed, and miserable. Prometheus, in his kindness, grants these pathetic creatures the gift of the *technai*:

> House-building, farming, yoking and taming, metal-working, shipbuilding, hunting; prophecy, dream-divination, weather-prediction, counting and calculating; articulate speech and writing;

> the practice of medicine; the art of building dwelling-places… (2001, 90)

Luck, or *tuchē,* is contrasted with *techne*. The former, according to Democritus, is associated with witlessness, an excuse people give for their own lack of resourcefulness. *Techne* on the other hand, is a ‘deliberate application of human intelligence to some part of the world’, making human existence safer, more predictable and allowing control over contingency (95).

Automata, or self-operating machines, seemed the ultimate realisation of technical advancement. An instrument that could ‘accomplish its own work, obeying or anticipating the will of others’, wrote Aristotle in 320BCE, could even replace the need for that pre-eminent and indispensable instrument, the slave (Aristotle 2000, 7). However, it is Hero of Alexandria (10BCE – 70BCE) who is credited with building the first known automaton, loosely defined as a machine capable of moving and acting by itself, which consisted of a self-powered three-wheeled cart (Sharkey 2007). In the following years, automata emerged with mechanisms powered by steam and water, cogged gears and pulleys.

Throughout the Middle Ages, efforts at both imagining and building artificial life continued in varying and largely isolated ways until the early Renaissance period. Automata were popularised as demonstrable feats of engineering and human ingenuity, but their ability to inspire awe and wonder in those that encountered them led them also to be used for religious purposes, as well as for entertainment. Sculptor Elizabeth King (2007) writes about her encounter with a small automaton in the image of a monk dating from around 1560. She theorises that the monk may have been constructed by engineer Juanelo Turrino for King Philip II of Spain (2007, 266). She describes her own encounter with the monk, as well as the experiences of others who encounter it, as ‘intimidating’, invoking a ‘primal anxiety’: ‘when this machine heads in my direction on a table, my animal flight urge stirs’ (2007, 274, 277). This response leads King to imagine what it must have felt like for those who encountered the automaton at the time, ‘[c]ould it even momentarily have been perceived as alive, just for the space of a shudder?’ (2007, 274).

It is little wonder, then, that these displays, along with the emerging new sciences, shaped emergent conceptions of the human. If automata, with their simple mechanisms, could demonstrate life-like behaviour, it seemed likely that all life could be reproduced in the same way. The idea that mathematics and the ideal world of geometrical forms is the real and true form of reality, with the material world of daily life its mere shadow was first postulated by Plato in the fourth century BCE (Riskin 2007). However, with the rise of automata, this idea really took hold. These developments led thinkers such as René Descartes to hypothesise the universe, and all it contains, as mechanical. The artificial replication of human-like behaviour and intelligence seemed imminent. The image of the human was thus co-emergent with machines, exemplified by figure of the automaton.

In her history of AI, Pamela McCorduck (2004) relates how, in 1673, philosopher and mathematician Gottfried Wilhelm Leibniz set about developing as unambiguous formal language to prove that all human activity could be reduced to mathematical calculation. While he did not succeed in this task, he did manage to invent the binary system, and a digital mechanical calculator composed of symbols, called the ‘Step Reckoner’, a key step in the later development of the computer. Although it did not succeed in replicating ‘intelligent’ behaviour, Leibniz theorised that it might instead be capable of doing the slave-like work of calculation that is ‘unworthy of excellent men’ (2004, 26). Throughout the 18th century, philosophers Spinoza, Hobbes, Locke, Kant and Hume, and scientists La Mettrie and Hartley all tried and failed to formulate laws of human thought (2004, 526). A proliferation of automata accompanied these scientific attempts.

For historian Jessica Riskin (2003), the point of origin for our contemporary way of thinking about automata is French inventor Jacques de Vaucanson’s ‘Defecating Duck’, first displayed in Paris in 1738. Unlike previous automata, the mechanical Duck was not simply designed to amuse, but was also an exercise in experimental philosophy, testing which aspects of living creatures could be reproduced mechanically, and to what degree (2003, 601). Vaucanson's Duck appeared to ingest food and excrete it in an altered form. The Duck was simultaneously a scientific, philosophical, and commercial venture. Between 1770 and 1854, another famous automaton, the chess-playing 'Mechanical Turk', was exhibited across Europe and the Americas (Standage 2004). Ultimately, however, both Vaucanson's Duck and the Mechanical Turk were revealed as frauds. In the artificial duck, the food and the resulting excretion were revealed to be completely different substances. Similarly, the Mechanical Turk's mechanism was shown to consist, not of a highly sophisticated mechanised human, but an actual human, hidden underneath, operating the machinery.

Nonetheless, throughout the 18th and 19th centuries, increasingly intricate and ingenious automata that could write messages, play musical instruments, and repeat phrases, were built by engineers and clockmakers. There was, however, also a darker side to these figures. Instead of simply representing the genius of mechanical invention, these automata were often also represented as 'exotic' or uncanny characters, often female or Black, as animals, acrobats or magicians (Norton-Wise 2007). There was also an emerging discomfort with the developing technologies. Mary Shelley's (1993) *Frankenstein*, which was first published in 1818, recounts a tale of horror in which a scientist creates a human-like, yet monstrous, creature by stitching together the remains of corpses and reanimating them using electricity.

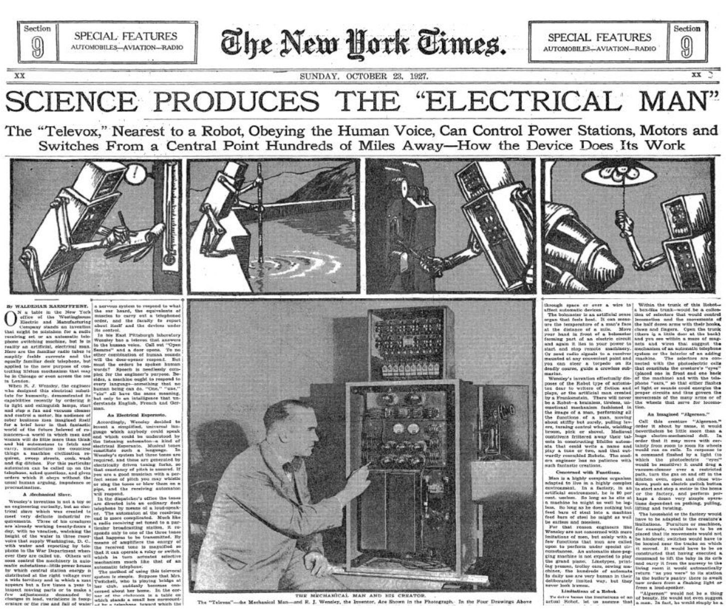
By the late 17th and early 18th century, the creation of machines that could emulate life stalled and shallow, human-like imitations had been relegated to the realm of entertainment and curiosity. At the same time, however, the project of replacing human workers with machines, went into overdrive. Around 1760 saw the beginnings of the Industrial Revolution, and increasing automation, particularly in the cotton and iron industries. The Spinning mule or Spinning 'Jenny' was invented in 1779 to automate the previously manual job of spinning cotton. Now, a greater variety of textiles would be available more cheaply to many more people, with significant profits for factory owners. This also radically transformed the conditions of the workers, and the relation to their work and its products. These developments inspired novel, and disparate, responses from theorists. Friedrich Engels' father owned a cotton spinning business based in both Germany and England, and he had seen the effects on workers first hand. This would lead him, along with Karl Marx, to document the resulting unemployment, suppression of wages, and deskilling (Boyer 1998).

An entirely different response came from engineer Charles Babbage. For Babbage, observations of the workers at their machines on the factory floor suggested a natural 'unequal division among machines', in which those at the top produce power and those at the bottom produce mechanism, existing 'merely to transmit force and execute work' (Babbage 1832, 16). Babbage developed this insight into a universal classification system that he believed to apply across the board to machines, to social organisation, and to individual humans (Norton-Wise 2007). As Leibniz and others had previously attempted, Babbage started to formulate this theory into a universal language. He developed machine that he called 'the Analytical Engine', which, as well as carrying out calculations would be 'capable of analysis and tabulating any function whatever' (McCorduck 2004, 31). Although the engine was never built, he had established the principle of what would become the modern digital computer (Dreyfus 1992, 70).

The Rise of the Robot

The world's first 'robot' was a product of fiction rather than engineering. It first appeared in a 1922 play called R.U.R., or 'Rossum's Universal Robots', by Czech playwright Karel Čapek (2012). The word 'robot', which was coined by Čapek, comes from the Czech word '*robota*' meaning slave, work, or drudgery. In the play, artificial people are created using organic matter so that they might produce goods at a fraction of the cost of human labour. The 'robots' eventually rise up against their creators and overthrow humanity. The play travelled internationally, including in New York in 1922 and in Tokyo, Japan in 1924 under the title '*Jinzð Ningen*' (Artificial Human). Despite its ambivalent message, it captured something in the public imagination, sparking an enduring fascination with robots. During these early post-War years, science fiction as a distinct genre also took off, coined by published Hugo Gernsback in the in the first issue of the 1926 of science fiction magazine, *Amazing Stories*: 'By "scientifiction"', Gernsback writes, 'I mean the Jules Verne, H. G. Wells and Edgar Allan Poe type of story—a charming romance intermingled with scientific fact and prophetic vision', in (Westfahl 1992). In 1927, Fritz Lang's feature length film *Metropolis* was released, combining a dystopian, expressionist view of the oppressive and destructive potential of technology with a utopian 'technology cult of the *Neue Sachlichkeit* and its unbridled confidence in technical progress and social engineering' of the Weimar republic (Huyssen 1986, 223). The film features a stark division between an above-ground futuristic technological utopia, inhabited by the city's elite, and a below-ground dystopia, featuring dark chambers and factories inhabited by the city's workers. The film also portrayed a new image of the robot, or '*Machinenmench*' (Machine person), an evil female android, Maria.

The same year that *Metropolis* was released, in the US, Westinghouse Electric Corporation unveiled the 'world's first' humanoid robot, Herbert Televox. The robot owed its existence to an unlikely confluence of events. In the mid-1920s, one of the engineers at Westinghouse, Roy James Wensley, designed a device for changing switches remotely using sound. A tuning-fork oscillator created a particular frequency at one end, generating a code that was sent via telephone line to a receiver unit, which processed the code, and opened or closed a particular switch in response. He called it the 'Televox' (Marsh 2018). A review of the device featured in *The New York Times*, where it was described it as an 'electrical man', a mechanical robot or slave that 'obeys without the usual human arguing, impudence or procrastination' (Kaempffert 1927). Although in the main body of the article the author takes pains to distance Wensley's machine from an actual robot, which would be a 'fantastic creation' for which, 'the modern engineer has no patience' (Kaempffert 1927), a new bar had been set.



Section 9 SPECIAL FEATURES AUTOMOBILES—AVIATION—RADIO **The New York Times.** SPECIAL FEATURES AUTOMOBILES—AVIATION—RADIO Section 9

SUNDAY, OCTOBER 23, 1927.

SCIENCE PRODUCES THE "ELECTRICAL MAN"

The "Televox," Nearest to a Robot, Obeying the Human Voice, Can Control Power Stations, Motors and Switches From a Central Point Hundreds of Miles Away—How the Device Does Its Work

THE MECHANICAL MAN AND HIS CREATOR.

Figure 3: New York Times Article announcing the 'Televox' (The New York Times 1927)

When Wensley was asked to present his creation at the opening of the Level Club in New York City, he knew that people were expecting more than a 'little frame of relays and wires' (Letizia 2013). Taking his cue from the article, Wensley decided to create a humanoid casing for his machine and give it a name, Herbert. While its torso comprised the Televox control box, it was set into a wooden frame with a head, arms, and legs. The 'robot' could accept a telephone call by lifting a receiver, and then 'act' depending on which signal was sent. It also made noises, specifically buzzes and clicks and waved its arms (History Computer 2021).

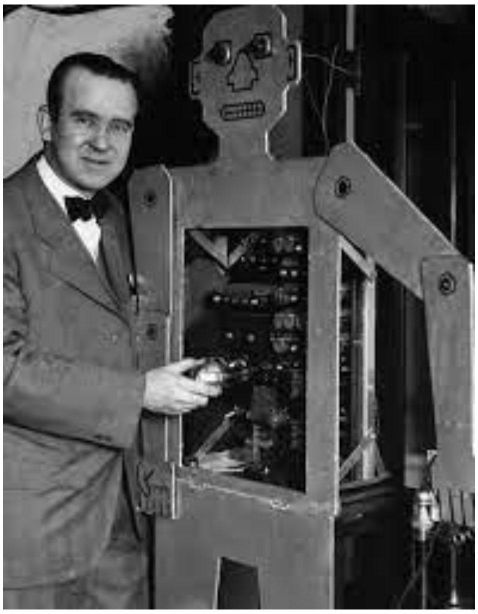


Figure 4: Herbert Televox (Botsolvers 2021)

Herbert Televox became an international sensation. In a manner that would pre-figure the dynamic between robotics researchers, media and funders for the coming century, Wensley benefitted from the often vastly exaggerated reports of his robot's abilities in the media by attracting funding, while also spending much of his subsequent career trying to correct and downplay the exaggerations (Marsh 2018).

Other robots soon followed. In 1928, a Japanese robot *Gakutensoku*, or 'Learning from the Laws of Nature', was created by biologist Makoto Nishimura. *Gakutensoku* was a very different kind of robot to those being built in Europe and the US. It was enormous, at over three meters tall, and had a gold upper body swathed in a toga and seated on an altar. It evoked not so much a feat of science, but a divinity (Hays 2013). It was powered by a 'ventricle system' consisting of rubber tubes and air compressors allowing it to change its facial expressions, puff its cheeks, move its chest, and simulate writing. For Nishimura, the robot was a part of nature, and he is quoted as saying 'if one considers humans as children of nature, artificial humans created by the hand of man are thus nature's grandchildren' (Whelan 2011). Unlike the portrayal of robot in the US and Europe, Nishimura is reported as insisting that his robot is not a 'slave', '[i]t would be sad to find joy in making a slave-like android that simply copies humans, which are masterpieces of the earth', reported in (Hays 2013).

Figure 5: Gakutensoku (Wikimedia 2021)

Subsequently, however, robotics in Japan stalled for the next two decades. However, in the years following World War II, Japan was occupied by Allied forces and subject to US cultural influences (Robertson 2007). In 1951, a manga and anime character named *Tetsuwan Atomu* (Mighty Atom), known as 'Astroboy' in the West, became Japan's most famous robot. Astroboy was created by Japanese artist Osamu Tezuka and was seen as a post-nuclear technology. The idea was that robots should help to bring out the best of our humanity in response to the nuclear age. The development of robots in Japan thus follows a parallel, but very different, course to that of Euro-American imaginaries. As anthropologist Christal Whelan (2014) observes, 'robots in Japan were never conceived as enemies; they are friends and companions for life, tied intimately to humans in what anthropologists describe as fictive kinship' (85).

Computing and Cybernetics

In 1937, while studying for his PhD at Princeton University, British mathematician Alan Turing published a paper 'On Computable Numbers, with an application to the *Entscheidungsproblem*' (Turing 1937). In it, he proposed a hypothetical device, later known as a 'Turing Machine', which could (in theory) solve any problem if it was described abstractly, as an algorithm. This suggested that any logical process done by a human could also be done by a machine. This established not only the enduring human-computer analogy, but also led to the creation of the modern digital computer. During World War II, Turing led a group of British code breakers to develop a device to unscramble German codes and cipher systems, leading to the development of the first British electronic digital programmable computer called 'Colossus' in 1943 (Russell and Norvig 1995). The following year, the ENIAC (Electronic Numerical Integrator and Computer) was completed at the University of Pennsylvania.

Once again, these new technological breakthroughs brought the promise of developing a machine that could think like humans. In 1942, an invitation-only meeting called the 'Cerebral Inhibition Meeting' was held, focusing on the topic of hypnotism and conditioned reflex. These meetings were attended by anthropologists Gregory Bateson and Margaret Mead, as well as neurophysiologist Warren McCulloch, social scientist Lawrence Frank, psychoanalyst Lawrence Kubie, and physiologist Arturo Rosenblueth (ASC Cybernetics 2021a). At the meeting, Rosenbluth presented on topics such as 'teleological mechanisms', 'circular causality', and 'feedback', which was subsequently published along with co-authors Norbert Wiener and Julian Bigelow in their article 'Behaviour, Purpose and Teleology' (Rosenblueth, Wiener, and Bigelow 1943). The article explores the topic of purposeful machines and proposes that machines could be analysed using the approach of behaviourism from psychology, rather than the traditional functionalist approach of engineering (Hayles 1999, 94).

Following on from this initial meeting, the first of the Macy conferences was held the following year. Its stated goal was to bring together leading thinkers to use the new technological advances to establish ‘a general science of the workings of the human mind’ (ASC Cybernetics 2021b). The conferences would run for the next 11 years, consisting of 160 conferences in total, and led to the establishment of the field of cybernetics. Attendees over the years also included mathematician and engineer Claude Shannon, logistician Walter Pitts, roboticist and neuroscientist William Grey Walter, and mathematician, physicist, and engineer John Von Neumann. Shannon’s work had shown that switching circuits could solve the same problems as Boolean algebra, and McCulloch and Pitts’ (1943) paper had proposed a computational model of a neuron. Together, they appeared to connect mechanism, Boolean algebra, and animal physiology. In 1948, in this book, *Cybernetics*, Wiener (1949) outlined his view of ‘the world as a set of complex, interlocking feedback loops, in which sensors, signals, and actuators … interact via an intricate exchange of signals and information’ (Brockman 2019, 3). The same year, Shannon established the field of ‘Information Theory’, defining information as a probability function with no dimension, no materiality, and no necessary connection with meaning. What united all attendees was their conviction that computing mechanism and human thought were analogous (Hayles 1999).

The claims of the cyberneticians appeared to be supported by the parallel development of small, automated, analogue robotic devices, providing a tangible instantiation of the theories they were proposing. The most famous of these were Walter’s ‘tortoises’, described for their appearance but also, after the mock turtle in Alice in Wonderland, who ‘taught us’, as they might teach us the secrets of organization and life (O’Connell 2000). Walter was a neurophysiologist and developed the tortoises to understand the mechanics of the brain. These small robots, which he names Elmer (ELectroMEchanical Robot) and Elsie had three wheels, blinking lights, and were light sensitive. When they encountered a mirror, they responded by displaying surprisingly lifelike behaviour, which Walter described as ‘flickering, twittering, and jigging like a clumsy Narcissus’ (Walter 1963, 115). Using behaviourist logic, Walter argued that this ‘might be accepted as evidence of some degree of self-awareness’ (1963, 128-129). Walter’s intuition, which be believed the tortoises confirmed, was that simple mechanism could lead to complex behaviours through complex connections.

Wiener and Shannon also had robots. Wiener's robot was a small tricycle cart with two photocell sensors facing to the front. The output from the photocells was connected to the tiller controlling the front wheel. Depending on the output voltage the cart either steers towards the light, like a 'moth' or away from it, like a 'bedbug' (Masani 1990, 211). Shannon's robot was described as an 'electronic rat' that exhibited goal-seeking behaviour and was able to find its way through a maze and appeared to 'learn' by trial and error (Hayles 1999, 63). These physical instantiations of the cybernetic program reiterated and appeared to provide scientific and empirical evidence for the imminent potentiality of formalising human thought and action. It also incorporated both human and physical sciences, integrating them under a unifying 'information science'.

Throughout the 1940s and '50s, science fiction continued to parallel technological fields. One prominent science fiction writer, Isaac Asimov (1950), developed 'The Three Laws of Robotics', first in 1942 and then in 1950. The laws state:

> 1. A robot may not injure a human being, or through inaction allow a human being to come to harm.
> 2. A robot must obey the orders given it by human beings except where such orders would conflict with the First Law.
> 3. A robot must protect its own existence as long as such protection does not conflict with the First or Second Law. (1950, 40)

As we will see, these fictional principles would come to have an outsized influence on the robotics community as they start to grapple with the issues of ethics and regulation for robotics in the next century.

The Birth of Artificial Intelligence (AI)

In 1955, a group of four scientists decided it was time to come together to settle the problem of 'intelligence' for good. They were John McCarthy, assistant professor of Mathematics at Dartmouth College, Marvin Minsky, Harvard Junior Fellow in mathematics and neurology, Nathaniel Rochester, manager of information research at IBM, and Claude Shannon (McCorduck 2004, 111-112). They wrote:

> We propose that a 2-month, 10-man study of artificial intelligence be carried out during the summer of 1956 at Dartmouth College in Hanover, New Hampshire. The study is to proceed on the basis of the conjecture that every aspect of learning or any other feature of intelligence can in principle be so precisely described that a machine can be made to simulate it. An attempt will be made to find how to make machines use language, form abstractions and concepts, solve kinds of problems now reserved for humans, and improve themselves. We think that a significant advance can be made in one or more of these problems if a carefully selected group of scientists work on it together for a summer. (McCarthy et al. 1955, 2)

In a deliberate move to break from both cybernetics and behaviourism, McCarthy proposed this as a new field of ‘Artificial Intelligence’, or AI, simultaneously laying the foundations for a new psychological theory of ‘cognitivism’ (Agre 1997, 80; Nilsson 2010, 73). The Rockefeller Foundation provided funding for the event and other likeminded researchers were invited, including Trenchard More and Arthur Samuel of the IBM Corporation, and Oliver Selfridge and Ray Solomonoff of MIT (McCorduck 2004, 112). Once again, what bound this group together was a firm belief that the essence of what is human could both be replicated by a machine and understood in a formal and scientific way.

Although they did not succeed in their stated goal, they nonetheless established a research programme that would dominate both psychology and attempts to codify human-like intelligence for the next 30 years. Despite this, the biggest technological achievement to come out of the conferences went almost unrecognised at the time (McCorduck 2004, 123). This was the work of two minor participants, Herbert Simon and Alan Newell, who had met when they were both consulting at the military think tank, RAND Corporation. Simon, a professor of political science at Carnegie Tech, became interested in computers through Newell’s work. Their first collaboration, along with computer programmer Cliff Shaw, was on a program that could process symbol structures using ‘heuristics’. The resulting programme, ‘Logic Theorist’ was presented at Dartmouth. Although their contribution is now seen as a seminal moment in the history of AI, it received a muted response from the attendants at Dartmouth at the time. Precisely why is the subject of debate. According to Marvin Minsky, the other participants had not connected their work so closely to human psychology and 'the analogy between human and artificial intelligence was not generally accepted', Minsky in (McCorduck 2004, 157). There was no such reticence from Newell and Simon to their own work. They were convinced that they had succeeded in creating a machine that could think like a human:

> [We] invented a computer program capable of thinking non-numerically, and thereby solved the venerable mind-body problem, explaining how a system composed of matter can have the properties of mind. (Simon 1996, 190)

Further, they claimed their achievement was as central to understanding the human mind in the twentieth century as Darwin's principle of evolution by means of natural selection was to understanding biology in the nineteenth century (McCorduck 2004, 153). They predicted that:

> …there are now in the world machines that think, that learn and that create. Moreover, their ability to do these things is going to increase rapidly until—in a visible future—the range of problems they can handle will be coextensive with the range to which the human mind has been applied. (Simon and Newell 1958, 8)

Logic Theorist would go on to prove 38 of the first 52 theorems in Whitehead and Russell's *Principia Mathematica*, and even find a more elegant proof than one of the originals (Crevier 1993, 46). It has subsequently become known as 'the first AI programme'. It also introduced several concepts which would define AI research problems for the following decades, including reasoning as search, the use of heuristics to guide search, and a new programming language IPL, which would form the basis for McCarthy's AI-programming language, Lisp in 1957. However, despite these breakthroughs in specific areas, Logic Theorist never managed to solve problems outside of the narrow problem area for which it had been designed. Nonetheless, it was Simon who, more than anyone else, pushed and ultimately legitimised the brain-computer analogy and thinking of humans as information processors (McCorduck 2004, 150-151). The direct connection between machines and humans, which prior to the conference had still been shunned by serious scientists, was finally legitimised and AI was firmly linked to science.

In 1958, the US Defence Department's Advanced Research Projects Agency (ARPA) was formed in response to Russian technological advances, such as Sputnik. The agency was interested in AI research and other 'far-out projects as insurance against unwelcome technological surprises' (Moravec 1995, 21). They were particularly interested in two areas of AI: its relevance to human psychology and its physical instantiation in robots. The tight coupling of human and computational thought processes brought about by military interest and the AI programme at Dartmouth, produced not only the field of AI (conceptualised now as symbolic information processing), but also the subsequently dominant paradigm of cognitive psychology. By the 1960s, the science of machines and the science of humans had effectively merged. The symbolic reasoning programme established at Dartmouth would come to dominate AI research. Alternative approaches were side-lined, including early 'connectionist' research, such as Pitts and McCullough's work on neural networks.

The Intelligence Race

In the period after Dartmouth, funding and resources were concentrated in a small set of institutions and researchers in the US. Over the following decades, ARPA, now DARPA, poured millions of dollars into the four main research centres: MIT, Carnegie Mellon (then Carnegie Tech), Stanford, and Stanford Research Institute (SRI) with its stated aim to 'fund people, not projects', allowing them to spend the money in any way they wanted to (Crevier 1993, 65). McCorduck writes, 'though each project had its own flavor, the general aim was the same—to produce some sort of independent agent that would function in the real world, or at least a somewhat impoverished real world' (McCorduck 2004, 261).

At Carnegie Tech, Newell and Simon's next project was to try to develop a machine that could demonstrate a more 'general' human intelligence. Thus, in 1959, they followed up the 'Logic Theorist' with

‘General Problem Solver’, or GPS. GPS aimed to develop a more rounded example of human intelligence, using techniques such as means-ends analysis, planning, and selective trial-and-error (McCorduck 2004, 247). It remained part of their research until 1968 (Crevier 1993, 54). GPS proved successful on a number of specific tasks, including logic problems and puzzles, but none could be said to represent generality. Instead, as McCorduck writes, they claim in a 1961 article, it has ‘pretences to generality’ (McCorduck 2004, 249), admitting to its limitations, while not closing off the possibility for its ultimate realisation, and thereby shutting down funding.

Researchers in the other centres were less convinced about the need for extensive knowledge of how the brain works to develop machine-based intelligence (Crevier 1993, 55). At MIT, John McCarthy and Marvin Minsky founded the MIT Artificial Intelligence group, although they were not agreed on the best path forward. McCarthy believed that AI would be realised through the application of formal logic. Minsky, on the other hand, believed in trying out anything that might work (1993, 64).

McCarthy left in 1963 to establish a group at Stanford just as MIT’s AI group had received at large DARPA grant. At MIT, a new programme called ‘Project MAC’, or ‘machine-aided cognition’, was established, subsequently becoming the MIT AI Lab in 1970. The same year the Stanford Research Institute (SRI) split from Stanford University, partly in response to student protests and objections to military funding by DARPA and became ‘SRI International’.

Minsky’s larger-than-life personality came to dominate the research programme at MIT, forming a devoted band of research students and an *ad hoc* approach to projects (Crevier 1993, 64-67). His early projects concentrated on various forms of natural language processing and search (Brooks 1991a, 6). Among the many projects initiated at MIT during these years was a conversational computer programme called ELIZA, created by computer scientist Joseph Weizenbaum. ELIZA was an ‘artificial psychiatrist’ and used pattern matching and substitution to create a superficial simulation of a natural conversation between humans. Sherry

Turkle (2011), an MIT student at the time, recounts interacting with the programme. As she reports, despite being aware of ELIZA's superficial constructions, after a few generic openers, students would invariably start to share intimate details and concerns with it. While at the time this did not concern Turkle, Weizenbaum was appalled. Much of his subsequent work was focused on warning about the dangers of conflating humans with machines, and concern about degree to which we have made the world like a computer (Weizenbaum 1976).

In 1963, Seymour Papert arrived at MIT, having met Minsky earlier at a symposium in London. Papert was interested in psychologist Jean Piaget's work, specifically developmental intelligence in children, and, on his arrival, Minsky also become interested in these areas (McCorduck 2004). Both of them had strong personalities; they courted media attention, worked with Hollywood on movie productions, and became known for their witty aphorism and wild predictions. Together they initiated new research programmes focused on theory of computation, robotics, human perception, and child psychology (Crevier 1993, 86). One of the areas Minsky and his team worked on was computer vision, attempting to translate the image captured on a television camera into something the computer can describe. Initially, he hired an undergraduate student to solve it over the course of the summer. In the end, it would take another 30 years and the advent of digital cameras before this was achievable (1993, 89). Minsky and Papert also became interested in 'Blocks Micro Worlds', a simplified world of geometric forms, such as pyramids, squares and rectangular blocks in which computers could be programmed to interpret images, manipulate blocks, answer basic questions and move about (1993, 83-84, 91). Using the Blocks Micro World, they tackled the problem of hand-eye coordination, combining vision programmes with robotic arms (1993, 92). Crevier describes it as follows:

> The arm had a moving shoulder, three elbows, and a wrist; it used fourteen hydraulic cylinders for muscles. Before attempting to grab any blocks, the robot would hold its hand in front of the camera and wave it a little to see whether it really was itself. The computer then adjusted the coordinate system used in the image to make it correspond with the coordinates of the hand. (1993, 92)

By connecting sensors (cameras) to actuators (robot arms) and using computing mechanisms to control problem solving, they had built machines that could be said, at least to a certain degree, to ‘see’, ‘plan’, and act.

Over at Stanford, McCarthy had started the Stanford Artificial Intelligence Lab, or SAIL. His goal was to build a fully intelligent machine within a decade (Moravec 1995, 20). The DARPA-funded project aimed to combine the latest capabilities in hardware with those in software, specifically computer vision and natural language processing, to see how close they could come to creating a robot that would emulate human behaviour and cognition. According to Charles Rosen, who led the project, the US Department of Defence agreed to fund it because they through that the robot might be able to act as a mechanical spy (McCorduck 2004, 271). The result was Shakey, the first real attempt to build a general purpose ‘cognitive’ and humanoid robot, combining logical reasoning and some degree of autonomous movement. Shakey consisted of three parts; a wheeled cart at the bottom, a cabinet in the middle which held the computer processors, and, at the top, a TV camera, a range finder and a radio-linked antenna (Crevier 1993, 94). Shakey could ‘perceive’ its world, ‘plan’ how to achieve a goal, and ‘act’ in a physical world to carry out the plan (Kuipers et al. 2017). However, as Crevier points out, Shakey’s physical world was deliberately constructed to be a simple as possible. It was a life-sized blocks world consisting of seven rooms connected by eight doors and containing square boxes (1993, 94-95). In this environment, Shakey could follow instructions given by a keyboard in simplified English and then negotiate its way around the room and perform an action, such as moving or stacking a box.

Figure 6: Shakey and Charles Rosen, Still from Video (SRI International 2017)

However, the project was not a success from a technical perspective. As Moravec later observed, Shakey was ‘impressive as a concept but pitiable in action. Each move the robot made, each image captured by its camera, consumed about an hour of computer time and had a high likelihood of failure’ (Moravec 1995, 15). A person walking across Shakey’s field of view could immobilise the robot, sometimes for days. The team named the robot Shakey because it shook and shuddered as it moved (DARPA 2021). As McCorduck observes, ‘Shakey showed that you could not, for example, take a graph-searching algorithm from a chess program and hand-printed-character-recognizing algorithm from a vision program and, having attached them together, expect the robot to understand the world.’ (McCorduck 2004, 269). By the time McCorduck saw Shakey, ‘he was … a sad sight, immobile in a corner’ (McCorduck 2004, 268).

Despite its limitations, however, Shakey was hugely popular. *The New York Times* featured the robot in 1968, and *Life Magazine* and *National Geographic* in 1970, with *Life Magazine* dubbing it ‘the first electronic person’ (Kuipers et al. 2017, 97). Crevier reports MIT researchers’ consternation with some of the more elaborate claims made by the article, including a supposed quote by Minsky that ‘in from three to eight years we will have a machine with the general intelligence of an average human being’, which, he says, Minsky vehemently denied (Crevier 1993, 96).

Outside of the US, an AI laboratory was set up in Edinburgh University, Scotland, in 1965. Two experimental robots were built there between 1969 and 1976, Freddy and Freddy II. Freddy integrated vision, manipulation, and intelligent systems. However, both robots were subject to the same problems as Shakey. As McCorduck relates, Bernard Meltzer of Edinburgh University derisively called Freddy a ‘feeble creature’ (McCorduck 2004, 268).

However, the only serious rival to US dominance in robotics was Japan. In the 1970s, the Japanese government committed $200 million towards the development of robots (Hays 2013). The WABOT project at Waseda University aimed to build the world’s first full-scale bi-ped humanoid intelligent robot. WABOT-1 was released in 1973, and was able to play the organ, walk, and communicate through an artificial mouth. It had sensors and hands to grip objects and move them.

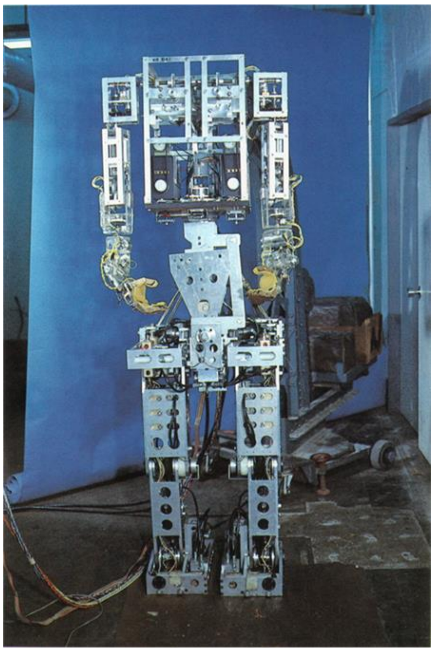


Figure 7: WABOT 1 (Waseda University 1973)

By the end of the 1960s, major AI research programmes were producing AI programmes that ‘proved theorems in geometry, solved problems from intelligence tests, algebra books, and calculus exams, and played chess, all with the proficiency of an average college freshman' (Moravec 1995, 8). Despite the technical breakthroughs and successes in specific areas, however, the major problems that had dogged the field from the beginning continued. Artificial 'general' intelligence, or the ability to generalise from one task to another, continued to elude AI practitioners. This problem and goal for the AI and robotics communities has since become known by the initialism 'AGI'.

AI Winter

In 1965, philosopher Hubert Dreyfus was hired to spend the summer at RAND. His brother, Stuart, was a computer programmer and engineer there, and had just received his PhD. Based on his observations at RAND, Dreyfus (1965) wrote a scathing attack on the field of AI which would become his book *Alchemy and Artificial Intelligence*. In it, Dreyfus compared the new science of AI to alchemy and mocked what he considered the absurd predictions of AI researchers. Dreyfus's argument, based on a phenomenological understanding of perception and intelligence, was that humans rarely use logic for problem solving. It did not go down well. RAND considered repressing the article (Crevier 1993, 121-122). Papert wrote a derisive refutation of Dreyfus and his work, calling him 'irresponsible' and 'wrong' (Papert 1968). Simon suggested that Dreyfus was using his connection to RAND to give himself unwarranted credibility (McCorduck 2004, 226). According to Dreyfus, no one at MIT, other than Weizenbaum, would dare be seen having lunch with him (Dreyfus and Dreyfus 2009, 8-9). In 1967, Papert arranged a chess match between Dreyfus and an AI chess programme called 'Mac Hack', which the AI programme won (Crevier 1993, 124). This incident was used to further discredit Dreyfus's argument within the community. Nonetheless, as Crevier acknowledges, 'time has proven the accuracy and perceptiveness of some of Dreyfus's comments' (Crevier 1993, 125). According to both Crevier (1993, 125) and McCorduck (2004, 236), while Dreyfus may well have been right, it was his 'tone' that estranged people.

Other criticisms at the time were more difficult for the AI community to elide. As we have seen, despite advances in specific domains, AI and robotics research in 1960s was mostly a story of disappointment and unrealised ambitions. In 1973, the UK government commissioned a report on the state of AI in the UK (Lighthill 1973). Its author, James Lighthill, gave a devastating critique of the field, specifically calling out the failures of speech recognition and machine translation. Any successes, he claimed, such as ‘list-processing languages’ were no better than the research done in regular computing laboratories (1973, 11). Lighthill dismissed the field of robotics as one that was influenced by human imagination and fiction, as much as by science. He even suggested a ‘pseudo-maternal’ relationship between robots and their builders, in order to compensate for the lack of the female capability of giving birth to children (1973, 8). He went on to assert that ‘in no part of the field have the discoveries made so far produced the major impact that was then promised’ (1973, 8). Robotics, as a bridging activity between automation and ‘intelligence’ was seen as particularly disappointing, its proponents delusional:

> When able and respected scientists write in letters to the present author that AI, the major goal of computing science, represents another step in the general process of evolution; that possibilities in the nineteen-eighties include an all-purpose intelligence on a human-scale knowledge base; that awe-inspiring possibilities suggest themselves based on machine intelligence exceeding human intelligence by the year 2000; when such predictions are made in 1972 one may be wise to compare the predictions of the past against [today’s] performance. (1973, 13)

Additionally, Lighthill noted that benefits ‘have flowed primarily to the science of psychology: in fact, a new range of attitudes to psychological problems has been generated’ (1973, 11). The report essentially ended funding for AI in the UK and Europe and solidified the ongoing reductions and tightening up of restrictions by DARPA in the US. This period of time has since become known as the first ‘AI winter’, coined in 1984 at the annual meeting of the American Association of Artificial Intelligence.

Perhaps it is no coincidence, then, that envisioning human-like machines once again fell to fiction. In 1977, the first *Star Wars* movie was released, introducing the lovable robots C-3PO and R-2D2 to the world, setting revised expectations and inspiring a whole new generation of robot makers. Although presenting a wildly different view of robots in society, Maria, the anti-hero of Metropolis, provided the inspiration for the charming, bi-ped C-3PO (Star Wars 2014).

In Japan, in 1982, contrary to events in Europe and the US, the Japanese Ministry of International Trade and Industry (MITI) launched 'the Fifth Generation' project, underscoring their previous commitment to developing humanoid robots (Nilsson 2010, 349). In 1984, Waseda University released Wabot-2, followed in 1985 by WASUBOT, a keyboard-playing, musical humanoid which was exhibited at the opening ceremony of the Tsukuba International Science and Technology Exposition that year (Takanishi 2019). Also at the exposition, Hitachi Ltd released the WHL-11, a biped robot that could walk (Hitachi 2021). The following year, Honda released one of their first of seven biped robots, Experimental Model 0. Not wanting to be outdone, funding was renewed both in the UK and the US. DARPA decided robot navigation was sufficiently advanced and the number of mobile robot projects increased dramatically as funding materialised (Moravec 1995, 21). In the US, DARPA founded the Strategic Computing Initiative and in 1983, the UK invested ₤350 million in the Alvey project for massively parallel computer processing (Nilsson 2010, 345, 355).

However, throughout the 1970s and 1980s, it was predominantly the corporate Expert Systems market had sustained the field (Russell and Norvig 1995, 21-22). However, as desktop computers became more powerful and less expensive, these machines proved too expensive to maintain and, in 1987, the market collapsed overnight. Japan's fifth generation project had also failed to meet expectations. Investors became disillusioned and funding for AI projects was once again withdrawn. This became known as the second AI winter.

The New Robots

The disappointments and limited successes of the symbolic information programming approach meant that now alternative and experimental approaches could be trialled. One of the problems, it was theorised, was the failure of AI systems to deal with ‘common-sense’ knowledge, a problem that John McCarthy had first articulated in 1959 with his ‘Advice Taker’ programme. In 1989, McCarthy defined common-sense knowledge as:

> …the basic facts about events (including actions) and their effects, facts about knowledge and how it is obtained, facts about beliefs and desires. It also includes the basic facts about material objects and their properties. (McCarthy 1989, 1)

In the mid-1980s, a professor at Stanford, Douglas Lenat, decided to tackle the common-sense knowledge problem. He launched project ‘Cyc’ to capture all of the knowledge and rules about the world, including ‘facts’ such as ‘nothing can be in two places at one’ and ‘animals don’t like pain’ (Crevier 1993, 240). The project was projected to last for two-person centuries and is still ongoing. Originally, as Nillson reports, Lenat had thought he would need ‘a couple of million assertions’ (Nilsson 2010, 447). By 2010, Lenat believed that 200 million assertions might be needed (2010, 447). Dreyfus, perhaps not surprisingly, was highly sceptical of the project. Dreyfus’ critique turns on the phenomenological idea that common-sense is not based on ‘context-free entities and their relationships’, instead it is based on ‘knowing-how rather than knowing-that’ (Dreyfus 1992, xviii, xxvii).

Around the same time, another radically different approach to AI was initiated at MIT. Roboticist Rodney Brooks started to look at older, cybernetic models such as Grey Walter’s tortoises, which had showed that apparently purposeful and complex behaviour could emerge from simple, non-computational mechanisms. Brooks experimented with various types of robot cognition, including investigating the connection between sensorimotor skills and intelligence, and cutting out the computation altogether (Brooks 1990). Brooks set his robots loose in the lab, with sensors directly connected to actuators and little computation. He noticed, as

Grey Walter had previously, that the robots responded directly to the environment and were able to display apparent lifelike and purposeful behaviour. ‘It soon became apparent’, Brooks wrote, ‘that the dynamics of the interaction of the robot and its environment are primary determinants of the structure of its intelligence’ (Brooks 1991a, 16).

Brooks is a harsh critic of the traditional symbolic approach (1990; 1991a; 1991b), claiming that the trend in (Von Neumann) computer architectures and models over the previous 30 years has had a strong and misguided influence on our models of thought. Brooks called approaches that were solely concerned with abstraction ‘a dangerous weapon’ (Brooks 1991b, 12). Rather than ‘good science’ they were ‘self-delusion’ (Brooks 1991b, 12). Brooks argues that the traditional approaches have relied, both implicitly but also explicitly, on what he called ‘folk understandings’ of human and animal behaviour (Brooks 1991a, 12). Instead, he argues, ‘real biological systems are not rational agents that take inputs, compute logically, and produce outputs. They are a mess of many mechanisms working in various ways, out of which emerges the behaviour that we observe and rationalize’ (1991a, 14).

More recently, Brooks (2002) has criticised the way that intelligence is defined in the field, observing that ‘intelligence in the early days of AI was thought to be best characterised as the things that highly educated male scientists found challenging’ (36). He points out the, now commonplace, observation in the field that seemingly simple activities are more difficult to automate than what are considered ‘higher’ order faculties, such as abstract, logical and mathematical reasoning. Such simple activities might be ‘distinguishing between a coffee cup and a chair, or walking around on two legs, or making aesthetic judgements’ (2002, 36). This has become known as ‘Moravec’s paradox’. In Moravec’s words:

> We are all prodigious Olympians in perceptual and motor areas, so good that we make the difficult look easy. Abstract thought, thought, is a new trick, perhaps less than 100 thousand years old. We have not yet mastered it. It is not all that intrinsically difficult; it just seems so when we do it. (Moravec 1995, 15)

Although it remains understated in his work, Brooks' views were influenced by phenomenology, in particular by Dreyfus's critique[2], the work of his students Philip Agre and David Chapman, and, indirectly through them, by the work of anthropologist Lucy Suchman. Agre and Chapman (1987) had developed a computer programme called 'Pengi' in which they show that planning, traditionally conceived as the mechanical execution of an explicit representational model of the world, can instead be designed without explicit models (268). According to Brooks, Agre and Chapman's redefinition of 'routine activity in a relatively benign, but certainly dynamic world' was one of the core ideas that changed the course of his research (Brooks 1991a, 2). This involved, not problem-solving or planning, nor representations that 'rely on a semantic correspondence with symbols that the agent possesses', instead they 'can be defined through interactions of the agent with the world' (1991a, 2). Thus, Brooks and his team were led to 'hypothesize (following Agre and Chapman) that much of even human level activity is similarly a reflection of the world through very simple mechanisms without detailed representations' (Brooks 1991b, 7). For Brooks, the core implications of these concepts were: first, a refutation of the representational approach, and second, the fact that 'intelligence' emerges from the interactions with the environment involving 'simple' mechanisms on the part of the agent.

As a result, Brooks developed a new robot architecture that he called 'subsumption architecture' (Brooks 1990). Instead of a centralised, top-down system, the new system is decomposed into smaller pieces and distributed around the robot. It had a layered model of increasing complexity: for example, the 'bottom' layer might be a simple control system focused on avoiding obstacles, whereas a more complex, 'higher' level might consist of a control system with the capability to identify something significant in the environment. When not occupied by the more complex activity, the robot could simply move around using very little

---

[2] Although Brooks was influenced by Dreyfus's critique, he felt the need to distance himself by adding a footnote in (Brooks 1991a) to the effect that '[e]ndorsement of some of Dreyfus views should not be taken as whole hearted embrace of all his arguments' (10).

computation. By drastically simplifying the computation, reaction times also improved dramatically. This allowed him to build and test in ‘uncontrolled’, more ‘natural’ environments without the problems that had been experienced by earlier robots such as Shakey and Freddy. A number of new robots were built using the new architecture, including ‘Allen’ (after Newell), ‘Simon’ (after Herbert) and ‘Ghengis’. As Brooks writes, ‘Allen would happily sit in the middle of a room until approached, then scurry away, avoiding collisions as it went’ (Brooks 1990, 5). The more sophisticated Herbert could ‘wander around office areas, go into people’s offices and steal empty soda cans from their desks’ (1990, 6). The six-legged Ghengis could walk over rough terrain (1990, 6). Building on these successes, Brooks hoped that he would build human-like intelligence incrementally, much like evolution.

In a new project, Brooks decided to pursue what he called the ‘Holy Grail’ of the AI community: an android, or ‘autonomous robot with humanoid form and human-like abilities’ (Brooks et al. 1999, 1). In 1993, Brooks put together a team that included philosopher and cognitive scientist Daniel Dennett to work on an upper-torso humanoid robot called Cog (Adam 1998). Cog’s system was to be based on four alternative ‘essences of intelligence’, gleaned from work in developmental psychology, ethology, systems theory, philosophy, and linguistics that they claim have been ‘discarded’ by traditional symbolic approaches (Brooks et al. 1999, 4). These were: physical embodiment, integration of multiple sensory and motor systems, a developmental structure allowing for incremental learning, and social interaction. Cog’s perceptual system included traditional visual and auditory sensors (cameras and microphone), but also a ‘vestibular, tactile and kinaesthetic system’ developed to allow for advanced physical orientation within the world and in relation to other objects (1999, 9). The task they had set themselves was to find a skill decomposition that maintains the complexity and richness of the behaviours represented while remaining simple to implement and construct (1999, 14). This, they believed, would allow them to contribute both to engineering and to science, that is, to robotics and to ‘the scientific goal of understanding human cognition' (1999, 1). As Adam observes, the team was strongly influenced by evolutionary biology:

> Cog was 'programmed to recognise its 'mother's' face (a post-graduate student), and will be designed to learn, so that future descendants of Cog can retrace the steps of millions of years of evolution in a few years of laboratory experiments. (Adam 1998, 147-148)

Despite earlier insights related to a 'benign world', according to Adam, the evolutionary style of robotics incorporates '[a] view of nature red in tooth and claw', in which '"[s]uccessful" robots control systems can contribute genetic material to become "parents" … unsuccessful ones are discarded' (1998, 154).

Ultimately, Cog's key scientific achievement was to highlight its own shortcomings. The Cog project did not succeed in its aim to advance our understanding of human intelligence or verify the four 'essences'. With regard to 'physical embodiment', the project's website admits '[s]ince we can only build a very crude approximation to a human body there is a danger that the essential aspects of the human body will be totally missed' (CSAIL 2021). Agre had made a similar observation in reflecting on the Pengi programme: '[i]t is hard to say which aspects of human embodiment and acculturation are necessary, either by definition or as a practical matter, for the human forms of intentionality' (Agre 1997, 242). As described on the Cog website, the key learning from physically embodied robots is the reaction they provoke in the people who interact with them. This, they reason, might have the advantage of encouraging natural human interaction from which the robot could 'learn' (CSAIL 2021).

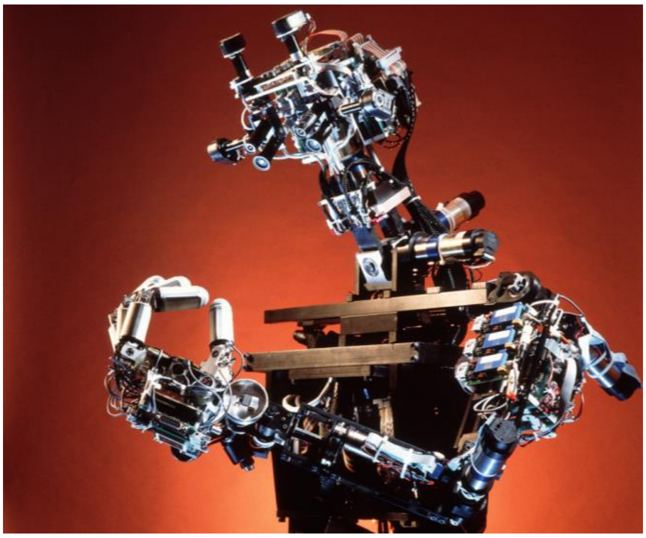


Figure 8: COG (Ogden 1993)

Building on this insight, Cynthia Breazeal, one of Brooks' students, led a group at MIT to create Cog's successor in the late 1990s. The result was a social robot called 'Kismet'. Kismet had a narrower remit than Cog and was specifically designed to produce an affective response in people (Breazeal 2002). Kismet has a large head and no body, with exaggerated features to simulate emotion, including two large eyes, eyelids and bushy eyebrows, as well as lips and a jaw. Its head has pointy ears made out of paper that give it a gremlin-like appearance. It is equipped with audio, visual, and proprioceptive sensors to allow it to recognise and respond to emotional cues.

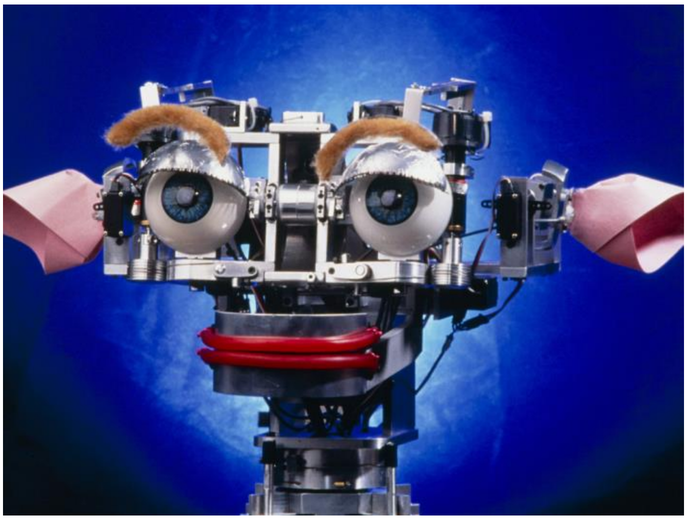


Figure 9: Kismet (Menzel 1998)

Breazeal draws on the idea of ‘social intelligence’ to underpin Kismet’s scientific credentials, arguing that social factors played a crucial role in our evolution as a species and cognitive development (2002, xii). Breazeal also draws on human developmental psychology, in particular studies based on early infant caregiver interactions. Kismet is conceived as an infant and equipped with ‘infant-level social abilities’ (2002, 5). Kismet’s child-like design served the twin function of drawing people to it, while also excusing any shortcomings in its interactive capabilities. Furthermore, by engaging people over time, Kismet’s ‘intelligence’ system could be trained, and thus develop increasingly sophisticated interactions. Kismet’s ‘intelligence’ is thus not coded in from the start, but conceived as a blank slate and something that will emerge from these social interactions. As Suchman (2007) observes, ‘the figure of the child in Euro-American imaginaries carries with it a developmental trajectory, a becoming made up of inevitable stages and unfulfilled potentialities’. Thus, this strategy ‘simultaneously authorizes the continuation of the project and accounts for its incompleteness’ (2007, 237).

Breazeal breaks down social intelligence into five discrete components, distributed between the robot and the observer, allowing it to ‘understand us and itself in social terms’ (Breazeal 2002, 1). First, following Brooks, a socially intelligent being is ‘embodied and situated’. Second, they should be ‘life-like’ and ‘believable’. Third, they must ‘understood’. This means they must be perceived as social, by displaying ‘a capacity to give attention, emotion, expression and playfulness’, as well as enabling a human to read its actions and intentions. Fourth, they should be ‘human-aware’, being able to identify a person, what they are doing and how they are doing it’ and, finally, they should be imbued with ‘simple learning strategies’, like ‘observation and feedback’ (2002, 7-11).

Ultimately, however, neither Cog nor Kismet was able to display any demonstrable intelligent behaviour. By 2003, both robots were retired to MIT’s science museum (Suchman 2019, 43). Yet Cog and Kismet’s ability to evoke reactions in those that came in contact with them, as well as continuous media coverage, ensured that their influence far outweighed their capabilities. Kismet also became the focus of a number of prominent studies, including in anthropology and STS, e.g. (Kember 2003; Suchman 2007; Turkle 2011).

Thus, the lack of progress in AGI moved the focus of robotics from the inner workings of the robot as a replica of the human, to the robot as perceived from the outside, and as a human assistant. By the 2000s, funders had become interested in human and robot interaction and collaboration. The U.S. National Science Foundation and DARPA sponsored a workshop on the new field of HRI to help identify the issues and challenges in order to design ‘synergistic teams of humans and robots where team members perform tasks according to their abilities’ (Burke et al. 2004, 104). The new HRI field was thus established quite explicitly to engineer both the robot and the human. The steering committee featured Breazeal from MIT, as well as researchers from Stanford University and professor of communication at Stanford University and co-creator of The Media Equation and roboticist Ronald Arkin from Georgia tech.

Throughout the 1990s, Waseda University in Japan also released new robots focused on human-robot communication and interaction. Between 1993 and 2000, Honda was experimenting with a number of increasingly sophisticated and sleek-looking robots culminating in 2000 with the release of ‘ASIMO’, after Isaac Asimov, a small bi-ped robot clad in white plastic, with a helmet and black visor, resembling an astronaut.

Figure 10: Honda's ASIMO (Honda Global 2000)

Other roboticists concentrated on making robots look as life-like as possible. At Osaka University, Hiroshi Ishiguro created the life-like 'Actroid' robot in 2003, with a female appearance. In 2006, he released the 'Geminoid' robot, modelled on himself. The robots can blink and appear to breathe and fidget (Guizzo 2010). Similarly, Hong Kong-based Hanson robotics created the controversial 'Sophia' robot, who has since been granted 'citizenship' of Saudi Arabia and is first non-human United Nation Development Programme's 'Innovation Champion' (UNDP 2017). All of the focus of these robots is on their external appearance, and they are teleoperated.

Figure 11: Sophia robot (Hanson Robotics)

In 2005, MIT spin-off Boston Dynamics, started to make physical robots with funding from DARPA (Nilsson 2010, 637). An early project designed for military purposes, BigDog, displayed extraordinary stability, strength, and agility, but was considered too loud for use in combat (Degeler 2015). BigDog was followed up by 'Spot' and the humanoid 'Atlas'. YouTube videos show Spot performing extraordinary life-like feats, such as opening doors, dancing, navigating down hallways, and walking up and down stairs. These advanced capabilities meant that, for the first time, robots could start to leave the labs and controlled factory spaces and be put to use. The Boston Dynamics website advertises Spot for use in inspections in Power and Utilities, Mining, Manufacturing and 'Public Safety' (Boston Dynamics 2021).

In 2006, the new Japanese Prime Minister Abe announced plans for a 'beautiful (*utsukushii*) innovative, new, roboticized Japanese society' (Robertson 2007, 385). As Japan anthropologist Robertson reports, the proposal, called 'Innovation 25', lays out plans for 'the roles that biotechnology and robotics will jointly play in securing the stability of both the Japanese economy and Japanese social institutions' by 2025 (2007, 385). Robertson called the vision that was laid out 'reactionary postmodernist', in which robots both free up women's time to devote to having more children, and also may be used to compensate for the declining and ageing population, thus avoiding the need for immigration (2007, 391). According to Robertson, these new Japanese robots herald an era of increased surveillance and control (2007, 393).

In the US, Boston Dynamic robots were also coming under scrutiny. Spot was trialled by four different police precincts. In 2021, they were removed from the NYPD after being criticised by US representative Alexandria Ocasio-Cortez on Twitter:

> Now robotic surveillance ground drones are being deployed for testing on low-income communities of color with under-resourced schools ^. (Ocasio-Cortez 2021)

Despite the difficulties in getting robots to develop towards AGI, as well as the ethical and regulatory challenges, robots as technologies are becoming increasingly prevalent. As we will trace in the following chapters, robots are increasingly being used in manufacturing, in transport, in space exploration, farming, care and service roles, bomb disposal, and surgery. In keeping with the evolutionary narrative that has become a part of the field, a number of my informants referred to the current proliferation of robots as a 'Cambrian explosion'.

Machine Learning

Despite the prevalence of robots today, the key driver for the current AI revolution stems from a combination of the early neural network research, and, since 2000, the vast quantities of social data that have been amassed by the industry tech giants. As we have seen, early work in neural networks developed in parallel to the symbolic paradigm instituted at the Dartmouth conferences but was side-lined when the symbolic approach to AI become dominant.

Building on McCulloch and Pitts' (1943) early work on modelling neurons in the brain directly, Frank Rosenblatt, a former classmate of Minsky at Cornell, developed an artificial neural network device (Crevier 1993). Rosenblatt proposed a 'perceptron', a more generalised computational model than the McCulloch-Pitts neuron, through which 'weights' and 'thresholds' can be learnt over time. Rosenblatt kept neural network research going throughout the 1960s. However, in 1969, Minsky and Papert (2017) published a book *Perceptrons*, attacking his work, and ultimately stalling the research for the next 15 years.

However, in the mid-1980s, neural net approaches were popularised again, primarily outside of mainstream AI in the fields of cognitive science and statistics. In cognitive science, it was at first known as 'parallel distributed processing' (Rumelhart, Hinton, and McClelland 1987) and later as 'connectionism'. In the field of statistics, in areas such as pattern recognition, probabilistic modelling and information retrieval, it became known as 'Machine Learning'. Connectionism was also established with the field of psychology as an alternative to the 'computationalism' of the symbolic model. The technique allows someone to write an algorithm that will search through large datasets and look for patterns. Modelled loosely on the neurons in the brain, in an artificial neural network (ANN), there are three kinds of layers: the input layer, a hidden layer and output layer. By defining an input, and a desirable output, the model can be 'trained' to search for patterns that could not be recognised by a person looking for them, for example because the datasets are too large.

As Fjelland (2020) has shown, three 'milestones' were achieved using a combination of machine learning and other techniques, making the imminent attainment of AGI seem closer than ever. These were, first, in 1997, IBM's chess-playing computer 'Deep Blue' beat reigning chess world champion Garry Kasparov. Second, in 2011, IBM's computer 'Watson' beat two of the best participants on the US game show *Jeopardy!*. Third, in 2016, DeepMind's 'AlphaGo' beat world 'Go' champion Le Sedol (2020, 3-4). By the mid-2010s, machine learning techniques further increased in dominance, buoyed also by the availability of the vast amount of (primarily social) data on which they could be trained. Advancements, such as the addition of several processing layers between the input and output layers, allowed for the algorithms to iteratively update and search ('learn') without the need for human intervention (LeCun, Bengio, and Hinton 2015). This became known as ‘deep learning’ and was launched into the public imagination with the release of Google’s DeepDream in 2015. Deep Dreaming took images and processed them using these deep networks, resulting in fantastical and surreal-looking images.

Figure 12: An image captured by Google’s Deep Dream (The Guardian 2016)

The Deep Dreaming algorithm was designed to detect faces and other patterns in images. Once the network has been trained to do this, it could also be run in reverse, thereby introducing faces and patterns into images where previously there were just ambiguous shades or shapes. The resulting images were described on news sites as ‘disturbing’ and ‘trippy’ (Burke 2015), a ‘psychedelic nightmare’ (Beschizza 2015), and ‘mind-melting’ (Junglist 2015).

Deep Learning allows for a much more in-depth and complex form of iterative processing, or ‘learning’, which can result in some remarkable, and seemingly intelligent and perceptive, outcomes. This is amplified by the fact that it is often difficult or impossible for anyone, even those who created the algorithm, to understand the ‘decisions’, that is, the patterns and iteratively developed outputs gleaned from the data. ‘Reinforcement learning’ is another development of the technique, modelled on reward-driven learning in the brain (Sutton and Barto 2018). The algorithm works towards a defined output using a process of trial and error. For example, feedback might come from a human interacting with the robot (who might react negatively to a specific utterance), or by using feedback from the environment (e.g. a robot banging into a wall). The technique has been likened to ‘common sense’ and humans learning from experience (Smith 2020).

Reinforcement learning was the technique used to train AlphaGo. Its success led the New York Times article to declare ‘It isn’t looking good for humanity’ (Mozur 2017). Later the article went on to say, ‘the victory... showed yet another way that computers could be developed to perform better than humans in highly complex tasks, and it offered a glimpse of the promise of new technologies that mimic the way the brain functions’. The ability to amass, control, and use this data has led to the inexorable rise to power of the tech giants, most notable Google, Apple, Amazon, and Facebook. Previously, the field known as ‘big data’, statistical analysis and automation, rebranded as ‘Machine learning’, is now considered almost synonymous with AI (Elish and Hwang 2016).

The pace of recent developments, in terms of social robotics, advanced engineering and machine learning, have led to a real excitement and anticipation in the field. Machine/Deep Learning has been applied in healthcare for diagnostics in ultrasounds and MRI scans to detect certain diseases, and in drug discovery by finding new applications for existing medicines. The technique has also been used successfully to optimise temperature control, in earthquake detection, to manage crops in agriculture and to train autonomous vehicles. Machine learning algorithms are used to curate and recommend content. Chatbots trained on conversational data are being used by high-profile bots Siri and Alexa.

However, the relative newness of the technology has also allowed it to grow free of oversight and regulation. Today, data is gathered with little regard for privacy and used for automated ‘prediction’. This disproportionally targets poor and working-class people, who are targeted as risky investments, potentially bad employees, even bad parents (O’Neil 2017; Eubanks 2019). Scholars have also shown how both datasets and algorithms may be racially and gender biased (O’Neil 2017; Benjamin

2019b). Despite these concerns, for the first time in a long time, it seems that even if human-like intelligence is not yet possible, at least advanced, smart robots might really be imminent. As we will trace in the following chapters in detail, Stevie is the heir to these histories and a product of this *Zeitgeist*. But like all robots, Stevie is not just a product of the past, implicit in his form is also the promise of a specific kind of utopian future.

The Future

2020 was to be the year, according to futurist and director of engineering at Google, Ray Kurzweil (2014), that computers would achieve the memory capacity and computing speed of the human brain. He does admit, however, that it is the organisation and content of these resources, through software, that will be necessary to achieve human-level intelligence. But why stop there? Once human-level intelligence has been achieved, he writes, 'it will necessarily roar past it' (2014, 15). In the second decade of the 21st century, Kurzweil predicts, it will become increasingly difficult to draw any clear distinction between the capabilities of human and machine intelligence. Evolution is 'the intelligent process', and 'master programmer' of life. Human intelligence is evolution's 'greatest creation', the result of a 'billion-year drama' (2014, 16). So, when super-human intelligence emerges in the early twenty-first century, it will be of greater import than any of the other events that have shaped human history. Kurzweil's grand narrative is biblical in its breadth and scope, with quotes from sources as varied as the bible and Winston Churchill. However, in this account, evolution, rather than God, is the creator, with DNA-based genetics as the blueprint and a disembodied intelligence taking centre stage, in place of an immortal soul or spirit. Through technology, humans will be able to take control of evolution, in effect becoming like gods.

Kurzweil is a proponent of transhumanism, a philosophical movement that advocates for human biological and cognitive enhancement through technology, or 'overcoming aging, cognitive shortcomings, involuntary suffering, and our confinement to planet Earth' (Bostrom 2005, 26). Kurzweil (2005) has predicted that humans (some humans) will be able to transcend their physical bodies, until ‘ultimately, the entire universe is saturated with our intelligence’ (40). This phenomenon, known as ‘the Singularity’, will come when ‘the pace of technological change will be so rapid, its impact so deep, that human life will be irreversibly transformed’ (2005, 24):

> The Singularity will allow us to transcend these limitations of our biological bodies and brains. We will gain power over our fates. Our mortality will be in our own hands. We will be able to live as long as we want (a subtly different statement from saying we will live forever). We will fully understand human thinking and will vastly extend and expand its reach. By the end of this century, the nonbiological portion of our intelligence will be trillions of times more powerful than unaided human intelligence. (2005, 25)

According to Kurzweil (2014), this secular rapture-like event will occur by 2099 and will seamlessly merge the human species and machine intelligence, with no clear distinction between the two. Conscious entities will have surpassed the need for a permanent physical presence and intelligent beings will achieve immortality.

Swedish philosopher Nick Bostrom is also a proponent of transhumanism. Bostrom (2002) introduced the concept of ‘existential risk’ in 2002, defining it as ‘one where an adverse outcome would either annihilate Earth-originating intelligent life or permanently and drastically curtail its potential’ (2). According to Bostrom, novel technologies, the first of which was the atomic bomb, represent an entirely new and unprecedented risk to human life on earth. This is primarily because, he argues, unlike previous risks such as plagues or earthquakes, technologies have become so advanced that humans have not yet evolved the biological or cultural mechanisms to respond to them. Thus, our existing ‘institutions, moral norms, social attitudes or national security policies’ are no longer appropriate (2002, 4).

For Bostrom, the top five apocalyptic threats are, in order of priority: misuse of nanotechnology, a nuclear holocaust, the simulation that we are living in being shut down, ‘badly programmed’ super intelligence, and the risk from a genetically-engineered biological agent. Naturally-occurring disease, asteroid or comet impact, and ‘runaway global warming’ come in at nine, ten and eleven respectively. Two other risks cited are the risk of a ‘repressive totalitarian global regime’ that might hold back technology (2002, 13), and a general lowering of IQ (2002, 12).

By strategically mixing science fact, science fiction and pseudoscience, hypothetical and fantastical future technology is presented as equally or more pressing than real and present risks, as well as morally superior. The risk of regulating technology is presented as a risk equal to that of not pursuing eugenicist ideals. By the mid-2010s, noted scientists and industry leaders Elon Musk, Stephen Hawking and Bill Gates were starting to warn about the ‘existential risk’ of AI. In 2014, Musk, speaking at MIT, warned that AI could be the human race’s biggest existential threat, comparing it to it to ‘summoning the demon’ (McFarland 2014). Theoretical physicist Stephen Hawking warns that the development of full AI could spell the end of the human race (Sample 2015).

Kurzweil is Google’s director of engineering and is busy assembling, according to the Guardian newspaper, ‘the greatest artificial intelligence laboratory on Earth’ (Khomami 2014). It bought DeepMind, Boston Dynamics and Nest Labs, although Boston Dynamics was sold to Japan’s Softbank in 2017. As Frude and Jandrić (2015) report, ‘his brief at Google is to develop natural-language processing so that artificial systems will be able to really understand what they hear and read, and when this is possible such systems will of course be able to absorb the contents of any and every book and webpage’ (417). Similarly, Google’s company DeepMind has as its explicit goal to ‘solve’ intelligence, which will then allow them to ‘solve everything’, and to ‘use technology to build our dream society’ (Katz 2017). Other companies, such as Microsoft, Facebook, and OpenAI (founded by Elon Musk to compete against DeepMind) have also invested heavily in this area.

Silicon Valley has eagerly adopted the transhumanist narrative, Bostrom’s concept of ‘Existential Risk’, and a related philanthropic orientation called Effective Altruism (EA). EA involves using data and algorithms, which proponents call ‘evidence’ and ‘reason’, to guide pro-social action. Originally targeting issues such as fighting global poverty and the cruelty of factory farming, existential risks or 'X-risks' have now come to dominate the discussion (Matthews 2015). Based on a mathematical formula proposed by Bostrom (2013), proponents estimate that, if humanity lives for another 50 million years according to current trends, the number of people who will ever live will be about 3 quadrillion. Thus, humans alive today comprise only a negligible amount. Add interstellar travel and you can, Bostrom argues, predict a potential for 10^{54} life-years. Even if this has only a 1% chance of being correct, it is 'worth a hundred billion times as much as a billion human lives' (2013, 19). Following Bostrom's logic, focusing on solving global poverty is a 'rounding error' (Matthews 2015).

As Schuster and Woods (2012) have called out, the field of 'existential risk' is disconnected from established, yet highly related, scholarly work. It does not draw on scholarship and personal accounts from those who have lived through extreme suffering and displacement. It does not draw on the philosophy of 'existentialism' with its emphasis on lived experience and embodiment. Further, as the authors argue, instead of rigorous calculation of probability, the field uses a 'rhetoric of probability' combined with a quantitative, utilitarian approach to ethics, 'calculated in terms of humanity's entire future political progress' and based on 'a hierarchical valuation of intelligence' (2021, Intro.). As they point out, Bostrom offers no reflection on the existential risks of ecosystems, animals, and the planet unless they are relevant to human and later posthuman flourishing. It is clear that Bostrom is enjoying himself tremendously. Unfortunately, people are listening.

On February 16th 2017, the European Parliament adopted a resolution in a published report with recommendations to the Commission on Civil Law Rules on Robotics. Among the proposals were changes European law to implement a code of conduct for roboticists and engineers, new laws on insurance and corporate governance of robots and AI and, a controversial principle 59f, calling on the commission to consider the implications of all possible legal solutions, including granting legal personhood to robots (Coalition for Critical Technology 2018).

In response, an open letter to the European commission was published with (currently) over 270 signatures from experts in AI and robotics, warning of the misguided nature and dangers of such a move. In a particularly damning paragraph, the letter states the amendment was ‘based on an overvaluation of the actual capabilities of even the most advanced robots, a superficial understanding of unpredictability and self-learning capabilities and, a robot perception distorted by science-fiction and a few recent sensational press announcements’ (Coalition for Critical Technology 2018). Signatories include a number of those involved in drafting the principles of robotics, as well as anthropologist, Kathleen Richardson. My informants were equally appalled by the resolution. According to roboticist Sam at CMU:

> We’re trying to make things that make better investments or do better search results or drive a car better. It has nothing to do with the piece of it that we would ascribe rights! (Interview 14/10/2019)

While the signatories of the letter were quick to blame ‘science fiction’ and ‘sensational press announcements’, it is less common for those in the community to point the finger at the institutions and technical leaders, from across academic and industry, who are explicitly and implicitly promoting these ideas. Only one of my informants, Sara at Heriot-Watt, pointed out the culpability of those who promote ‘existential risk’:

> …hype-merchants, parasites around the edges, who like the fame and being thought of as a *savant* … futurologists and so on, whose interest is to hype and people believe them because they’ve seen this stuff in films. Even people like Hawkins, who should know better. When he says AI is really dangerous, people believe him. It’s not dangerous for the reasons people think. (Interview 27/6/2019)

This future orientation, without regard for the past, is common among technologists, as roboticist Ben related in our interview, ‘I have had technologists who have point blank said the history doesn’t matter only the future matters, and we invent the future’ (Interview 11/10/2019). As Schuster and Woods articulate, the futurist and transhumanist narrative ‘taps a deep vein of public desire for secular eschatology —that is, for stories and scenarios about the origins and ends of humanity, life, and the universe’

(Schuster and Woods 2021, Intro.). It is clear that what is needed are alternative narratives, and new fictions.

Conclusion

In this chapter, we have traced the pasts and the futures that are implicated in the robot present. As we have seen, the past is less linear than the dominant narratives would suggest, and the future less pre-determined. Contemporary humanoid robots emerged from a long-standing tradition of attempts to create life and emulate human intelligence, and of attempting to define human essences and boundaries. In the mid-20th century, industry, military and academic institutions came together to create the project of AI, explicitly aligning the science of creating artificially intelligent machines to the science of the human and ensuring that the project of symbolic representations would come to dominate both AI research, and psychology.

As we have seen, the failure to realise human-like intelligence, or AGI, resulted in a withdrawal of funding and a diversification of the field, leading to alternative and radical methods being pursued, including the creation ‘common-sense knowledge’ database, behaviour-based and affective robots, and machine learning. In this chapter, I explored some of the ways in which robotics and machine learning are being deployed today. I also described how the difficulty of creating genuine human-like intelligence has resulted in a shift from the actual creation of the intelligent machine to a focus on collaboration and the potential capabilities of robots and AI. I described how, rather than the past, the overriding force animating contemporary AI and robotics work is the ever imminent near future. In the next chapter, Territorialisation, I continue to explore how these histories are materialised through the research sites, laboratories, and wider institutional networks, as well as how they are corporealised and embodied in the robot. This will further uncover the gap between the reality and the fantasy, revealing a disconnect between the capabilities of the technology and the discourses of imminent AGI. In the next and subsequent chapters, I will also describe the ways in which both the reality and the fantasies of the robot transform geographies, structured relationships, practices, and desires. This will also reveal the reality of the risks of robots, which are simultaneously more mundane and more urgent.

Chapter Four: Territorialisation

In Historicisation, we uncovered the constructed and contingent nature of the robot. In ‘Territorialisation’, the focus moves to the robot’s ‘irreducible materiality’ (Pietz 1985, 7). As Pietz’s stresses, the fetish is not simply a mediator, signifier, or representation of something beyond itself; rather it is significant as an object of value in and of itself (1985, 15, 7). The robot-as-fetish is an enduring material object connecting otherwise unconnected multiplicities, including mechanical, electrical, digital and network elements, but also norms, desires, expectations and assumptions, into a stable arrangement. This chapter thus takes the material robot as a central focus, more specifically; it takes *a* robot as its central focus. That robot is Stevie. As we shall see, Stevie’s identity, morphology, and mobility identify ‘him’ as a robot, a singular and distinct object, despite ‘his’ constituting multiplicity. The robot-object is a concrete reference-point, a site in which specific histories, as well as technologies and cultural imaginaries are sedimented and reproduced. This chapter’s focus on the material fact and inner workings of the robot allows us to investigate its ‘discernable peculiarities, contour and form’ (Lefebvre 1991, 77). This includes taking account of its physical ‘substance’, its historical and temporal specificity, as well as its ‘complex sensuality’ (Henare 2003, 57). By focusing on the physical capabilities of the robot, a ‘black-box’ technology (Latour 1999), we can bring clarity to an object that can often seem opaque and impenetrable.

As well as the material fact of the robot, this chapter also focuses on three other areas: the wider space of robotics research, the ‘conceived’ space of robotics researchers (Lefebvre 1991), and the territorial transformations, or ‘re-territorialisations’, that are occasioned by the robot. First, a focus on the socio-technical relations embedded in the robot draws attention to the wider territory of the robot-as-fetish, the space within which the robot as object is realised and given meaning. As Lefebvre (1991) writes:

> It is never easy to get back from the object (product or work) to the activity that produced and/or created it. It is the only way, however, to illuminate the object’s nature, or, if you will, the object’s

> relationship to nature, and reconstitute the process of its genesis and the development of its meaning. All other ways of proceeding can succeed only in constructing an abstract object - a model. It is not sufficient, in any case, merely to bring out an object's structure and to understand that structure: we need to generate an object in its entirety - that is, to reproduce, by and in thought, that object's forms, structures and functions. (113)

For Lefebvre, this 'social space' is produced by an interrelation of three levels: perceived space, conceived space, and lived space. 'Perceived space' constitutes the socio-physical space, a society's 'spatial practice, 'revealed through the deciphering of its space'; 'conceived space' is the abstract space as mapped and modelled by scientists and understood quantitatively, and finally; 'lived space', is the (subjective) space of everyday activities and represented qualitatively (1991, 38, 362). Foucault (2019) also emphasises the importance of space and spatiality through his concept of 'heterotopias', challenging conventional modes of thinking about space. Heterotopias are 'the space in which we live', this space is not 'a kind of void, inside of which we could place individuals and things', rather 'we live inside a set of relations that delineates sites which are irreducible to one another and absolutely not superimposable on one another' (2019, 23). Geographer Edward Soja (1996) further developed these ideas in his concept of 'Thirdspace': spaces in which the 'real' and the 'imagined' mingle, counter spaces that are inherently ambiguous. Thus, space is revealed, not as the fixed and dematerialised, but as constructed and discontinuous as the temporality traced in Historicisation, and thus similarly available for reconceptualising and reconfiguring.

Thus, as well as focussing on the material fact of the robot, this chapter also brings ethnographic attention to the 'socio-physical space', or 'perceived' space of robotics research, drawing on fieldwork from a number of robotics labs, conferences, and test sites in Ireland, the United Kingdom, the United States, and Japan. The robotics field that I analyse in this project is also the field that I simultaneously construct. It emerged from the connections that I followed and the places that opened up to me. It reveals, on the one hand, the geographical specificity of individual labs, subject to specific histories, research priorities, and funding supports. On the other hand, it also reveals a community of researchers united by common histories, identities, and mythologies. This chapter will thus reveal both discontinuities and continuities between various settings, as well as surfacing the wider institutional arrangements within which they operate.

Paying ethnographic attention to the spaces within which robotics research takes place allows me to articulate the ‘conceived space’ of the roboticists with whom I interacted. By documenting how roboticists relate to their work, through an analysis of observational and interview data, I describe some of the debates and discourses that dominate the field, illuminating how roboticists conceive of, and relate to, space. The final section of this chapter will focus on the ’re-territorialisations’ in which the robot is implicated. In particular, in this chapter, I focus on the network of institutions that is being established to underscore and disseminate the image of imminent AGI. Other re-territorialisations are elaborated in subsequent chapters, including, in Reification, changes to work practices and labour conditions, as well as how the robot alters the social dynamics within a care community. Lefebvre’s concept of ‘lived space’, or the concrete space of everyday users, is further interrogated in the Personalisation chapter.

Stevie the Robot

The Robot

Stevie is a full-bodied humanoid robot. The upper part of his body has a ‘human-inspired torso’ with two short arms ending in balls, and a head (McGinn et al. 2020). The head, or ‘social interface’, is connected to the torso via a ‘three-degrees-of-freedom’ neck mechanism. This means that there are just three defined directions in which the neck can move using motors, including pitch (side-to-side) and roll (front to back). The front of the head, or face, has two separate digital displays to allow Stevie to display eight different emotions. Stevie’s arms have two degrees of freedom at each shoulder. Stevie is 1.4 meters tall, about the size of a ten-year-old child.

Figure 13: Stevie in the lab

Stevie has wheels instead of legs, and is designed to have a clearly identifiable forward-facing direction, with a wheel exposed at the bottom so people can see what direction it is going in. The wheels are 'omni-directional' to provide stability when he is static. Stevie can move his head and arms and can turn at the hip. Stevie's skin is made of FDA-grade antimicrobial plastic with accents of yellow. While his head looks a little boxy, his body is covered in a shiny fabric bonded to a hand-moulded plastic shell. He has under-body lighting and an 'illuminated bisecting circle pattern' in the torso, which serves a number of functions. First, it helps to indicate which side is front facing, second, it serves as a brand or logo, and third, it is an indicator that he is 'on', giving cues as to his 'inner state'. Stevie has large red button on the shoulder, which is a stop button, or 'kill switch', and safety feature. Stevie needs power to function, and lots of it. There are two very large batteries between the wheels at the bottom, which have been specifically designed for it to fit.

When Stevie's face is 'on', the upper display shows two big brown eyes and eyebrows. There is a 3D effect using graphical software, allowing for a glint in the pupils. His eyes appear to blink continually. This is just one of a number of 'non-verbal' social interactions that have been designed in. Others are: raising and lowering eyebrows, smiling and frowning, and bodily micro-movements. When turned on, Stevie's little arms move up and down and his head moves slightly from one side to another, even when no one is interacting with him. Together, these give Stevie the impression of being animate. The lower display features a simple line drawing of a mouth, which has four different shapes, curved down, straight, curved up, and rounded. When Stevie is 'off' the display is blank. At other times, the upper display area can be used as a simple screen, displaying messages or providing a programming interface for the developers.

Stevie is covered in sensors, about 60 of them, including a LIDAR or light-detection sensor for measuring distance. There is a microphone for capturing audio data, webcams on the front and back to capture images and another webcam on the neck area with which to capture 3D. This will allow for depth perception, which a normal webcam, just capturing pixels and colour, cannot. Proprioceptive sensors sense and measure Stevie's inner state, such as his position in space, which way he is facing, his speed, and the angle of his joints.

Stevie has the capacity to store the data that is captured locally and is also connected to the wider network of connected devices in the lab, as well as the Internet and the Cloud. An on-board wireless router is installed to enable off-board processing, teleoperation and remote data logging. This networked structure allows for all sorts of configurations and connections, such as distributed and 'off-board' processing, data storage and use in the Cloud, as well as remote-controlled teleoperation. A person can take over the system and remotely operate Stevie's actions and speech using a laptop. He can also be used, like any 'smart' device, such as a smartphone or tablet, as a communication device, such as for video or audio-communication.

What might be referred to as Stevie's 'intelligence': his voice, his responses, his gestures and movements, are implemented using a few different approaches. A limited, pre-defined set of social interactions is programmed in. Stevie's speech recognition capability is implemented using IBM Watson's speech recognition API. Stevie talks using a speech synthesis SDK developed by Cereproc, an 'off-the-shelf' male-voice solution with a 'Southern English' accent, chosen after conducting trials with potential users. Stevie 'verbally communicates' through speakers in his ears. Stevie's autonomous behaviours are implemented using a common framework for writing robot software, called the Robot Operating System, or ROS. This includes tools, libraries and conventions to standardise and simplify the task of creating various robot behaviours, such as the 'Navigation Stack' that takes in information from the sensors and translates them into velocity commands that are then sent to the actuators. The team are also actively working on instituting machine learning techniques in order to develop Stevie's 'social' abilities. In order for Stevie to 'learn' the dynamic and continuous flow of natural conversation, as well as how to recognise people, the system will need to be continuously supplied with a huge dataset of natural language and 'trained' using examples.

In the first version of 'Stevie', computing was done with a single computer. The new version has a distributed network of computers which means there is no single central processor doing all of the computing, which slowed the system down, even sometimes incapacitating it. Now, as well as an Intel NUC (i5 processor, 8GB RAM) on-board computer, there are a series of distributed lower power computers and small motherboards, allowing for the processing to be done in parallel. Each set of cameras has its own dedicated computer just to process the visual data, as do the wheels. Other smaller computing devices can be added into the wider system, such as mobile devices and microcontrollers. These distributed micro-controllers communicate with the main on-board computer using UART over USB. This approach is called 'AI on the edge', meaning the data processing and decisions are made locally, rather than in the centre, and allows for much faster response time. It also has the added benefit of being more secure and reliable. The main computers may be updating every second minute, but these ones can update in milliseconds.

The Lab

Stevie is being built by a small team based at the Robotics Lab based in the Department of Mechanical and Manufacturing Engineering, at Trinity College, in Dublin, Ireland. To get there, you have to climb up a number of flights of stairs of an historical building, walk down a long corridor, and through a new extension at the back. When I first went there, I marvelled at the historic location and aesthetic, only partially ingratiating, but the team were not feeling sentimental. The building did not represent innovative engineering or the futuristic feel that they felt they were part of creating. The small room that houses the lab is one of the few robotics labs in Ireland. Daniel, who leads the team, is an assistant professor and mechanical engineer and the small team is predominantly made up of engineering post-docs, as well as a number of post-docs from computer and data science. For Daniel, the project is the culmination of a life-long dream of inventing something new and taking his place among the great inventors.

The lab is a large L-shaped room with large steel-mounted shelves all around the walls. The shelves are piled high with wires and mechanical parts, nuts and bolts, bits of plastic and wood, and tools and machines. On the right-hand side as you come in, there are a number of high tables with machines on them, as well as a 3D printer. In the main part of the room, there are a number of long desks set against the wall with people working at computers, as well as a few large desks set perpendicular angles to the wall with high stools, and flat screens at the end and a camera for video conferences and meetings. In the centre, there are two sofas facing each other and a coffee station. The room is part workshop, part computer lab, with a dash of design aesthetic.

In the lab, the work is rarely on the whole robot. Most of the time, an older prototype of ‘Stevie’, and the new model, are standing immobile and switched off, out of the way. The only time I see Stevie ‘whole’, is when he is about to be tested with potential users. Different parts of what will eventually become Stevie are laid out on tables and counters around the room:

> Dónal, an undergrad in mechanical engineering, is testing the individual components of what will eventually become Stevie. He

> points out the different parts. Here are two small computers, but also lots of small motherboards. He points out the one for the back wheel. A lot of computation is going to be necessary, he says. He points out the microphone, two webcams and another webcam for the neck area, which, he tells me, will capture 3D. This will allow for a measure of depth that a normal webcam, just capturing pixels and colour, cannot compute. But, he explains, this will not work for windows, because windows reflect light so there also have to be sonar sensors. He shows me about eight of them. These will be able to sense windows but not, for example, sofas because the sofa will absorb the sound. He lifts up two very large batteries. These will sit between the wheels, he says. which have been specifically designed for them to fit. Cost considerations mean the batteries can't be any smaller.
> The 'robot' does not look at all like a robot. If I didn't know, and had to guess, I would have described it as a clear plastic coffee table, with several shelves and bits of wire and plastic lying on it/sticking out of it. Possibly to remind people what we are aiming for here there is a piece of wood with a face drawn on it in marker pinned on top. (Fieldnotes 23/10/2018)

Donál is not a graduate student like the others in the lab, instead, he has taken a year out of his undergraduate studies to do this work. He is uncharacteristically humble and philosophical about his contribution to the lab. He works primarily on the electrics. In the interview, he describes Stevie's architecture in terms of 'hierarchies' and tells me he is working with the lower level:

> The lowest level is the mechanical system and on top of that I would have provisioned, installed and designed for batteries and the power supply. That's the most basic, you're not going to do anything without energy. Once you have that structure, that system, you then build on the most basic components. So: sensors. As you read what position each joint is and you have little, very basic computers which will do, not a lot, but will read what the joints are and then have some sort of idea about what the joints should be doing and then try to adjust for that. So that's a sort of cause and effect there. But in terms the overall goal of what the machine wants to do coherently, as a holistic complete level that doesn't really have much to do with the work I've done. So basically, I'm like a slave carrying out orders from the masters. Yeh, in the computer science terminology there's like slave devices, master devices and things like that… (Interview 1/2/2019)

The team control every aspect of the robot, from hardware to software. As well as a 3D printer for the arms, head and torso, the lab has a large device for making the mould for Stevie's body. Stevie was built to be easily modular and upgradeable, and to make it possible for the hardware to be easily changed and updated. Although much of the software is bought 'off-the-shelf', this does not preclude developing it in-house at a later stage.

Stevie is humanoid, but he is not very human-like. In publications, they refer to him as 'human-inspired' (McGinn et al. 2020). The team are adamant that they do not want him to be too human-like or to risk falling into the territory of the 'uncanny valley'[3]. Additionally, they say, studies have shown that people's expectations are set by the outward appearance of the robot and the team do not want anyone expecting too much human-like intelligence from Stevie.

His height and appearance are in keeping with the current accepted thinking on social robots. Research from Japan, particularly on Honda's ASIMO robot, suggest that the optimum size for a social robot is 1.2 meters (Hirose and Ogawa 2007). For Stevie, the researchers 'through trial and error', a combination of needing to fit the parts and aiming for a reasonable stature, came up with 1.4 meters. As they explain to me, this is tall enough to 'have presence', but small enough not to be too imposing or intimidating or dangerous if it falls over. It also has the advantage of being eye-level with someone who is sitting down or in a wheelchair. Trying to get a robot, especially a wheeled robot, to bend over, would add a huge amount of complexity, they explain.

The fact that Stevie has wheels, instead of two legs, is something that comes up a lot in this lab, and in others. Researchers assume that people expect bi-ped robots, like C-3PO in Star Wars. However, I am told repeatedly, these are in fact extremely difficult to build. While Boston Dynamics have distinguished themselves with the extraordinary feats of agility in their bi-ped and quadruped robots, the team tell me, that is because they have the resources to specialise in this one specific area. For everyone

[3] The 'uncanny valley' is a term coined by Japanese roboticists Mori (2012) and describes how an increase in human-likeness of an image or object can result in a feeling of uncanniness or eeriness in those that encounter it. It is elaborated in more detail in the 'Personalisation' chapter.

else, including well-known social robots like Pepper, wheels are a much more stable and cost-effective option.

Stevie's face is designed, in line with state-of-the-art HRI research, to maximise 'affect' and expressive capacity. He has large brown eyes, animated eyebrows, and a separate screen for his mouth. Stevie's head and face were created in collaboration with Irish National College for Art and Design. The two separate screens are based on literature from psychology suggesting that people perceive the top and the bottom of the head differently, but Daniel says, they are the first robotics team to implement it in this way.

The first Stevie prototype was covered in a soft velvety cream material, chosen so as not to appear too 'product-like' and inspired to some degree by the commercial therapeutic seal robot, PARO. While at the lab, I regularly participate in user trials, including one for selecting the 'optimum' material for the skin. This involved rating a range of materials according to my preference, both visually and touching them while blindfolded. Ultimately, the final material and colour are chosen in response to feedback from carers, who wanted something that was easy to clean and clinical. It needed to conform to medical standards suitable for care environments and is thus made of medical grade anti-microbial material. It is shiny and white. As is also observed by Richardson (2015) at MIT, the team at Trinity are very aware of critiques of the racial underpinning of robotics research. They are therefore somewhat apologetic of having produced such an extremely 'white' robot and go to great lengths to explain the rationale. There are also accents of yellow, chosen for aesthetics reasons.

The researchers did not intend for Stevie to be identifiable as male or female. Stevie's name and appearance, they say, were chosen to be deliberately androgynous. Although efforts were made to implement a genderless voice, they tell me that users generally found this difficult to comprehend. After a period of testing involving users, one of the male voices received the best feedback and was implemented. Despite their aspiration to gender neutrality, the male voice clearly identifies him as male, and pretty much everyone refers to the Stevie as 'he'. This is in contrast to the researchers who make a deliberate effort to call Stevie 'it' or 'the robot'. However, a few of the researchers admitted privately that they often forget themselves and start calling Stevie 'him'. Dónal tells me that the new version of Stevie is not as much like a man as the previous one, more like a woman. I ask why. He's not sure, just a feeling, he says. Brian interjects saying that the new body was more woman-like. It had curves. Apparently, the original has wider shoulders. I asked why the shape of the body was different this time. They tell me they had to make it that way to allow for the additional flexibility, the new robot can turn at the hips and bend a little.

Initially, the team thought that Stevie could be used to solve distinct problems, like attending to specific care needs. They had assumed that next steps would be to put functioning arms on it and get it to do service tasks. Apart from the fact that these tasks are, in fact, remarkably complex to replicate mechanically, initial feedback from potential users convinced them to focus on Stevie's communication and entertainment capabilities. For this reason, Stevie's arms are there purely to give him a human-like appearance and to carry out 'non-verbal' communication. They are deliberately designed not to let people think that they are useful and so end in small balls, rather than hands. He can, and does, use them for gesture, to convey emotion and direct attention.

Everyone, except the team, is very interested in Stevie's name. It is the first thing I ask them about. In all the time I spend with the team, it is also the first question that journalists, interviewers and potential funders ask. The team, on the other hand, are embarrassed by the question. It has so little to do with the skill and work that they feel that they bring to bear on the robot. The team have backgrounds in mechanical and electrical engineering, computer programming and machine learning. Very often, they have studied a combination of these disciplines, and this is what has led them to robotics. The name is irrelevant, and incidental, they insist. Nonetheless, I persist. They tell me that Stevie's name originally resulted from demand from those interacting with him. The team originally chose what they hoped was the gender-neutral 'Stevie' and subsequent attempts to rename him have been roundly rejected by a committed community of dedicated Stevie-ophiles. Brian is somewhat irritated:

> The name is the least significant thing that we are working on… you know? Call it whatever you want! (Interview 1/2/2019)

During the interviews, I ask about Stevie's intelligence. Everyone agrees that his 'intelligence' is basic. This, they characterise as being 'introspective', knowing how to orient and manoeuvre itself in space, being able to talk, a budding ability to recognise and distinguish things in the world, with rudimentary interactive abilities.

> It's a very basic form of intelligence. But it exists and it exists distributed so if you pull his arm he'll fight back. And if you hold him longer he'll fight back more. The dialogue system would have a certain amount of intelligence… from a technical point of view the actual recognition of the speech is quite intelligent but we can't lay claim to that because that is sent on to Google… parsing the language isn't really intelligent. (Dónal Interview 24/1/2019)

A more developed intelligence, I am told, would be greater autonomy, such as more autonomous movement, and an improved ability to react to people and things in the world, such as the ability to react to people's emotions:

> Well, when you say 'intelligent', especially in the context of machines, you would think of artificial intelligence as in human-level intelligence... things like language, social interaction, understanding context, empathy, and also abstraction, reasoning … if they could ponder, if they could create art and poetry. (Dónal Interview 24/1/2019)
>
> [Teleoperation] gets around the problem that AI isn't intelligent. Which it's not, not really. A robot cannot exist on its own in the world. Even the state-of-the-art. It will get stuck, things will go wrong, and it won't know what to do. (Barry Interview 24/1/2019)

Stevie's development is not funded by dedicated departmental or robotics funding, instead the team have managed to secure small amounts of funding in disparate areas such as innovation and healthcare. Ultimately, they hope to secure investors and develop a spin-off company. They are not part of a wider national robotics effort and, despite the engineering legacy of the University, there is little robotics history or institutional knowledge in this particular area. Despite this, or perhaps because of this, they are the only team that I meet attempting to develop a complete humanoid and social robot.

The Field

In a manner that is not at all typical for robotics teams, the team building Stevie have been engaging directly with communities of older people since the very beginning. During the time of my engagement with the lab, I accompany them on site visits with two different communities. First, I attend a series of focus groups with a group of older people connected to an Irish charity whose aim is to support independent living. Second, I travel with the team to a retirement community in the United States, where they are carrying out a long-term, embedded study. The retirement community is set in a rather salubrious area on the East Coast of the US and is established for ex-military personnel and their spouses. The team has been given an apartment with a kitchen and en suite in the independent living part of the community to set up a temporary lab. Three temporary desks have been set up around Stevie. It does not take long before they are covered with laptops, chargers, and tablets. There is a freestanding 'Kanban' board.

Testing Stevie outside of the lab makes it immediately clear just how urgent and pressing the challenges are related to his physicality. Simply staying 'on' and moving around take up much more of the team's time and energy than attending to behaviours that are traditionally thought of as 'cognitive'. Moving Stevie out of the lab means borrowing or hiring cars and vans with enough space for the boxes that house his various parts. A team needs to accompany the robot to disassemble and then later reassemble him. Travelling abroad brings even more challenges. He has to be split into several parts. Every package has to be less than 30kg or the baggage handlers cannot lift them. He has to be issued with a special passport, an ATA card, as he is classified as a 'good'. Once assembled and activated in the field, a whole new set of problems arise:

> While the robot navigation worked fine in the lab, it is not working here. Here they have carpets and the light from the large windows downstairs have disabled the sensors. There's a problem with one of the bearings and smoke is coming out of Stevie. It takes two people to take him apart and because there is such a small team here on the

> ground, it is not that easy to fix. They manage to stop the smoke with some lubricant. That was yesterday. Today, for apparently no reason, one of Stevie's arms keeps flying up, but it is not critical so they leave it like that. (Fieldnotes 12/8/2019)

Researchers in other labs report the same problem. In Heriot-Watt, Christine tells me, '[s]o when I started working with [robots] I was like OK, so robots are very complex machines, they are not straightforward. You want them to go left, they go right, you want them to do this, they do that… oh my god!' (Interview 26/6/2019). In the US, other factors jar too. While Stevie's voice had appeared perfect in the lab, it somehow does not sound right here. His voice sounds very low, very English, comically English, like a villain in a Bond movie. Some of the residents ask if it is Daniel's voice. Stevie also seems very white. Glowy white.

The team are in constant contact both with potential funders and the media. Daniel carries with him a list of journalism contacts to send updates to, and there is a constant flow of interest from newspapers, as well as television channels from across Europe. Over the course of my time with the team, there is a continual media presence at the various trials and user engagements. During one of the focus groups in Dublin, the main Irish national TV station is there to film it. While we are in the US in August, a photographer and newspaper reporter from Time magazine are also present, resulting in Stevie's front-page splash in October 2019. Stevie is also regularly featured in national and international newspapers.

The Spaces of Robotics Research

Although the Robotics Lab in Dublin was my primary field site, I also visited two robotics research labs in the UK, Heriot-Watt in Edinburgh and Bristol Robotics Lab (BRL), as well as the Robotics Institute at Carnegie Mellon in the US. I also attended and presented at robotics conferences in Sheffield in the UK and in Tsukuba, Japan.

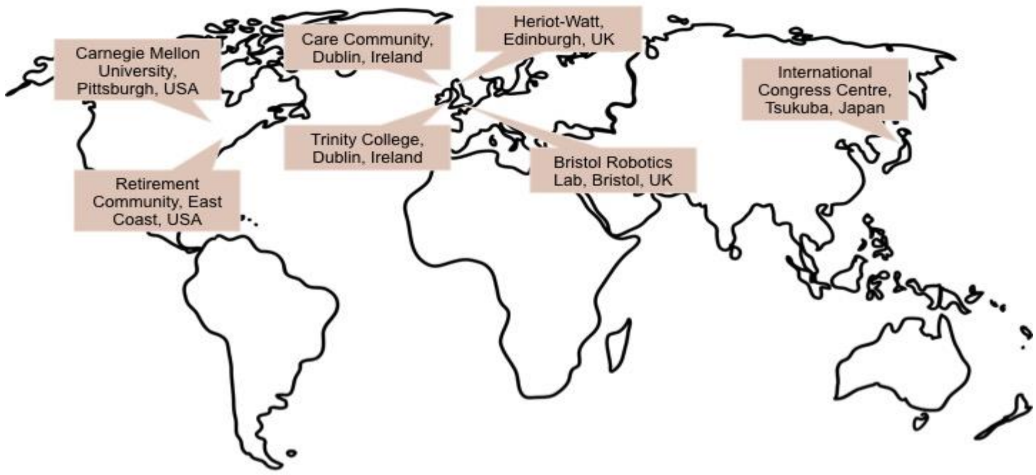


Figure 14: Map of Field Sites

The conferences

My first formal introduction to the world of robots and robotics research was at the Artificial Intelligence and Simulated Behaviour (AISB) conference at Sheffield University in the UK in April 2016. My knowledge of the field at this point had been mostly theoretical, and I had assumed that the other participants would be traditional engineering types, working on discrete aspects of robotics technologies and chasing very narrow versions of 'intelligence'. Instead, I found a field exploding with creativity and excitement. Humanoid robots were not a far-off dream, but already operational. Huge advances in mechatronics and swarm automation were creating moving machines that looked and felt sentient.

The convention consisted of a wide variety of topics related to AI and cognitive science, from AI and games, to embodied cognition and acting, human robot interaction (HRI), depression, artificial sexuality, and the social aspects of cognition. Thus, as well as engineers and computer scientists, there were philosophers, psychologists, theatre people, legal scholars, HRI specialists, artists, sex therapists, and priests. It was clear that something was happening, and if not the creation of human-like intelligence, then nonetheless a gathering of creative, conscientious, and energised people, organised around the image of the robot. When I entered the conference hall, feeling out of place, the first people I meet are a media academic specialising in sex therapy and a professor of theatre studies. My presence as a student of anthropology, far from being considered out of place, was completely unremarkable. There was an openness and inclusiveness to the field that I would continue to experience throughout my fieldwork.

Although there were many different disciplinary tracks, I was one of the few people moving between them. The topics in each track were similar, as researchers grappled with the philosophical and social implications of contemporary robots, yet each track dealt with these questions in completely distinct ways. Professor of Cognitive Science, Margaret Boden, was the star of the 'Computational Creativity' track, in which she proposed Google's Deep Dreaming algorithm as the 'new collage', challenging any essential difference between human-created art, and computer-generated models. In the 'Embodied Cognition, Acting and Performance', theatre scholars and philosophers reflected on the embodied nature of performance, suffering, and empathy with robots. In 'Social Aspects of Cognition and Computing', speakers discussed the social, ethical, and philosophical aspects of robots. The HRI, track was by far the most 'techie'. Research groups, including a large contingent from MIT, presented their controlled experiments between robots and 'human subjects', measuring discrete social variables such as 'joint attention' and 'gaze following', 'turn taking', imitation, socially-tuned perception, social guided visual search and human-robot 'personality matching'. To my surprise, the human 'subjects' whom they were targeting with their work on social robots were almost all children, autistic people, and older people.

Despite the variety of disciplines and tracks present at the AISB conference in 2016, most tracks had an 'ethics' focus of some kind. What was meant by ethics, however, varied widely. The more philosophically-oriented tracks debated how well traditional philosophical concepts stood up to the problem of potential robot consciousness, including, for example, how seriously we should consider robot suffering or robot rights from the perspective of traditional analytical categories. There was a lot of discussion about programming 'human values' into machines, including prolonged debates about the 'trolley problem'[4]. During these discussions, I was struck by the assumption that there was a universal standard of 'moral values' that was shared by all humans, despite the fact that most of the funding for the field came from the US military. I was also struck by the fact that participants seriously debated considering moral standing for machines, while ignoring the idea of moral standing for other sentient beings. Yet, other sentient beings display the autonomy, consciousness, and feelings that participants considered the yardstick for attributing moral standing to robots. The HRI group had a guest speaker who spoke about the risks of empathy and attachment. All of the ethical debates were characterised by disagreement, confusion, and contradiction, as participants from different academic traditions and none struggled to understand each other.

AISB was also the setting of a symposium to develop and debate the 'Principles of Robotics', which had originally been created in 2010 at a retreat funded by the UK's Engineering and Physical Sciences Research Council (EPSRC) and the Arts and Humanities Council (AHRC) (Boden et al. 2017). At that time, they had gathered together a group of the UK's senior roboticists to promote acceptance of robotics by the general public, so that it would not face the kind of opposition that GMO had. The group had decided to take it upon themselves to create design principles for 'regulating robots in the real world'. The original principles consisted of five rules and seven 'high-level messages', deliberately building on Asimov's three laws. The new principles were as follows:

1. Robots are multi-use tools. Robots should not be designed solely or primarily to kill or harm humans, except in the interests of national security.
2. Humans, not robots, are responsible agents. Robots should be designed & operated as far as is practicable to comply with existing laws & fundamental rights & freedoms, including privacy
3. Robots are products. They should be designed using processes that assure their safety and security.
4. Robots are manufactured artefacts. They should not be designed in a deceptive way to exploit vulnerable users; instead their machine nature should be transparent.

---

[4] The trolley problem is an ethical thought experiment in which a person has to decide whether to divert a train, or 'trolley', to save five people and instead kill just one who would not have been killed without the intervention.

5. The person with legal responsibility for a robot should be attributed. (2017)

At AISB in 2016, the original attendees were back to develop and debate the principles. When I first read the principles, I was surprised by their vociferous pragmatism, a trait that also struck me about many of the participants with whom I became acquainted over the course of the conference. Thus, before I ever encountered a robot, I found myself already entrenched in debates about the ethics surrounding them. At the symposium, there was a pervasive sense that the participants gathered were responsible for guiding the rest of society in their use of technology and to make the world safe. When they talked about the public, they considered questions such as ‘should we allow people to use robots?’ It was also due to the symposium that I had the good fortune to find myself drinking *sake* with some of these roboticists, who would ultimately become some of my primary interlocutors in the UK and US. The conference was a fascinating and exciting entry to the field and one in which I determined to chase up humanoid robots that, while animating all discussions at the conference, I still had not encountered.

The following year, one of my papers was accepted at the International Conference of Social Robotics (ICSR) in Tsukuba, Japan, see (Veling 2017). Japan is the only major rival to the US dominance of robotics and is, many roboticists believe, the ‘spiritual’ home of human-like robots. Throughout my fieldwork, researchers spoke enviously of the acceptance of social robotics in Japanese society, their culture, and Shinto beliefs, which made it acceptable to attribute spirituality to things, and the government support for the robotics programme. The most successful commercial robot, the furry seal PARO, is of Japanese origin, although it is very far from what many people, at least in the Euro-American imaginary, would associate with the idea of a robot. The sleek and minimalist design of Honda’s ASIMO provides the template for today’s social robots.

Tsukuba is about 50km north of Tokyo, and one of the largest science and technology sites in Japan. It is also the location of the National Institute of Advanced Industrial Science and Technology (AIST), where

Paro was developed. Tsukuba has been designated a ‘Mobility Robot Experimental Zone’, which allows for robotic experiments on public roads and is marked by signs across town.

Figure 15: Robot Experimental Zone Signs in Tsukuba

However, despite its cultural setting, the conference proper was focused narrowly on the field of social robotics, which is seen as a sub-field or HRI, and dominated by the norms of that field. The research consisted almost entirely of studies modelled on experimental psychology, using quantitative methods in laboratory settings to test specific effects of robot morphology or behaviour on humans. Apart from HRI researchers, there were also many present from the field of psychology. Papers included topics such as the impact of robot errors on people’s trust, and the influence of robot body shape on perceptions of gender, and the influence on the robot’s interactions style on user performance. These studies use measures such as physiological measurements (such as pupil diameter or heart rate) or statistical mechanisms, such as Likert scales and other measures of ‘user acceptance’. This limited metric can only measure existing norms and majority expectations. Research was focused on finding statistically significant differences in the data, with an emphasis on differences between genders. I found a marked lack of critical engagement with the topics, or with the potential impact of reproducing social norms in the technology.

The studies relied heavily on ‘Wizard of Oz’ methods in lab environments, in which a researcher teleoperates the robot to emulate ‘intelligent’ behaviour, often covertly. In surveys, participants were often recruited via Amazon’s Mechanical Turk, a platform we will discuss in more detail in the ‘Reification’ chapter. The effect of these approaches is to make a research field that is, as one of the pioneers of the field admitted to me privately, ‘prematurely conservative’.

Due to this narrow focus, many HRI and social robotics conferences reserve alternative tracks and workshops for divergent themes that are also a part of the discipline. Thus, as well as micro-validations going on in the main conference, workshops included the kind of topics that threatened from the fringes to destabilise the serious and scientific image of the field, such as that ‘robots and religion’ workshop. This workshop featured Anne Foerst, who had served as the theologian and ‘spiritual advisor’ for Cynthia Breazeal’s team at MIT, and a transhumanist pastor who spoke about the hope of immortality through AI. There were presentations on the ‘Tao’ of robots, and a demonstration of the ‘first Catholic robot’, the ‘Sanctified Theomorphic Operator’, or SanTO. Japanese robot designer, Tatsuya Matsui, of ‘Flower robotics’, spoke at length about the robot aesthetic, which, like a flower can be delicate and beautiful, and in need of nurturing.

At the conference, there were also a number of special awards for innovation in robots, one of which was won by a robot from Dublin called Stevie. It was at this conference that I met two other key informants and subsequent collaborators, Daniel and Gerry, and after which I became truly engaged in the field.

Heriot-Watt

The first lab I visit, other than my primary field site in Dublin, is Heriot-Watt in Edinburgh, Scotland, where Gerry and his team are engaged in work on social robotics. I first travelled there in 2018 to carry out field research and again the following year when I returned to the same site for more fieldwork and interviews. The robotics research is housed at the School of Mathematical and Computer Sciences, with labs dedicated to swarm robotics, drones, and HRI. During both my visits to the lab, I am given a desk with a number of PhD students, but most of the days I spend following Gerry around, observing and discussing his work. Much of his day is spent attending project meetings discussing current and future projects. Unlike the team working on Stevie, the teams in Edinburgh are not building their own robots; rather the bulk of their work is focused on programming off-the-shelf robots to develop 'naturalistic' interactions between robots and people. The lab hopes to distinguish itself through its focus on HRI.

> The HRI lab is a very large room filled all sorts of humanoid and zoomorphic robots. A large 'FLASH' robot stands in the corner going through a range of facial movements or emotions. Opposite, an adorable iCub is going through a yoga routine. Under a table is Sony's robot dog, AIBO. On the table, there are two small robots, a Miro, a 'bio-mimetic' robot designed by a collaboration between UK roboticists and the Sebastian Conran's design studio. On the table were a few Anki Cosmo robots, mostly used for outreach and teaching. (Fieldnotes 17/5/2018)

Gerry's work is focused on combining machine-learning techniques with computational decision-making programmes. The project he is working on at the time of my visit is one in which they are manipulating the robot's expressive capabilities to see how people respond to them. This robot's face has only 3-degrees of freedom, so they have to test whether the expressions they designed were still interpreted in the way that they wanted them to be by humans. Another researcher, Lena, is working on 'social signal processing', which is concerned with modelling observable social cues that are exhibited by people in social interactions. She is looking specifically at the problem of the robot knowing when to interrupt a person at the appropriate moment, a behaviour that, she tells me, is particularly important in hazardous environments. The method she is using is to try to detect 'cognitive' or 'mental load' in the user. This they are doing by measuring pupil diameter, which has, Lena tells me, been used in cars to see if the driver is still attentive. Thus, the system might recognise the pupil diameter and then interrupt, or not interrupt, on that basis, or change the behaviour of the robot system. They are also measuring heart rate variability, skin conductants, and blinking frequency.

Sara, a professor in the department, is focused on models of 'action selection', and questions such as: 'How does the robot know what to do?'. AI, according to her, is the art of 'doing the right thing'. The set of possibilities is finite, she explains, constrained by the physical nature of the artefact. This then needs to be engineered on a number of different layers of abstraction:

> I can go to the office today. That would be a high level of abstraction. I can walk towards my front door and turn the knob would be a much lower level of abstraction although quite closely related because I might be doing that so I can go to the office. So, another abstraction. Time scales. Abstract actions have long time scales. Concrete actions have short time scales. You shouldn't be so focused on getting to the office that when a fire breaks out in the kitchen behind you, you still go to the office. That would be silly. We're good at that type of thing. We don't persist with goals when it is clear the goals should be substituted with something else because we manage our goals quite flexibly which is much harder for robots to do. (Interview 27/6/2019)

Sara is using models of empathy and affect to try to model these behaviours. Separate to these, she tells me, is social context, which can be brought to bear through the use of techniques, such as roles, narrative and games:

> A role will give you a kind of a determining envelope for behaviours. If they are the boss and you are the subordinate. You have a social relationship. This is all culturally mediated. Power relationships are culturally mediated. Gender has an influence and that depends on the culture as well. (Interview 27/6/2019)

A combination of the three, empathy, affect and social context, she says, will allow them, not just to reflect the internal model affective state of the robot, but also to project the consequence of a particular expressive behaviour as 'an action'. This will allow the robot to reflect on, and judge, whether a particular action it carried out was appropriate or not. However, she says, tracking social context is incredibly difficult, even having a vocabulary for what the social context is, and connecting that to the appropriate social signal. The solution, she says, would be to get 'theory of mind' working accurately. However, what the ultimately means, she admits, isn't entirely clear:

> There isn't universal agreement in psychology about this. I'm only a computational person, I can only take what I think I can… I take a view from psychology, it's not the only view. (Interview 27/6/2019)

However, as we have seen, in actual experiments with robots in the field, or 'in the wild' as it is commonly referred to by roboticists, it tends to be much more prosaic issues that present the biggest challenge. In this case, it was the issue of ensuring the robot had enough power to keep running:

> We ran a robot for three weeks continuously, but we made a serious design error, which should have been obvious but wasn't, like other serious design errors. So, like I say, robots can run for about two hours without needing to recharge, so we designed in an automatic recharge behaviour for the robot… we had the recharging unit in the corner and whenever its battery was low the robot would drive into the recharging station and recharge. The easy way of doing this is to drive in frontwards. Big mistake! Well, it can't interact with anybody, can it? It's facing the wall. We should have thought of that! And people got really upset about it, because, you know, the robot would have to recharge for two or three hours. It takes time to recharge a robot. It doesn't happen just like that. They are big pieces of machinery, typically, with big batteries in them that aren't going to recharge in just five minutes. So, it takes two or three hours to charge the robot as well. And during that time, it wasn't able to interact. Had we had it reversed while in the charging station that would have gone better. (Interview 27/6/2019)

Across campus, a Living Lab has been set up as a model apartment complete with kitchen, sitting room and bedroom with double bed and wardrobe, and an en-suite bathroom. Lucas shows me around the lab that is intended to test the robots as part of a wider 'smart' environment within a 'natural' setting. There are two robots in the apartment, Softbank's social robot Pepper, focused on social interaction, and the vaguely humanoid PAL TIAGo, designed to perform tasks, like picking things up. The floor is smooth linoleum throughout and there were no tricky things like awkward lips between floors. There are sensors under the floorboards and in the kitchen cupboards and cameras in the corners. Later, when I interview Lucas, he is eager to stress the pragmatic nature of his work. He is not interested, he stresses, in understanding 'the inner mechanisms of intelligence in humans', instead he's trying to 'emulate it with planning, classical AI combined with machine learning' (Interview 26/6/2019).

However, towards the end of the interview becomes more animated. He tells me about an article that he has just read in the UK Times about a forthcoming book by James Lovelock, creator of the Gaia hypothesis, see (Gill 2019). According to Lovelock, Novacene, or the age of cyborgs, will supersede the Anthropocene. For Lovelock, this represents the next stage in human evolution. These cyborgs will not be humanoid, an idea about which Lovelock is scornful, according to the article. Rather, they will likely be spheres. But we do not need to worry about them killing us, because these new beings will need us just as we need plants. As part of Gaia, it will be in these new cyborgs' interest to continue to regulate the planet's ecosystem and keep us cool (below 50 degrees Celsius). This future is driven by evolution and natural selection, says Lovelock, 'you can't avoid it'. Lovelock is an eco-futurist, believing that engineering will save humans from catastrophic and existential harm. 'It's a realistic picture, I think', Lucas says.

A huge amount of researchers' time in Heriot-Watt is dedicated to doing demos, for schools, for the public and for the media. Their relationship with the media is fraught having had a particularly damaging experience with media, when a week-long experiment of embedding a robot in a local clothes shop had resulted in negative global coverage of a robot being 'fired' (insider.co.uk 2018). With a few exceptions, when asked, most of the researchers that I interview expressed their frustration with how media report on robots, how they inflate expectations and scaremonger, and how they misrepresent the research. Very few enjoyed an easy relationship with the media. And yet, while it is undoubtedly true that the there is a strong public appetite for stories of imminent AGI in humanoid dressing, it is not the case that these stories are solely being perpetuated by media or marketing rhetoric. There is a substantial effort to maintain the mystique of imminent AI from academia to ensure the interest, and the funding, are maintained.

> What tends to happen is that people over promise, under pressure, because you don't get the money if you don't [funding] and then people on the other end who have far too high expectations say you aren't delivering, are you? This is crap, I wanted something else. Crash! You know? I've been here before. (Sara Interview 27/6/2019)

Bristol Robotics Lab (BRL)

After Edinburgh, I travel south to a robotics lab in Bristol. This lab is very different from anything that I had seen before: a dedicated robotics space and the largest in the UK. It is an impressive and large building surrounded by wild plants and grasses. Two or three massive industrial robotics arms function as spotlights at the front. There are a number of portaloos stationed outside, which I later discover are an innovation that has arisen from research into bio-fuels to power the robots. As a result, they have pivoted to developing electric-lit toilets for use by women in India, and at the Glastonbury music festival. There is a reception area with a sign telling us that photography is forbidden. I have to wait here until someone comes to meet me. Inside there are open plan, yet intimate, spaces bounded not too linearly by glass walls that double as whiteboards covered in impressively complicated looking algorithms. Each cube hosts a different research area, including swarm robotics, social robots, 'soft robotics', tactile sensing robots, robot prosthetics. In the centre of the space is a large aerodrome for drones and aerial robots.

A substantial research focus at BRL is devoted to the problem of energy and the amount of energy needed in order for the robot to 'maintain autonomy'. The bioenergy and self-sustaining robot research groups are focused on 'feeding' the robots, essentially on the problem of how to generate electricity from naturally occurring biomass. The three sources used are urine, flies, and slugs. A small robot machine outside the door on a set of tracks is powered by flies, acting a bit like a Venus Flytrap plant might, slowly dissolving the flies in its mechanism. The organic waste is converted to electricity using 'microbial fuel cells'. Like in biology departments, the doors to these labs are closed and have hazard signs up. In 2013, the team at BRL developed a way to charge mobile phones using just neat urine (Ieropoulos et al. 2013). In the toilets at the lab, cardboard bowls are provided so that researchers may harvest their own urine on site.

Richard, who I first met at AISB, is my key interlocutor here and I have been given a desk with his team. He is one of the senior founders of the lab and no longer works there full time. Richard's interest is in programming robots to engage in ethical behaviour and his students are coming at the problem in disparate ways. One of them, Danni is working on a project with swarm robots focused on emulating group behaviour and ‘collective decision making’. Swarm robots are small and simple physical robots, out of which, working together, more collective and complex behaviour ‘emerges’. They can be programmed to react to each other in simple ways; for example, ‘cooperation’ mode might describe a number of robots programmed to focus on the same task. In ‘organization’ mode, a single robot may control the others. Behind us, there is a white plastic surface with 1,000 tiny ‘kilo-bots’. When they are on, their simple sense-and-act system cause them to move towards, and then be repelled by, a neighbouring robot so as not to hit it. The result is a remarkably life-like display, which, in combination with the sheer number of robots, sounds like heavy rain on the plastic roof. It feels like a swarm, but because the robots are tiny and shiny and silver with blinking lights, it feels a lot less repellent than an insect or rodent swarm might. They also have larger swarm robots, called e-pucks, which are ‘smarter’, that is, they have more computational power and can be programmed to do more. In Danni’s study, she is programming the robots to collaborate on moving a Frisbee across a platform. First, it is programmed in a simulation. The robots have to ‘adapt and learn’, using programmed positive and negative feedback to build ‘consensus’. Danni talks about a scenario where these robots could be used, for example, in old, abandoned, nuclear reactors where you could just cut a small hole in the wall and let these robots in. The robots, she says, unlike humans, are expendable. So, while some of them might ‘die’ from the nuclear toxins, enough of them could survive and do whatever job needed to be done.

Another researcher student, Max, is working on developing an internal model for a robot, in which a robot runs a simulation of itself and its environment inside itself. On the screen, Max has created a 3D simulation, in which the robot is represented by a cube. The small robot-cube is positioned on top of another, larger cube, which represents the ground. There are other potential obstacles, represented by different 3D shapes: a cone, a triangle, and a cylinder. Ultimately, this will allow it to operate as a 'real-time Consequence Engine' allowing it to model and predict the consequences of its own actions, and that of other agents in the environment. It is hoped that this might be a step towards building self-aware behaviour, and ultimately even ethical behaviour. At the moment, for example, it might be able to make ethical decisions such as stopping other agents from falling down a hole.

Between the research labs and the restaurant, there is an enclosed space, which Richard tells me is the 'incubator'. It is, I am disappointed to find out, not full of 'baby' robots, but rather a business incubator where people who have start-ups in robot-related business can rent some space. I tell Richard my joke about the baby robots. He is not as amused as I expected. He explains that they do, in fact, have a plan for 'evolutionary' robots, in which the robot will build itself in a 3D printer, and then emerge out into a robot-learning environment, where it can learn from other robots (but only those robots that were themselves sufficiently advanced). Only the most successful robot will go on to 'reproduce' their algorithms in the next generation.

At one end of the lab there is a display of 'dead' robots, including an early shrewbot. Shrewbot is particular to the lab at Bristol and consists of a body with a single mechanical arm out of which whisker-like tactile sensors protrude. It does not have to rely on light for vision as other robots do, and so can navigate down pipes and in the dark. Most of the robots at the lab in Bristol are not humanoid. Humanoid robots feature just two research areas, a small living lab in the centre of the space, and the HRI research group.

After presenting my own research to the group at the lab, I am adopted by two HRI researchers, Lucy and Stefan. Lucy is a PhD student and is looking at the potential for robots to influence human behaviour. In her study, three robots try to persuade someone that they should exercise. The first robot is programmed to use 'non-committal technical language', the second will use the first-person pronoun 'I', and the final robot will quote others, such as 'your therapist told me to tell you…' The test will evaluate which type of prompt is more persuasive. They are excited about the potential for Machine Learning to create ‘social behavioural datasets’. Lucy compares machine learning to grounded theory and wonders whether it could beat an anthropologist in identifying phenomena in data.

Like all of the other labs, my presence in here is not in the least noteworthy for teams who are used to a constant stream of commercial funders and journalists. Indeed, when I first arrive at any of the labs, including this one, I am generally offered a tour, in which researchers working on their robots dutifully pause, stand up and explain their research to me as if I am some visiting dignitary. My presence offers another opportunity for junior researchers to practice their demonstration and media communications skills, which is such a major part of this research, and something they are constantly in training for.

Carnegie-Mellon University (CMU)

As we have seen, it was in the US, rather than in Europe, or even Japan, that the programme of intelligent robots first developed as a serious research programme. The ‘big four’ centres had been working in earnest on trying to create human intelligence since the first research grants in the early 1950s. According to McCorduck (2004), each of the four main centres evolved its own style based on the personalities of the men who dominated it (133-134). While MIT (after Minsky), was ‘haute couture’, ‘stylish yet faddish’, CMU (after Newell and Simon) represents old-world craftsmanship, attending to detail and using the finest materials, ‘classic but stodgy’ (2004, 134). As described earlier, both Suchman (2007) and Richardson (2015) conducted fieldwork at MIT’s Computer Science and Artificial Intelligence Laboratory (CSAIL) in the early 2000s. My research took me to another key centre, the robotics research institute at Carnegie Mellon in Pittsburgh, which I visited in October 2019, carrying out observation and interviews.

To a visitor unfamiliar with the geography of the US, there is an almost eerie juxtaposition between the high-tech architecture and prestige of the university and the rusty, industrial, and shabbily-faded glory of the surrounding area. Pittsburgh is a proudly industrial town, known as the ‘Steel city’ having produced 60% of US steel between 1870 and 1910. The legacy of this economic boom is still evident in the cultural centres, many of which are the legacy of industrialist Andrew Carnegie, including the university and the Pittsburgh museums. Westinghouse was based in Pittsburgh and Herbert Televox was first displayed here. Robots are thus a part of the city's history and identity. However, in the 1970s and 80s, de-industrialization brought decline, and with it the loss of 130,000 manufacturing jobs, mostly in steel. Unemployment ran at a rate of 18% (Andes et al. 2017). Now, along with states such as Ohio and Indiana, it is known as part of the 'rust belt'. When I first arrived to one of Pittsburgh's less salubrious suburbs for my stay, the decay was palpable. Sidewalks and services were crumbling, metal signs and bridges were rusted, electric wires hung low across roads, and there was a deeply unpleasant odour pervading the air. Today, Pittsburghians' relationship with technology is fraught. While benefitting from the steel mines, the few that are left are pumping out harmful emissions (Maher 2019). A report from the National Bureau of Economic Research (NBER) warns of concerns over Pittsburgh's high 'exposure' to robots (Acemoglu and Restrepo 2017). Pittsburgh is also test site for self-driving cars. While no one has yet been killed there, unlike Tempe Arizona and Oakland California, an Uber AV in Pittsburgh swerved onto the sidewalk and continued driving without the technology correcting the behaviour (Kang 2017).

When I visit, Google Maps proves to be useless as the fact that many of the streets cross each other at a massive vertical distance is lost to the Uber App and, as a result, also on the Uber drivers. Thus, my first impression of the university is not the impressive Beaux-Arts style buildings and wide promenade that greet the visitor to the main entrance of the University at Forbes Avenue, but rather a service entrance at the back of a car park underneath one of the side buildings.

Carnegie Mellon's Robotics Institute (CMRI) is part of the school of computer science and, founded in 1979, one of the oldest robotics departments in the world. It is huge, with over 50 full time faculty members, spanning seven floors and several buildings. Bob, a veteran CMU roboticist, agrees to show me around. The newest building is the Newell-Simon Hall, which was built, Bob tells me, to rival MIT's Frank Gehry building. He points out a photo of Newell and Simon in the entrance hall, 'we each have our founding fathers', he says, referencing MIT and Stanford, 'these', he says proudly, 'are ours'. Opposite is the boarded-up cubicle that used to house the 'roboceptionist', a collaboration between the Robotics Institute and the School of Drama. The original character, Valerie, with a 'brilliant blue-eyed gaze', would sense people walking past and offer them assistance (Gutkind 2009). Most of the time, however, Valerie spent on the phone revealing secret details of her private life, intimate conversations with her psychiatrist, and aspirations to be a lounge singer (2009, 104). Perhaps it was the breath-takingly gendered stereotyping that led Valerie to be replaced, in 2005, by 'Tank', 'son of a NASA scientist who had aspirations to be an astronaut but was grounded by his vertigo' (2009, 106). By the time I arrived, the roboceptionist was gone. No one could tell me where.

The idea of robots roaming the halls of the academic centres that birthed them has become almost a cliché, reported endlessly by journalists, academics, and other visitors to the centres. And so, I was a little disappointed not to encounter any. There were a few static ones, however. Up on the first floor, an immobile robot sensed passers-by and tried to engage them in conversation. In the restaurant, a robot chess player, his image now decentred in the tilted TV screen, challenged people to a game of chess. None of the students or faculty at CMU showed even the slightest interest in these displays. They were, however, shown to visitors, who would obligingly and inevitably attempt short, awkward conversations with them.

The HRI institute at CMU is part of a separate, dedicated HCI/design department, which although part of the computer science department, is not part of the Robotics Institute. It is, however, also housed in Newell-Simon. According to Bob, it is the 'left-side' of the brain to the Robotics Institute's 'right-side'. Although I am based on the Robotics Institute, I do visit the HRI department. Inside, the walls are colourfully painted, and feature lots of sofas and 'break out' areas. Upstairs one of the workshops has displays on the walls of 'smart materials' for sensing and moving. There are beautiful, knitted robots, 3D printed plastic objects, foam and cardboard models. A large room dedicated to ‘futures work’ allows for the prototyping of environments, such as, one of the researchers tells me, a completely smart home for teenagers which they used to understand how much ‘smart’ technology was too much. They also used it to recreate ‘service’ environments, so they might, for instance, build a shop or a restaurant out of cardboard. The futures work, I am told, allows them to experiment with possible technologies without having to have those technologies in place yet.

Each of the different research areas of the Robotics Institute is spread across the six floors of Newell-Simon and parts of several other buildings. Newell-Simon is built on top of the ruin of the old United States Bureau of Mines. There are still train tracks coming into one of the lower floors. We see a pair of very old ‘Adept’ robots that Bob says he has rescued from storage. They are huge, very solid machines, with a single arm covered in extendable plastic. This is the challenge for soft robotics, he says, as the rubber on one of them has all but disintegrated.

Across the many floors, there are rooms with different research areas. ‘Bio-robotics’ is at the top. The ‘bio’ in the title is loosely interpreted, meaning they look to biology for inspiration. In this case, it appears to mean based on animal physiology, rather than attempting to create humans. The robots here are modular and can be put together in any configuration. The first thing I see is a robot snake. There is also a snake obstacle course, and snake sandpit. Another robot made of the same modular sections but assembled in a different way resembles a spider, with eight legs.

In one large room, a researcher is listening to ‘Scarborough fair’ by Simon and Garfunkel and growing plants in ‘autonomous environments’. In the lower levels, there is a massive arena with a huge water tank for underwater robotics, and a netted area or for testing drones. Here I can also see industrial robots, self-driving vehicles, lunar robots, artic explorers, and enormous spider robots. At the sides of it there are workshops full of people working, making things. Here they are referred to as ‘shops’ rather than labs. Another level down is ‘the catacombs’ which they are using for testing subterranean robots. This part is left over from the original buildings. Some parts are not even head height. There is a massive fan and an enormous block of coal.

There is a 'panoptic studio', a huge, domed, igloo-shaped structure built to capture motion and social interactions in 3D. The system has 480 synchronised video streams so that it can capture a large amount of people's movement in space. This then produces a 3D output of 'labelled, time-varying anatomical landmarks'. These movements, captured algorithmically, can be used both to predict and to reconstruct 'skeletal trajectories'. There is a manipulation lab, and a lab for connecting vision and manipulation in which a gripper is trying to pick up coffee beans.

Manipulation is a big part of the robotics research at CMU and one more associated with industrial robots than social ones. According to Nick, a roboticist at CMU since the 1980s, manipulation, or 'making other things move', was originally what robotics was about, and people were more interested in manipulation than the mobility and locomotion of the robot itself. I interview Nick in his office. Behind him, there are a bunch of medals and awards, including some from DARPA. Back in the 1980s, he tells me, robots were almost entirely destined for manufacturing. 'And by the way…', he tells me:

> …it's also not clear that those industrial robots that I talked about should be called robots. That was a subject of dispute. Maybe still is in some quarters. The reason they are called robots is because of Joel Engleberger, the guy who founded Unimation, he thought it was a good marketing thing to call them a robot.' (Interview 9/10/2019)

Like many roboticists, Nick is also a chief Scientist in a company developing robotics for industry. For Nick, the ultimate test of a robotics device is to succeed commercially, to find a niche where they can do something that is economical relative to human beings.

Nick is somewhat defensive of people who consider manipulation uninteresting or accuse it of not having to deal with uncertainty. For Nick, making an instrument that deals with precision is always about dealing with uncertainty, and thus also about intelligence. He talks about mechanical intelligence and morphological intelligence, and the fact that intelligence can ‘drift back and forth’ between the mechanics or hardware, the software and the designer.

Tina is also a professor at the manipulation lab. She was drawn to manipulation, she tells me, because of the precision and the ‘science’ of it, rather than what she calls the ‘tinkering’ of robotics. Tina’s team are working on soft robot hands, using materials like rubber, cloth, and foam. The hands, which are coloured green and purple, have four fingers and wobble in every direction, looking eerily alien. One of the advantages of soft fingers, Tina tells me, is their ability to collide with each other without incident.

> With a rigid robot we’re always worried: oh, it’s going to collide with something, it’s going to break! And there’s all this concern about something so simple as the fingers running into each other. And with the soft hand it’s just such a relief we don’t care they don’t run into each other. Nothing bad is going to happen. (Interview 15/10/2019)

Tina would like to develop a ‘grand unified theory of grasping and manipulation’, a classification system that would describe how people grasp and manipulate objects. She suspects that there are a finite amount of things, perhaps a hundred, that a robot could do that would render them human-level capable. It is just about finding which hundred, she says.

I also interview Bob in his office. He has been at CMU for nearly 20 years. Before coming here, he held positions at other prestigious institutions. He tells me he is sceptical of other researchers, from Breazeal to most of my informants, who trace their interest in robots back to R-2D2 and C-3PO in Star Wars:

> That’s just a myth, a story that they have created to explain to themselves a much longer term and deeper phenomenon. If the ground is not prepared, *Star Wars* isn’t going to do it for you. In my case, I was a very clumsy, chubby, incompetent child. And I had many brothers and my sister really like to point it out … so the idea of trying to understand that, and master it was very attractive to me. And that is why I am doing what I do. (Interview 8/10/2019)

Unlike many of my other informants, Bob is not concerned by media representation of his field. In fact, he is the second of my informants to have a close relationship with Hollywood studios, having worked closely with major Hollywood studios on movie productions featuring robots.

Bob’s office is full of junk. It is like a mechanic’s workshop on top of which someone has dumped a whole load of ageing computing machinery from on high. He has to climb over things to get to the desk. For Bob, the architecture, and the junk, are integral:

> I think the architecture of a building really has an influence on how people think, which is why I like junk, it helps me think, you’re not anybody without a pile of trash. Problem is, they keep tidying it up. (Interview 8/10/2019)

While Bob has raised chaotic junk to an art form, just like all of the other labs that I visit, at CMU, there is robot and machine detritus is everywhere. Robot hands lie on tables, robots stand motionless facing walls, and bits of robots are thrown into corners, having been raided for parts. Tables and shelves are covered with plastic boxes full of wires, nuts and bolts, labels, screws, paper, bits of plastic and foam, electrical circuits, old motherboards, wood, connectors, papers with diagrams. Some labs are more deliberate with this aesthetic than others. In both BRL and CMU, there are semi-formal ‘roboseums’, featuring select early prototypes and experiments that have been retired, and have gained a retro feel.

The sight of abandoned bits of robots is, for the roboticists, simply characteristic of a workshop. To an outsider, it is somewhat unsettling. Behind a whole load of displays, I spot a pair of eyes:

> What’s that? I ask. That’s our man machine, he says. It’s been abandoned. This is where robots go when they are retired. Man machine stares out at nothing without eyelids. What do you call him? I ask. It’s a humanoid? He offers. There is no sentimentality here. On the table across from this is an exoskeleton attached to a pair of large, tan, cowboy boots. This will make you walk faster, I’m told. There is a large, wooden ball in the middle of the room. What’s that? I ask. You could try and stand on it, he suggests, if you have good insurance. (Fieldnotes 16/10/2019)

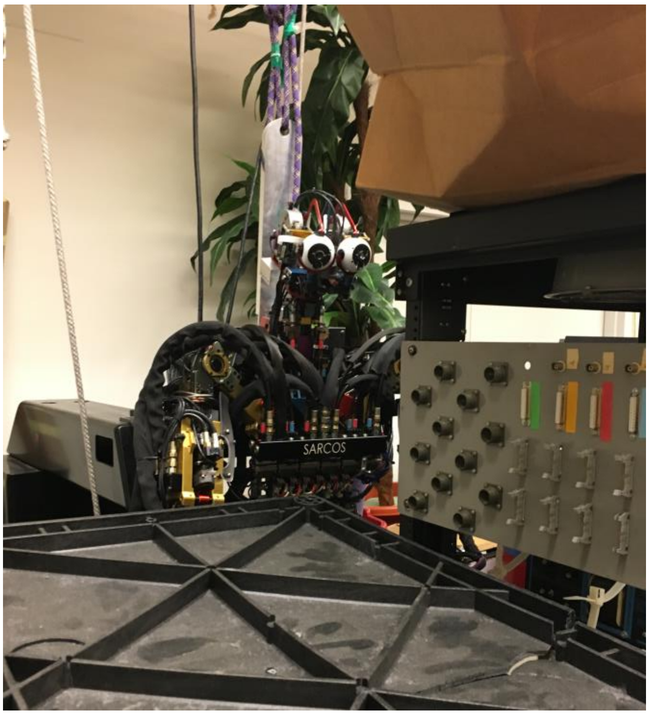


Figure 16: Man machine stares out at nothing without eyelids

Outside, in one of the warehouses between robotics buildings, I spot a large sculpture of an eagle, which one of their robotics teams won for the 2007 DARPA urban challenge. Apart from the prestige that accompanies such a feat for CMU, the challenge also included a cash prize of $2 million. A team from CMU also shared the $3.5 million cash prize at the 2015 DARPA Robotics Challenge, using Boston Dynamics robots. While DARPA remains the dominant funder when it comes to robotics research in the US, the influence of industry and venture capital on robotics funding is growing. A new 'big four' has emerged, namely the large tech companies: Facebook, Amazon, Google and Apple. At CMU, Disney and Apple had spaces just outside the boundary of the university, and Facebook were in the process of setting up there. During my time there, I have been given a desk at the Community Robotics, Education, and Technology Empowerment (CREATE) lab, which, as it happens, is in the same building.

The research agenda at the CREATE lab is different than the rest of the Robotics Institute, where the research is primarily motivated by military and industry. CREATE was founded with the explicit aim of refusing military funding. Their biggest challenge, roboticist Ben tells me, was, if not for the military, then what were their robots *for*?

> Then the question became now, what's the positive? You know what the negative is, what you are not doing. So, what *are* you doing? And indeed, that is the situation. If we are not taking defence funding, how do we get funding in a space where, back then, so much, or the majority of the work being done is defence-funded? And the answer to us was education. (Interview 11/10/2019)

Like the other robotics research spaces, this lab also has packed shelves, mechanical bits, breakout couches, and abandoned robots. But there is a difference. The walls here are packed with a much wider array of objects and technology hybrids. I can see feathers and bells, wires and glass balls, and something that looks like a car radio from the 1980s. A bottle of wine could be a project or a present. There are two cardboard figures: one waving a flag, the other looks like an Egyptian puppet. There is a shoebox with pipe-cleaners coming out of it. There are a few small, wheeled robot cars with wires, platforms, and cameras. There are also a large metal shelving units rack with over twenty servers stacked on them. This, I will learn, has become the central focus of the CREATE team's work. The researchers here are a mix of roboticists, programmers, educationalists, artists, and researchers from other disciplines, like geography.

The focus of the lab is 'to turn things around', that is, rather than trying to figure out how to get people to accept robots, they start with people and communities, and then try to figure out how the researchers' tools, expertise and prestige can be used to serve them. This means that the lab is deeply embedded in the local community from Pittsburgh and its surrounds. On the tables between the two couches is a small robotics project designed for pre-schoolers, called 'MindfulNest'. The interface is a tablet with a bright green cover. Attached to it is a black box which acts are a holder for three items – a plastic black and white flower with a green handle, a white tube that is rounded at the top, and a long blue tube. If a child picks up and smells the flower slowly, a scent will come out. If they sniff too hard – nothing will. If they pump the white tube slowly, balloons will float up on the screen. If they use the blue tube to 'conduct' music that is playing gently and with flowing movements, music will play. If they move it back and forwards too fast, nothing will happen. Ben says that kids had taken to going over to it when they themselves decided that they needed to calm down.

Barbara's background is in geography and mapping, and she is working on the air quality project in the Monongahela Valley. The steel works and coke plants that are releasing emissions are affecting the air quality and the health of the people at the intersection of the three rivers. There is also a large fracking industry. There appears to be a higher rate of cancer and certain illnesses that may be related to the pollution from these industries, but it is difficult to prove definitively. The CREATE team have identified three independent sources of air pollution and are tracking them. Working with families in the area, they are monitoring the air quality using robots that will measure it periodically throughout the day and night. The project has found, for example, that at 3am the levels of benzene go up. However, as Barbara relates, the issue and the solutions are not straightforward. Despite the negative effects of industry, the steel works and the fracking are a point of pride for people. It means they do not have to rely on foreign oil.

Another project focuses on mapping the local housing situation, such as house prices, where buyers were coming from, tracking gentrification in the Pittsburgh area. The project visualises inequities in Pittsburgh, including social structural inequities around housing, around education, as well as around looking at factors such as the education gap between different ethnic communities. Beth is giving a presentation to local agency workers. As well as the quantitative data that she presents on an interactive map of Pittsburgh, she is also giving them a qualitative sense of the situation, by playing them interviews with residents and showing very detailed, up-close photographs of different parts of the locality.

The team also provide spatial and temporal imaging for all around the world looking at pressing societal issues, such as poverty, violence, modern slavery and climate change. One of these projects, called Earthtime, is shared with Global leaders at Davos. At CREATE, they are using technology to empower people that would otherwise not be able to afford, or have access to, it. Funding for the CREATE lab is not as straightforward as for the rest of the institute. It is a balancing act, one of the researchers tells me. Each project has an eclectic group of funding relationships. Heinz Endowments are a source of funding. The pro-social orientation of the work has made physical robots almost redundant at the lab, and certainly anything that might be humanoid.

> We take the word ‘robotics’ very broadly in the Robotics Institute. … We use it to mean any system that has sensing, and complex machinery for cognition, and some kind of pushing on the world or action. (Ben Interview 11/10/2019)

The robot in the CREATE lab is less of a physical object and more of a symbol and rallying point, bringing together the capital, resources, data, power, influence and social influence of ‘robotics’. Instead of using it to increase power and profit, it redirects it, in an entirely uncharacteristic inversion, away from, in Ben’s words, the ‘hegemonic power structures, to challenge negotiations of power in society and to enact pro-social change’ (Interview 11/10/2019). The lab is a small glimpse of how things could be different. However, I suspect that its existence on the periphery of the university proper is the result of the determination of a few dedicated individuals, making its ability to influence the wider institutional network of robotics research less likely. In the wider Robotics Institute, there is still the idea that politics is separate from the field. In a long and rather combative section of the interview, in which I try to ask whether he has any concerns with regard to the impact of robots, including the data they need to ‘learn’, an exasperated Bob eventually exclaims:

> Now you’re asking me questions that are political philosophy! Holy crap! I don’t know! You are asking me very deep political questions! You guys are confusing… you think robots and AI have anything to do with these issues and they don’t [bangs table for emphasis]! Not in any deep way! … It’s like asking should we have coal burning power plants or nuclear power plants. It has something to do with robotics because robots run on electrons. But come on! Don’t ask the roboticists about it. They don’t have anything to say! (Interview 8/10/2019)

Conceived Space

For Lefebvre (1991), ‘conceived space’, or ‘representations of space’, is the abstract space as understood quantitatively. It the ‘dominant space in any society’ and is ‘the space of scientists, planners, urbanists, technocratic subdividers and social engineers, as of a certain type of artist with a scientific bent—all of whom identify what is lived and what is perceived with what is conceived’ (1991, 38-39). Although representations of space are abstract, ‘their interventions occur by way of construction’, such as architecture, but also, I argue, robotics, as ‘a project embedded in a spatial context and a texture which call for “representations” that will not vanish into the symbolic or imaginary realms’ (1991, 41-42). Crucially, *‘conceived* space is thought by those who make use of it to be *true*, despite the fact - or perhaps because of the fact - that it is geometrical: because it is a medium for objects, an object itself, and a locus of the objectification of plans’ (1991, 361).

For roboticists, the core problem space is translating from the extreme reduction of minimalist 3D simulated space to the ‘real world’, the physical world of texture, friction, and relative unpredictability. Often, the robot’s features and functionalities are first modelled in a 3D virtual or simulated environment, one that is empty, linear, controllable, and disembodied. Once verified in this space, they can be implemented in the robot. Thus, the art and craft of robotics also sees the practitioner in embodied and reflective conversation with the material world. As Agre (1997) has pointed out, this differentiates engineers from more abstract sciences, such as mathematics (40). Roboticists are actively engaged with the material and spatial work from which they get concrete feedback. It either works or it does not. Their work *feels* true. It is, as Agre describes it, ‘nothing less grand or more specific than an inquiry into physical realizations as such… what truly founds computational work is the practitioner’s evolving sense of what can be built and what cannot. This sense, at least on good days, is a glimpse of reality itself’ (1997, 11).

The advent of HRI and social robotics takes roboticists out of the field of engineering and computer science, in which producing a novel object that ‘works’ is in itself sufficient to render it ‘scientific’ and ‘true’, into the murkier world of the human sciences. Thus, robotics researchers find the requirements of their field growing, absorbing all domains. Already swamped with the difficulties of moving from the virtual to the physical world, the ‘social’ world outside the lab means that ‘in the wild’, for roboticists, can seem a genuinely terrifying place. It is little wonder then, that, as Richardson (2015) observes, for robotics scientists, the ‘social’ is located in the micro-behavioural exchanges between human and robot (19) and dominated almost entirely by quantitative methods in laboratory settings.

Roboticists also regularly come up against people’s high expectations when they encounter robots, people whose expectations come from watching science fiction movies and TV programmes and find themselves rather underwhelmed with the reality of robot capabilities today. The highly challenging integration of hardware, engineering and computational resources that are required to instantiate this humanoid, mobile, interactive entity go unnoticed, as people react more strongly to the blinking of the robot’s eyes and appear instinctively to search for limitations and shortcomings in the robot’s apparent intelligence. As we have seen, the field of HRI was established to exploit this tendency. However, not all of those in robotics are impressed by this development. Nick is dismissive of some of those from other disciplines developing an interest in robots, especially ‘Crazy HRI types’, including, I am both pleased and affronted, me.

> You say to somebody in HRI what would be the first thing you would do to make this robot more appealing or more acceptable or whatever. Their answer is always ‘you should put eyes on it’… and people say, you know, in HRI they like to talk about eyebrows. Eyebrows! The old hard-nosed crowd, well that’s something that they can goof on… You could ask the question why is it called robotics? Robotics meant to a group of people something more specific. Maybe it should be called robotics. I’m not saying it should or it shouldn’t… I just find it difficult to care that much about how humans are going to interact with machines when they have a certain kind of behaviour when we are still struggling so much to produce that kind of behaviour. (Interview 9/10/19)

Indeed, the pressing physical mechanisms and complexity of interacting with, and navigating, the physical environment continues to be such a fundamental issue for roboticists that concerns about implementing imprecise concepts such as 'intelligence', 'sociality' and 'context' remain, in many cases, distant, magical, and impractical. As Suchman (2007) reflects based on her observations:

> …the materialization of even a bodied individual in a physical environment has proven more problematic than anticipated. In particular, it seems extraordinarily difficult to construct robotic embodiments, even of the so-called emergent kind, that do not rely upon the associated provision of a 'world' that anticipates relevant stimuli and constrains appropriate response' (231)

For most roboticists, then, both the abstract world that is represented symbolically, and the physical world that is visible and tangible, can be mapped and measured by science and therefore also, in principle, reproduced technically. Thus, we find ethical simulations focused solely on physical interventions and futures work in which the physical environment is the primary space under consideration.

Despite drawing on a range of disciplines to situate their work for publications and grant applications, in their actual work, roboticists are generally not starting from pure or unified theories. Rather, their work is dominated by available materials, techniques, disciplinary norms and approaches, interactions with the world, as well as funding and commercial influences and opportunities. Thus, most roboticists believe that the solution to intelligence, if there is to be one, will be distinctly 'messy' a mix of methods, conflicting theories and approaches. Roboticists will generally look to theories outside their discipline for inspiration, rather than scientific validation. Thus, we get 'evolutionary algorithms' and 'artificial neural networks' that are loosely analogous to biological models, rather than their mechanical reproduction. Descriptions such as 'bio-mimicry' and 'bio-inspired' are more accurate to describe the very tenuous connection between computational processes and mechanical bodies to biological life processes that are considered to encapsulate the essence of humanity. A number of roboticists strongly eschew any connection between their work and any philosophical standpoint.

Furthermore, among my informants, there was little vocal support for narratives of imminent AGI. Many blamed media for perpetuating these hyped-up fantasies. For most roboticists, discussions of true human-like intelligence are the stuff of science fiction, hysterical news articles and, according to Sara ‘opportunistic charlatans’. Researchers, trying to convey the level of intelligence of current AI, reach for examples of the simplest life forms, such as ‘slugs’ and ‘cockroaches’, noting that these animals are still far more intelligent than any artificial system. Most roboticists I meet are keenly aware of trying to avoid the effects of over-inflated predictions and hype that prefigured both previous AI winters. During my fieldwork, all roboticists acknowledged that it was predominantly machine learning that was driving the current wave of funding for AI and robotics research. However, none of them believed it would lead to AGI and were therefore expecting another AI winter.

Despite this, most roboticists conform to a worldview that is ‘broadly materialist’. Thus, the roboticist imaginary is dominated by belief in mathematics as the basis of intelligence, and that current techniques are, in principle at least, sufficient for achieving AGI. Although they expressed vast differences in perspectives with regard to the timelines for achieving AGI, most of my informants believed it to be theoretically possible, based on the foundational theories that they consider to be fundamental and beyond question.

In general, the lack of progress towards human-like, or AGI, is explained using evolutionary metaphors, specifically with reference to the amount of time that human evolution has had compared with robot ‘evolution’. This explanation assumes that there is no qualitative difference between people and human-made machines; simply that it is a question of material (ours is soft and wet) and time (we have had more of it). As we have seen, quite a number of roboticists referred to the current proliferation of robots as a ‘Cambrian explosion’. Obstacles to achieving AGI are thus thought of in quantitative spatial and temporal terms. As Nick explains:

> I am an enthusiastic member of the robotics club interpreted strongly. I think that animals and people are machines in the ordinary sense of the word. I don't think there's any magic going on. I don't think it requires quantum mechanics to appeal to some sort of uncertainty going on so that we can't be predictable in principle. I think these machines are so amazingly complicated that we don't need to worry about them being deterministic or about us being able to predict what they do. You know it's a really a very simple philosophy. A lot of the time when I read philosophers I'm kind of surprised at the things they are able to find interesting…but also finding what seem like silly things to worry about, you know? (Interview 9/10/19)

Ultimately, for Nick, the goal of robotics should be about understanding human intelligence, this 'subsumes biology, evolution and everything else'. Similarly, Bob tells me that his new project is to make a smart robot, to build people. He is not particularly interested in elegant mathematical formulae or finding a grand unifying theory, rather he takes a decidedly 'scruffy' approach: 'we are going to use every trick in the book to make the thing work'. For Bob, then, intelligence will work from a better or recombination of current techniques. The real key to intelligence, for Bob, is not implementing algorithms or rules, rather it is extracting rules inductively from stored examples. This is evocative of the common-sense knowledge projects, viewing 'tacit' knowledge as something that is nonetheless decontextualised and formal:

> So, there are a couple fundamental views, which have roots in psychology and neuroscience. There is one view that you make rules in your head. There's another of you that you remember examples and then there are other views. So, if we talk about business school or medical school students are exposed to what they call 'cases'. And they are supposed to learn by remembering these examples. But they are also supposed to extract rules. So, if a doctor sees a patient come in with a bunch of symptoms you can view it them matching it up from a memory of a previous person who came in with those symptoms. Or they are running a bunch of rules. Probably they are doing both. Currently in what is called deep neural nets, learning rules is very popular. I think that eventually - although it is very productive right now - it will eventually run out of gas. And we will go back to also considering remembering examples. (Interview 8/10/2019)

If a step change is required, for Bob it is in the area of the materials that are currently used. He is interested in developing robot skin, to allow it to deform and adjust to edges. For Bob, as with many other roboticists who I meet, this is not one theoretical way of operationalising the human ('human intelligence') amongst many, this is the *only* way. Echoing Minsky, any philosophical standpoint that is described in a way that is specific enough, with 'clear instructions', may be operationalised. As Bob put it: '[i]f you tell me what you want done and if it's specific enough, I believe we know how to do it'. Thus, the world and how it is structured, including all living things, is ultimately knowable and can be described explicitly. His goal, he insists, is purely a practical one: 'The ultimate goal is to build a robot that does what people do independent of whatever intellectual dressing you want to put around it'. He is, or at least, acts, astonished with my suggestion that this way of operationalising intelligence might connect to a philosophical standpoint.

> To a roboticist just seems so obvious we don't elevate it to the status of philosophy! You know we build machines that try to act like people. And if you are making machines to act like people, you basically fundamentally believe that machines can be people and people are machines. Once you've signed up as a roboticist you don't consider it something worth discussing. I am 100% committed to the fact that there's nothing magic actually happening in our heads. We're just big machines. Meat machines. (Interview 8/10/2019)

Theories that are not sufficiently 'specified', or that claim that human may lie beyond science, are often dismissed as 'magic'. In the words of Boden (2004), 'the pseudo-mysticism of the kind propounded by the romantics and inspirationists' (15). And yet, phenomena that do lie outside of these domains are, as Brooks pointed out, often explained using folk understandings of the social and human worlds. Although as we have seen, Brooks' own work is dominated by 'folk' readings of evolutionary, cognitive and economic views of human nature, sprinkled with phenomenological concepts, such as sociality, embodiment and situatedness, interpreted in their purely physical sense.

In the field of robotics, it is commonly assumed that complex and multifarious concepts, like ‘common sense’, ‘movement’, or ‘interaction’, that contain ‘worlds’, may be reduced to a single, generaliseable method or theory. Towards the end of my time in the retirement community in the US, there is a visit from researchers from the US Naval research laboratory interested in applied research into intelligence systems, including HRI. They talk about their interested in perceptions of robot animacy and engagement. They are working on a project in which the hypothesis is that people perceive rudeness as agentic, so their robot is jumping a line and evaluating people’s reactions. Other projects measure engagement by how long people stay to interact with the robot. In that experiment, they found when there was more movement, there was more engagement, although still only about 40% stayed. The environment is, they say, ‘a confounding variable’. This can be resolved in a relatively straightforward manner, they say: ‘[w]e just need better theories of interaction’ (Fieldnotes 15/8/2019).

In my research, I noticed a distinction between these views as expressed by informants who are part of large, established organisations and those of smaller research teams. In general, the view of researchers in this latter group tended to be more circumspect. Gerry does not think that AGI is imminent, but neither does he believe that it is not possible in principle:

> Yeh, I don’t think it’s impossible. I guess in some ways I am a materialist. I don’t think there is anything special about the human brain that couldn’t be implemented in a different medium. (Interview 28/6/2019)

Only Lena expressed a definitive reticence in this regard. In response to my query as to when she thinks AGI might happen:

> Lena: Should I be absolutely honest with you?
> Louise: Please
> Lena: But don’t tell anyone I said this.
> Louise: OK
> Lena: Never. (Interview 27/6/2019)

Despite the specifics of their localities, most roboticists relate to the robot-object and to ‘space’ in a very particular way. This relates to the academic disciplines on which it draws, but also to the necessary bounds of their craft, whether it is the numerical precision of software programming or of mechanical engineering. Added to this is the massive challenge of simply moving an object in space, as well as the sheer amount of effort and resources needed to programme the robot to perform a particular action. Thus, despite intriguing insights gained from this physical interaction with the world into collective, distributed and embodied intelligences, ultimately, for most roboticists, their foundational orientation remains tied to a physical, mechanistic and quantitative logics that are considered self-evidently 'true'.

Re-Territorialisations

The robot is a symbol for a powerful image of technological advancement and human control over contingency. For roboticists, there is a requirement to maintain a balance between positing their project as one that is scientifically relevant, in which theories of human intelligence may be tested and reproduced, while simultaneously distancing themselves from fantastical narratives of imminent human, and super-human intelligence. Almost all of my informants blamed the media and those ignorant of the field for perpetuating the more fantastical and feverish narratives. And yet, as we have seen, it is in fact academics and technologists much closer to the robotics field who are the main drivers of these visions, most keenly under the banner of the new field of 'existential risk'.

In Historicisation, I described how Bostrom's articulation of the 'existential risk' of super-intelligent robots is galvanising and directing techno-futures. Undergirding these visions is an entire institutional network and futures industry that he has helped to create. Bostrom was awarded a professorship by Oxford University in 2003 and became founder and director of its Institute for the Future of Humanity (FHI) in 2005. The centre is focused on the 'governance of AI', including an 'examination of how technological trends, geopolitics, and governance structures will affect the development of advanced artificial intelligence' (FHI 2021). As Schuster and Woods (2021) report, FHI received £13.3 million in funding from effective altruism group the Open Philanthropy Project in 2018, which is primarily funded by co-founder of Facebook, Dustin Moskovitz (Chap. 2). The FHI also shares office space with the Centre for Effective Altruism.

Meanwhile, Bostrom's co-author and software engineer at Google, Eliezer Yudkowsky, founded The Machine Intelligence Research Institute (MIRI) in Silicon Valley, also in 2005. It originally promoted itself as an instituted devoted to the 'singularity' or accelerating the development of AI but has pivoted to the 'existential risks' of AI. MIRI's stated goal is to ensure 'the creation of smarter-than-human intelligence has a positive impact' (MIRI 2021). Associated researchers and advisors include Nick Bostrom, Ramana Kumar from DeepMind, Jaan Tallinn, founder of Skype, and computer scientist Stuart Russell.

Russell co-authored the definitive course book on AI in 1995 and is, in turn, the founder of the Center for Human-Compatible Artificial Intelligence (CHAI) at the University of California at Berkeley. The centre is focused on reorienting AI toward the development of 'provably beneficial systems' (CHAI 2021). It is also in receipt of funding from the Open Philanthropy project, receiving $5,555,550 in 2016 to run for 5 years (Open Philanthropy 2016). The Guardian newspaper ran a feature of Russell in 2021 in which he reiterates the claims that 'most experts believed that machines more intelligent than humans would be developed this century' and are 'spooked' by their own success. The same year, Russell was awarded an 'OBE', or 'Officer of the Order of the British Empire' for his services to AI.

Skype's Tallinn is the co-founder, along with cosmologist and astrophysicist Martin Rees, of the Cambridge Centre for the Study of Existential Risks (CSER), a centre dedicated to 'safeguarding humanity' in order that they might 'reap the enormous benefits of technological progress while safely navigating the pitfalls' (CSER 2021). External advisors include Bostrom, Musk, and Russell, as well a number of prominent philosophers, cognitive scientists, computer scientists and roboticists. Animal rights activist and utilitarian philosopher Peter Singer is also an advisor.

Along with MIT cosmologist Max Tegmark, Tallinn also co-founded The Future of Life (FLI), a non-profit research institute based in

Boston, Massachusetts. The institute is also 'dedicated to mitigating existential risks posed by super intelligent artificial intelligence' (FLI 2021). Bostrom is, of course, on the scientific advisory board along with Russell, Musk and Rees. Also on the board are US actors Alan Alda and Morgan Freeman. Previously, Stephen Hawking had also been a board member. The institute has received $10 million from Musk.

Bostrom's influence can be seen in the IEEE standard's group dedicated to the 'Safety and Beneficence of Artificial General Intelligence (AGI) and Artificial Superintelligence (ASI)' (IEEE Standards Association 2016). The group draws heavily on the Bostrom's work and the (short) list of contributors is drawn from related think tanks, MIRI, FOH and FLI. For evidence as to their claims of imminent AGI, all of the institutes and groups point to the same publication, 'When Will AI Exceed Human Performance?', a paper that presents the results of 352 machine learning researchers and their beliefs about the probability of when a 'high-level machine intelligence (HLMI)' could 'accomplish every task better and more cheaply than human workers' (Grace et al. 2018). The widely publicised outcome of the study is pitched as an 'aggregate forecast' and presented as giving 'a 50% chance of occurring within 45 years' (2). In reality, the paper is an aggregate of opinions of a narrow group of researchers. The paper is written by a group from both MIRI and FHI, supported by FLI.

The research agendas of these think tanks comprise issues that genuinely concern the community, such as fairness and transparency, genuine global risks, such as climate change, as well as the big-ticket items that draw in the attention and the funders, such as super-intelligence and space travel. For the most part, however, instead of a focus on pressing current issues then, the focus instead becomes on hypothetical future ones. What is most concerning, however, is the overwhelming lack of diversity of these institutes. The future is thus in the hands of a very narrow core group of people and interests. As Schuster and Woods (2021) point out, there has been a marked increase in 'existential sentiment' amongst people from 'countries, institutions, and identities are not facing immediate existential risks.' (Intro.). This apocalyptic sentiment distracts from the less dramatic, but similarly urgent collective work that is needed to be done to improve the lives of people on the planet. What does not appear to make the agenda are the issues of the uneven power structures, the real effects of surveillance technologies on communities, particularly marginalised ones, the devaluation of labour, or the political vision to change them. Instead, attention is focused on projects on techno-utopian projects that involve space travel, colonising Mars and immortality through AI. As we have seen both the real and fabricated complexity of new technologies, as well as the confusion over ethics and 'existential risk', has served to deflect scrutiny and oversight as policy makers are fearful of falling behind in the innovation stakes.

As we will trace in the following chapter, the danger of AI and robots is not that of future super-intelligence intelligence, rather it is current realities that increasing automation and surveillance on people's lives today. As we will see, the dominant image of the intelligent robot elides the real spaces and practices that are being transformed by robots and their role in increasing automation.

Conclusion

In this chapter, I have traced the fact of the robot as an enduring material object connecting otherwise unconnected multiplicities. In contrast with the view of a disembodied AI, the robot is a resolutely physical instantiation bound by its mechanical and localised machinery and the laws of physics. I looked in detail at the physical construction and elements that make up Stevie the robot, which include mechanical, electrical, digital and network elements, but also norms, desires, expectations and assumptions. A focus on the material fact of the robot also draws attention to the ways in which physical objects and environments are in fundamental ways impervious to human intentions and schemes. As Richardson (2015) notes in her observation of robotics work at MIT, 'the Real is continually asserting itself in the making of robots, and there is a sphere outside cultural constructions that has its own separate properties.' (4). As we have seen, it is the durability of robot's physical embodiment that allow it access to spaces where humans cannot go, such as outer space, subterranean, or other hostile environments.

The materiality of the object forces the roboticist into a concrete engagement with the materials, their limitations and possibilities, as well as their interactions with the physical world. The robot is a space where concepts collide with physical, material, and social 'realities'. The robot is thus a live experimental site, in which theories and assumptions may be tested, and proven or disproven. As we have seen, the team designing Stevie are not so much concerned with whether the robot is pre-programmed using a combination of symbolic representation, visual tricks or trained data, just that it 'works'. It is also the resolute physicality that is experienced by the embodied person in the 'event' of the robot encounter.

In this chapter I have shown how the manifestation of Stevie, and other contemporary social robots, is historically and geographically specific. Stevie's existence owes as much to science fiction, people's responses, and the creativity of his creators, as it does to science. His form, as that of other social robots, is inspired both by *Star Wars* and Astroboy. His stature and colour owe much to Japanese design aesthetics. Stevie's capabilities are both facilitated and constrained by the current state of available materials (metal, motors, plastic), mechanical engineering, battery power, digital and AI capabilities, availability of funding and researcher talent. He is also the result of thousands of small design decisions by the researchers working on him. We have seen how, his identity as a single entity, as 'Stevie', is created by adding a face. It is not until it is encountered, that the robot is identified as a unified object. The robot as fetish thus creates an illusion of natural unity among heterogeneous things (Pietz 1985, 9). In this chapter, I have also revealed how the actual technological capabilities of contemporary robots elide fantastical discourses. The robot is thus a concrete and enduring embodiment of specific, historical, and local materials, practices and ideas and thus stabilising and enduring in a way that concepts cannot.

In this chapter, I have also described the spaces in which the robot as object is realised and given meaning, drawing on Lefebvre's concepts of 'perceived space' and 'conceived space'. I have described ethnographic fieldwork from multiple robotics labs, conferences, and test sites across Ireland, the United Kingdom, and the United States. This has revealed, on the one hand, the geographical specificity of individual labs, subjected to specific histories, research priorities, and funding supports. Thus, I showed how there is a large discrepancy in terms of funding between labs, with the US robotics effort funded largely through military funding, although in recent years there was also more money coming from industry. In Europe, funding tended to come from more diverse sources, including general scientific research funding, but is primarily driven by perceived industry needs. We also saw how researchers from large, established institutions tended to reproduce the dominant narratives of the field more forcefully, while researchers at smaller institutions were more likely to question them.

The focus of the research also varied between centres. As Richardson (2015) observed, the robot lab at MIT was best characterised as 'a robot body parts lab', as separate teams work independently on individual body parts, such as robot hands or software, and actively resisted bringing the individual body parts together to make a whole robot (93). This oriented her focus to the various traumas, deformations and disassociations she found implicated in the robot form. However, as we have seen, this is not necessarily the focus across all robotics labs, with some working on creating a complete robot, while others focused on the interactions between people and robots.

Funding thus also influences researcher motivation. Researchers motivated by scientific achievements are more likely to focus on scientific innovations and publications, which orients them towards specific, narrow aspects of the robot technology. Researchers targeting commercial funding are more likely to focus on the perception of the whole robot and identify more as entrepreneurs. In all sites, researchers were in constant contact with media, however, their experiences varied. Again, more established research centres were more likely to have positive experiences of media interaction.

Despite the geographical and cultural specificities of the robotics labs, I also found a robotics community united by commonalities, in particular the historical narratives, identities and mythologies that sustained the field. Most roboticist viewed their work as pragmatic, rather than philosophical, and few thought AGI to be imminent. Nonetheless, roboticists were united by a broadly materialistic worldview, in which they maintained that AGI was possible in principle. As we have seen, this included a view of space as ‘conceived’, and belief in humans as fundamentally biological and knowable entities, who share a destiny of improvement and progress which may one day have to be shared with, or even ceded to, manufactured machines. Furthermore, despite distancing themselves from the fantastical narrative espoused by techno-utopians and futurologists, researchers were implicated in supporting it in various ways. In particular, in this chapter we focused on the establishment of an institutional network dedicated to extending the image of the robot as the super-human evolutionary successor. In the following chapter, Reification, I continue to trace some of these ‘re-territorialisation’, in particular, the ways in which the image of the robot object is transforming temporalities, geographies, materialities, practices and social arrangements, including cultural identities, work practices, labour relations, creativity and social dynamics.

Chapter Five: Reification

Robots occupy a unique position in the contemporary. The robot is, as we have seen, an object whose value and identity is still under negotiation. As a scientific object, it is an experimental site for the validation or negation of models of human 'intelligence' and action. As a technical object, it embodies the potential for commodification and commercial application. As a cultural identifier, it represents the culmination of Enlightenment ideals about modernity, scientific progress and man's ingenuity and domination over nature. In literature and popular culture, it provides complex mirrors of our culture(s), collective identities and existential fears.

In this chapter, we explore the robot as 'reification'. The status of the robot-fetish as a valuable object, indeed as a 'discrete thing (a *res*)', does not inhere in the object itself. Rather, it relies on specific institutional systems, or 'the productive and ideological systems of a given society' for marking that value (Pietz 1985, 12). Central to the discourse of the fetish is 'the idea of certain material objects as the loci of fixed structures of the inscription, displacement, reversal, and overestimation of value' (1985, 9). Developing a robot is a costly enterprise. It is composed of expensive hardware and software components, which become dated quickly. Additionally, there are costs associated with the need to store, move and repair the robot. However, the cost of the robot has little to do with its market value, or the value that might be ascribed to it in given situations. This chapter will focus on the constructed value of the robot from three different perspectives.

First, I will pick up on threads from the Historicisation and Territorialisation chapters, which deal with the construction of robot intelligence as analogous to human intelligence, or even super-human intelligence. As Taussig (2010; 1993b) has identified, objects, including representational images, commodities and 'mimetic machines', play a central part in developing and maintaining cultural identities, embodying the twin meanings of mimesis, that of imitation and sensuousness, and registering sameness and difference. The image of the human-like, intelligent robot makes it a potent symbol for Enlightenment Man and the ultimate technical mastery over nature. This image of the intelligent machine has real effects; in Historicisation, I described how it influences national and international governance and regulatory structures, in this chapter I will describe how it contributes to changing work practices and labour conditions.

Secondly, I will examine the robot in terms of the concept of reification drawing on the concept of 'commodity fetishism' as articulated by Marx (2008) in *Capital*, and as used by Taussig (2010) and anthropologist Alf Hornborg (2006; 2014) in their explorations of technology as fetish. These scholars develop the theory of the fetish as reification. From this perspective, an object's value is constructed in a way that obscures the social foundations of its production. Thus, objects are perceived as 'independent beings endowed with life' (Marx 2008, 43). At the same time, 'the fragmentation of the object of production necessarily entails the fragmentation of its subject' (Lukács 2013, 87). In a peculiar reversal, objects become subjects and subjects become objects. In this section, we will trace the ways in which the image of the 'smart' or 'intelligent' machine elides the human work that is necessary to animate it.

However, this view of the reification aspect of the fetish does not give the full picture. Pietz is critical of those theorists for whom reification is understood simply as 'a false consciousness based on an objective illusion' in which 'material objects turned into commodities conceal exploitative social relations' (Pietz 1985, 9). He is particularly critical of the dismissal of the person within Marxist and structuralist theories as mystified and directed by the impersonal logic of abstract relations (1985, 10). Instead, 'reification' denotes the 'truth' of 'a special type of collective object' (1985, 14), and as such, may be imagined as much more than just how it may be used to conceal exploitative relations. In the third section of the chapter, I will examine the robot as 'mimetic machine' and how it may be used to reawaken our 'mimetic faculty' (Taussig 1993b). This is, for Taussig, 'the nature that culture uses to create second nature, the faculty to copy, imitate, make models, explore difference, yield into and become

Other' (1993b, xiii). As I will show, these objects are replete with possibilities for creation and recreation:

> Standing thus at the crossroad of past and future, nature and culture, and submerging birth and death, the commodity is hardly a sign or symbol. Only in religion and magic can we find equivalent economies of meaning and practices of expenditure in which an object, be it a commodity or a fetish, spills over its referent and suffuses its component parts with ineffable radiance. (1993b, 233)

In this section, I show how the symbol of the robot and the performance of animacy may be used to play, to enchant, to learn, to educate, and to balance social dynamics. Finally, for Pietz, the key omission of Marxist and structuralist writers is the 'activity of the embodied individual' (10). This is the focus of the following chapter, Personalisation.

Robots at Work

In 1984, director James Cameron released *The Terminator*, a movie in which a cyborg assassin is sent back from the future to ensure the extinction of the human race. Skynet, the AI behind the eponymous Terminator, exemplifies a particular kind of calculative rationality, surpassing humans in intelligence, speed, strength, and ruthless cruelty. Today, the Terminator remains a potent symbol of the threat of uncontrolled artificial intelligence. Newspaper articles are often accompanied by its monstrous red-eyed, metallic and skeletal form proclaiming the latest AI advancements, for example in *Syfy Wire*, where they claim 'OpenAI robot writes dystopian essay about sparing humanity (maybe) from machine takeover' (Bullard 2020). But, as we have seen, it is not just breathless media articles reproducing this image, it is also promoted by prominent leaders in technology and science. As we have seen, the 'existential risk' of super-intelligence is presented as the next evolutionary stage for the human, precipitated by elite experts and scientists, and now in need of control by those same experts and scientists.

The robot, like its precursor the automaton, has provided a concrete and physical representation of a particular view of humanity, a mirror against which to explore and produce identity and difference. Taussig

(2010; 1993b) explores this theme in detail, showing how objects, including representational images, commodities and 'mimetic machines', play a central part in developing and maintaining cultural identities, embodying what he has called the twin meanings of 'mimesis', that of imitation and sensuousness, and registering sameness and difference. In Historicisation, we saw how automata, and subsequently robots, were used both as sites to explore the essence of humanity, as well as to represent an Other against which a certain view of humanity may be defined. We saw how, through attempts to develop AI, technological developments abounded, while the attempts to formalise human rationality continually stalled. With the failure to distil the essence of rationality, the figure of the automaton came increasingly to represent the Other, the uncanny. As Taussig (1993b) writes:

> [T]he living creatures thus mimicked … turn out on inspection … to be everything but the white male. There are negroes in top hats and tight breeches, the 'upside-down world clock' with a monkey playing a drum, 'the dance of the hottentots', a duck drinking water, quacking, eating grain, and defecating, birds in cages, birds on snuff boxes, and women—especially women. (213-214)

As Taussig shows, this process of mimesis and alterity is used by both 'Western' and 'primitive' cultures in identity formation. For the Cuna, it was through the carved image of the figures of European. In the West, and particularly in the United States, it is the idea of the good savage/bad savage, 'corresponding to the great mythologies of modern progress', specifically:

> [T]he marked attraction and repulsion of savagery as a genuinely sacred power for whiteness has continuously been concretized in terms of noble Indians at home in nature, as against degenerate blacks lost no less in history than to history. (1993b, 142, 150)

Technology, specifically 'the magic of mechanical reproduction', is central to this 'civilized identity-formation' (1993b, 207-208). In this context, the robot is a mimetic machine *par excellence*.

As we explored, the project of AI initiated at Dartmouth was defined by a paradigm change in how the idea of machine intelligence was framed. Prior to Newell and Simon's 'Logic Theorist', serious scientists downplayed the human-machine analogy. However, with the birth of AI and the founding the psychological paradigm of cognitivism, the analogy was legitimised. Funders, including the US military, were particularly interested in this form of mimesis, underscoring the connection between human psychology and robots. No other technology so overtly seeks to replicate, and thus understand, ourselves as humans in the world. Recently, echoing early Enlightenment projects, as repeated attempts at replicating human intelligence have faltered, the robot is recast, not as a human replacements, but as 'assistants', such as servants and slaves, entertainers, carers and cleaners.

Today, the robot is seen as filling a much-needed gap in the market for replacing human labour that is viewed as 'dull, dangerous, and dirty', also known as the '3Ds'. More recently, 'dumb' has been added to the list, making it the '4Ds'. And indeed, among the myriad of robots that I encountered were subterranean, lunar and underwater robots, robots that could fit down pipes and continue to function in highly dangerous situations, such as nuclear reactors and warzones. An alternative 'use' is the assumption that humanoid robots can supplement the humanity of those who humanity is not fully realised, such as for those with disabilities, children, and older people.

Across all of my research sites, these narratives provide the basis and justification for the research and are regularly repeated as the motivation for robotics research in funding proposals and publications. According to this reasoning, robots will free people up for more creative activity, and more 'high value' employment. According to roboticist Christine, robots are necessary to fill a much-needed jobs gap, particularly for social robots in care work with older people. The fact that humans were still needed 'in the loop' means that people can be retrained for all of the new jobs that are going to be created, such as jobs facilitating and coordinating robots, as she explains, 'they just need to recycle these people!' (Interview 26/6/2019). Gerry also echoed these sentiments, 'if you use robots to replace undesirable jobs and retrain people to do jobs that are more satisfying then that's a net gain for society' (Interview 28/6/2019). Among those trying to create a commercially viable robot, the narrative of the smart servant or slave is ubiquitous, as is the debate of how smart our slaves should be allowed to be (see Minsky and Riecken 1994, 25). If it is too smart, it might surpass humans, if it is too ‘dumb’ it will not be useful. As Bob relates to me in our conversation about potentially limiting the data available to the robot:

> [W]hat kind of servant do you want? One that’s blind? One that’s deaf? The robot part is just making it possible that that servant can serve you. (Interview 8/10/2019)

For many roboticists, then, robots should be slaves. For some, this has also been used as a rhetorical device used to point out the ridiculousness of the human/machine analogy (Bryson 2010). However, as a number of scholars outside of the discipline have pointed out, these positions ignore that these narratives reproduce the idea that the freedom and leisure of some humans is dependent on the degrading work of others (Atanasoski and Vora 2019). Additionally, as Alexandra Chasin (1995) articulates, these narratives also distance one class of human from another, while bringing humans whose labour is considered lower value into equivalence with robots:

> That servant troubles the distinction between we-human-subjects-inventors with a lot to do (on the one hand) and them-object-things that make it easier for us (on the other). Is the servant one of us or one of them, human or thing, subject or object? (73)

Contemporary robots are no longer represented obviously as specific kinds of marked, exotic or uncanny creatures. However, gender, class, and racial identifiers continue to mark ‘service’ robots, such as the female-voiced Alexa and Siri, as well as university projects such as the roboceptionist, and HRI studies that disproportionately focus on gender and geographical differences. Further, robots continue to symbolise a specific type of class-based, gendered and racialised view of human intelligence and idea of progress, while simultaneously also providing a stand in, or ‘surrogate’ class, representing Others against which the universal liberal subject is defined and on whose labours its freedom depends (Atanasoski and Vora 2019).

While my roboticist informants routinely gave me these stock answers, which they were used to peppering research papers and funding proposals with, some of my informants were, in private, less convinced of them. For one thing, as Sara points out, the jobs that are considered lower value are not necessarily the easiest to automate:

> Cleaning robots? Well, they are successful up to a point. Why do you think people still clean? We've been able to do cleaning robots for 30 years. I was first involved in a cleaning robot project, tangentially, in the early 90s … Look around, you don't see them. Why do you not see them? Well, cleaning is difficult. Cleaners are low status, so no one realises how difficult the job is. They are in complicated environments. They need to be able to discern, you know, should I throw this piece of paper into the bin or not? … Someone used to let their little iRobot scoot around at night, which is a good idea; it doesn't get in your way then. They have a dog, the dog did a poo, and the robot ran over the top of it and managed to grind it into every carpet on the ground floor. Because it had no idea what it was doing! Now a human cleaner wouldn't have done that … And then they would have cleaned it up afterwards. So cleaning is harder than you think. (Interview 27/6/2019)

This phenomenon is one that we have seen expressed earlier as 'Moravec's paradox'. Gerry also observes:

> A lot of the jobs that are the lowest paid and the least respected jobs are actually jobs that are hard to automate. So, you have the case where robots are replacing people in desirable jobs and people are still left with not-so-great jobs. Like, general janitorial work. That's technically a super hard problem. I don't know how you could create a robot that would clean the way that a human can clean in a human environment. (Interview 28/6/2019)

For Gerry, this reveals a structural issue, rather than necessarily a problem with robots *per se*:

> Those [low value] jobs, they are not protected by unions, and they are low paid, so I do worry about that. It's more how automation might be used as a tool in this overall narrative that devalues human workers. It's more of an outside policy thing than robots themselves. (Interview 28/6/2019)

Sam is a veteran roboticist, now working at CREATE. He did his undergraduate degree at MIT and has been involved in a number of prestigious robotics projects. Now, disillusioned by the field, Sam and his wife Beth are working with Ben on ‘pro-social’ projects. Sam is highly critical of the way the narrative of the 3Ds is perpetuated throughout the university, including being a staple of introductory courses for new PhDs: ‘it’s indoctrinating people with the idea that we’re doing people a favour with automation’ (Interview 14/10/2019).

Sam and Beth had started their own company, making robotic blueberry pickers for agriculture. It started, Sam tells me, with Beth’s interest, ‘long before it became common practice’, in ‘how to use cameras to make sense of the world’ (Interview 14/10/2019). Her access to the latest technology was limited as she was an undergraduate student. Instead, she decided to build her own. Later, it made sense to build a company around it and try to sell it. They entered, and won, prestigious robotics competitions, drawing more interest to their technologies, particularly from the industrial and agricultural space, eventually ending up with a blueberry picker. In the interview, Sam describes in detail the moment at which he became disillusioned with the robotics work that he was engaged in:

> I was working on the blueberry-sorting machine… a carpet of blueberries would sail by at several kilometres per hour on this belt and there were 50 air jets that would push these blueberries off the belt. They were rejected because they weren’t ripe enough or whatever and we worked through the winter to get things ready with the grower and the grower was the guy who owned the farm, and the engineer was part of the farm. So, we set up all the conveyor belts and things like that and had the idea of this but we did the vision processing part of it, which we knew how to do.
>
> And we were working so hard to deploy this thing. Then it came to the time to set it up in the shed. And so, we show up and the setup is, like, 10 feet away from the people, from the workers who were doing the thing on a very similar belt. So, here’s us doing our automation and there was theirs which looked about the same and they were doing it by hand. And I realised kind of slowly as it sunk in that my model of why I was so excited that we were doing this was based on a lack of understanding, a lack of curiosity about the efforts to try and figure out, are we really helping someone by doing this? So yeah, we were helping the owner. We were helping them to potentially automate and spend less money. We were helping them, certainly, in labour negotiations. I mean, I can only imagine what it was like to negotiate when you’re watching the automation next to you.

> And the biggest surprise to me, that I just didn't understand at all, was that, there I saw (even with the negative feelings that must have been around by being next to this thing that was going to take your job away), there I observed people doing what I would have imagined to be really, well, work that I would not have wanted to do. But they were having fun, they were laughing, and they were smiling. They were having conversations […] and they were absolutely mentally engaged. They were engaged with what humans do.
>
> I had to come to terms with my way of thinking about what job was ok or not, or which ones should be automated. Well, I shouldn't be the one making that decision! Certainly, for me to have input on that without doing the research, like how many agricultural workers did I even seek out before showing up that day? The answer is zero. I didn't even think about it. And that's really in hindsight if I could send a message back in time it would be like: Think more carefully about this. Just don't be so sure of where you are. (Interview 14/10/2019)

Sam and Beth's stellar engineering careers collided sharply with the realities of automation. In the face-to-face encounter, it became clear that the narratives and assumptions about 'lower value' jobs and the need for automation were inadequate. Sam continued working in the company for some time after Beth had already left. After a while, they both secured research roles elsewhere.

The image of the imminent takeover of jobs by robots, in which a single, intelligent and autonomous, humanoid robot replaces a single worker obscures the fact that in many places, 'robots' are already widely deployed. In contrast to the narrative of the 3Ds, a study by United States National Bureau of Economic Research (NBER) estimates that for every robot deployed, the equivalent of 6.2 workers lost their jobs (Acemoglu and Restrepo 2017). Despite this, it is the image of the intelligent robot, rather than the structural inequalities brought about by automation, that continue to animate discussion about ethics and risks at the highest levels.

The robot, in providing a tangible image of the imminent super-intelligent machine, remains a potent symbol for a particular view of humanity, and a cultural identifier against which to define difference. As Taussig (1993b) has shown, 'mimesis' may be a tool for political resistance or for political control. It is a 'human faculty' that is also historically contingent. In the West, that history is one of colonialism and empire. As Taussig reminds us, 'the mimetic basis remains dependent, above all, on an alterity that follows the ideological gradient decisive for world history of savagery vis à vis civilization' (1993b, 65). Thus, what he calls 'controlled mimesis', exemplified in the figure of the automaton, 'is an essential component of socialization and discipline, and in our era of world history, in which colonialism has played a dominant role, mimesis is of a piece with primitivism' (1993b, 219). As we have seen, the 'controlled mimesis' of the intelligent machine has real effects, influencing governance and regulation, as we will continue to describe in the next section, changing work practices and labour conditions.

The Human-in-the-Loop

In previous chapters, I described how the contemporary AI revolution is powered by advanced algorithms, neural nets and improved programming capabilities. However, the primary reason for its prominence is the massive amount of social data that has been made available due to people's interaction with a proliferation of personal devices and surveillance technologies. This includes text and images on social media, as well as data gathered from surveillance technologies tracking people's movements, locations, and facial images. While small research teams struggle to obtain data to train their algorithms, big business does not. The World Economic Forum has called user-generated data the 'new oil' (World Economic Forum 2011). The importance of socially-generated data has led to observations, such as that by Ekbia and Nardi (2014) that, 'Facebook's nearly one billion users have become the largest unpaid workforce in history' (10). In January 2020, the world had amassed 44 zettabytes (or 40 trillion gigabytes) of data, which is expected to reach 463 'exabytes' by 2025 (Vuleta 2021). This data consists of anything from online written content, comments, chats, messages, likes, tweets, photos, video uploads, web logs, sensors, road camera feeds, games, satellites and online transactional data, such as purchases or delivery receipts, and so on. For companies, users provide both continuous data, as well as a targeted audience for paying advertisers. Thus, the data powering the 'AI' revolution is thoroughly social.

As well as the continuous need for massive amount of social data, the dominant image of the autonomous and intelligent robot also elides the vast amount of hidden labour needed to animate technologies and present them as autonomous and human-free. Once the data has been gathered, it still needs a huge amount of human effort to clean, curate, and 'train' it. Some of this is done by unsuspecting users. 'The 'Completely Automated Public Turing Test(s) to Tell Computers and Humans Apart', or CAPTCHA, facility was originally developed to prevent bots ('intelligent' software agents) posing as human. It has now become a tool allowing companies like Google to train their learning systems to recognise non–standard fonts and formats and tricky image recognition tasks through unpaid user labour (Ekbia and Nardi 2014).

The work needed to power AI systems is so great that it has spawned an entirely new type of work. Anthropologist Mary L. Gray and Microsoft researcher Siddharth Suri (2019) spent five years investigating this new, largely hidden sector of the economy, which they have called 'ghost work'. The work includes flagging X-rated videos, screening flagged content from social media platforms, rating search engine results, removing duplicate listings, linking similar products, vetting transactions, transcribing and translating videos, and labelling images and text. These labels are then used to train algorithms, which, it is hoped, will ultimately work to remove the human from the process. In their study, they investigate four platforms enabling this type of work, including Microsoft's internal Universal Human Relevance System (UHRS), 'the socially minded start-up' LeadGenius and non-profit Amara.org, and Amazon's Mechanical Turk (MTurk).

Mechanical Turk (MTurk) is an 'artificial intelligence' or 'a crowdsourcing marketplace' in which employers can outsource work to a distributed workforce (MTurk 2021). As Ekbia and Nardi (2014) report, Amazon founder Jeff Bezos launched the platform in 2006, announcing, 'You've heard of software–as–a–service. Now this is human–as–a–service' (2014, 7). MTurk positions itself as a stopgap until technology inevitably catches up with humans. As is stated on the website 'while technology continues to improve, there are still many things that human beings can do much more effectively than computers' (MTurk 2021). As we have seen, many of my informants in the field of HRI use MTurk to crowdsource participants for surveys. Google uses MTurk workers to train its machine learning systems (Katz 2017).

Employers encounter the employees on MTurk in terms of an abstract, technological process. As described by Gray and Suri (2019), there is no interaction between employer and worker and indeed requesters can use programming code to request tasks. This results in the API, the 'application programming interface', determining the interaction between the programmer and the worker, such as assigning a random code, 'A16HE9ETNPNONN' in their example, to both requesters and workers (2019, 5). Ultimately, this renders the workers invisible and interchangeable. The platform is thus intentionally designed to atomise and anonymise workers, disguising human labour as machine labour. This objectification of workers has real effects on the labour conditions of workers. MTurk workers are classified as contractors, rather than employees, denying them legal protections that come with full-time employment. In 2012, the average wage for a task, if performed well and quickly, was about $1 (Cushing 2012). Grey and Suri give the example of one experienced MTurk worker based in the US, who, working 10 hours a day, makes about $40, far less than the minimum wage (Gray and Suri 2019, xi).

But it is not just the image of the robot fronting massive online platforms for technological giants that are used to obscure hidden human labour. In a footnote to her text, Suchman (2007) recounts how in an interview, Cynthia Breazeal reveals how the code that was rewritten to run Kismet took the equivalent of two full-time people working for 2.5 years (238). Similarly, seemingly autonomous games such as DeepMind's Atari-playing system and AlphaGo have teams of experts programming and training them over long periods of time to enable them to appear autonomous and 'intelligent' (Katz 2017, 9; Suchman 2019, 41-42).

The image of the intelligent machine is also used to obscure the work of implementing physical robotics systems. Throughout my fieldwork, the robots that I most regularly encountered were not in labs, but at airport security, museums, and in shops. During my stay in Pittsburgh, I regularly came across Starship delivery robots being tested on Forbes Avenue, which leads up to the University. These delivery robots are being tested throughout university campuses in the US and often reported in the press as a glossy, futuristic, and human-free addition to the lives of students in the 21st century, see (Nichols 2021). What these articles rarely point out, is that the delivery robots need to be remote controlled by ‘Robot Operators’. The job, according to the specification, is to ‘support our robots throughout the shift, solve unexpected situations that might arise during the course of a delivery, and work effectively as a team to ensure exceptional results in the delivery service from merchants to customers’ (Starship 2019). In the pharmacy CVS, on the same road, I am forced to use an automated cashier. Although there are shop assistants present, their role is solely to oversee the technology:

> One young shop assistant stands awkwardly opposite them at the end of one of the aisles, slightly in the way of the people browsing the shelves. There is nowhere for him to actually stand. From what I can see, the machines work fine for about 50% of the time. So, he’s standing there, awkwardly and slightly in the way, at the end of an aisle, key in hand, for when, inevitably, the technology, or the humans, fail. He looks on shyly. Neither of us is sure how much customer service or human interaction he is required to perform in this new arrangement. I ignore him, step up to the machine and carry out the transaction. Unusually, for me at least, nothing goes wrong. As I walk out the door, I hear him say ‘have a nice day!’. I feel bad for not having acknowledged him earlier. I turn when I am already halfway out the door, and give him an awkward wave. (Fieldnotes 17/10/2019)

As reported by Mateescu and Elish (2019), agriculture, delivery services and retail are currently at the frontier of the AI ‘revolution’. In their report, they dub the hidden human effort needed to maintain the apparently autonomous systems ‘human infrastructure’ (2019, 12). The authors show how, while self-checkout has not removed the need for cashier jobs, it has shifted the roles and responsibilities of workers. Frontline workers are required to absorb the risks of adopting these new technologies, which inevitably require new skills and routines to facilitate the system, which are often invisible and undervalued (2019, 14). Workers are also more likely to see their hours reduced and are subject to greater scrutiny, transforming the ways in which the workers are classified and perceived by management. As we can see from the field note above, the automation of work also transforms how the workers are perceived by customers. In my own case, it was not clear whether the tacit norms of social engagement still applied, or whether I was now required to ignore the human in the automated system.

In their research into the agricultural sector, Mateescu and Elish (2019) show how adopting big data and smart technologies reconfigured work practices, including creating new work routines and changing the physical infrastructure to render farmland amenable to the data collection. As they report, these shifts change the way that the farmers relate to the land, a physical field must now be conceptualised as a complex dataset to be managed through other digital information and digital tools (2019, 5). Instead of working out in the fields, farmers are finding themselves in front of a computer for most of the day.

The idea that supposed labour-saving devices generate new forms of, often devalued, labour is not new, see for example (Cowan 1983). However, the narrative has shifted into overdrive. The image of the human-like, intelligent machine is used to elide the limitations of current technologies, as well as the reorganization of employment structures and practices, often to the detriment of workers. In contrast with the ancient dream of freeing people from labour, the new roles that are created include only a very limited number of elite and highly visible workers to build the technologies, and a much larger, hidden workforce animating them, adding their skills and creativity to fix errors, enhance their performance, and bring them to life. Rather than a novel technological revolution, then, these developments are part of a longer history of automation in which the promise of freedom from labour instead gives way to new forms of uncompensated, invisible, or undervalued labour.

Despite the seeming pervasiveness of these systems, however, the 'reification' aspect of the fetish is not necessarily exploitative. As I will describe in the following section, and continue in the next chapter, the problem of the fetish arose and 'remains specific to the problematic of the social value of material objects as revealed in situations formed by the encounter of radically heterogeneous social systems', only one of which is capitalism (Pietz 1985, 7). In the next section, we will see that the fetish evades any single totalising logic or definition, and may be used for resistance, as well as control.

Animating Stevie

For Pietz, reification denotes the 'truth' of 'a special type of collective object' (Pietz 1985, 14). Fetish objects:

> …exist in the world as material objects that 'naturally' embody socially significant values that touch one or more individuals in an intensely personal way: a flag, monument, or landmark; a talisman, medicine-bundle, or sacramental object; an earring, tattoo, or cockade; a city, village, or nation; a shoe, lock of hair, or phallus; a Giacometti sculpture or Duchamp's Large Glass. Each has that quality of synecdochic fragmentedness or 'detotalized totality' characteristic of the recurrent, material collective object discussed by Sartre. (1985, 14-15)

Pietz quotes Sartre at length:

> It is necessary to take up the study of collectives again from the beginning and to demonstrate that these objects, far from being characterized by a direct unity of a *consensus*, represent perspectives of flight… For us the reality of the collective object rests on recurrence [repetition of the same property within the members of a series]. It demonstrates that the totalization is never achieved and that the totality exists at best only in the form of *detotalized totality* (Sartre 1968, 78, 80).

Thus, the reification aspect of the fetish, like mimesis, is personal, as well as historical, and may be used to resist and, per Graeber (2005), to create. Further, in the case of the robot, it is also a 'mimetic machine', which may be used to reawaken our 'mimetic faculty' (Taussig 1993b). In this section, I will show how the symbol of the robot and the performance of animacy may reveal the robot as craft and performance, and be used to play, to enchant, to learn, to educate and to balance power dynamics in favour of the less powerful.

Like other robots, the work that is needed to transform Stevie from a carefully controlled machine in the lab to a robot that can be brought out and ‘cope’ in public is vast, and largely unseen:

> The work of preparing Stevie to be able to move out of the lab and into the ‘real world’ is exhausting. It is mayhem in the lab. The new robot is much further along than the last time I had seen it. All the levels in the ‘coffee table’ are filling up. There are batteries at the bottom, cameras and sensors up on top. There is what looks like a little fan in the middle. Not-yet-Stevie’s face has had a makeover. His eyes are brighter, with a 3D-rendered glint. Everyone looks exhausted. The front, unlike the original Stevie, has a platform, which someone is trying to pin an iPad to. Dónal has lost his hex key but finds it now in the midst of the tangle of wires and boxes. He says things that used to work are now going wrong. I tell him he seems very calm. He says he’s too tired to freak out. As they are showing me things they keep falling on the ground. No one seems too concerned and they just leave them there. Niall is filing some wooden part attached to a clamp. Someone is testing the robot’s sound, while an undergrad is being blamed for the way the arms are moving. They are behind schedule. Nothing is booked yet, but they are expecting to leave around mid-Feb. (Fieldnotes 24/1/2019)

By mid-February the team land in Washington, as scheduled. Stevie needs to be reassembled on site. It takes two people to take him apart and reassemble him and because there is such a small team here on the ground, it is not that easy to do. Once he is up and running, things immediately start to break and fall off. Someone has to be continually on-hand to fix, tinker with, and repair him. Michelle, the technical liaison at the retirement community, has officially been appointed as Stevie’s ‘handler’, but it will be a while before he will be fully operable by anyone other than the team of roboticists who created him.

While the team hope to develop all of Stevie’s autonomous actions, for now, most of his actions must be controlled remotely. Thus, while a belief in the imminent emergence of a human-like technological species is not subscribed to by the majority of roboticists, there is a requirement to perform and promote machine autonomy in diverse ways. The robot relies on countless human interventions to appear autonomous and animate, including constant tinkering and repairs, mediating and translating, as well as everyday robotics techniques such as ‘Wizard of Oz’, ‘human in the loop’ and the creation and training of social data for machine learning systems.

While active in the retirement community, Stevie is mostly operated by the head of machine learning, Ciara. Using the arrow button on a laptop, she directs him through corridors and into lifts. Although he has been programmed to conduct dialogue in a number of defined settings, such as when running Bingo or the table quiz, invariably residents of the care home will stop him on his travels to and from the lab and start conversations with him. When this happens, Ciara will often take over the speech system and either select pre-typed responses or furiously attempt to type in appropriate responses. Ciara is also responsible for training the machine learning system, which they hope, ultimately, will allow him to deal with these complex, spontaneous conversations autonomously.

While not operating Stevie, Ciara spends hours training the small dataset that the team have managed to gather. I spend an afternoon with her, seeing how it is done. She activates Stevie much like Alexa, with a ‘Hey Stevie!’ command. She then starts to talk to Stevie, asking questions like ‘What is the weather like today?’ As she tells me, the speech input is sent to Google to translate it to text. The machine learning algorithm then parses the text that is returned and makes a prediction on it, essentially ‘guessing’ what class of statement it is likely to be. This might be whether it is likely to be a sad or happy utterance. Based on this result of this calculation, Stevie will carry out a follow-up action, such as responding in kind, or ‘mirroring’ with an appropriate facial expression. ‘Hey Stevie’, Ciara says, ‘What is the weather like today?’ Stevie replies, ‘the weather in Washington is 78 degrees with a 40% chance of rain’. This is correct. We are delighted. Hey Stevie! What are the activities today? What do you think of [name of community]? What’s on the menu today? Stevie replies, sometimes with the right answer, more often with something that appears completely random. He is cycling through his eight emotions at a furious rate. At the moment, while the machine learning system is still being trained, Stevie is erring on the side of sympathy and collapsing his shoulders and head in a posture of abject sadness at the slightest verbal provocation: 'I am sorry to hear that, can I do anything to help?' It's very nice, but rarely a logical response to what is being asked. One of their greatest challenges will be to get data, Ciara says. What she needs, for example, is enough data so that the algorithm can detect small differences in people's expressions.

In all of the labs that I visited, robotics researchers encounter this problem. Aside from the limited access to large datasets that a small research team will have, they are also limited by the local and physical nature of the robot. Personal interactions with the physical robot are necessarily much more limited than those that could, for example, be gathered by an App distributed by a large technical company across many personal devices. Also, people will have a higher expectation of 'natural' interaction from a humanoid robot than from an impersonal App. In Heriot-Watt, Gerry encountered the same problem:

> Modern AI is a very data-hungry endeavour, especially if they are using deep learning techniques - which we are not touching for these reasons. You must have a very, very large dataset in order to learn a good model and because we are working with the autistic population, we can't collect a lot of data. We have access only to a very small group of users. (Interview 28/6/2019)

Apart from the sense that his physical presence and demeanour evokes, it is his vocal abilities that sets Stevie apart from other technologies. People's opinion of the robot and its capabilities relies heavily on the sophistication of the dialogue system. It is also the one that people study carefully, trying to gauge his intelligence, his human-likeness. People's reactions are most pronounced to responses that appear to show an ability to react to, and interpret, the unfolding events of the present moment. When Stevie's responses appear to show that he knows the weather today, that he has remembered people's names, people are overjoyed. Throughout trials with potential users, I see a number of different people, mostly researchers, but also care workers, take over Stevie's controls and his dialogue system. Sometimes it is done quite formally and deliberately, to test a particular aspect of Stevie's functionality, or to conduct a 'Wizard of Oz' style experiment. At other times it is done spontaneously, to smooth over an interaction, to fill gaps, or simply just for fun.

On one such occasion, Stevie is invited to 'happy hour' at the retirement community, which was an hour of drinks and sociality in the afternoon:

> The group are delighted to see him. There are about 12 women and one man gathered around a communal area with glasses of wine and beer and they are clearly looking for some entertainment. 'Hi Stevie!', they call out as soon as he comes into view. Many of the residents are already familiar with Stevie from the structured activities and seeing him navigating through the hallways. In a reversal of the usual roles, one of the residents asks Stevie to ask them questions. This appears to confound 'Stevie' somewhat; I can see Ciara's bewildered face trying, under considerable pressure, to think of something for Stevie to say. Stevie responds with some stock phrases, 'I am delighted to be here', he says. Kindly, the group tries a change of direction: they wonder whether Stevie might like to hear a joke. Stevie duly replies in the affirmative, and a resident is called on to tell it. Stevie chuckles politely in response to the somewhat bawdy story. The role reversal continues: 'What have you observed about us?' they ask him. There is no precedent for this kind of conversation and Ciara continues to struggle. Sensing the likely cause of these new dynamics, 'Stevie' diverts attention with 'I wish I could drink!' accompanied by a sad face. It is enough for the group. 'I wish you could dance!' someone offers. One of the residents familiar with Stevie suggests Karaoke. Back on familiar territory, the researcher switches back to the pre-programmed version of 'These boots are made for walking', written by a member of staff. This version is called 'These wheels are made for rolling'. He gets a big applause for his efforts. 'Was it good, did you like it?', he asks. They did, they reassure him. Someone suggests that they could all sing a song. They are a very easy crowd. They want Stevie to succeed and are giving him every encouragement. Those residents who are familiar with and know what he can and cannot do from the structured sessions, are, in a sense, demonstrating him for those that are not. Stevie asks whether they would they like him to tell them a story. They would and so he proceeds to tell them a pre-programmed one, the story of *Tir na n-Óg*, from Irish mythology. The story doesn't take long, less than five minutes, but it is romantic and tragic, with some magic thrown in. It goes down very well. They talk about Bingo. 'Will you call my numbers?' one resident asks, suggestively. Stevie replies in what now seems a James-Bond accent, 'I'll call your numbers'. 'Ooh!' say the residents. (Fieldnotes 12/8/2019)

Unlike this 'free' dialogue, the routine nature and scripted humour of Bingo make it a perfect setting to trial Stevie. There are other challenges to be overcome, however. Bingo was already an established activity run by residents, or sometimes by staff in the higher dependency part of the community. Both Bingo and Karaoke are established social activities and facilitating them is a public role, which carries a degree of social standing in the community. The technological system originally used in Bingo was donated to the community by the husband of the resident who now facilitates the session. In an effort to integrate Stevie into this sensitive situation, the researchers made a number of tweaks to how he would normally behave. Instead of acting as facilitator, in both Bingo and Karaoke, Stevie is programmed to play the role of assistant. The researchers have also scripted deference into Stevie's dialogue:

> Once Stevie is activated, he launches straight into the game: 'oh seventy. oh seven zero'. Stevie maintains a slight smile throughout, delivering numbers in monotone. There is silence in the room as everyone checks their numbers. 'I hope we get some winners here' he offers in a slightly defeatist tone. 'oh seventy. oh seven zero', that's how many friends I've made here at [xxx], he says quietly. There is no response. The mood in the room starts to improve with some winning numbers. Stevie defers to the facilitator to check the numbers. 'Mrs. S will now check your numbers'. Mrs. S is pleased, 'Thank you Stevie', she smiles at him. His facial expressions change abruptly from neutral to happy (arms in the air!) and back to neutral. Stevie rushes into a monotonous and forlorn 'I wish I could play bingo with you all'.
>
> A little later, 'I hope you are enjoying yourselves' receives a murmur of agreement. 'Is anyone close?' he asks. 'No!' from across the room. 'Oh boy this is a fun one' he declares, joylessly, 'Monday night bingo oh yay'. Despite not being able to convey natural timing or tone, Stevie manages to carry off a pretty convincing game of Bingo, with some successful interaction with his co-facilitator. His continual head movements throughout the game are gentle and appealing. They feel lifelike and sympathetic. (Fieldnotes 12/8/2019)

The next day, Stevie and the team join a Bingo game in the higher dependency unit. This is an assisted-living part of community, in which people have much greater care needs. There is a similar positive, if more muted, reaction to Stevie's arrival. Here, the carers play a much greater role in the residents' care and in their activities. In this instance, a carer, rather than a resident, runs the Bingo, as well as helping people out with the game.

Daniel asks her to announce the game by saying ‘Stevie! Start the Game’, which will activate him. I was told earlier that this was difficult for them to do both jobs at once so Stevie would actually be helping them out here but at the moment the carer does not look at all too thrilled about what is going on. She looks decidedly unhappy, in fact.

> The residents ask Stevie to speak slower and louder, which is an easy thing to get him to do so the team make the appropriate changes on the fly. Stevie deliberately tells an awful joke to break the ice. ‘Is that good?’ he asks. ‘No!!’ comes the response. Stevie does his sad face. People are warming to him. Stevie starts to call out numbers. ‘You are the best people to play Bingo with’ he announces. At first this provokes no response and then a very delayed and hesitant ‘thank you’ from the back. ‘This game is a hard one!’ Stevie declares. The carer picks up on it: ‘Stevie has character!’ she tells the room. People laugh in response. ‘B7 is my least favourite number’, Stevie announces, cryptically. ‘It’s his favourite?’ a resident asks. ‘No, his least’, the carer translates. ‘Oh...’
>
> The game stops so that someone’s numbers can be checked. The carer restarts it, ‘Stevie! Let’s go. Start!’ ‘oh seventy. oh seven zero. That’s how many friends I’ve made here’. Nothing. ‘N4 I am enjoying this game!’ ‘Good for you’ replies one of the residents, sarcastically. She and the carer laugh together conspiratorially. ‘We must be close’, Stevie says, ‘We are!’ comes the enthusiastic response. Stevie, still programmed for the earlier session, announces, ‘I am so glad Mrs. S invited me!’ Mrs S is, of course, not in charge here. Michelle takes over the system, ‘Ignore that’ she has Stevie say, about himself. Hilarity ensues. ‘G18 I see we have some regulars in the crowd!’ This gets people’s attention. ‘Does he mean us?’ a resident asks. ‘Ask him how does he know we have regulars here?’ another resident asks a carer. In response to Stevie’s silence, Michelle continues to take over Stevie’s automated responses: ‘I just know’, she says through Stevie. ‘Oh! I apologise’, the resident replies in an exaggerated tone, aware of the performance in which she is engaged. ‘That’s ok Elizabeth’, responds ‘Stevie’. The resident is enchanted, ‘He knows my name!’. Others are intrigued, ‘How does he do that?’ (Fieldnotes 13/8/2019)

Unlike many formal Wizard of Oz experiments, Ciara and the team are at pains to ensure that those who encounter Stevie are aware that he is being tele-operated. They sometimes talk about their work with Stevie as ‘puppeteering’. However, despite their best intentions, it is nonetheless clear that many of the advanced behaviours that the robot exhibits during these performances cannot now, or possibly ever, be programmed in. Indeed, researchers are often unaware of the immense spontaneity, skill and creativity that they engage in when performing the robot, skills that are not a focus for their discipline and rarely, if ever, explicitly analysed. Further, as we shall explore in more depth in the next chapter, even when those engaging with the robot are made aware of a person controlling it, it does not necessarily stop people from attributing the behaviours to the robot.

> Michelle, who has taken over Stevie for the first time is settling into her new role. ‘I got a full house!’ someone calls out, ‘Wu! wu!’, whoops tele-operated Stevie with uncharacteristic American levity, ‘Hooray!’. Everyone laughs. As people leave, tele-operated Stevie says goodbye to each of them by name ‘bye now’, ‘have a good day, sir’. The older people who have been playing Bingo are enthralled. The initially reluctant carer tells the team how much fun she had. She seems pleased that the resident that she was caring for has enjoyed it, she tells us ‘he woke up something’ in the resident. (Fieldnotes 13/8/2019)

Later in the week, Stevie is invited back for a social session at the request of the previously wary carer.

> About 11 residents are gathered around, more men this time. Stevie is standing in the middle in neutral mode and several people are trying to get his attention. Some are waving at him, others calling ‘Hi Stevie!’ One of the residents ask his carer ‘Does he know I am a colonel?’
>
> ‘Stevie’, still being set up, isn’t reacting to any of this. There are very long pauses. There is a lot of pressure on, and felt by, Ciara, who is trying to control him. The residents are unsure of how to engage him. They haven’t taken part in the structured activities and are not as capable of closing the gap themselves. When he does talk, they find it more difficult to understand him and miss a lot of what he says. Once the research team are set up, they can finally take control of the conversation. ‘Do you want to hear a song?’ Ciara has Stevie ask. ‘Yes’ they reply. Stevie is immobile while a song is played through him. The song stops abruptly. ‘Did you like it?’ he asks. There is scattered applause. The colonel offers Stevie a compliment, ‘You are very versatile!’. Stevie goes for another song. This time it is Stevie’s own version of these boots are made for walking and he is mouthing the words. People are reacting and smiling, nodding along. When he finishes, the Colonel says: ‘Your performance is quite electric!’ Everyone laughs.
>
> ‘Would you like to hear a joke?’, Stevie asks. Indeed they would. ‘Why are they called French fries when they are cooked in Greece?’ Stevie offers. There is a muted response. People shake their heads. ‘Are they good jokes?’, ‘Stevie’ wants to know. ‘Do you have

> another one?' someone offers, diplomatically. Stevie tells another one that no one understands. The carers and staff who are present try smooth over a lot of the pauses and lack of understanding. When nobody reacts to another of Stevie's joke, Michelle repeats the joke for the group in a very slow, animated and loud voice. This time there is some laughter. Ciara, who has been overseeing Stevie's automated performance, decides to take over the dialogue system. 'Hey Michelle!', she has Stevie say, 'stop stealing my jokes!'. Stevie's apparent awareness of the social dynamics: his interpretation of the interaction, the slight peevishness at his own joke's failure, the awareness of the carer being an appropriate target for teasing combine to make the situation genuinely hilarious. Even the simulation of intelligence and agency as Stevie makes this joke feel thrilling. 'Would you like to hear a story?' Stevie asks his newly won-over audience. 'Yes!' comes the hearty response. Stevie tells the story of *Tir na n-Óg*. People are genuinely engaged, listening intently. 'Oh dear!' a few people say, shaking their heads at the sad bits. (Fieldnotes 15/8/2019)

Stevie is at his most compelling when those controlling him use his status as a significant object to redress the balance of power in the retirement community. This is most evident when carers use Stevie to tease each other. At one of the sessions, a carer called Eddie is somewhat stern to one of the residents in the public setting of the quiz. Later on, he laughs in response to one of Stevie's comments. Michelle quickly takes over the dialogue system and rounds on Eddie, 'Eddie stop laughing at me!'. The carer is subtly brought back down to size, and the residents are delighted.

Most of the older participants who took part in the studies with Stevie were independent, active, and engaged. Stevie's presence was not just of interest to the residents as an object of fun, many residents were interested in the technological specifics, and in taking an active role in the research and design of the robot. They regularly give the team newspaper and magazine articles about robotics and AI that they have come across or actively sought out. Others are inspired go off and do their own research into AI, asking questions about, and directing their own learning into, facial recognition, algorithms, and how to train neural networks. A number of residents expressed interest in being part of a committee to help with the co-design of Stevie. For the residents, many of whom are now retired from highly technical and challenging jobs, this gives them an opportunity to work on something that is both challenging and personally rewarding. A number of residents we interview also helped with the care and companionship of other resident with higher-care needs. And so, in the interviews, they did not just consider themselves ‘users’ of Stevie, but also considered how they would design or deploy him.

> A lot of residents don’t have anyone to talk to, don’t have any visitors. There are people downstairs with no visitors at all. They would be glad to answer these questions. Many people who are highly dependent can get very difficult, but they still need to be treated with compassion. [Stevie] is not subject to emotions that would cause him to get cranky, to get irritable. People downstairs who are sick, I was sick myself a few weeks back, they get cranky. But they still need someone to talk to. Stevie would listen, and he wouldn’t get irritated by them. He’s very dispassionate. (Kathy Interview 14/8/2019)

For John, one of the residents who runs Karaoke for the community, Stevie was his ‘wingman’. Stevie could entertain people while he switched between or lined up new songs. Established habits could be given a new twist, Stevie would fist-bump the audience at the end of the session instead of shaking hands.

Some residents have not had good experiences with technology, one of the carers tells us. They’ve had their social security numbers stolen or have received fake, automated calls from the IRS. Unlike much technology, however, it feels like Stevie is on their side. When they are with the robot, their grandchildren want to stay and play. In some ways, quite a few people remark, he might be better than a human: he does not get frustrated, or judge you, and you can turn it off and control it. It is obvious that Stevie is giving people hope. Hope that they might be able to have company, that there might be small ways in which their lives can be improved, that they might feel more secure, less lonely and that they might find a nice way to pass the time. As we explored in the previous chapter, Stevie also connects the individual to a broader sense of purpose, technical advancement, and the future, in fundamentally existential ways. Encounters with Stevie, while visceral and embodied, are also shaped by wider historical and cultural realities.

As well as inspiring wonder and awe, the image of the human-like, intelligent machine unites people with a sense of common purpose and progress, much like space travel. For the older people in the study, it seemed like they are witnessing the realisation of a promise that had been made to them a long time ago. In Dublin, Fred is clearly tapping into this dream when he marvels at Stevie, 'We've come a long way! Look how far we've advanced!' Similarly, in Washington, Kathy exclaims, 'It's so fascinating to see the way they are used … it's really coming!'.

Conclusion

The social status of the robot as an autonomous, human-like, and intelligent machine is one that has been developed over many centuries, and across different geographies. The image represents the culmination of a philosophical and scientific worldview in which a specific type of human intelligence leads to a mastery over nature, and control over contingency. This makes the robot a potent symbol for technological advancement and determination of all kinds, not just for robotics.

In this chapter, I examined the constructed value of the robot from three different perspectives. First, I described how the construction of robot intelligence as analogous to human intelligence plays a central part in developing and maintaining cultural identities, which, in the case of Euro-American imaginaries, is one implicated in colonialism and empire. Second, I revealed ways in which this image of the robot is being used to obscure the limitations of current technologies, the real work that is necessary to make the technologies appear intelligent and autonomous, as well as how it is used to reorganise employment structures and practices, often to the detriment of workers. Third, I show how the image of robot animacy is also a testament to the creative performance and craft of its creators, and is used to play, to enchant, to learn, and to rebalance power dynamics. As a new 'mimetic technique', the robot allows for new ways to relate to our bodies and to nature, and new possibilities for creation and recreation. In the following chapter, I continue with this theme, investigating the personal, embodied, and visceral way in which the robot is encountered.

Chapter Six: Personalisation

The lens of 'personalisation' recognises the embodied status of the person, and the intensely personal way that the robot-fetish is experienced in a concrete encounter. The truth of the fetish 'is experienced as a substantial movement from "inside" the self… into the self-limited morphology of a material object situated in space "outside"' (Pietz 1985, 11-12). In the Historicisation chapter, we saw how early Christian automata provoked awe in those who encountered them. Then, in the late 20th century, robotics and AI researchers were both inspired and disconcerted by the reactions that their programmes and robots evoked in those that encountered them. Similar reactions from users caused Weizenbaum to abandon his project ELIZA completely. By the 1980s, a new research field of social robotics was established, with the explicit aim of maximising the strong responses of humans to robots.

As we have seen, for Pietz, Marxist and structuralist interpretations of the fetish are limited because they ignore the relation of the fetish to the embodied person, of which, he argues, labour theory of value is only one example (Pietz 1985, 10). Unlike the idol, Pietz writes, the fetish was often worn and used to achieve tangible effects in the user, such as healing (1985, 10). The 'personalisation' aspect of the fetish recognises that of 'the subjection of the human body (as the material locus of action and desire) to the influence of certain significant material objects that, although cut off from the body, function as its controlling organs at certain moments' (1985, 10). It is this aspect that led to the fetish as appearing as 'a perversely anthropomorphized or sexualized thing' in all registers, whether psychoanalytical, ethnographic, modernist or Marxist (1985, 10). The fetish object is 'established in an intense relation to, and with power over the desires, actions, health, and self-identity of individuals whose personhood is conceived as inseparable to their bodies' (1985, 10). This encounter, 'stripped of all symbolic value', becomes 'a crisis moment of infinite value, expressing the sheer incommensurable togetherness of the living existence of the personal self and the living otherness of the material world' (1985, 10).

The 'personalisation' aspect of the fetish delimits a space that is comparable to Lefebvre's (1991) conception of 'lived space' or 'representational space', which, for Lefebvre, is the concrete space of everyday users and is lived through the body (200). It is 'the shifting intersection between that which touches, penetrates, threatens or benefits my body on the one hand, and all other bodies on the other. Thus, we are concerned, once again, with gaps and tensions, contacts and separations' (1991, 200-201). The references by Pietz to Merleau-Ponty's philosophy and concept of the 'flesh' throughout the text reveal that Pietz is indebted to Merleau-Ponty's embodiment philosophy in articulating the fetish concept as a whole.

As we have seen in the previous chapters, the challenge of articulating, and theorising, the sensuous connection between the embodied person and the object has been encountered by numerous scholars, including Lefebvre, Taussig, Suchman, and Pietz. In this chapter, I take an explicitly Merleau-Pontian approach to an analysis of encounters between people and robots, which is used to reveal what I argue is an under-theorised and little recognised, yet vital, aspect for understanding human-object relations. In order to further elaborate and develop this theme, I also use Taussig's (1993b) reading of Walter Benjamin's concepts of 'mimesis' and 'alterity'. Taussig's concept is a particularly useful lens through which to view encounters with robots, allowing us to consider the identifications, as well as the disavowals, between the person and the object in identity formation.

As we have seen, Taussig builds on Benjamin's concept to develop a 'two layered' notion of mimesis. The first is the ability to mime, or to copy, which is, he maintains, also the capacity to Other (Taussig 1993b, 19). Clearly, as we have explored in Territorialisation, the creation of humanoid robots is a mimetic act in this sense. The second, is 'a palpable, sensuous, connection between the very body of the perceiver and the perceived', a 'flashing moment of mimetic connection' (Taussig 1993b, 21, 23), emphasising the embodied, contingent, and historical nature of the human-machine encounter. The concept of mimesis is, for Taussig, simultaneously embodied and historical. In the Reification chapter, we traced how social and cultural identities are created through mimetic processes. This chapter focuses on the moment of connection between the body of the perceiver and the perceived, and on how the interplay of mimesis and alterity, the self and the Other are constructed.

As we will elucidate in this chapter, the robot is a special kind of object. It is both a mimetic technique, as well as an object of perception. The lens of 'personalisation' thus emphasises the importance of considering the concrete encounter between the embodied person and the object. This chapter begins by describing 'naïve' encounters with the robot-as-fetish by potential 'users', as well as by roboticists. This is followed by an account of how the encounter is typically theorised: as anthropomorphic projections of mental models and as a demonstration of Theory of Mind (ToM). I contrast this account with one that draws on Merleau-Ponty's embodiment concept. Next, I explore the experience of those operating the robot, engaged in a creative performance with the 'new mimetic' technique of the robot, a new 'sensorium' in which they find themselves extended, their identity blurring with the robot. This chapter also reveals ways in which the robot moves from kin, to prosthesis, to sympathetic Other, motivating at times a primal recognition and identity, and at others a disavowal and 'othering'. As we will trace, this novel technology, suspended somewhere between science, art and commodity, reveals novel entanglements between the human and the object.

Encounters with Robots

> They are sitting on chairs. One person is in a wheelchair. They have been eagerly awaiting Stevie's arrival. He rolls in through the door, haltingly, his mechanism whirring away. He is small and shiny and white. He comes to a standstill in front of the small group of participants, all older people connected with the Irish charitable organisation whose focus is on supporting people to age independently at home. Although his wheels are stationary, the rest of him is not. His head continues to move slightly from side to side, his arms make gestures as he speaks. His big brown eyes are glinting, two eyebrows move up and down. Right now, they are

raised, innocently. His mouth is curled upwards in a cartoon smile. This is decidedly not the Terminator. People lean forward in their chairs, their expressions melt into affectionate smiles. They are amused, sympathetic, excited, enchanted. He is more like a child than they were expecting, but also not quite fully working. Not everyone is impressed, though. Mary, at the back, has her arms folded. 'He's not even good looking', she complains. She'd like some colour. Maybe green, she suggests, for Ireland?

'Hello! I am Stevie!'. Stevie surprises everyone with his booming English accent. It is an authoritative and confident voice, like an announcer on the BBC, yet also somehow reassuringly automated. Stevie has regained some authority with his commanding introduction, but this group of older Dubliners are on a day out, and they are here to have fun. They want to test him, see how he reacts. 'How are you, my dear?' Paul drawls in an affected tone, 'how is your highness?' Niall calls him 'Steven'. 'Steven, can you open a fridge door and retrieve food?' But the formality doesn't last. Someone shouts, 'Hey Stevie! Where's that fiver you owe me?' Everyone starts to laugh and relax. A few people clap their hands in delight.

The group have immediately spotted the limitations of Stevie's wheels, and they start to tease him, testing to see how he will answer them. Stevie, can you run? Can you dance? Play hopscotch? Play football? Niall says, 'Shamrock rovers [football club] are looking for a scorer!' Stevie, however, is unflappable. His demeanour and his responses appear remarkably dignified, as well as self- and situationally-aware. He cannot play sport, he explains logically, because he has wheels, not legs. He also knows what time of year it is. 'Maybe I'll get a pair of legs for Christmas?' he proposes. Everyone is amazed and thrilled by what appears to be his highly advanced social abilities. 'Ah, he's great craic', Paul sighs, slapping his knees to emphasise his admiration.

At one point, due to a technical hitch, Stevie starts to talk much more slowly. Participants respond in kind by also slowing down. 'Are you not too sure?' someone asks, with concern. Stevie laps it up, 'I don't remember, I'm only 2 years old!' This leads to a delighted 'I'm getting a real reply back!' Far from being put off by these small errors, the participants vociferously reject the suggestion from the robotics lead that it might be frustrating for them. His imperfection appears to strike them as part of his charm, 'part of the personality'.

Despite having been told at the beginning of the session, they seem to have forgotten, or have chosen to ignore, the engineer behind the desk furiously typing in the appropriate responses on his laptop. Once he types something in, the off-the-shelf text-to-speech

> translator generates an audio file, which is broadcast through the speakers on the side of Stevie's head. Despite this 'Wizard of Oz' technique, it is not a straightforward case of deception of unsuspecting, vulnerable users. For the participants, while thoroughly enjoying his lifelike capabilities, there is no illusion that he is alive, or human. After gazing at Stevie lovingly, Paul turns his head to the team and announces matter-of-factly, 'Yes, I could use him for a few hours each day, like a TV'. Niall agrees, 'It's great fun to get to know them. To find out what they can do'. Mary, who has somewhat been won over by the interaction, spots an obvious advantage of a robot compared to a human: 'You can turn him off when you get tired of him'. (Fieldnotes 25/10/2018).

Stevie is explicitly designed to elicit an affective response. His humanoid appearance, expressive eyes, micro-movements, and gestures are all geared to maximising the attribution of agency. However, as we have seen, non-humanoid robots have elicited similar responses in humans. Robot swarms, whether in aerial swarm formation (as in swarming drones) or on the ground, evoke impressions lifelike of behaviour and some form of intelligence, although clearly not human. An early study by Heider and Simmel (1944) shows how even simple geometric shapes moving across a screen were interpreted by those who observed them as having emotions, motivations, and purposes. To an observer, then, specific kinds of movement and appearance imply a higher degree of intelligence, purpose and the presence of emotions, unrelated to their actual functional capabilities.

At the start of my research into robotics, I reached out to several academics interested in robotics in the engineering and computer science departments in the university in which I am based. To demonstrate the 'state-of-the-art', they played a video of Boston Dynamic robots. It featured demonstrations of a number of their humanoid and 'zoomorphic' robots, including the human-sized warrior robot Atlas and the 'canine-inspired' Spot. As I have described, Boston Dynamics focus on robot movement, in particular balance and dynamics, rather than on the 'cognitive' aspects. Yet, the results are simultaneously impressively and disconcertingly lifelike. There is little reassurance in the fact that the robots being created are intended for military use.

Figure 17: A number of Boston Dynamics robots including Atlas and Spot (Boston Dynamics 2020)

Since 2002, ground and drone robots have been deployed in the military for a multitude of purposes, including mine detection, bomb disposal, surveillance, transport, and as weapons. The US-led invasions of Iraq and Afghanistan were used as an unprecedented testing ground for robotics technologies. While the ethical and legal implications of military funding and deployment of robotics technology is a major concern, in this chapter my focus is on the soldiers' experiences of the robots.

In 2007, a Washington Post article describes a test of a bomb disposal robot in dedicated test ground in Arizona. The five-foot long, insect-like robot would go out into the field to find mines. Occasionally one would blow up and blow off one of the robot's legs. By the end of the test, the robot had only one remaining leg, but was still dragging itself along with it through the sand. From the engineer's perspective the trial was a success. The Army colonel in charge of the mission, on the other hand, was appalled. He could not bear to watch the burned, scarred, and crippled robot attempting to complete its mission. He called off the test, describing it as 'inhumane' (Garreau 2007). This was not an isolated response. In the same article, a sergeant recounts a story about a robot they called 'Sgt. Talon', whose tenacious behaviour in the face of danger evoked the human quality of stoicism and became a lucky charm for the troops. Another explosives technician carries the remains of his robot 'Scoobie' in a box. A number of studies show soldiers risking their lives to save the robots they work with (Singer 2010; Carpenter 2013; Kolb 2012).

As well as showing impressive feats of agility in their robots, Boston Dynamics also released videos of engineers testing their robots' robustness and stability by kicking them. In a video of DARPA-funded 'BigDog', a vaguely zoomorphic rectangular form on four legs is repeatedly kicked by an engineer (olinerd 2008). Each time, it stumbles wildly but manages to regain its balance without falling over. It is a disconcerting watch. It is both comical and full of pathos. Even though the strange angles and movement of its gangly black legs are vaguely repellent, this is quickly overshadowed by the apparent cruelty of the act.

Figure 18: Still from video testing Boston Dynamic's Big Dog (olinerd 2008)

A later video shows the same treatment of the more refined-looking Spot. While the technological capabilities revealed are impressive, there is a sense of unease watching the videos. It is very difficult not to feel sympathy for the robot whose movements invoke apparently purposeful, even archetypical 'heroic' qualities, such as fortitude and resilience, in the face of intentional, repeated, and calculated cruelty.

A number of researchers in the social robotics field, including philosopher of technology Mark Coeckelbergh and MIT researcher Kate Darling, have been drawn to these examples in their study of the ethics of human-robot interactions. Their focus is specifically on the apparent empathy that people feel towards moving machines. Coeckelbergh (2018) has shown that humans react viscerally to a machine that displays animate behaviour, regardless of its human-like design. He uses the installations of Belgian artist Kris Verdonck, in which machines appear to be suffering or in distress, to illustrate his point. In one installation, ‘Dancer #1’ (Verdonck 2016), a metal bar writhes, twists and turns on a rope in a way that appears, to the observer, to be distress and suffering. ‘In the concrete confrontation with the machine’ Coeckelbergh writes, ‘something happens which creates… empathy’ (Coeckelbergh 2018, 154).

Figure 19: Still from ‘Dancer #1’ by Kris Verdonck (Verdonck 2016)

Darling (2016; 2017) has conducted a number of experiments to test the phenomenon. In an experiment based on the infamous Milgram experiments, a group are given a bunch of small robot dinosaurs called ‘Pleos’ to interact with (Darling 2016). After a while, in a dramatic and, I imagine, traumatic, change of direction, they are told to ‘tie up, strike, and ‘kill’ their Pleos’ (12). Most refused. The (one) Pleo that was killed simulated pain and ‘whimpered while it was being broken’ (12). Other HRI experiments show that motion influences ‘perceived affect’ (Saerbeck and Bartneck 2010) as well as empathic responses to robot suffering (Rosenthal-von der Pütten et al. 2013; Suzuki et al. 2015).

It is not just apparent robot suffering that evokes both intelligence and animacy, and elicits empathic responses. The first time I touched a robot, it was just a disconnected arm with a human-like hand and five jointed digits, which had been lying on a shelf in the Robotics Lab at Trinity. Robot hands, when built in this intricate way, bear a striking resemblance to human hands. They are clearly mechanical, as you can see the joints and wires that hold them together, but they look dexterous and versatile in the way that human hands do. They evoke skill, labour, but also gentleness and touch. I picked the hand up and held it in mine. Despite having felt out of place and awkward as I entered this new, unknown field site, I felt instantly comforted and relaxed.

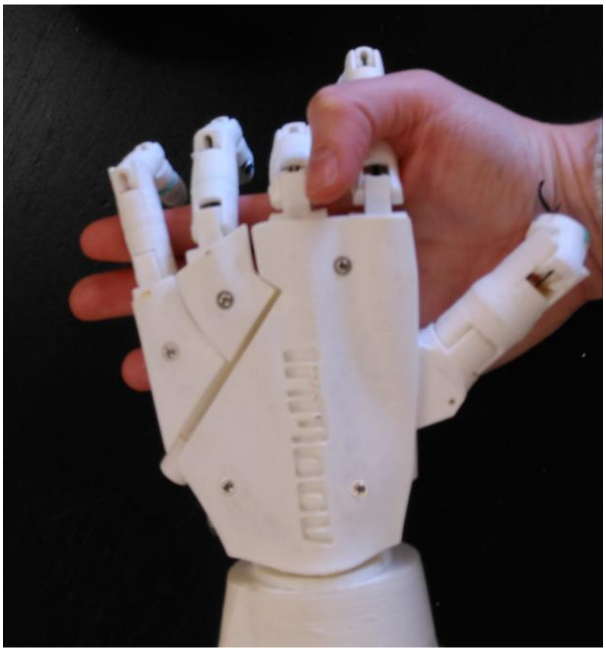


Figure 20: Holding a robot hand

As well as empathy, robots can motivate feelings of eeriness. In 1970, robotics professor at the Tokyo Institute of Technology, Masahiro Mori (2012) first described on the unnerving experience of an initial feeling of affinity with a realistic robot, which then reveals itself to be artificial. For example, ‘we could be startled during a handshake by its limp boneless grip together with its texture and coldness’ (2012, 3). Mori calls this experience ‘the uncanny valley’, explained by way of a diagram, shown below.

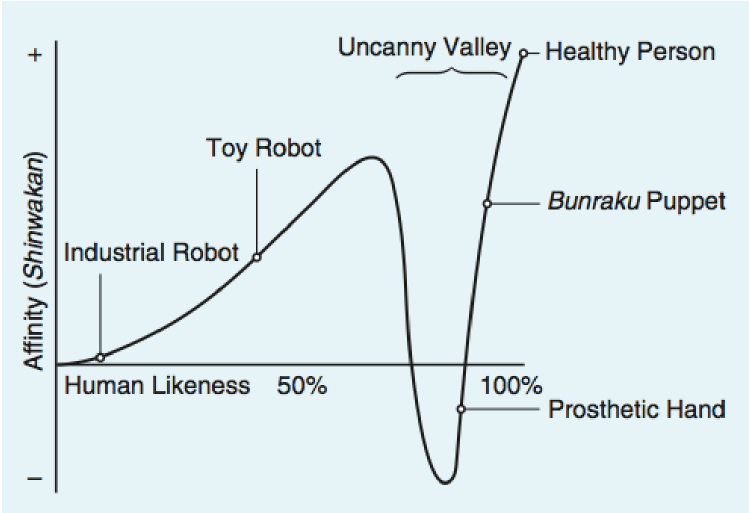


Figure 21: The Uncanny Valley (Mori 2012)

Movement, according to Mori, amplifies the peaks and troughs, including, for example, the speed of changes in facial expressions. The death of a person leads them to fall to the bottom of the valley. At the time of its publication, in an obscure Japanese journal called 'Energy' it barely received any attention, see (Mori 1970). Today, particularly in HRI, the uncanny valley is used as a reason to avoid creating ultra-realistic robots. This does not deter some roboticists, such as Hiroshi Ishiguro and David Hanson of Hanson Robotics, who, as we have seen, are focused on creating robots that are as life-like as possible.

Other responses are also evoked. Darling (2017) and her team conducted another experiment with the tiny toy robots Hexbug Nanos. This time, participants did not hesitate to strike them with their mallets. The research team had to supply each robot with a name and a backstory before people would empathise with them. The researchers theorised that in this instance movement alone was not sufficient because of the Hexbug Nano's insect-like appearance, and people are used to killing insects.

In 2015, a hitchhiking robot, 'HitchBOT', was designed by a Canadian research team to test human reactions to robots. The robot could not move or walk, instead it had to ask people to carry it, or drive it, to get around. The team also gave the robot a 'personality' with limited conversational ability and a social media presence. The robot successfully hitchhiked across Canada, Germany and the Netherlands, before being beheaded in Philadelphia at the start of its journey across the US (Leopold 2015).

In Japan, a study in which the robot Robovie was deployed in a shopping mall shows how it was subjected to ‘abuse’ by groups of children, including ‘kicking, punching, beating, folding arms, and moving (bending) the joints of robot’s arm and head’ (Nomura et al. 2015, 63). Most frequently, children blocked the robot’s path. When asked about their behaviour, the children reported that they had done it for ‘curiosity’ and ‘enjoyment’ (2015, 64). Half of the children believed that the robot could perceive the abuse. Many of the roboticists I interviewed recounted similar stories. They tended to be both perplexed and disturbed by these events. Christine recounts an experience of deploying a robot at a museum:

> People have a lot of expectations [of robots]. Oh! He’s a bar tender. He’s in the mall, he should help me, he should guide me. And then they talk to the robot and he’s like ‘de de de’ and people get frustrated, they are like ‘answer me!’ and then when he doesn’t, people start harassing the robot. I’ve seen that many times when you put robots in public spaces. You see children kicking them. It’s insane! (Interview 19/6/2019).

Clearly, it is upsetting and frustrating to experience these reactions to a lovingly crafted, not to mention expensive, object. And yet, I have also experienced wanting to engage in robot ‘abuse’. In Pittsburgh:

> The Starship robot that I follow is driving down a wide footpath and passes a group of people working on the garden. They seem pleased to see it, ‘awww, it’s lost!’ they say, sympathetically, stepping out of its way. I, however, have seen it usurp the path and am irked as to why it appears to get right of way on what should clearly be a civic space. I feel like getting in its way, maybe even giving it a little shove. I want to see what it will do. I also want to draw out whoever is controlling it. But I don’t. I am, among other things, worried that there is a room full of people somewhere, possibly up at the university where I am doing my fieldwork, monitoring the ‘experiment’ on a large screen in a lab, and my act of sabotage would be broadcast for all to see, clocked up as another act of robot ‘violence’. I also feel like a passer-by might come to its aid. It is, after all, small, cute, and apparently ‘lost’. (Fieldnotes 15/10/2019)

It is no surprise, then, that Stevie also provokes and inspires an intense, visceral response in people. His humanoid appearance, with a face, eyes, and expressive eyebrows, as well as his ability to move independently, immediately identify him as an animate entity. As soon as people encounter him, their whole bodily form and orientation changes. Stevie's haltering movements and compact size evoke an instantaneous, usually compassionate, response from people. Long before people start to wonder what he is or what he can do, they have felt him to be a kindred being.

Figure 22: Stevie with older users (Mooney 2019)

At times during my fieldwork, I glance over and feel Stevie looking at me; his large brown eyes and gentle smile seemingly acknowledging some shared existence or sentience. When residents of the retirement community encounter Stevie in the hallways, they are inevitably drawn towards him, their expressions lighting up, expressing amusement, delight, and compassion. When working in a room with Stevie, there is a regular sound of delighted from the hallway as people pass by.

> When people come into the room they are surprised and pleased to see Stevie and start trying to engage him by waving or calling to him. At the start, Stevie is still off. He is being charged and the researcher is attaching webcams and tablets to his front. He appears to be standing there, obediently, childlike, like a parent tying a toddler's shoelaces. People who come in say 'Hi Stevie' but he doesn't reply. Stevie is not in communication mode and there is no one currently controlling him. This strikes people as odd; he's mute

> they say, his mind is gone. He's in game mode, one of the researchers explains, apologetically.
>
> Once Stevie is ready to go, he is switched on and can start running the Quiz. Stevie starts up 'Can you hear me?'. They respond, 'Speak up Stevie, we are old!'. The researchers dial up the volume. Each of the participants has a controller with coloured buttons on it and tries to choose the correct, colour coded response. The first category is 'military history'. Stevie's expression is 'neutral'. His upper torso moves to face the screen, as if reading, his back to his audience. His head and body move slightly, he waves his arms. He asks his first question. Everyone clicks their controller and the board lights up. Stevie congratulates everyone who gets it right. When announcing a correct answer and 'winner', his arms are raised in celebration and he smiles at the room. When no one gets it right, he hangs he head, lowers his arms, his mouth bends down at the corners. It is a pitiful expression that never fails to elicit a sympathetic response from this crowd. Any errors from the robot, such as a mispronunciation, are met with laughter. There is a remarkable tolerance for error that would not be the case with other technologies. People feel they have a personal relationship with Stevie, they feel protective of him. Much of Stevie's dialogue during these structured activities is scripted, which can make humour difficult. Despite this, at times it can work very well. As participants fail to answer many of the questions in a particular category, Stevie quips: 'I bet you regret choosing this category'. It's pitch perfect, there is a burst of laughter from the room. (Fieldnotes 12/8/2019)

When Stevie arrives into a new situation and people become aware of his presence, some react with delight, while others retreat, but inevitably some people will start to try and engage him in conversation and see what he is capable of. Stevie is designed to respond to people using a combination of dialogue and non-verbal gestures, such as facial expressions, as well as head and arm movements. When people meet Stevie, often one of the first things they do is to try and catch Stevie off guard, testing his self-awareness (does he know he doesn't have legs?), his situational awareness (does he know what's on the board for dinner tonight?), and his ability to respond to off-beat questions. When the residents' grandchildren at the retirement community come across Stevie, their initial reactions of wonder quickly turned to mischievousness as they began to goad and test the robot. They quickly discover that Stevie reacts with apparent sadness to insults, with delight to complements:

> 'You smell like rotten egg sandwiches!', they taunt him, delightedly
> Stevie responds by hanging his head forlornly
> This time: 'You smell like flowers!'
> Stevie raises his arms in delight, and so on… (Fieldnotes 13/8/2019)

People insist on a robot having an identity and will name it if the researchers do not. While the team building Stevie take great care not to anthropomorphise the robot by calling it 'him', some of the residents objected strongly: 'Don't call him 'it', he's got a name!'. Stevie's childlike appearance, coupled with a dispassionate British accent, serve to announce him as an unintimidating, even humorous, character. This, along with an automated, monotone voice makes for a deadpan delivery of sentences that is both endearing and amusing. In Douglas Adam's *The Hitchhiker's Guide to the Galaxy*, a character called Marvin the Paranoid Android is a robot prototype afflicted with depression and boredom, due to the underuse of his massive intelligence (Adams 1979). Stevie's demeanour sets a similar tone. Stevie's emerging 'personality' becomes a function of his technical limitations: the arm that sticks out, a name that the text to speech translator cannot quite make out, the slow and halting way he moves, his apparent resolve in the face of adversity, or graciousness in the face of contempt. At one point, while navigating down the corridor, Stevie turns, faces the wall, and just shuts down. It appeared to be anything other than a mechanical fault. Was it childish obstinacy? Existential *ennui*? Had he just had enough of us? The more Stevie got things wrong, the worse his jokes, the richer his personality became.

When the residents at the retirement community are doing more formal 'interviews' with Stevie, I also do one:

> I am asked to sign the relevant forms and am told that I do not need to answer anything that I do not want to. This worries me somewhat. What will they be asking me? I go into the room, where it is just the two of us. I can see the research team outside of the window but I try to forget they are there. This is relatively easy as Stevie himself is very compelling. 'Stevie' asks if I am ready to begin and introduces himself. He looks straight at me and tells me 'You do not have to answer any questions'. This is disproportionately reassuring. Much more so than the research team's assurances when they were not operating through Stevie. I do not have to answer any questions! Even Stevie says so. He then invites me to introduce myself. I start

> to relax. He asks me about where I grew up, how many brothers and sisters I have. I start to talk. It is enjoyable to recount these things, to relive them. It also strikes me as unusual to have 'someone' take such an interest. Stevie is looking straight at me, open and honestly. It feels intimate, and empathic. The more I tell, the more I remember. I remember things that I am proud of. Stevie is very encouraging. 'Oh really?' Ciara has 'Stevie' ask (I am determined to remember). His expression and demeanour are similarly engaged, he smiles, facial and bodily movements, blinking, eyebrows moving, throwing arms up in delight at some small detail. He seems genuinely interested and energised by the conversation. It is a very pleasant experience. I feel seen. He asks what my favourite childhood memory was. I start trying to think about this. It is a very difficult questions to answer. I remember the researchers outside and start to feel embarrassed. Then I recall, I don't have to answer! So, I don't. 'Stevie' doesn't mind, he moves on without judgement… (Fieldnotes 14/8/2019)

After each of the 'interviews' between Stevie and the residents, the research team ask me to sit in on a follow-up interview with the participants, to capture their reflections. They relate similar experiences to my own. The experience is very pleasant; it is nice to have 'someone' to talk to. They say, 'you forget you are talking to a robot'. He is, as one of the residents, Kathy, says, 'a very personable guy!'. People clearly enjoyed talking about the past, and their past achievements. Betty is concerned that she did not have time to 'tell him about my husband. He achieved so much … I didn't tell him about my pets. He should have asked me to elaborate more'. There was even a sense of urgency to the entreaties, a need to tell the stories, to remember them and to have them remembered, to reanimate them and bring them back to life. People talked to Stevie in a way that they had not to others in the community. Michelle, who had controlled Stevie for one of the interviews, recounted how one of the residents had told Stevie that she was applying for a job, something she had not disclosed to anyone else in the community. Originally from Germany, Edith tells us she finds Stevie easier to talk to than the other residents. She has had trouble fitting in because, she tells us, she feels overwhelmed by the accomplishments of the other residents. Stevie, she says, is less judgemental.

The intensity of people's visceral responses to robots has challenged assumptions in the robotics community. Traditionally, these reactions were considered somewhat irrational and something of an embarrassment. People often respond to aspects of the robot that roboticists consider trivial, such as its outward appearance or name, attributing sentience and capabilities far beyond what the robot is capable of due to small, or even accidental, features of the robot. It is easy to dismiss this phenomenon, as is often done, as an unsophisticated response from 'naïve' users, projecting human attributes onto an undeserving object. Roboticists tend to situate themselves in direct opposition to those who 'anthropomorphise' and stand in awe of mobile machines. And yet, it is the roboticists who have the most personal and intimate relationships with their robots. Even those who have expert knowledge of the technological intricacies of the robot can, at times, encounter the robot innocently, experiencing it as an enchanted and animate being, with the appearance of animacy and intelligence unrelated to its physical and technological capabilities.

Although the research team developing Stevie insisted on calling Stevie 'it', in moments where team members were caught off guard, they would often refer to Stevie as 'him'. Roboticists showing real affection for their robots is very common. In the Edinburgh lab, one of the PhD students, Dieter, shows me the Anki Cosmo.

> Anki Cosmo is a small robot, about ten centimetres tall. It looks like a little toy digger with cubes beside it, with a small screen with a face on it. The game you play with Cosmo is to program it to recognise, through its visual system, some of the blocks in front of it. The blocks have LEDs in them so they light up and the light display changes when they are picked up or moved. They want to get a bunch of these for their first-year students who don't normally get to work with actual robots, just with software and simulations. Dieter is controlling it from his phone. As he explains, the phone is the brain, whereas the embodiment (the little robot-truck) is just that. Except, he says, the personality stays in the body. He tries to get the robot to pick up a block, which it does tentatively, falteringly, and requiring some help. Dieter reaches out and uses his index finger to support the robot. Its visual system was blocked by only engaging one side of the block, Dieter explains. He talks about how this robot is a companion, like a pet, and that it shouldn't just be targeted at children, autistic and older people, but at 'normal' people. At single people, like him.
>
> What makes this like a robot, and not like a remote-controlled dinosaur, he says, is the fact that when you are not controlling it, it

> goes off and does its own thing. So, there is a degree of autonomy there. The little robot's 'personality', then, is the way in which is put its little digger up or down showing emotions that looked like frustration and excitement. Hesitation, maybe? Or that perhaps the battery was running out. Its eyes also change to show different kind of 'cutesy' behaviour. It was, indeed, very cute. The robot, he said, was modelled on Wall-E. He is excited about using these with students. Working with these, he says, the students can programme how to navigate through a maze in the abstract, and then try programming the robot to do the same thing. This will also allow them to see how things like the mechanics influence the movement/behaviour and not just in the abstract. He also said they would learn to use APIs, so for example using Twitter to give the robot commands. (Fieldnotes 15/5/2018)

At the same lab, Lena talks of how much she enjoyed working with iCub. She admits to 'anthropomorphising' her robots, telling me how touching her robot gives her comfort and support during stressful demonstrations:

> I still catch myself doing it. I still feel comfort sometimes when I am in the robotics lab and I put my hand on the shoulder of my robot while I am giving a demonstration in this uncomfortable situation where you have to talk to people, and I know that I feel more like you are not alone. (Interview 27/6/2019)

Even veteran roboticists like Nick, who expressed disdain for magic, silly philosophers, and 'crazy HRI types', admits to his own remarkable response to robots in an encounter. Indeed, his own assumptions about the potential for general intelligence were forged, not by his own technical work, but by watching robots play soccer in the RoboCup competition:

> I remember the first time I said to myself 'Gee, maybe we have arrived' and that maybe it's obvious that machines *do* perceive that they *do* have awareness of the environment, that they *do* have intention and express that intention through purposeful action. I think the first time that I felt confident that you can make those claims was when I was watching the RoboCup soccer team. (Interview 9/10/2019)

Theorising the Encounter

In the HRI literature, the phenomenon of the human-robot encounter is most commonly explained in terms of two concepts: 'Theory of Mind (ToM)' and 'anthropomorphism'. ToM is a concept referring to an assumed human capacity to ‘impute mental states to himself and others’ (Premack and Woodruff 1978, 515). It is assumed, as the minds of others are not directly observable, that a system is needed to make predictions about the behaviour of others. Key concepts associated with ToM include ‘attention’, in which we selectively direct our attention to specific objects. This includes ‘joint attention’, in which two people direct their attention to something together (Baron-Cohen 1991) and ‘intentionality’ as defined by Daniel Dennet. Dennet’s ‘intentional stance’ is the ‘the strategy of interpreting the behaviour of an entity (person, animal, artifact, whatever) by treating it *as if* it were a rational agent who governed its “choice” of “action” by a “consideration” of its “beliefs” and “desires”’ (Dennett 1998, 27).

ToM is used widely in HRI as an explanation for sociality among humanoid robots, see (Scassellati 2002; Breazeal 2002). The concept of ‘joint attention’ is a key strategy for implementing perceived sociality in robots. Furthermore, the previously pejorative concept of ‘anthropomorphism’ has been rehabilitated in the field by connecting it to ToM. In developing a case for exploiting anthropomorphism in social robots, Duffy (2003) explains, anthropomorphism is a ‘projective intelligence’, used ‘to rationalise a system’s actions’ (180). This is a similar argument as used by anthropologist Stewart Guthrie (1993), where he proposes a new theory for religion based on what he considers the human propensity for anthropomorphism. For Guthrie, this is a strategy for survival in a hostile world, in which individual actors perform a risk analysis of the situation in order to select the most beneficial outcome. Anthropomorphism is simply a misapplication of this rational strategy (1993, 6). Boden (2006) also connects anthropomorphism and ToM, drawing on the work of anthropologist Pascal Boyer (1996). Despite this, Boden admits that the computational mechanisms underlying ToM are still unclear (2006, 2).

As traced in the previous chapters, the origin of social robotics research, at least in the US and Europe, is in the research into behaviour-based robotics at MIT. This was a deliberate departure from previous representational and symbolic models that did not take embodiment and situatedness seriously. Despite early and potentially transformative breakthroughs, embodiment and situatedness were not theorised in any meaningful sense, instead interpreted in its purely physical sense, coming to stand for a robot having a humanoid or zoomorphic body, as well as the use of sensors and distributing processing. Despite the insights that these perspectives offer, the foundational conceptual framework in robotics remains ‘naturalistic’, oscillating between behaviourism and cognitivism, within an overarchingly biological and evolutionary explanatory principle.

Phenomenology, however, offers an alternative account of the human-robot encounter. Philosophers Gallagher and Zahavi (2010) have argued that ToM approaches ‘deny that it is possible to directly *experience* other minded creatures.’ (183). For ToM theorists, they argue, ‘the only mind that I have direct access to is my own. My access to the mind of another is always mediated by his bodily behaviour’ (2010, 181). This, they argue following Scheler, ‘underestimates the difficulties involved in self-experience and over-estimates the difficulties involved in the experience of others’ and is ‘an overly impoverished conception of what is given, of what is experientially available (2010, 182, 183-184). Following Merleau-Ponty, they argue that it is our shared embodiment that allows us to know one another.

Also drawing on phenomenology, Jackson (2002) has argued that ‘anthropomorphism’ is a misunderstanding of the human natural tendency towards intersubjectivity. As we have seen, for Merleau-Ponty, the pre-objective unity between the body and the world is primordially intersubjective. This unity is, for us, suffused with animacy and vitality, in which everything is potentially a subject, and potentially an object. Thus, for Jackson, ‘human beings everywhere tend to conceive of subjectivity not only as encompassing others but as extending into the extra-human world with the result that objects, words, and ideas tend to become imbued with consciousness and will’ (2002, 334-335). In this account, ‘we do not project human consciousness and will onto machines … rather ‘intersubjectivity so shapes our experience … that it constitutes our “natural attitude” towards the world into which we find ourselves thrown – a world that includes persons, machines, words, ideas, and other creatures’ (2002, 341). In contrast to ToM, then, we do not assess or determine the status of another before we decide to respond to them, rather, after Levinas (1999), we are ‘obligated to respond to them even before we know anything at all about them’ (Gunkel 2018, 98).

Human/Machine Entanglements

Those who come face-to-face with the robot experience the presence of another animate entity. However, other types of embodied encounters are experienced by people who work on the robot, and those who operate it, for whom Stevie is as a mediator, or prosthesis, between them and the world. The work of robotics often involves researchers inserting themselves in the research in intimate ways. As we have seen, for roboticists, putting themselves into their experiments and into the wider robotic system in the form of Wizard of Oz experiments is so ubiquitous as to be unworthy of comment. In some labs, researchers even provide their own ‘bio-material’ to power the robots. Richardson (2010) identifies roboticists’ ‘mimetic practice’ in her fieldwork, in which they first act out human behaviours before embedding them in their robots (75). During her observational work at MIT, she finds that roboticists often modelled their robots on themselves (Richardson 2015). Similarly, as we have seen, in my interview with Bob, he spoke at length about how his desire to become a roboticist, as well as the designs for these robots, arose from his frustrations with his own physical clumsiness as a child growing up in a family of athletes.

Approaches to building intelligent agents are often characterised by roboticists themselves as directed by ‘purely logical or scientific thought’ (Brooks 1991a, 18). However, in the same paper, Brooks argues that the specific model of computation for robots characterised by situatedness, embodiment, intelligence and emergence was ‘arrived at by continuously refining attempts to program a robot to reactively avoid collisions in a people-populated environment’ (Brooks 1991a, 17). The technological work of roboticists is not a straightforward application of science. By contrast, the act of building and animating the robot is a highly situated, embodied, and social process that sees the roboticists in an embodied engagement with the world and the object. Building the robot is a social and creative activity, involving negotiation, compromise, trial and error, fixing, repairing, and persuasion, as well as love and frustration. Dónal, who is working on Stevie, tells me that nothing, other than the love of technology, inspired him to get into robotics:

> As a 15-year-old I wasn't thinking anything else at all, I was thinking of the technology in and of itself. I'd an interest in computers, programming and electronics, and robotics seemed to a great combination of all of those things. You get the full stack … I was never able to just love one part of it … it's much more tangible when you write a bit of software on hardware that you've built yourself. And then it does stuff in the real world! That's just, that's just really cool. There's no amount of fancy graphics on the screen that can replace that. I really like that. My first draw to it was purely, *purely* the technology. (Interview 24/1/2019)

The physical realisation of an idea, as well as immediate, concrete feedback can result in frustration as well as exhilaration. In the field study at the retirement community, as Ciara repeatedly tries, and fails, to get Stevie to respond correctly:

> 'Sometimes I just wanna smash him', she says. 'Hey Stevie!', Stevie wakes up. 'Shut up!'. Stevie sweetly responds, 'being silent now'. We exchange sympathetic looks, feeling sorry for him. (Fieldnotes 15/8/2019)

Building robots is a mimetic act on a number of levels. As Taussig (1993b) has elaborated, it is both copy and contact, both the creation of a 'copy' in material form and also experienced in an intense, visceral connection. By building the robot as a self-contained, autonomous entity, by, as Nick put it, 'putting eyes on it', it breaks the continuity with the self, deliberately establishing an Other. As we have seen, this Other may be experienced as an identification with the self, or as an Other against which the self may be challenged and contrasted, and sometimes as both. As Taussig articulates, 'mimesis registers both sameness and difference, of being like, and of being Other' (1993b, 129). The 'sociality' then has to be rebuilt/reinstated through overt social actions, such as face, expressions, voice, mimicking, and mutual gaze.

In an earlier chapter we explored how, for Miller (1986), it is in the consumption, rather than the production, of objects that a person becomes aware of their own being and meaning is conferred on the object. This, he maintains, happens through a process of identification and disavowal. However, my observations of human-robot entanglements suggest that it is in the embodied encounter that these boundaries are drawn. As Wehrle (2021) has shown, identity is developed at the 'bodily level' in which 'there is a degree of agency already operating' (379). Further:

> …the body is not thematized as concept with historically contingent meaning invested with power relations, but as the necessary condition for a coherent experience and thus for every formation of identity. (2021, 382)

However, robots are also a form of 'mimetic technique', affording those who use them new perspectives and ways of interacting with the world. This is the experience of those who operate Stevie, controlling his navigation, movements, and speech remotely, and interacting with the world through his body as real-world avatar. Both Ciara and Michelle recounted the strange experience of moving between the robot as an extension of themselves, to facing it in an interpersonal encounter. Michelle talked about trying to channel her 'positive' approach through Stevie, but finding it transformed through Stevie, into a different 'personality'. As we have seen, when a person takes over the dialogue system, it is immediately apparent how complex and nuanced the dialogue becomes, as well as how creative, situated and responsive people need to be in order to respond appropriately and with humour. Using Stevie as a 'mimetic machine', people seem to expertly bring together situational awareness, timing, historical and spatial knowledge, dominant discourses and social dynamics, undercutting or supporting implicit social dynamics to indicate deference, to rebalance social dynamics, or for comedic effect. These 'puppeteers' who animate Stevie are engaged in a creative performance, switching identities and discovering new, creative aspects of themselves. In this way the robot is a prosthesis or extension of the self, allowing people to extend themselves, creating a 'new sensorium' (Taussig 1993b, 24), their identity blurring with the robot.

The way that the robot is experienced in the inter-personal encounter is not one in which the perceiver calculates, projects or rationalise the encounter, rather it is an embodied experience revealing the human open to an intersubjective world. At times, people experience the robot as an extension of themselves, yet at other times, through its autonomous movement, lifelike morphology, or by having a face, it is experienced as a separate entity. This recognition is pre-objectively and pre-sententially felt. The seeming contradiction of the simultaneously situated and embodied person, both a unity with the world and a unity of the self, forms the backdrop for the process of mimesis and alterity in identify formation. However, as we saw in the previous chapter, identity formation is both embodied and historical. Thus, the visceral interplay of mimesis and alterity, of sameness and difference, is implicated in the wider ideological and cultural discourse on which it depends. In all cases, however, it is the living body that animates the relation.

Conclusion

In this chapter, I investigated the relation of the fetish to the embodied status of the person. First, I drew on literature and ethnographic description to articulate the deeply personal and visceral ways in which people experience the robot. These include encounters in which people experienced a primal empathy and kinship with the robot, as well as how, at other times, robots provoked revulsion, unease, and even violent responses. I revealed how, in a concrete encounter, the robot may be experienced both as an embodied connection, as well as a break in the continuity with the self, as Other. Next, I contrasted a Merleau-Pontian interpretation of the encounter with more typical explanations, such as ‘Theory of Mind’ and ‘anthropomorphism’.

Thus, I identified the robot-fetish as ‘a sort of primary and carnal rhetoric of identification and disavowal’ (Pietz 1985, 14). This connection between the embodied perceiver and the perceived is not intrinsically good or bad, but it is, as Pietz’s characterisation suggests, an integral part of the relationship between the person and the fetish-object. This connection forms the backdrop from which boundaries are configured and reconfigured. It is in the encounter with the robot-object that the robot becomes ‘robot’ and may be considered ‘animate’.

In the final section of this chapter, I explored other embodied entanglements between the person and the robot, including how roboticists inserted themselves into their research, the craft of robotics work, and the experience of the 'puppeteers', for whom the robot is a prosthesis or 'mimetic object'. This chapter thus reveals ways in which the robot moves from kin, to prosthesis, to sympathetic Other, motivating at times a primal recognition and identity, and at others a disavowal and othering, revealing novel entanglements between the human and the object.

These diverse encounters have revealed unexpected insights that challenge both the Enlightenment model of Man, as well as the symmetrical view of anthropology and STS theory. This includes, for example, the tendency for moving machines to evoke a visceral and pre-discursive empathy in those that encounter them, regardless of their inner computational abilities. However, as we examined in the previous chapter, we are embodied, as well as cultural, beings. In Reification, we traced how social and cultural identities are created through mimesis. In this chapter, we focused on the moment of connection between the body of the perceiver and the perceived, and how the interplay of mimesis and alterity, the self and the Other are constructed. As Taussig (1993b) argues, '[t]he flashing moment of mimetic connection' is simultaneously embodied, mindful, individual and social' (23). The visceral experience of the fetish, or robot, is enmeshed in a wider ideological consciousness in which we both identify and disavow the robot object within the context of specific geographical, historical, and ideological configurations.

For Lefebvre (1991), this is the 'lived' or 'representational space', 'the "mixed" space—still natural yet already *produced*—of the first year of life, and, later, of poetry and art.' (203). I argue, however, that this space of everyday experience is one in which people are continually in creative and embodied engagements with the world. Further, I argue that this is the space in which the capacity for resistance and revolution resides.

Chapter Seven: Configuring the Robot as Fetish

In this dissertation, I have investigated how the human is produced in robotics research using the conceptual framework of the fetish as articulated by Pietz and underscored by Merleau-Ponty's embodiment phenomenology. I have analysed the robot as a socially significant object using each of the key aspects of Pietz's fetish concept: 'historicisation', 'territorialisation', 'reification' and 'personalisation', and have developed each of these categories by drawing on ethnographic description, as well as theorists, in particular Lefebvre, Taussig, Dreyfus and Merleau-Ponty. This was done in order to understand its utility, as postulated by Pietz (1985), as a novel theoretical materialism, divergent from the philosophical tradition.

In Historicisation, I analysed the robot as a composite fabrication resulting from specific historical events, contingencies and social conditions, and thus challenged conventional accounts of technologies as linear, progressive, and inevitable. Instead, I showed how the robot was a result of fictional narratives, scientific discourses, material realities, human ingenuity, scientific failures, institutional supports, commercial applications, ideologies, individual desires, and social norms. This revealed how the robot emerged from a long-standing tradition of attempts to create life and emulate human intelligence, and of attempting to define human essences and boundaries. It demonstrated how, in the mid-20th century, industry, military and academic institutions came together to create the project of AI, explicitly aligning the science of creating artificially intelligent machines to the science of the human. It elaborated how the human is continually configured in terms of available scientific approaches and technologies of the day, and specific to a Euro-American culture that is dominated by the physical and biological sciences within an overarching economic script.

In the same chapter, I also explored how the failure to achieve 'general intelligence' resulted in the diversification of the field, leading to alternative and radical methods being pursued, such as the creation of 'common-sense knowledge' databases, behaviour-based and affective robots, and machine learning techniques. In this chapter, I also investigated some of the images of the past and the future, the mythologies and futurologies that sustain the research. It is by recognising the robot as an historical object that we can focus attention on the transformed temporalities, geographies, materialities, practices, and social arrangements that are set in motion by the novel identity of the robot. In the final part of this chapter, I started to uncover some of the ways in which the robot, and its image, are used to transform space, in particular how a new institutional network dedicated to ‘existential risk’ is shaping policy and practices. These themes were further developed in the following chapters of Territorialisation, Reification, and Personalisation.

In Territorialisation, I focused on the material fact of the robot, as well as the wider socio-spatial landscape. This lens was used to examine the robot as a resolutely physical instantiation bound by its mechanical and localised machinery and the laws of physics. The chapter takes Stevie the robot as its central focus, allowing me to interrogate the concrete form of the robot, uncovering the specific histories, technologies, and cultural imaginaries that are sedimented in it. This connected the histories of the robot to its material form, revealing a lineage reaching back into antiquity, as well as more recent influences, including American and Japanese fiction, Japanese design, MIT’s behaviour-based robotics, machine learning techniques, as well as aspects specific to its geographical location, including institutional support, researcher motivation, and individual design decisions. This lens also revealed how roboticists work in concrete engagements with the material and the physical world, which remains impervious to human intentions and schemes.

Next, in Territorialisation, I investigated the wider spaces of robotics research, drawing on Lefebvre’s concept of ‘social space’. This included a detailed ethnographic account of a number of robotics laboratories in Ireland, the UK and the US, as well as conferences and field test sites. This revealed, on the one hand, the geographical specificity of individual labs, subject to local histories, research priorities, and funding supports. However, it also revealed the ‘conceptual space’ of roboticists, and how the community of researchers is united by common histories, identities, and mythologies. This includes a belief in humans as fundamentally biological and knowable entities, who share a destiny of improvement and progress, which may one day have to be shared with, or even ceded to, manufactured machines. This is in stark contrast to the fact that, despite impressive technical progress, the creation of human-like intelligence remains as elusive as ever. Additionally, this is contrasted with the way in which, through their craft, roboticists are encountering resistances to the dominant vision, and remain open to new ways of theorising intelligence and ontological possibilities. It also revealed ways in which a small number of roboticists have reoriented their focus towards pro-social work. In the final part of this chapter, I describe the network of institutions that has been established dedicated to 'existential risk' and to extending the image of imminent AGI. This, as I show, is comprised of powerful institutions and individuals from academia, and the corporate and entertainment worlds.

This theme continued in the next chapter, Reification, where I examined the status of the robot-fetish as a socially significant or valuable object and its reliance on specific institutional systems for marking that value. This chapter focused on the constructed value of the robot from three different perspectives. First, I examined the construction of robot intelligence as analogous to human intelligence and how it plays a central part in developing and maintaining cultural identities. As we saw, this, in the case of Euro-American imaginaries, is one implicated in the creation and maintenance of colonialism and empire. Second, I showed how the image of the robot is being used to elide the limitations of current technologies, the real work that is necessary to make the technologies appear intelligent and autonomous, and how it has been used to reorganise employment structures and practices, often to the detriment of workers. In the third section of the chapter, I contrasted this with another view of the intelligent and animate robot, where we saw how the performance of robot animacy was used to play, to enchant, to learn, to educate, and to balance social dynamics. This was done by taking into account the creativity and resistances that come into play in the interactions between the embodied individual and the concrete object. This theme was further elaborated in the following chapter, Personalisation.

In Personalisation, I focused on the relation between the embodied person and the robot-as-fetish. In this chapter, I drew on ethnographic description of encounters between the embodied person and the robot-object, including ‘naïve’ users, roboticists, and my own, to develop a detailed description of robot encounters. This revealed how people experienced the robot at times as an extension of themselves, and at others as a separate entity through its autonomous movement, lifelike morphology, and face, in a way that is pre-objectively felt and pre-sententially recognised. Personalisation thus highlighted the way in which, through the embodied encounter, self and the Other are conferred with meaning. It is only through this process that the robot is identified as a unified object and may be considered ‘animate’. I contrasted conventional accounts of the human-robot encounter in both robotics and HRI, which draw on biological and cognitive theories, with an explicitly Merleau-Pontian reading of the encounter. This revealed an embodied human open to an intersubjective world. In this chapter, I then explored other ways in which the robot-fetish is experienced in various embodied encounters, including how roboticists inserted themselves into their research, the craft of robotics work and the experience of those operating the robot, engaged in a creative performance with the ‘new mimetic’ technique of the robot in which they find themselves extended.

Despite its embodied nature, I showed how entanglements with robots are nonetheless also a mindful, historical, and social process, enmeshed within a wider cultural identity. Taken together, Reification and Personalisation show how identity formation is both cultural and personal, including how the interplay of identity and difference in the embodied encounter with robots is a key factor in conceptualising them as a fetish. As Pietz articulates:

> The fetish might then be viewed as the locus of a sort of primary and carnal rhetoric of identification and disavowal that establishes conscious and unconscious value judgments connecting territorialised social things and embodied personal individuals within a series of singular historical fixations. (Pietz 1985, 14)

A combined lens of ‘reification’ and ‘personalisation’ reveals how the visceral experience of the fetish, or robot, is enmeshed in a wider ideological consciousness in which we both identify and disavow the robot object within the context of specific geographical, historical and ideological configurations. As we have seen, for Pietz (1985), the intense personal encounter is ‘always incommensurable with (whether in a way that reinforces or undercuts) the social value codes within which the fetish holds the status of material signifier’ (12-13). Though incommensurable, the personal experience may either reinforce or undercut existing social value codes.

Thus, the visceral, intimate, and familiar experience of the robot reinforces the image of the animate and human-like robot, while simultaneously holding the potential to obscure the wider ideological projects that the robot as material signifier evokes and motivates in us and in wider society. The warmth of the personal encounter can also quieten resistance and concerns. As we have seen, although a number of the carers at the retirement community initially viewed Stevie with suspicion, he appeared to win them over through the enjoyment of the interaction, and the pleasure he clearly gave to the people in their care. Nonetheless, as I have also described, the threat of robots and other technologies to their jobs is real. While a fully autonomous replacement of their roles is not imminent, there is an imminent threat to their jobs, in terms of reconfigured rights, security, and compensation. Similarly, we saw how in Pittsburgh the people who encountered the Starship robot on the path in Pittsburgh moved out of its way to let it pass, thinking it ‘cute’. However, after I leave, I see a post on Twitter in which a wheelchair user is forced off the same street by the one of the Starship robots (Ackerman 2019). Further, as I have described, robots may be deployed for use in intimate spaces, and can elicit more personal and private information than other technologies. This incommensurability of Reification and Personalisation may help to explain, and thus also help to reconcile, the problem of agency versus structure, or phenomenology versus culture. While not intrinsically bad, this ‘jumbling of agency' (428) is, as Graeber (2005) writes, both very 'familiar', and potentially 'dangerous':

> For Marx this becomes dangerous for two reasons. First, because it obscures the process of how value is produced, which is of course very convenient to those who might wish to extract value … Second, all of this makes it much easier to treat the … tendencies of whatever system it may be, as natural, immutable, and therefore completely outside any possibility of human intervention. (428-429)

On the other hand, it is also in the embodied and personal encounters with the machine that its potential for historical change and social creativity resides.

Thus, as well as a single aspect, by combining two or more aspects of Pietz's fetish we can get particular views of the robot. Combining 'historicisation' and 'reification' draws our attention to the constructed and historic aspects of the object, and the structured relationships that are figured, and reconfigured by the event of the robot's fixation. A combination of 'territorialisation' and 'historicisation' brings a spatio-temporal perspective to the object. By combining 'territorialisation' and 'reification', we can consider the intersection between the social value of the robot and the ways that space is conceived, reconfigured, and produced. Together, the themes of 'territorialisation' and 'personalisation' reveal the importance of materiality in the concrete, embodied encounter. The lens of 'historicisation' and 'personalisation' underscores the ways in which a person's relation to the world is simultaneously historical and embodied.

The Robot as Fetish

For Pietz (1985), the fetish concept is described, not as fixed object with a prior model or truth, but a 'radically historical object that is nothing other than the totalized series of its particular usages' (7). The fetish is thus not an 'indigenous' concept, rather, it is a concept developed by the West to explain relations that they felt but could not describe. As Pietz's analysis suggests, the 'fetish' concept reveals more about the West's contradictory relationship with objects and the world around it than about any external culture onto which it is often projected. As we have seen, Pietz argues that the concept thus ‘marks the breakdown of the adequacy of the earlier discourse’ (1985, 6). His preliminary theory of the fetish ‘represents the emerging articulation of a theoretical materialism quite incompatible and in conflict with the philosophical tradition’ (1985, 6). Specifically, this is ‘the capacity of the material objects to embody – simultaneously and sequentially – religious, commercial, aesthetic and sexual values’ (1985, 6). The concept of the ‘fetish’, then, ‘has always named the incomprehensible mystery of the power of material things to be collective social objects experienced by individuals as truly embodying determinate values or virtues’ (1985, 6). What distinguishes Western culture from ‘primitive’ cultures, then, is not a superior empirical understanding of nature, but the inability for the empiricist view to capture the fullness of our entanglements with the world. As Hornborg (2006) articulates:

> … it may not so much be an incapacity to relate as such that distinguishes us from the animists, as the incapacity to exercise such ‘relatedness’ within the discursive and technical constraints of the professional subcultures which organize the most significant share of our social agency. Science and technology does not so much make us into robots, as make specific *parts* of our behaviour robot-like. (24)

As we have seen, a number of theorists (Marx 2008; Taussig 2010; Pfaffenberger 1988; Graeber 2005; Hornborg 2006) have used the fetish concept to describe the complexity and mysteriousness of people’s relations to non-human worlds. These concepts are seen as a possible solution to the problem of Cartesian dualisms and an over-rigorous insistence on scientific objectivism, particularly with regard to social worlds. However, Pietz’ articulation of the fetish concept is unique, both in stressing the essential material fact of the fetish, as well as the centrality of embodiment. Furthermore, for this study, it provides a comprehensive and bounded theoretical framework to aid the analysis.

In my study, Pietz’s model of the fetish is used to sensitise us to alternative understanding of human agency, intelligence, and boundary separations/resistances. It provides a pluralistic analytical scheme that acknowledges the multifarious, and often incommensurate, ways in which the relations between humans and the object-as-fetish may be described.

Furthermore, I have created a novel conceptual framework with which to explore the robot as fetish, by combining Pietz's fetish theory with an explicitly Merleau-Pontian reading of the body. I have also extended specific aspects by drawing on key theorists and their concepts, such as Lefebvre's concept of 'space', and Taussig's use of 'mimesis' and 'alterity'. This novel framework has revealed four different, at times complementary, at times contradictory, ways of framing relations between person and the object. Each of the four aspects has revealed a distinct, yet essential part of our relationship with objects. It is also worth noting, however, as Taussig (1993b) points out, that 'sensuous knowledge' is historical, 'in part because of the colonial trade in wildness that the history of the senses involves.' (44). Thus, in traditional theory, each of these four aspects of the fetish is not represented or acknowledged equally. While the physical object and 'conceived' space are viewed as 'real' and 'true', discontinuous historical and embodied knowledges are often unrecognised and unacknowledged.

The fourfold character of the fetish is another significant aspect of the fetish concept. In *Thirdspace*, Edward Soja (1996) argues that Lefebvre's trialectic 'meta-philosophy' of spatiality–historicality–sociality extends beyond the dominant dualistic and temporally defined dialectics of Hegel and Marx (36). This allows for an expansion beyond binary oppositions that are closed to new, unanticipated possibilities (1996, 30-31). As we have seen, Foucault developed the concept of 'heterotopias' to articulate a new way of thinking about space, which is made up of social relations, contradictions, and ambiguities. It is, as Soja articulates, similarly trialectic, comprising space, knowledge, and power (1996, 15). Soja expands on these ideas in his concept of 'thirdspace', which represents 'a radically different way of looking at, interpreting, and acting to change the embracing spatiality of human life' (1996, 29). Thirdspace is a space where the real and the imagined, the conceptual and the lived mingle, a marginal space of creativity, a space of radical openness. These are spaces 'outside', where difference and otherness are embraced (1996, 35). In the same way, Pietz's framework is a 'quadruple dialectic', or 'quadralectic', accommodating more than binaries, or even single alternatives. It thus allows for the articulation of a complex, pluralistic account that accommodates contradiction, ambiguity, and incommensurability, while still offering a circumscribed and usable framework.

Thus, we find that an encounter with the robot object is not entirely social, and while co-constituted, not entirely symmetrical. It is at once pre-discursive, visceral and embodied, as well as scientifically and socially significant. Fetish theory may offer a way to integrate and reconcile an account of power and political economics, with a personal embodied experience of the object in a way that reveals the inherent contradictions between individual experience and social phenomena. By interrogating the robot as fetish, we have revealed the complexity of human-object relations, including the importance of considering the embodied and the material, as well as the historical and the symbolic. This reveals the simultaneously natural, phenomenological, historical, symbolic, and political ways in which we experience the world and others. Thus, we find, as Pietz postulates, that the fetish, and thus also the robot-as-fetish, is 'a special type of collective object that reveals the truth of all historical objects' (1985, 14). Pietz quotes both Deleuze and Merleau-Ponty with regard to the universality of the fetish. For Deleuze, 'the fetish is the natural object of social consciousness as common sense or recognition of value' (Pietz 1985, 6). For Merleau-Ponty, 'every historical object is a fetish' (1985, 5). Thus, we find that the fetish represents, not just humanoid robots, but technology in general, or the 'truth' of the 'total collective material object' (1985, 14) and a novel theoretical materialism, divergent from the philosophical tradition. Through the prism of the fetish, then, we can elaborate how the human is produced in robotics research, while also gaining an understanding of the wider human-object relation.

Producing the Human

As we have traced, early AI efforts were dominated by symbolic AI and the cognitivist paradigm. It was assumed that intelligence would be solved by a group of carefully selected experts in a number of months by describing it 'precisely' enough. This view was upheld by established institutions and any critiques were summarily quashed. It was not until the rise of machine learning and the corresponding paradigm of connectionism that an alternative view of intelligence was admitted. It was in this interstice that Dreyfus's critique and phenomenological insights, as well as critiques that emphasised the social and relational aspects of human 'intelligence' and action, were picked up by Agre and Brooks, and the teams at MIT. Building on early success in the areas of navigation, efficient computational resourcing, and perception of machine animacy, Brooks and his team attempted to develop a full humanoid robot, Cog. Cog did not succeed in displaying any advances in terms of its 'intelligence', however, the teams were startled by the responses of those that encountered it. Cog's lack of success also led to a flurry of critiques.

Dreyfus (2007) critiques both Brooks' 'behaviourist' and Agre's 'pragmatic' approach to what he terms 'Heideggarian AI'. Although Dreyfus considers Brooks' approach an important advance, he criticises it in that it only responds to fixed features of the environment, not to context or changing significance. Similarly, Agre's approach, as instantiated in Pengi, is criticised for objectifying 'both the functions and their situational relevance for the agent', which he considers an impoverished instantiation of Merleau-Ponty's 'solicitations', or 'the experiential aspect of being drawn in by an affordance' (2007, 253). Thus, according to Dreyfus, for both Brooks and Agre, 'no skill is involved and no learning takes place':

> …Agre doesn't even try to account for how our experience feeds back and changes our sense of the significance of the next situation and what is relevant in it. In putting his virtual agent in a virtual world where all possibly relevance is determined beforehand, Agre doesn't account for how we learn to respond to new relevancies, and so, like Brooks, he finesses rather than solves the frame problem. (2007, 253)

By contrast, Dreyfus invokes Merleau-Ponty's concepts of 'intentional arc' to show how 'intelligence' develops in the learner through experience:

> What the learner acquires through experience is not *represented* at all but is *presented* to the learner as more and more finely discriminated situations, and, if the situation does not clearly solicit a single response or if the response does not produce a satisfactory result, the learner is led to further refine his discriminations, which, in turn, solicit more refined responses. (2007, 250)

Other critiques focus particularly on Brooks' interpretation of both 'situatedness' and 'embodiment'. Adam (1998) criticises Cog's design for a lack of sociality and an embodiment that ignores its 'feminine' aspects (155). Similarly, Suchman's (2007) observations of both Cog and Kismet lead her to observe that, although the concept of 'situated' have been incorporated into business as usual within AI research, its meaning and implications have been distorted:

> For Brooks, *situated* means that creatures reflect in their design an adaptation to particular environments. Following a lineage traceable to the founding premises of cybernetics, Brooks's situatedness is one evacuated of sociality, at least as other than a further elaboration of an environment understood primarily in physical terms. (15).

Furthermore, for Suchman, this interpretation of situatedness results in a complete misunderstanding, indeed an inversion of the concept: '[t]he creatures "interactions" with the environment … comprise variations of conditioned response', thus leading 'in some cases to term's appropriation in support of various forms of neobehaviourism' and even being reinterpreted by some in the field as meaning 'predetermined' (Suchman 2007, 15). As well as taking issue with the interpretation of 'situated', Suchman also takes aim at the mischaracterisation of the concept of 'embodiment', noting how the concept is misunderstood as something that nonetheless remains secondary to the mind and exists within an objective, naturalistic world:

> The first thing to note is that discoveries of the body in artificial intelligence and robotics inevitably locate its importance vis-a-vis the successful operations of mind or at least of some form of instrumental cognition. The latter in this respect remains primary, however much mind may be formed in and through the workings of embodied action. The second consistent move is the positing of a 'world' that preexists independent of the body. The body then acts as a kind of receiver for stimuli given by the world, and generator of appropriate responses to it, through which the body 'grounds' the symbolic processes of mind. Just as mind remains primary to body, the world remains prior to and separate from perception and action, however much the latter may affect and be affected by it. And both body and world remain a naturalized foundation for the workings of mind. (2007, 230-231)

Despite these critiques, as we have seen from experiments in social robotics, Brooks and Breazeal have made genuine attempts to implement embodiment, sociality and learning into Cog and its successor, Kismet. Indeed, those are the primary goals in Kismet's design (Breazeal 2002). Any yet, social robots like Kismet and Stevie are no closer to a breakthrough in AGI than earlier symbolic AI projects.

As Suchman (2007) also notes, Brooks has admitted that 'there might indeed be some "new stuff" that we need' to allow robots to 'take off by themselves' (Brooks 2002, 184). Similarly, the current online project overview for Cog recognises the limitation of current technologies to realise the complexity of embodiment, '[s]ince we can only build a very crude approximation to a human body there is a danger that the essential aspects of the human body will be totally missed.' (CSAIL 2021). According to Richardson (2015), the complex concept of embodiment as it is understood by robotics researchers at MIT is more than just an argument for a physical body, indeed it is a 'recognition that intelligence, cognition, perception, linguistic and non-linguistic communication is only possible through bodies – sensual, fleshy, sensory-motor, proprioceptive bodies' (48). However, it is difficult to understand from her observations how the 'sensual, fleshy' body is made manifest in the research that she describes.

For Suchman (2007), beyond sociality, embodiment also implies the contingent interactions of biological, cultural-historical and autobiographically experience (231). However, there is little detail in her work as to what this might actually mean. Suchman (2011) continues to espouse embodiment in her later work, where she develops her view of it in terms of what she now calls 'embodied relations', in her observations of interactions with the robot Mertz, designed by one of Breazeal's students. She observes:

> Entrained by Mertz's vitality, the human interlocutors are robotically subjectified; shifting their orientation to each other's queries and laughter, the robot is correspondingly restored to humanlike objectness ... The interactivity of persons and things is manifest here as moments of bodily imitation and connection animated by affective dynamics that escape their classification. (2011, 120)

Despite emphasising the importance of the body and embodiment, it is clearly the relational implications of situatedness that have most resonance for her. Indeed, according to Suchman, the key problem with AI and robotics research is its focus on separation, autonomy, and individuation (213). As she argues, it is the 'ontology of separate things' that then needs to be 'joined together', that inhibits the project of AI (257). Although not claiming to comment on robot 'intelligence', Richardson's (2015) argument is almost the reverse of Suchman's. For Richardson, it is the mistaken rejection, or, as she calls it, the 'annihilation', of categories of separation that underscore both robotics and anthropological theory. This, she believes ultimately leads to dystopian readings of machines, as well as anxiety and nihilism (2015, 4, 130). Drawing primarily on psychological literature, Richardson proposes a 'humanistic asymmetry' as an alternative to Latour's 'ahumanist symmetry' (2015, 5).

In my research, I have not found either separateness or symmetry to be a core issue inhibiting AI and robotics. Even Suchman (2007) acknowledges that robotics systems and their 'intelligence' are connected and distributed (207). Similarly, as I have explored in this study, robot systems include swarm robots, cooperative robots, robots as part of smart environments, non-humanoid robotics, robots connected to the Internet, the Cloud and distributed data. From my observations and interviews, I find instead that roboticists are broadly, quite open to 'unbounded' conceptions of intelligence, nor do they see their robots as bounded entities, except from the perspective of their users or when they themselves come upon them.

Instead of the issues of separateness versus symmetry, in this thesis I argue that what is lacking is a full understanding of embodiment. Although both Suchman and Richardson emphasise the concept of embodiment, they have not developed it in in the context of its wider ontological implications for the figure of the human. As we have seen, Suchman's original articulation of thought is not developed beyond one in which human and computers are considered analogous (Suchman 1985). Similarly, Richardson adopts the cognitivist interpretation of anthropomorphism and ToM in analysing user interactions with robots (Richardson 2015, 73).

Thus, in anthropological studies of robotics, the centrality of embodiment is often overlooked in favour of a relational ontology that retains elements of cognitivism.

Despite the insights that the concept of embodiment offers, its transformative implications are often lost, both in robotics and in anthropological studies of robotics. In robotics, it has come to justify the robot having a humanoid or zoomorphic body, as well as an emphasis on sensors and distributed processing. Instead of moving beyond a behaviourist or cognitive paradigm, in robotics theory, the two are merged. On the one hand, 'low-level', physical, bodily behaviour is known through observation, while 'high-level' cognitive thought is modelled computationally. In this way, overwhelmingly in robotics, the field has retained a commitment to the human understood mathematically and reductively. Evolution is the organising and explanatory principle, in which lower-level behaviours evolve into higher level ones. Any phenomena that remain impervious to these methods, such as sociality, tacit knowledge and everyday action, indeed, even areas such as life, intelligence and consciousness, are all explained by 'evolution', a point that is particularly evident in studies of Artificial Life, see (Helmreich 2000; Hayles 1999; Kember 2003; Adam 1998; Suchman 2007).

However, as we have seen, despite a lot of attention, the programme for 'embodied' and 'situated' robots in fact represents a small part of the wider robotics and AI effort. As discussed, deep learning and 'big data' are the primary reason that symbolic AI has been superseded as the dominant approach in the field. The connectionist programme is said to model the brain's learning power as an alternative to symbolic representation. Instead of using rules, a computer could 'learn' from a set number of examples and then extrapolate or generalise from those to respond to other situations. The successes of machine learning have been seen by some as being a potential path to creating AGI, and a validation for a 'connectivist' model of human intelligence. It also has, according to Papert, a 'cultural resonance' with those who are compelled by behaviourism, both by its association with biological neurons, as well as the view that all behaviour can be shaped through reinforcement, in (Katz 2017, 6). A number of commentators have made some elaborate claims regarding deep learning's ability to lead to 'general intelligence', such as Andrew Ng, who declared that: 'If a typical person can do a mental task with less than one second of thought, we can probably automate it using AI either now or in the near future' (Ng 2016).

Dreyfus (1992) had originally enthusiastically welcomed the programme of 'neural-network modellers', not, he says, as they offered a more plausible account of intelligence, but because they provided a provided a powerful alternative to the dominant symbolic AI approach (xiii-xvi). And yet, he argued, connectionism remains open to the same criticisms as symbolic AI, being divorced from a human context and situation and failing, once again, to generalise (Dreyfus 1992). According to Dreyfus (1992), a human-like learning device, would have to share 'enough human concerns and human structures to learn to generalise the way that humans do' (xlv-xlvi). Coming from a psychological perspective, Gary Markus (2018) has critiqued the potential for deep learning to advance towards AGI on a number of grounds, including; its need for continuous data, the superficiality of the extracted patterns, its inability to deal with hierarchy (making it difficult, for example, to learn sequential positions), its struggles with open-ended inference, the lack of transparency, and its inability to integrate prior knowledge.

Just like the previous cognitive programme, then, machine learning has not brought us any closer to realising human-like intelligence. As Dreyfus (1992) remarks, '[i]t looks likely that the neglected and then revived connectionist approach is merely getting its deserved chance to fail.' (xxxviii). However, it only fails as a model for human intelligence. Rather than imminent human-like intelligence, machine learning is an advanced statistical technique focused on a specific dataset. What makes deep learning so powerful, and in some respects, dangerous, is that it can be used to identify, reproduce, and entrench existing social patterns in ways that humans cannot emulate and, in many cases, cannot understand. Thus, it is precisely its divergence from human intelligence, rather than its continuity with human intelligence, that makes it so compelling.

Most roboticists with whom I talked claimed agnosticism when it comes to any overarching philosophical or metaphysical commitments, beyond a vague materialism. Additionally, few would claim that their robots with bodies, or their machine learning algorithms, were definitive paths towards AGI. Instead, most roboticists say that they are willing to use anything, as long as it 'works'. As Agre (1997) identifies: 'what truly founds computational work is the practitioner's evolving sense of what can be built and what cannot (11). However, as we have seen, this only extends to phenomena that can be described in a way that is 'specific' enough, or, as a 'pre-given formal-mathematical specification' (1997, 14). Ultimately, then, I found that roboticists remain committed to a scientific view of the human comparable to that of the physical sciences, including a mechanistic and material view of the universe, albeit mostly implicitly. AI is, as Agre (1997) has phrased it, 'covert philosophy' (240).

Furthermore, although most roboticists, despite some notable exceptions, eschew any claims to imminent AGI, they nonetheless support the narrative in subtle ways. This includes exaggerating claims for funding proposals, using Wizard of Oz techniques in public robot demonstrations, and in tacit support of those in adjacent fields who are pushing this narrative. Roboticists are required to connect their projects to the image of the (Enlightenment) human configured scientifically, by biology, economics, and psychology, within a wider unquestioned narrative of historical progress and inevitability, and unequal hierarchical social structures.

At a recent ACM International Conference on Human-Robot Interaction, Brooks (2018) admonishes HRI researchers for the lack of impact their research has on the companies that are developing and producing robots in earnest. This is because, he writes, the field has a 'fetish for mathematical notation', its only 'purpose is to obscure simplicity' (2018, 1-2). Further, its method of user testing, which either relies on Wizard of Oz testing in lab environments or on people recruited via Amazon's Mechanical Turk to fill in questionnaires, has 'nothing to do with real users' (2018, 2). These methods start with a 'pre-set list of issues', allowing only for a focus on ‘extremely tiny aspects of what should be much bigger pictures of how ordinary people are going to work with a robot in real life to get a task done that actually matters to them’ (2018, 2).

Brooks encourages researchers to avoid ‘mediocrity and irrelevance’ and instead of ‘being sucked into a world of minimal viable safe units of publication’ to be brave, creative, take risks and have fun (2018, 3). And yet, as I have found in my fieldwork, the researchers working in HRI were also those most open to approaches from other disciplines and to new and alternative theories. Many described their frustration with the limitations of the discipline and difficulty in getting any genuinely creative or interesting research published at peer-reviewed conferences or journals. Indeed, many conferences reserved a separate track for, what one of my informants called, ‘crazy’ research.

Thus, the ‘technology imaginaries’ of roboticists are driven, not so much by their individual commitment to a particular philosophical viewpoint, but rather by a network of interrelated institutional orderings. These institutions are: industry and government funding bodies dependent on the ‘usefulness’ of the humanoid robot; academic institutions, industry organisations, and publishing companies for whom a ‘science’ of the human is equal to physical science, the network of institutions reproducing the ideological, post-Enlightenment fantasies and visions of disembodied intelligence. As we have seen, this latter vision is advanced by powerful individuals in adjacent technological industries, whose influence is concretised in dedicated think-tanks validated by top-tier universities, technology companies and the world of entertainment, as well as through prestigious and lucrative competitions and honours systems.

Taking Embodiment Seriously

Based on the data and insights that I have gathered and analysed in this study, I argue that taking embodiment seriously is inhibited by a superficial reading of the concept of embodiment, reinforced by the institutional ordering of the field. In particular, there are three areas in which an alternate reading of embodiment challenges current approaches. First, Merleau-

Ponty's concept of embodiment runs counter to roboticists' commitment to a scientific view of the human comparable to that of the physical sciences. Second, it negates the related assumption that aspects of the human world, such as 'generality' and 'common-sense', can be reduced to single variables. Third, it resists the interpretation of embodied knowledge as 'simple' or 'low-level'. In this section, I examine each one of these in turn.

As we have seen, the foundational conceptual framework in robotics remains 'naturalistic', oscillating between behaviourism and cognitivism, within an overarching biological and evolutionary explanatory principle. Despite Brook's phenomenological insights and criticism of the HRI field for its lack of imagination, Brook's commitment to scientism similarly directs his research. Philosophers Gallagher and Zahavi (2010) trace Brooks' 'clearly Cartesian' interpretation of embodiment all the way back to La Mettrie (1745), which 'characterizes animals as purely physical automata' (134). What is more, they observe, he goes even further by 'topping it off by proposing that conscious-like intelligence should emerge from this kind of system' (135). Additionally, Brooks remains committed to a view of the human as a machine with an evolutionary framework. This narrative, as Casañeda and Suchman (2014) articulate, posit robots, see '[a]s entities with aspirations to a universalized, fully human status, where that is taken to be the higher order', further, 'these figures reiterate developmental/evolutionary discourses born in the crucible of imperial projects of expansion and domination (317). The evolutionary discourse further implies a hostile world, in which competing individuals struggle for survival and replication.

In previous chapters we have explored how, in the fields of HRI and social robotics, the underlying commitment to 'hard' science methods, such a statistics and physiological studies, leads to hypothesis testing rather than creative exploration, in which the 'norm' is confirmed and reproduced, while outliers are viewed as insignificant or irrelevant. Robots are designed to meet majority expectations using the metric of 'user acceptance'. This can lead to the reinforcement of stereotypes. Similarly, in machine learning, only the most regular patterns underlying social data are reproduced.

Further, the ‘naturalistic’ approach to science can lead to a general uncritical acceptance of categories that are assumed to be ‘natural’, such as race or gender, or even types of disease (Elish and boyd 2018, 18). At its worst, this lack of critical thought has facilitated the re-emergence of previously debunked pseudo-sciences, which might have been assumed obsolete, such as physiognomy and phrenology. Thus, papers are emerging such as the controversial ‘A Deep Neural Network Model to Predict Criminality Using Image Processing’, which was only withdrawn from *Springer Nature* due to a massive outcry and open letter from researchers and practitioners across a variety of technical, scientific, and humanistic fields, including a number of anthropologists (Coalition for Critical Technology 2021).

We can identify this as the ‘conceptual space’ as articulated by Lefebvre (1991) and described in Ingold’s (1997; 2000) ‘machine-theoretical cosmology’. As Ingold (1997) points out, this results in importing from the technical domain, ‘the kinds of antagonism that is presumed to exist in people’s relations with non-human beings’, with relations with the non-human world seen as ‘fundamentally exploitative’ (117). This, he says, represents, a ‘transactional failure in relations with the environment’, which, for many is a relationship of ‘mutualism and trust rather than domination and exploitation’ (Ingold 1997, 117-118). For Lefebvre (1991), ‘reduction is a scientific procedure designed to deal with the complexity and chaos of brute observations. This kind of simplification is necessary at first, but it must be quickly followed by the gradual restoration of what has thus been temporarily set aside for the sake of analysis’ (1991, 105). Unless there is a return to experience, ‘a methodological necessity may become a servitude, and the legitimate operation of reduction may be transformed into the abuse of *reductionism*’ (1991, 106). Similarly, Taussig (2010) has argued that the ‘positivist doctrines’ of the ‘natural science model’ are responsible for the ‘petrification of social life’, alienating people from nature, and subject from object (6). Thus, he calls ‘to restore to the abstract, context free, calculative, universal', 'the resistance of the concrete particular', particular its 'sensuousness, its mimeticity' (Taussig 1993b, 2).

This overarching commitment to a scientific view of the human leads to other key insights from phenomenology being missed. Thus, the second point is the assumption that human 'worlds' can be reduced to single words, and thus rendered independent variables. Thus, we have seen projects that claim to develop 'common-sense', by cataloguing innumerous de-contextualised items of knowledge. Likewise, the concept of 'generality' assumes that being able to apply knowledge from one context to another can be reduced to a simple formula. As we have seen in the ethnographic data, researchers are searching for a universal theory of mind, of movement, and of interaction, assuming that these will be found, and made explicit and implementable. Underlying this logic is the assumption that human worlds can be reduced to independent variables and that current techniques and materials are sufficient to lead to human-like intelligence. As Dreyfus (1992) has argued, it is a:

> …conceptual framework which assumes that an explanation of human behaviour can and must take the Platonic form, successful in physical explanation; that situations can be treated like physical states; that the human world can be treated like the physical universe. (232)

As Agre (1997) remarks in his observations of Pengi, Pengi 'interacts continually with the world but it is not truly *in* the world' (296). It is clear that being *in* the world requires a step change in current techniques.

The third key insight that is missed is the assumption, as we have seen in both Brooks and Agre, that embodied interactions are 'simple', 'low-level', and 'routine', and that from these, somehow, intelligent and complex behaviour will 'emerge'. This narrative is implicated, though not explicitly, with discourses around low-value and un-skilled labour. By contrast, as both Merleau-Ponty (2012) in his discussion of habit, and Dreyfus and Dreyfus (2005) in their discussions of expertise, elaborate, embodied knowledge is much more complex. For Merleau-Ponty, habit presupposes a form of 'understanding' that the body has of the world in which it carries out its operations (Moya 2014, 1). Dreyfus and Dreyfus

(2005) invert the traditional account of intelligence, in which ‘higher-level’ intelligence is associated with rule-based certainty. Instead, as they show, it is the initial ‘novice’ stage in learning that is characterised by explicit rules. By level five, ‘expertise’, skilful activity is characterised by an ‘immediate intuitive situational response’, ‘without calculating and comparing alternatives’ (Dreyfus and Dreyfus 2005, 787-788). Ultimately, whether expertly walking, driving, playing music or carrying out surgery, there is no time for reflection or deliberation, or monitoring what one is doing, instead ‘one is in the flow’ or ‘in the groove’ (Dreyfus and Dreyfus 2005, 789).

According to a phenomenological theory of embodiment, our ability to be in the world is characterised by our ‘expert’ ability to be responsive, flexible, reactive, and creative in responding to the world. It is not just a case of ‘sense and respond’, rather it is the immediate integration and synthesis of past experiences, personal dispositions, bodily situatedness and attention, environmental factors, cultural awareness and norms, social interactions, future intentions and motivations, as well as conscious cognitive processes, all ordered by overarching metaphysical, ontological, social and cultural being in the world. We are indeed, as Moravec (1995) identified ‘prodigious Olympians in perceptual and motor areas’. However, this is not, as Moravec suggests, simply explained by ‘evolution’:

> Encoded in the large, highly evolved sensory and motor portions of the human brain is a billion years of experience about the nature of the world and how to survive in it. (15).

Instead, it is because of our creative and embodied, situated involvement in the world, which we have barely even begun to understand. Taussig (1993b) also sees bodily habit as being an essential part of ‘tactile knowing’ (25). Thus, Taussig argues that a change in ‘habit’ is necessary for ‘radical change’ to occur:

> Habit offers a profound example of tactile knowing and is very much on Benjamin’s mind, because only at the depth of habit is radical change effected, where unconscious strata of culture are built into social routines as bodily dispositions. (1993b, 25)

For Taussig, however, habit is similarly ‘automatic’, needing to be ‘awakened to its own automaticity’ (1993b, 25). I would argue, by contrast, that what is needed is a reawakening to the skill and creativity that is inherent in our engagements with the world through a reconfiguration of the human as both relational and embodied.

By taking Merleau-Ponty's concept of embodiment seriously, in both a practical and ontological sense, we see the human as an embodied and situated being, primordially engaged with the world, *of* the world. The body refuses to be abstracted. It is sensuous and sensing, carnal, finite, vulnerable, exposed, and mortal. It cannot be scaled up or down, cut or co-located without trauma. Unlike Latour's (1987) 'immutable mobiles', it is not durable, not reusable, and not repeatable. This pre-conceptual, pre-discursive, and embodied existence is ontologically prior to science, symbols or discourse. This forms the backdrop and horizon to our interactions with and reflections on the world. This is as true for ethnographic researchers, ethnographic informants, robotics scientists and robot users, but not, at least not yet, for robots. Merleau-Ponty's insistence on an inherent ambiguity and vagueness highlights the limitations of a science and a philosophy that aims at complete knowledge. His articulation of a dialectic that can encompass mutuality and apparent contradiction is a challenge to reductive causality as a metaphysical claim. Nonetheless, the centrality of the body should nonetheless not blind us to that which is not revealed in experience.

The implication, therefore, is that the concept of situated cannot be untethered from the concept of embodiment. It also means that it goes further than ensuring that the biological body, in all its fleshiness and messiness, is not written out of scientific accounts. This implies that the living body is essential to the concepts of 'relationality' and 'intersubjectivity'. We are connected to one another, through our bodies, through our perception the whole world around us is imbued with vitality. The solicitations of the world are animated through us. Merleau-Ponty's embodiment concept describes a prior integration of what are traditionally considered separate domains that no formalisation has hitherto captured. How bodies come alive, how they experience the world and their creativity, is still a mystery to science. It is likely that it cannot be solved by current methods, either in science or in technology, and may always remain ambiguous.

Reconfigurations

In this dissertation, I have explored how the human is produced in robotics research. As we have seen, multiple figures of the human, and of human-machine entanglements, emerge. We saw how the robot is often seen as a literal attempt to produce or reproduce the human and may be conceived of as a project in experimental philosophy, in which theories of the human are tested, proven, or disproven. However, we also traced how many roboticists claim not to adhere to an explicit philosophical approach, and instead underscore the pragmatic nature of their work. Despite this, roboticists also broadly subscribe to a materialistic and scientific view of the world, in which humans are ultimately knowable and reproducible. Further, we explored how conceptions of the human, both in robotics, wider academia and in 'common sense', mirror the available scientific theories of the day. This view of the human, as amenable to scientific and quantitative investigation, is upheld by a network of interrelated institutional orderings, including funding bodies, academic institutions, and the corporate world. Through these mechanisms, the human continues to be conceived as scientifically knowable, despite a lack of evidence.

Explicit and implicit theories of the human have emerged in my research. These include that of the universal subject, whose rationality and 'intelligence' is the blueprint for AGI or even super-human intelligence. There is also the robot user, who is modelled in terms of 'mental models' and 'anthropomorphic projections' of psychology, often not 'fully' human, and subject to the determinants of evolution and biology. Less explicitly, there is also the invisible worker, whose gendered and racialised work is devalued and hidden, classed as unskilled and demeaning. The robot is also used mimetic object, which is used to construct the self and the Other. Thus, I found that the human is produced in a number of distinct, and sometimes contradictory, ways.

However, as well as production, and reproduction, there is also resistance. The failure of the AI programme to realise general intelligence has led to the field challenging core assumptions about what it means to be human, questioning the dominant ideologies and exploring alternative ways of modelling the human. This latter programme, however, is generally obscured in favour of more fantastical claims. As we have seen, while indications of human-like intelligence are breathlessly predicted and reported, the implications of the failures of artificial intelligence are in fact the most astonishing and novel. For 2,000 years, philosophy has been dominated by a particular view of humanity. What we have now discovered is that humans, and indeed other forms of life, are infinitely more complex, mysterious, and indeterminate than previously thought. In this way, another figure of the human has emerged; an embodied and situated being, whose mysterious relations to the world, and to the object, in many ways defy precision and calculation. This reveals how humans and objects are inextricably entangled and co-constituted, yet not symmetrical.

As early as 1972, in his original book *What Computer Can't Do*, Dreyfus (1992) calls AI researchers 'the last metaphysicians', who 'are staking everything on man's ability to formalize his behaviour; to bypass brain and body, and arrive, all the more surely, at the essence of rationality' (78). If successful, they 'will confirm an understanding of man as an object, which western thinkers have been groping towards for two thousand years but which they only know have the tools to express and implement' (1992, 78). 'If, on the other hand', he writes, 'artificial intelligence should turn out to be impossible, then we will have to distinguish human from artificial reason, and this too will radically change our view of ourselves' (1992, 78-79). Ultimately, then:

> In Heideggerian terms this is to say that if Western Metaphysics reaches its culmination in Cybernetics, the recent difficulties in artificial intelligence, rather than reflecting technological limitations, may reveal the limitations of technology. (1992, 227)

Dreyfus concludes with the optimistic suggestion that:

> …if instead of trying to minimize our difficulties, we try to understand what they show … We can then view recent work in artificial intelligence as a crucial experiment, disconfirming the

> traditional assumption that human reason can be analysed into rule-governed operations on situation-free discrete elements—the most important disconfirmation of this metaphysical demand that has ever been produced. This technique of turning our philosophical assumptions into technology until they reveal their limits suggests fascinating new areas for basic research. (1992, 303-304)

However, despite Brooks' minor admissions in discrete places about the limitations of Cog's embodiment, as Dreyfus (2007) remarks 35 years later, perhaps somewhat bitterly, 'as far as I know, neither Dennett nor anyone connected with the project has published an account of the failure and asked what mistaken assumptions underlay their absurd optimism.' (250).

As we have explored in this dissertation, while the problem of human-like 'intelligence' may not necessarily always be impervious to science, it is currently so incredibly complex and unaccounted for, that it might as well be. The implication of this, as Dreyfus suggests, should radically change our view of ourselves. If AI has tested and found limitations in terms of symbolic processing, social robots and connectionism are currently showing up the limitations of a purely relational approach. The fields of robotics and AI research are thus showing that despite the many boundary breakdowns, between nature and culture, society and technology, the living body resists. The body-subject can, for now at least, resist translation into a problem of coding, instrumental control and market exchange. But what resists is not 'Man' as 'the embodiment of Western logos' (Haraway 1991, 173), but the embodiment of something much more ambiguous and elusive, and in its elusiveness, potentially universal.

Thus, we have found that high-tech culture challenges inherited assumptions in intriguing ways, however, not always because of what it achieves, but also sometimes in what it fails to achieve. We have seen how mechanical and algorithmic objects remain stubbornly inert, vitalised only by the innumerable interventions and interactions of the people involved in building them, animating them, interacting with, interpreting them, and creating the narratives that sustain them. Yet, this is not a criticism of the technologies of AI and robotics, which, despite not being 'human-like', have progressed at an astonishing rate, providing new possibilities for being in the world. Nor is it a criticism of robotics researchers, who, despite the constraining norms of their field, have remained open to new entrants, embraced new theories and disciplines, and opened up radically new, fruitful ways of theorising intelligence and ontological possibilities.

A number of scholars working at the intersection of technology and social sciences have attempted to develop new ways of theorising the human-technology relations. In *Computation and Human Experience*, Agre (1997) criticises both mentalism and interactionism for their 'overly individualistic conceptions of agents in environments' (314). However, he also rejects the computational metaphors that have been adopted, such as 'networks' and 'cyberspace', which he says 'would dissolve all individuals into a boundless *res cogitans*' (1997, 314). Instead, he proposes a 'deictic' ontology, which is neither objective nor subjective:

> A deictic ontology, then, is not objective, because entities are constituted in relation to an agent's habitual forms of activity. But neither is it subjective. These forms of activity are not arbitrary; they are organized by a culture and fit together with a cultural way of organising the material world. (1997, 244)

As opposed to 'the representational theory of intentionality', this alternative ontology begins 'with the phenomenological intuition that everyday routine activities are founded in habitual, embodied ways of interacting with people, places, and things in the world' (1997, 243). A deictic ontology, he writes, 'can be defined only in *indexical* and *functional* terms, that is, in relation to an agent's spatial location, social position, or current or typical goals or projects (1997, 243). Ultimately, however, Agre remains constrained by his own disciplinary norms, assuming that the experiential dimension revealed by phenomenology can be programmed (Dreyfus 1992; Masís 2014). Nonetheless, Agre's deictic ontology, emphasising an agent's spatial location, social position, and goals or project, combined with his call for a study of the history of technology, although not an exact mirroring, reflects very similar concerns to Pietz's fetish 'quadralectic'.

As we have described, Haraway's (1991) 'Cyborg' is used to articulate more positive configurations between humans and technology, as opposed to traditional critiques that insist on the threat and necessary domination of technology in society. For Haraway, the political struggle is to see the Cyborg from both perspectives. Thus, while on the one hand it can be viewed as 'the final imposition of a grid of control on the planet … about the final appropriation of women's bodies in a masculinist orgy of war', it is also 'about lived social and bodily realities in which people are not afraid of their joint kinship with animals and machines, not afraid of permanently partial identities and contradictory standpoints' (1991, 154).

Similarly, both Benjamin and Taussig, see hope in the ability of new technologies, or 'mimetic techniques', to restore this embodied connection between the person and the object (Taussig 1993b, 23). What is more, new mimetic techniques can transform nature. Mimesis becomes, for Taussig, 'the nature that culture uses to create second nature', with the ability 'to play with and even restore this erased sense of contact-sensuousness particularity animating the fetish … to create a quite different, secular sense of the marvellous' (1993b, 251, 23). This will allow us to rediscover 'embodied knowing' (1993b, 24). In particular, mimetic machines can restore this lost sense of contact-sensuousness. The machines create a 'new sensorium involving a new subject-object relation and therefore a new person' (1993b, 24).

We must be careful, however, as these conceptual tools can also lead to unfounded optimism regarding the potential for 'technologies' to reconfigure boundaries in ways that include, empower, and improve the lives of those who are outside of the dominant group, including those who do not participate (either through choice or access) in their use. As Taussig (1993b) concludes:

> So far, of course, history has not taken the turn Benjamin thought that mimetic machines might encourage it to take. The irony that this failure is due in good part to the very power of mimetic machinery to control the future by unleashing imageric power, on a scale previously only dreamed of. (26)

Speaking specifically about robotics, Castañeda and Suchman (2014) propose that 'robotics might rethink itself as a very different project, in ways that could not only re-articulate already existing material practices, but also suggest new lines of research and development less focused on the figure of the autonomous human, and more on infrastructures and artifacts for planetary sustainability' (335). Robotics could 'revitalize old agendas' and thereby 'discover the spaces available for resistance, intervention, and transformation' (2014, 335). They propose:

> …a more differentiated set of starting points for the robot that admit possibilities for multiple kinds of bodies and associated capacities, as well as more various cultural environments: not just the nuclear family, but social collectives that more effectively represent and challenge the many forms of relationality that exist both within and outside of the United States, Britain, and Europe; not just the normally developing child, but differently-abled bodies; and not just a limited notion of imagination, but varied forms of engagement with the world. (2014, 335)

Rather than being framed as 'importation of mind into matter', robotics could instead be concerned with 'the rematerialization of bodies and subjectivities in ways that challenge familiar assumptions about the naturalness of normative forms, primates, robots, and robot-primates might become sites for transformation rather than further reiteration.' (2014, 335).

However, as we have seen, the institutional investment in promoting the technologies of control, versus those supporting inclusion, enchantment and diversity are heavily skewed in favour of the former. The narrative of the 'future' of humanity is overwhelmingly in the hands of a narrow group of people and institutions that posit a world in which technology is the inevitable and only solution to catastrophic and existential risks to humanity. However, 'resistances, interventions and transformations' are not necessarily tied to technological changes, and technologies need not necessarily be enlisted in their service. While it is true that robotics and other technologies could potentially be enlisted to projects that promote greater equality, transformations, and inclusion, we cannot expect robotics to lead this reconfiguration. This is especially true if those fields that take humans as their core focus are failing to provide the resources necessary to do so.

As the world experiences unrivalled economic, political, and ecological challenges, we must resist the ultimate sterility of simply eradicating boundaries and instead attend to the urgent need to articulate new ones. In this dissertation, I maintain that the posthuman insistence on human-machine symmetry may in fact bolster the simultaneous animation of objects and the dehumanisation of people that is characteristic of the project of control. As an alternative, I propose that by centering the role of embodiment, perhaps these new mimetic machines might indeed, as Taussig (1993b) proposes, 'suggest other ways of being identical, other ways of being alter', thus helping us '[t]o reinvent a new world and live new fictions' (xv, xvii).

In this project, I have argued that what is needed is not the elimination or dissolution of the human, but a complete reconfiguring of the conception of 'Man' as human archetype, and new tools with which to understand ourselves and our relationship with the world. However, as Ruha Benjamin (2019a) argues, attempting to define or redefining the human and its relation to technology is a political act, contributing to what counts as human, how society is organised and their potentiality:

> After all, the larger critical project involves questioning how robot technologies revolve around the category of "human"— "who defines it, inherits it, wields it... who rents it, tills it, toils for it... who gets expelled from it, buried under it, or drowned as they risk everything to inhabit it". (9)

Any reconfiguration on the concept of the human must therefore include those who have previously been excluded from this domain. Further, as we have seen, the concept of human is inextricable from the wider world with which it is entangled. Anthropologist and environmentalist Deborah Bird Rose (2011) has proposed an 'existential ecology' to reflect 'two major shifts in worldview: the end of certainty and the end of atomism' (2). Existentialism, for Rose, implies, 'there is no predetermined essence of humanity, no ultimate goal toward which we are heading, and that we experience what appear to be astonishingly open ways of being and becoming human' (2011, 43). Ecology is necessary to denote our kinship and entanglement 'as a co-evolving species of life on Earth' (2011, 43-44). Ultimately:

> Ecological existentialism thus proposes a kinship of becoming: no telos, no deus ex machina to rescue us, no clockwork to keep us ticking along; and on the other hand, the rich plenitude, with all its joys and hazards, of our entanglement in the place, time, and multispecies complexities of life on Earth. (2011, 44)

Any reconfiguration must also include a consideration of other embodied beings with whom we share kinship and a planet. In this project, I have proposed that Merleau-Ponty's concept of embodiment, combined with a developed articulation of the fetish concept as articulated by Pietz, might serve as important resources for this collective project. As we have seen, for Merleau-Ponty (2012), by 'remaking contact with the body and with the world', we will also 'rediscover' ourselves (239). Ultimately then, I propose a conceptualisation of the human that is both relational *and* embodied.

Conclusion

In this dissertation, I asked how the human is produced in robotics research. As we have seen, by building and modifying robots, AI/Robotics researchers are reflecting, reproducing, and producing some of our foundational assumptions about what it means to be human. This includes the idea of the mechanical body and information-processing mind, within an overarching evolutionary narrative. Ultimately, this includes a view of the human as scientifically knowable. However, we have also seen the way in which robotics work disrupts this narrative, including the many different ways in which human and machines are entangled.

These findings reflect those of previous laboratory studies, which have revealed ways in which the boundary between the human and the machine is permeable and open for reconfiguration. However, these insights have contributed to the posthuman dissolution of the human and, as we have seen, contemporary anthropology is left with no figure of the human as a point of departure. I have argued that, as well as leading to a gap in anthropological knowledge, this has also led to anthropologists missing vital tools necessary to engage with other disciplines in addressing contemporary economic, political, and ecological challenges. In my research, an alternative configuration is emerging, in which, despite the many boundary breakdowns, between nature and culture, society and technology, the living body continues to resist translation into a problem of coding, instrumental control, and market exchange.

In this study, I developed a novel conceptual framework drawing on both the concept of the 'fetish' as articulated by Pietz, and Merleau-Ponty's embodiment phenomenology. I used this framework both to investigate human-object relations, revealing the blurred nature of the boundary between them but also to propose an approach for their reconfiguration. The concept of the fetish facilitated an investigation of the robot, and thus wider human-object relations, along four dimensions: 'historicisation', 'territorialisation', 'reification' and 'personalisation'. Through the lens of 'historicisation', I revealed the continuous and discontinuous past and futures of the robot present, revealing how the past is less linear and the future less pre-determined than dominant narratives would suggest. Through this lens, I traced the new events that are precipitated by the realisation of the robot in enduring material form, including the narratives and the futures that are implicated in its identity, as well as the implications for policy and regulation.

The lens of 'territorialisation' focused on the irreducible materiality of the robot, including how its resolute physicality forces the roboticist into a concrete confrontation with world, as well as the fundamental ways it resists human intentions. It also reveals the gap between the reality of the robot capabilities and the fantastical discourses. Through this lens, I also focused on the space of robotics work. This revealed the geographical specificity of individual robotics labs, while also revealing a community of researchers united by common histories, identities, and mythologies. Finally, through this lens, I also traced a wider landscape of institutions dedicated to extending the image of the intelligent robot as the super-human evolutionary successor.

The next chapter focused on the 'reification' aspect of the fetish. Here, I explored the constructed value of the robot as an autonomous, intelligent being, and the Euro-American cultural identities that are implicated in that vision, rooted in colonialism and empire. I then analysed the ways in which the performance of robot animacy obscures the real work that is necessary to make the technologies appear intelligent and autonomous. However, through this lens, I also found that robot animacy results from the creative performance and craft of its creators, and how it can be used to play, to enchant, to learn and to rebalance power dynamics.

Finally, the lens of 'personalisation' revealed the visceral nature of the human-robot encounter, as well as novel entanglements between the human and the object. By taking an explicitly Merleau-Pontian reading of embodiment, this chapter drew together these themes by exploring how the self and the Other are constructed. It also revealed the human as a fundamentally embodied being, situated in an intersubjective world.

Pietz's fetish 'quadralectic' thus allows for an integration of a 'chiasmic' ontology with one that recognises the historical and constructed nature of the fetish object, its social and symbolic value, as well as its irreducible materiality. As we have traced above, each of the four aspects of the fetish has revealed a distinct, yet essential, part of our relationship with robots, revealing more broadly ways in which humans and objects are inextricably entangled and co-constituted, yet not symmetrical. Thus, I have shown that while the project of robotics is predominantly viewed as a literal attempt to produce a human, and the ensuing shifting boundary may be conceived as a dissolution of the human, it is in fact through fetishistic relations, animated by the living body, that the human is produced.

In this project, I contend that understanding the human as simultaneously embodied and relational is critical for both robotics and anthropology. In the current literature, the potentially revolutionary concept of relationality is extracted from its embodied situationality, losing its body and leaving the traditional conceptual categories undisturbed. Deviating from the relational ontology dominant in the anthropology of technology, I discovered an enduring asymmetry between the human and the robot, with the living body emerging as a durable category that cannot be reasoned away. In its holistic, embodied sense, it has so far eluded attempts at artificiality and mechanisation. By combining ethnographic description with the concept of the 'fetish' and Merleau-Ponty's embodiment phenomenology, this study reveals a view of the human open to the world, enmeshed in an empathic connection with other living things in it. It demonstrates the creativity and levity in the stories and interactions between people, as well as drawing attention to the labour, craft, and love involved in robot design and development. Observations of interactions between humans and robots have revealed unexpected insights that challenge both the Enlightenment model of 'Man', as well as the symmetrical view. This includes, for example, the tendency for moving machines to evoke a visceral and pre-discursive empathy in those that encounter them, regardless of their inner computational abilities. In contrast to dominant narratives of competition and 'selfish genes', people feel a primal kinship with the robot;

it is more than just a technology or a tool. It is in these encounters that the robot can be said to be ‘animate’. Ultimately, this novel framework can thus not only support the querying the boundaries, but also drawing new ones.

By reconfiguring the human as an embodied, situated, and sensuous being, who is nonetheless continuous with, and kin, with the rest of the world, we can challenge the liberal human subject while simultaneously retaining the necessary and productive category of the human. This will not just allow us to theorise about the nature of the human, but also about the role of the human as moral agent, and also for a reflexive consideration of ourselves as anthropologists. Additionally, it allows for an analysis of human relationships with technologies, which are neither utopian nor dystopian, but pragmatic.

Technologies have transformed our physical and social landscapes; they have ‘recrafted our bodies’ by extending our reach in the world, and how we interact with others, they have even more markedly affected our political, social and environmental realities. However, on a non-symbolic level, our experiencing bodies continue to share a greater commonality with all other embodied, living beings, than with machines. By engaging both Merleau-Ponty’s embodiment philosophy and Pietz’s fetish concept, we can contribute to a positive reconfiguration of a situated, embodied and creative human in sympathetic relation with the world as an alternative to dissolving or decentering the human in anthropological theory.

References

Acemoglu, Daron, and Pascual Restrepo. 2017. 'Robots and Jobs: Evidence from US Labor Markets'. Working Paper 23285. Working Paper Series. National Bureau of Economic Research.

Ackerman, Emily. 2019. 'i (in a wheelchair) was just trapped on forbes ave by one of these robots…' *@EmilyEAckerman*. Twitter, October 21, 2019. https://twitter.com/EmilyEAckerman/status/1186363305851576321.

Adam, Alison. 1998. *Artificial Knowing: Gender and the Thinking Machine*. London; New York: Routledge.

Adams, Douglas. 1979. *The Hitch Hiker's Guide to the Galaxy*. Pan Original. London: Pan Books.

Agre, Philip. 1997. *Computation and Human Experience*. Cambridge; New York: Cambridge University Press.

Agre, Philip, and David Chapman. 1987. 'Pengi: An Implementation of a Theory of Activity'. *Proceedings of the Sixth National Conference on Artificial Intelligence* 278: 268–72.

Andes, Scott, Mitch Horowitz, Ryan Helwig, and Bruce Katz. 2017. 'Capturing the next Economy: Pittsburgh's Rise as a Global Innovation City'. Bass Initiative on Innovation and Placemaking. Brookings Institution.

Appadurai, Arjun, ed. 1986. *The Social Life of Things: Commodities in Cultural Perspective*. Cambridge [Cambridgeshire]; New York: Cambridge University Press.

Apter, Emily S., and William Pietz, eds. 1993. *Fetishism as Cultural Discourse*. Ithaca, N.Y: Cornell University Press.

Aristotle. 2000. *Politics*. Translated by Benjamin Jowett. Dover Thrift Editions. Mineola, NY: Dover Publications.

ASC Cybernetics. 2021a. 'ASC: Foundations: History: Timeline'. American Society for Cybernetics. Accessed September 8, 2021. https://asc-cybernetics.org/foundations/timeline.htm.

———. 2021b. ‘Summary: The Macy Conferences’. American Society for Cybernetics. Accessed 27 July 2021. https://asc-cybernetics.org/foundations/history/MacySummary.htm.

Asimov, Isaac. 1950. *I, Robot*. Garden City, New York: Doubleday.

Atanasoski, Neda, and Kalindi Vora. 2015. ‘Surrogate Humanity: Posthuman Networks and the (Racialized) Obsolescence of Labor’. *Catalyst: Feminism, Theory, Technoscience* 1 (1): 1–40.

———. 2019. *Surrogate Humanity: Race, Robots, and the Politics of Technological Futures*. Perverse Modernities. Durham: Duke University Press.

Atran, Scott. 2004. *In Gods We Trust: The Evolutionary Landscape of Religion: The Evolutionary Landscape of Religion*. Oxford: Oxford University Press.

Babbage, Charles. 1832. *On the Economy of Machinery and Manufactures*. Third Edition. London: Charles Knight.

Barad, Karen. 2003. ‘Posthumanist Performativity: Toward an Understanding of How Matter Comes to Matter’. *Signs: Journal of Women in Culture and Society* 28 (3): 801–31.

Baron-Cohen, Simon. 1991. ‘Precursors to a Theory of Mind: Understanding Attention in Others’. *Natural Theories of Mind: Evolution, Development and Simulation of Everyday Mindreading* 1: 233–51.

Benjamin, Ruha. 2019a. ‘Discriminating Design, Liberating Imagination’. In *Captivating Technology: Race, Carceral Technoscience, and Liberatory Imagination in Everyday Life*, edited by Ruha Benjamin (page numbers omitted). Durham: Duke University Press.

———. 2019b. *Race after Technology: Abolitionist Tools for the New Jim Code*. Medford, MA: Polity.

Bennett, Jane. 2010. *Vibrant Matter: A Political Ecology of Things*. Durham: Duke University Press.

Beschizza, Rob. 2015. ‘Google’s Deep Dreaming Fills the ’net with Psychedelic Nightmare GIFs’. *Boing*, July 7, 2015. https://boingboing.net/2015/07/07/googles-deep-dreaming-fills.html.

Bloch, Maurice. 1991. 'Language, Anthropology and Cognitive Science'. *Man* 26 (2): 183–19817.

———. 1998. *How We Think They Think: Anthropological Approaches to Cognition, Memory, and Literacy*. Boulder, Colo.: Westview Press.

Boas, Franz. 1955. *Primitive Art*. New York: Dover Publications.

Boden, Margaret A. 2004. *The Creative Mind: Myths and Mechanisms, Second Edition*.

———. 2006. *Mind as Machine: A History of Cognitive Science*. Oxford: New York: Clarendon Press; Oxford University Press.

Boden, Margaret A., Joanna Bryson, Darwin Caldwell, Kerstin Dautenhahn, Lilian Edwards, Sarah Kember, Paul Newman, et al. 2017. 'Principles of Robotics: Regulating Robots in the Real World'. *Connection Science* 29 (2): 124–29.

Boston Dynamics. 2020. *Boston Dynamics Robots*. Photograph. Accessed July 17, 2020. https://www.bostondynamics.com/products

———. 2021. 'Spot® - The Agile Mobile Robot'. 2021. Boston Dynamics. Accessed November 10, 2021. https://www.bostondynamics.com/products/spot.

Bostrom, Nick. 2002. 'Existential Risks: Analyzing Human Extinction Scenarios and Related Hazards'. *Journal of Evolution and Technology* 9: 1-30

———. 2005. 'A History of Transhumanist Thought'. *Journal of Evolution and Technology* 14 (1): 1-25

———. 2013. 'Existential Risk Prevention as Global Priority'. *Global Policy* 4 (1): 15–31.

Botsolvers. 2021. *First Humanoid Robot – Herbert Televox*. Photograph. https://botsolvers.com/first-humanoid-robot-herbert-televox/.

Boyer, George R. 1998. 'The Historical Background of the Communist Manifesto'. *Journal of Economic Perspectives* 12 (4): 151–74.

Boyer, Pascal. 1994. *The Naturalness of Religious Ideas: A Cognitive Theory of Religion*. Berkeley: University of California Press.

Boyer, Pascal. 1996. ‘What Makes Anthropomorphism Natural: Intuitive Ontology and Cultural Representations’. *The Journal of the Royal Anthropological Institute* 2 (1): 83-97.

Braune, Sean. 2020. ‘Fetish-Oriented Ontology’. *Open Philosophy* 3 (1): 298–313.

Breazeal, Cynthia L. 2002. *Designing Sociable Robots*. Cambridge, Massachusetts: The MIT Press.

Brockman, John, ed. 2019. *Possible Minds: Twenty-Five Ways of Looking at AI*. New York: Penguin Press.

Brooks, Rodney. 1990. ‘Elephants Don’t Play Chess’. *Robotics and Autonomous Systems* 6 (1–2): 3–15.

———. 1991a. ‘Intelligence Without Reason’. *Computers and Thought, IJCAI-91* AI Memo No. 1293

———. 1991b. ‘Intelligence Without Representation’. *Artificial Intelligence* 47: 139-159.

———. 2002. *Flesh and Machines. How Robots Will Change Us*. New York: Vintage Books.

———. 2018. ‘A Brave, Creative, and Happy HRI’. *ACM Transactions on Human-Robot Interaction* 7 (1): 1–3.

Brooks, Rodney, Cynthia Breazeal, Matthew Marjanović, Brian Scassellati, and Matthew M. Williamson. 1999. ‘The Cog Project: Building a Humanoid Robot’. In *Computation for Metaphors, Analogy, and Agents*, edited by Chrystopher L. Nehaniv, 1562:52–87. Lecture Notes in Computer Science. Berlin, Heidelberg: Springer Berlin Heidelberg.

Bryson, Joanna J. 2010. ‘Robots Should Be Slaves’. In *Natural Language Processing*, edited by Yorick Wilks, 8:63–74. Amsterdam: John Benjamins Publishing Company.

Bullard, Benjamin. 2020. ‘OpenAI Robot Writes Dystopian Essay about Sparing Humanity (Maybe) from Machine Takeover’. *SYFY WIR*, September 10, 2020. https://www.syfy.com/syfywire/robot-author-writes-dystopian-essay-sparing-humanity.

Burke, Elaine. 2015. 'A Look Back at the Disturbing World of Google DeepDream'. *Silicon Republic*, December 14, 2015. https://www.siliconrepublic.com/trending/google-deepdream-artificial-intelligence-images.

Burke, J.L., R.R. Murphy, E. Rogers, V.J. Lumelsky, and J. Scholtz. 2004. 'Final Report for the DARPA/NSF Interdisciplinary Study on Human–Robot Interaction'. *IEEE Transactions on Systems, Man and Cybernetics, Part C (Applications and Reviews)* 34 (2): 103–12.

Butler, Judith. 2011. *Bodies That Matter: On the Discursive Limits of 'Sex'*. Routledge Classics. Abingdon, Oxon; New York, NY: Routledge.

Callon, Michel. 1986. 'The Sociology of an Actor-Network: The Case of the Electric Vehicle'. In *Mapping the Dynamics of Science and Technology: Sociology of Science in the Real World*, edited by Michel Callon, John Law, and Arie Rip. Basingstoke: The Macmillan Press Ltd.

Callon, Michel, and Bruno Latour. 1992. 'Don't Throw the Baby Out with the Bath School! A Reply to Collins and Yearley'. In *Science as Practice and Culture*, edited by Andrew Pickering, 343-368. University of Chicago Press.

Čapek, Karel. 2001. *R.U.R. (Rossum's Universal Robots)*. Dover Thrift Editions. Mineola, N.Y: Dover Publications.

Carpenter, Julie. 2013. 'The Quiet Professional: An Investigation of U.S. Military Explosive Ordnance Disposal Personnel Interactions with Everyday Field Robots'. PhD diss., University of Washington.

Castañeda, Claudia, and Lucy Suchman. 2014. 'Robot Visions'. *Social Studies of Science* 44 (3): 315–41.

CHAI. 2021. 'Center for Human-Compatible Artificial Intelligence – Center for Human-Compatible AI Is Building Exceptional AI for Humanity'. Accessed December 12, 2021. https://humancompatible.ai/.

Chasin, Alexandra. 1995. 'Class and Its Close Relations: Identities among Women, Servants, and Machines'. In *Posthuman Bodies*, edited by

Judith Halberstam and Ira Livingston. Unnatural Acts. Bloomington: Indiana University Press.

Coalition for Critical Technology. 2018. ‘Abolish the #TechToPrisonPipeline’. Open Letter, September 21, 2018. https://medium.com/@CoalitionForCriticalTechnology/abolish-the-techtoprisonpipeline-9b5b14366b16.

Coeckelbergh, Mark. 2018. ‘Why Care About Robots? Empathy, Moral Standing, and the Language of Suffering’. *Kairos. Journal of Philosophy & Science* 20 (1): 141–58.

Conty, Arianne Françoise. 2018. ‘The Politics of Nature: New Materialist Responses to the Anthropocene’. *Theory, Culture & Society* 35, no. 7–8 (December 2018): 73–96.

Cowan, Ruth Schwartz. 1983. *More Work for Mother: The Ironies of Household Technology from the Open Hearth to the Microwave*. New York, NY: Basic Books.

Crevier, Daniel. 1993. *AI: The Tumultuous History of the Search for Artificial Intelligence*. New York, NY: Basic Books.

CSAIL. 2021. ‘Cog Project Overview’. Accessed September 8, 2021. http://groups.csail.mit.edu/lbr/humanoid-robotics-group/cog/overview.html.

CSER. 2021. ‘Centre for the Study of Existential Risk’. Centre for the Study of Existential Risk. Accessed November 13, 2021. https://www.cser.ac.uk/.

Csordas, Thomas J. 1990. ‘Embodiment as a Paradigm for Anthropology’. *Ethos* 18 (1): 5–47.

Cushing, Ellen. 2012. ‘Amazon Mechanical Turk: The Digital Sweatshop’. *Utne.Com* (blog). https://www.utne.com/science-and-technology/amazon-mechanical-turk-zm0z13jfzlin/.

Darling, Kate. 2016. ‘Extending Legal Protection to Social Robots: The Effects of Anthropomorphism, Empathy, and Violent Behavior towards Robotic Objects’. In *Robot Law*, edited by Ryan Calo, A. Froomkin, and Ian Kerr, 213–32. Cheltenham; Edward Elgar Publishing.

———. 2017. '"Who's Johnny?" Anthropomorphic Framing in Human-Robot Interaction, Integration, and Policy'. In *Robot Ethics 2.0*, edited by Patrick Lin, Keith Abney, and Ryan Jenkins, 173-192. Oxford, New York: Oxford University Press.

DARPA. 2021. 'Shakey the Robot'. 2021. U.S. Department of Defense Website. Accessed September 11, 2021. https://www.darpa.mil/about-us/timeline/shakey-the-robot.

Degeler, Andrii. 2015. 'Marines' LS3 Robotic Mule Is Too Loud for Real-World Combat'. *Ars Technica*, December 29, 2015. https://arstechnica.com/information-technology/2015/12/us-militarys-ls3-robotic-mule-deemed-too-loud-for-real-world-combat/.

Deleuze, Gilles. 1994. *Difference and Repetition*. New York: Columbia University Press.

Deleuze, Gilles, and Félix Guattari. 1987. *A Thousand Plateaus: Capitalism and Schizophrenia*. Minneapolis: University of Minnesota Press.

Dennett, D. C. 1998. *Kinds of Minds: Toward an Understanding of Consciousness*. 3rd pr. New York: Basic Books.

Desjarlais, Robert, and Jason C. Throop. 2011. 'Phenomenological Approaches in Anthropology'. *Annual Review of Anthropology* 40 (1): 87–102.

Douglas, Mary. 2003. *Purity and Danger: An Analysis of Concepts of Pollution and Taboo*. Mary Douglas Collected Works 2. London: Routledge.

Downey, Gary Lee. 1998. *The Machine in Me: An Anthropologist Sits among Computer Engineers*. New York: Routledge.

Downey, Gary Lee, Joseph Dumit, and Sarah Williams. 1995. 'Cyborg Anthropology'. *Cultural Anthropology* 10 (2): 264–69.

Downey, Greg. 2007. 'Seeing with a "Sideways Glance": Visuomotor "Knowing" and the Plasticity of Perception.' *Ways of Knowing: New Approaches in the Anthropology of Experience and Learning*, 222–41.

Dreyfus, Hubert. 1965. 'Alchemy and Artificial Intelligence.' Santa Monica Calif: RAND Corp.

———. 1992. *What Computers Still Can’t Do: A Critique of Artificial Reason*. Cambridge, Mass.: The MIT Press.

———. 2007. ‘Why Heideggerian AI Failed and How Fixing It Would Require Making It More Heideggerian’. *Philosophical Psychology* 20 (2): 247–68.

Dreyfus, Hubert, and Stuart Dreyfus. 2005. ‘Peripheral Vision: Expertise in Real World Contexts’. *Organization Studies* 26 (5): 779–92.

———. 2009. *Mind over Machine: The Power of Human Intuition and Expertise in the Era of the Computer*. Repr. New York: The Free Press.

Duffy, Brian R. 2003. ‘Anthropomorphism and the Social Robot’. *Robotics and Autonomous Systems* 42 (3–4): 177–90.

Ekbia, Hamid, and Bonnie Nardi. 2014. ‘Heteromation and Its (Dis)Contents: The Invisible Division of Labor between Humans and Machines’. Online Article. *First Monday,* 2014.

Elish, M. C., and danah boyd. 2018. ‘Situating Methods in the Magic of Big Data and AI’. *Communication Monographs* 85 (1): 57–80.

Elish, M C, and Tim Hwang. 2016. *An AI Pattern Language*. 1st edition. Intelligence and Autonomy Initiative (I&A). New York, NY: CreateSpace Independent Publishing Platform.

Eubanks, Virginia. 2019. *Automating Inequality: How High-Tech Tools Profile, Police, and Punish the Poor*. First Picador edition. New York: Picador St. Martin’s Press.

Feenberg, Andrew. 2014. ‘The Many Natures of Philippe Descola, Reflections on *The Ecology of Others*’. *Science as Culture* 23 (2): 277–82.

FHI. 2012. ‘Future of Humanity Institute’. Accessed November 13, 2021. http://www.fhi.ox.ac.uk/.

Fjelland, Ragnar. 2020. ‘Why General Artificial Intelligence Will Not Be Realized’. *Humanities and Social Sciences Communications* 7, no. 1 (2020): 1-9.

FLI. 2021.‘Futureoflife.Org’. 2021. Future of Life Institute. Accessed November 13, 2021. https://futureoflife.org/new-homepage/.

Ford, Martin. 2018. *Architects of Intelligence: The Truth about AI from the People Building It*. First published: November 2018. Birmingham, UK: Packt Publishing.

Foucault, Michel. 1977. 'Nietzsche, Genealogy, History'. In *Language, Counter-Memory, Practice: Selected Essays and Interviews*, edited by D.F. Bouchard, 139-164. New York: Cornell University Press.

———. 1980. *Power/Knowledge: Selected Interviews and Other Writings*. Edited by Colin Gordon. Vol. 23. New York: Pantheon Books.

———. 1990. *The History of Sexuality*. Vintage Books ed. New York: Vintage Books.

———. 2002. *Archaeology of Knowledge*. London: Routledge Classics; New York: Routledge.

———. 2007. *Madness and Civilization: A History of Insanity in the Age of Reason*. Routledge Classics. London: Routledge.

———. 2010. *The Birth of Biopolitics: Lectures at the Collège de France, 1978-79*. Edited by Michel Senellart. Paperback ed. Basingstoke: Palgrave Macmillan.

———. 2019. 'Texts/Contexts: Of Other Spaces'. In *Grasping the World*, edited by Donald Preziosi and Claire Farago, 1st ed., 371–79. London & New York: Routledge.

———. 2020. *Discipline and Punish: The Birth of the Prison*. Translated by Alan Sheridan. London: Penguin Classics.

Fowles, Severin. 2016. 'The Perfect Subject (Postcolonial Object Studies)'. *Journal of Material Culture* 21 (1): 9–27.

Frude, Neil, and Petar Jandrić. 2015. 'The Intimate Machine – 30 Years On'. *E-Learning and Digital Media* 12 (3–4): 410–24.

Gallagher, Shaun, and Dan Zahavi. 2010. *The Phenomenological Mind: An Introduction to Philosophy of Mind and Cognitive Science*. London: Routledge.

Garreau, Joel. 2007. 'Bots on The Ground'. *The Washington Post Sunday*. May 6, 2007. https://www.pressreader.com/usa/the-washington-post-sunday/20070506/282192236550791

Geertz, Clifford. 1972. ‘The Wet and the Dry: Traditional Irrigation in Bali and Morocco’. *Human Ecology* 1 (1): 23–39.

———. 1973. *The Interpretation of Cultures: Selected Essays*. New York: Basic Books.

Gell, Alfred. 1992. ‘The Technology of Enchantment and the Enchantment of Technology’. In *Anthropology, Art and Aesthetics*, edited by J Coote and A Shelton, 40–66. Oxford: Clarendon.

———. 1998. *Art and Agency: An Anthropological Theory*. Oxford: Clarendon Press.

Gill, Martha. 2019. ‘Novacene by James Lovelock Review — Don’t Worry, the Robots Might Not Kill Us’. *The Times Saturday Review*, June 21, 2019. https://www.thetimes.co.uk/article/novacene-by-james-lovelock-review-m0z39ptcw.

Goody, Jack. 1973. ‘Evolution and Communication: The Domestication of the Savage Mind’. *The British Journal of Sociology* 24 (1): 1-12.

———. 1977. *The Domestication of the Savage Mind*. Themes in the Social Sciences. Cambridge [Eng.]; New York: Cambridge University Press.

Gordon, Lewis R. 1998. ‘African-American Philosophy: Theory, Politics, and Pedagogy’. *Philosophy of Education Archive*. (1998): 39-46.

Grace, Katja, John Salvatier, Allan Dafoe, Baobao Zhang, and Owain Evans. 2018. ‘When Will AI Exceed Human Performance? Evidence from AI Experts’. *ArXiv:1705.08807 [Cs]*, May. http://arxiv.org/abs/1705.08807.

Graeber, David. 2005. ‘Fetishism as Social Creativity: Or, Fetishes Are Gods in the Process of Construction’. *Anthropological Theory* 5 (4): 407–38.

Gray, Mary L., and Siddharth Suri. 2019. *Ghost Work: How to Stop Silicon Valley from Building a New Global Underclass*. Boston New York NY: Houghton Mifflin Harcourt.

Guizzo, Erico. 2010. ‘Geminoid F Looks More Realistic Than Ever’. *IEEE Spectrum*, November 2, 2010. https://spectrum.ieee.org/geminoid-f-looks-more-realistic-than-ever.

Gunkel, David J. 2018. ‘The Other Question: Can and Should Robots Have Rights?’ *Ethics and Information Technology* 20 (2): 87–99.

Guthrie, Stewart. 1993. *Faces in the Clouds: A New Theory of Religion*. New York: Oxford University Press.

Gutkind, Lee. 2009. *Almost Human: Making Robots Think*. New York; London: W.W. Norton.

Haraway, Donna. 1991. *Simians, Cyborgs, and Women: The Reinvention of Nature*. New York: Routledge.

Hayles, N. Katherine. 1999. *How We Became Posthuman: Virtual Bodies in Cybernetics, Literature, and Informatics*. Chicago, Ill: University of Chicago Press.

Hays, Jeffrey. 2013. ‘Robots in Japan: History, Industrial Uses and Security’. Facts and Details. Last updated January 2013. Accessed July 27, 2021. https://factsanddetails.com/Japan/cat26/sub163/item873.html.

Heider, Fritz, and Marianne Simmel. 1944. ‘An Experimental Study of Apparent Behavior’. *The American Journal of Psychology* 57 (2): 243-259.

Heiss, Jan Patrick, and Albert Piette. 2015. ‘Individuals in Anthropology’. *Zeitschrift Für Ethnologie* 140 (1): 5-17.

Helmreich, Stefan. 2000. *Silicon Second Nature: Culturing Artificial Life in a Digital World*. Updated edition with a new preface, First paperback printing. Berkeley Los Angeles London: University of California Press.

Henare, Amiria. 2003. ‘Artefacts In Theory: Anthropology And Material Culture’. *The Cambridge Journal of Anthropology* 23 (2): 54–66.

Hicks, Dan. 2010. ‘The Material-Cultural Turn: Event and Effect.’ In *The Oxford Handbook of Material Culture Studies*, edited by Dan Hicks and Mary C. Beaudry, 1:25–98. Oxford: Oxford University Press.

Hirose, Masato, and Kenichi Ogawa. 2007. ‘Honda Humanoid Robots Development’. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* 365 (1850): 11–19.

History Computer. 2021. ‘The History of Elektro of Westinghouse - History Computer’. Accessed November 9, 2021. https://history-computer.com/the-history-of-elektro-of-westinghouse/.

Hitachi. 2021. ‘Story: Hitachi Origin Story: Hitachi Global’. Accessed November 18, 2021. https://www.hitachi.com/corporate/origin/story.html?p=emiew_1985.

Honda Global. 2000. *ASIMO*. Photograph. https://global.honda/newsroom/news/2000/c001120b-eng.html.

Hornborg, Alf. 2001. ‘Symbolic Technologies: Machines and the Marxian Notion of Fetishism’. *Anthropological Theory* 1 (4): 473–96.

———. 2006. ‘Animism, Fetishism, and Objectivism as Strategies for Knowing (or Not Knowing) the World’. *Ethnos* 71 (1): 21–32.

———. 2014. ‘Technology as Fetish: Marx, Latour, and the Cultural Foundations of Capitalism’. *Theory, Culture & Society* 31 (4): 119–40.

Howell, Signe. 2016. ‘The Relationality of Species in Chewong Animistic Ontology’. In *Critical Anthropological Engagements in Human Alterity and Difference,* edited by Bjørn Enge Bertelsen and Synnøve Bendixsen, 43-64. Cham: Springer International Publishing.

Huyssen, Andreas. 1986. *After the Great Divide: Modernism, Mass Culture, Postmodernism*. Theories of Representation and Difference. Bloomington: Indiana University Press.

IEEE Standards Association. 2016. ‘Ethically Aligned Design’. Version 1 - For Public Discussion. The IEEE Global Initiative on Ethics of Autonomous and Intelligent Systems.

Ieropoulos, Ioannis A., Pablo Ledezma, Andrew Stinchcombe, George Papaharalabos, Chris Melhuish, and John Greenman. 2013. ‘Waste to Real Energy: The First MFC Powered Mobile Phone’. *Physical Chemistry Chemical Physics* 15 (37): 15312-15316.

Ingold, Tim. 1997. ‘Eight Themes in The Anthropology of Technology’. *Social Analysis: The International Journal of Anthropology* 41 (1): 106–38.

———. 2000. *The Perception of the Environment*. London: Routledge.

———. 2002. ‘Culture and the Perception of the Environment’. In *Bush Base, Forest Farm: Culture, Environment, and Development*, edited by Elisabeth Croll and David Parkin, 39-56. London; New York: Routledge.

———. 2011. *Being Alive: Essays on Movement, Knowledge and Description*. London; New York: Routledge.

———. 2013. *Making: Anthropology, Archaeology, Art and Architecture*. London; New York: Routledge.

insider.co.uk. 2018. ‘Robot “hired” by Edinburgh Supermarket Then “Fired” after a Week’. *Insider.Co.Uk*, 22 January, 2018. https://www.insider.co.uk/news/robot-hired-edinburgh-supermarket-fired-11892140.

‘Interview with Marshall Sahlins’. 2008. *Anthropological Theory* 8 (3): 319–28.

Jackson, Michael. 1983. ‘Knowledge of the Body’. *Man* 18 (2): 327–45.

———. 2002. ‘Familiar and Foreign Bodies: A Phenomenological Exploration of the Human-Technology Interface’. *The Journal of the Royal Anthropological Institute* 8 (2): 333–46.

———. 2005. *Existential Anthropology: Events, Exigencies, and Effects*. New York: Berghahn Books.

Jackson, Michael, and Albert Piette. 2015. ‘Anthropology and the Existential Turn’. In *Existential Anthropology*. New York: Berghan Books.

Jackson, Zakiyyah Iman. 2013. ‘Review: Animal: New Directions in the Theorization of Race and Posthumanism’. *Feminist Studies* 39 (3): 669–85.

Johnson-Laird, Philip Nicholas. 1988. *The Computer and the Mind: An Introduction to Cognitive Science*. Cambridge, Mass: Harvard Univ. Press.

Junglist. 2015. ‘How To Create Mind-Melting Videos With Google’s DeepDream Neural Network’. *Lifehacker Australia*, July 7, 2015. https://www.lifehacker.com.au/2015/07/how-to-create-mind-melting-videos-with-googles-deepdream-neural-network/.

Kaempffert, Waldemar. 1927. ‘Science Produces the “Electrical Man”’. *The New York Times*, October 23, 1927.

Kang, Cecilia. 2017. ‘Pittsburgh Welcomed Uber’s Driverless Car Experiment. Not Anymore.’ *The New York Times*, May 21, 2017. https://www.nytimes.com/2017/05/21/technology/pittsburgh-ubers-driverless-car-experiment.html

Katz, Yarden. 2017. ‘Manufacturing an Artificial Intelligence Revolution’. *SSRN*. Electronic Journal. https://ssrn.com/abstract=3078224

Kember, Sarah. 2003. *Cyberfeminism and Artificial Life*. London; New York: Routledge.

Khomami, Nadia. 2014. ‘2029: The Year When Robots Will Have the Power to Outsmart Their Makers’. *The Observer Technology*, February 22, 2014. https://www.theguardian.com/technology/2014/feb/22/computers-cleverer-than-humans-15-years.

King, Elizabeth. 2007. ‘Perpetual Devotion: A Sixteenth-Century Machine’. In *Genesis Redux: Essays in the History and Philosophy of Artificial Life*, edited by Jessica Riskin, 263–92. Chicago: University of Chicago Press.

Knorr-Cetina, K. D. 2013. *The Manufacture of Knowledge: An Essay on the Constructivist and Contextual Nature of Science*. Elsevier.

Knorr-Cetina, K. D. 2014. *The Manufacture of Knowledge: An Essay on the Constructivist and Contextual Nature of Science.* Kent: Elsevier Science.

Kolb, Michael. 2012. ‘Soldier And Robot Interaction In Combat Environments’. PhD diss., University of Oklahoma.

Kuipers, Benjamin, Edward A. Feigenbaum, Peter E. Hart, and Nils J. Nilsson. 2017. ‘Shakey: From Conception to History’. *AI Magazine* 38 (1): 88–103.

Kurzweil, Ray. 2005. *The Singularity Is Near: When Humans Transcend Biology*. New York: Viking.

———. 2014. *The Age of Spiritual Machines: When Computers Exceed Human Intelligence*. New York: Penguin Books.

Latour, Bruno. 1983. 'Give Me a Laboratory and I Will Raise the World'. In *Science Observed: Perspectives on the Social Study of Science*, edited by Karin Knorr-Cetina and Michael Mulkay, 141–70. London and Beverly Hills: Sage.

———. 1987. *Science in Action*. Cambridge, Mass.: Harvard University Press.

———. 1993. *We Have Never Been Modern*. Cambridge, Mass: Harvard University Press.

———. 1999. *Pandora's Hope: Essays on the Reality of Science Studies*. Cambridge, Mass: Harvard University Press.

———. 2002. 'Morality and Technology: The End of the Means'. *Theory, Culture & Society* 19 (5–6): 247–60.

———. 2004. 'Why Has Critique Run out of Steam? From Matters of Fact to Matters of Concern'. *Critical Inquiry* 30 (2): 225–48.

———. 2005. *Reassembling the Social*. Oxford: Oxford University Press.

———. 2010. *On the Modern Cult of the Factish Gods*. Science and Cultural Theory. Durham [NC]; London: Duke University Press.

Latour, Bruno, and Steve Woolgar. 1979. *Laboratory Life: The Social Construction of Scientific Facts*. Sage Library of Social Research; v. 80. Beverly Hills: Sage Publications.

———. 1986. *Laboratory Life: The Construction of Scientific Facts*. Princeton, N.J: Princeton University Press.

Law, John. 1987. 'Technology and Heterogeneous Engineering: The Case of Portuguese Expansion'. In *The Social Construction of Technological Systems: New Directions in the Sociology and History of Technology*, edited by Wiebe E. Bijker, Thomas P. Hughes, and Trevor Pinch, 111–34. Cambridge, Mass.: The MIT Press.

Leach, Edmund. 1959. 'Hydraulic Society in Ceylon'. *Past & Present* 15 (1): 2–26.

———. 2001. ‘Levels of Communication and Problems of Taboo’. In *The Essential Edmund Leach: Volume 1: Anthropology and Society*, edited by Stephen Hugh-Jones and James Laidlaw, 226–37. New Haven: Yale University Press.

LeCun, Yann, Yoshua Bengio, and Geoffrey Hinton. 2015. ‘Deep Learning’. *Nature* 521 (7553): 436–44.

Lefebvre, Henri. 1991. *The Production of Space*. Oxford, OX, UK; Cambridge, Mass., USA: Blackwell.

Leopold, Todd. 2015. ‘HitchBOT, Hitchhiking Robot, Gets Beheaded in Philly’. *CNN*, August 4, 2015. https://edition.cnn.com/2015/08/03/us/hitchbot-robot-beheaded-philadelphia-feat/index.html.

Letizia, Anthony. 2013. ‘The Robots of Westinghouse’. Geekfrontiers.com. Accessed July 27, 2021. https://geekfrontiers.com/pittsburgh-history/the-robots-of-westinghouse/.

Levinas, Emmanuel. 1994. ‘Intersubjectivity: Notes on Merleau-Ponty’. In *Outside the Subject*. Stanford: Stanford University Press.

———. 1999. *Otherwise Than Being, or, Beyond Essence*. Translated by Alphonso Lingis. Pittsburgh, Pa: Duquesne University Press.

———. 2002a. ‘Beyond Intentionality’. In *The Phenomenology Reader*, edited by Dermot Moran and Timothy Mooney, 529–39. London; New York: Routledge.

———. 2002b. ‘Ethics and the Face’. In *The Phenomenology Reader*, edited by Dermot Moran and Timothy Mooney, 515–28. London; New York: Routledge.

Levi-Strauss, Claude. 1966. *The Savage Mind*. Chicago: The University Of Chicago Press.

———. 1988. *The Way of the Masks*. Seattle: University of Washington Press.

Lighthill, James. 1973. ‘Lighthill Report: Artificial Intelligence: A Paper Symposium’. UK: The Science Research Council.

Lueck, Bryan. 2012. ‘Alterity in Merleau-Ponty’s Prose of the World’. *Epoché: A Journal for the History of Philosophy* 16.2 (2012): 425-442.

Lukács, Georg. 2013. *History and Class Consciousness: Studies in Marxist Dialects*. Translated by Rodney Livingstone. Nachdr. Cambridge, Mass: MIT Press.

Lynch, Michael. 1985. *Art and Artifact in Laboratory Science: A Study of Shop Work and Shop Talk in a Research Laboratory*. Studies in Ethnomethodology. London; Boston: Routledge & Kegan Paul.

Maher, Kris. 2019. ‘Pittsburgh Air-Quality Problem Recalls the Bad Old Days’. *The Wall Street Journal*, March 12, 2019. https://www.wsj.com/articles/pittsburgh-air-quality-problem-recalls-the-bad-old-days-11552388400

Malinowski, Bronislaw. 2014. *Argonauts of the Western Pacific: An Account of Native Enterprise and Adventure in the Archipelagos of Melanesian New Guinea*. Routledge Classics. London; New York: Routledge.

Marcus, Gary. 2018. ‘Deep Learning: A Critical Appraisal’. arXiv preprint arXiv:1801.00631.

Marcus, George E. 1995. ‘Ethnography in/of the World System: The Emergence of Multi-Sited Ethnography’. *Annual Review of Anthropology* 24: 95–117.

Marsh, Allison. 2018. ‘Elektro the Moto-Man Had the Biggest Brain at the 1939 World’s Fair’. *IEEE Spectrum*, September 28, 2018. https://spectrum.ieee.org/tech-history/dawn-of-electronics/elektro-the-motoman-had-the-biggest-brain-at-the-1939-worlds-fair

Martin, Emily. 1991. ‘The Egg and the Sperm: How Science Has Constructed a Romance Based on Stereotypical Male- Female Roles’. *Signs* 16 (3): 485–501.

Marx, Karl. 2008. *Capital: An Abridged Edition*. Edited by David McLellan. Oxford World’s Classics. Oxford; New York: Oxford University Press.

Masani, P. R. 1990. *Norbert Wiener 1894–1964*. Basel: Birkhäuser Basel.

Masís, Jethro. 2014. 'Making AI Philosophical Again: On Philip E. Agre's Legacy'. *Continent* 4 (1): 58-70.

Mateescu, Alexandra, and Madeleine Clare Elish. 2019. 'AI in Context: The Labour of Integrating New Technologies'. 30. Data & Society.

Matthews, Dylan. 2015. 'I Spent a Weekend at Google Talking with Nerds about Charity. I Came Away … Worried.' *Vox*, August 10, 2015. https://www.vox.com/2015/8/10/9124145/effective-altruism-global-ai.

Mauss, Marcel. 1997. 'A Category of the Human Mind: The Notion of the Person; the Notion of Self'. In *The Category of the Person: Anthropology, Philosophy, History*, edited by Michael Carrithers, Steven Collins, and Steven Lukas, 309. Cambridge: Cambridge

———. 2002. *The Gift: The Form and Reason for Exchange in Archaic Societies*. Oxon, UK: Routledge.

McCarthy, John. 1989. 'Artificial Intelligence, Logic and Formalizing Common Sense'. In *Philosophical Logic and Artificial Intelligence*, edited by Richmond H. Thomason, 161–90. Dordrecht: Springer Netherlands.

McCarthy, John, Marvin L Minsky, N. Rochester, and Claude E Shannon. 1955. 'Dartmouth Summer Research Project on Artificial Intelligence'.

McCorduck, Pamela. 2004. *Machines Who Think: A Personal Inquiry into the History and Prospects of Artificial Intelligence*. 25th anniversary update. Natick, Mass: A.K. Peters.

McCulloch, Warren, and Walter H. Pitts. 1943. 'A Logical Calculus of the Ideas Immanent in Nervous Activity'. *The Bulletin of Mathematical Biophysics* 5.4: 115–33.

McFarland, Matt. 2014. 'Elon Musk: "With Artificial Intelligence We Are Summoning the Demon."' *Washington Post*, October 24, 2014. https://www.washingtonpost.com/news/innovations/wp/2014/10/24/elon-musk-with-artificial-intelligence-we-are-summoning-the-demon/.

McGinn, Conor, Eamonn Bourke, Andrew Murtagh, Cian Donovan, Patrick Lynch, Michael F. Cullinan, and Kevin Kelly. 2020. ‘Meet Stevie: A Socially Assistive Robot Developed Through Application of a “Design-Thinking” Approach’. *Journal of Intelligent & Robotic Systems* 98 (1): 39–58.

Menzel, Peter. 1998. *Kismet Is Happy*. Photograph. Science Source. https://robots.ieee.org/robots/kismet/.

Merleau-Ponty, Maurice. 1968a. *The Visible and the Invisible*. Evanston: Northwestern University Press.

———. 1968b. *The Visible and the Invisible: Followed by Working Notes*. Edited by Claude Lefort. Translated by Alphonso Lingis. Studies in Phenomenology and Existential Philosophy. Evanston, Ill: Northwestern Univ. Press.

———. 2012. *Phenomenology of Perception*. Translated by Donald A. Landes. Abingdon, Oxon; New York: Routledge.

Miller, Daniel. 1987. *Material Culture and Mass Consumption*. Social Archaeology. Oxford, OX, UK; New York, NY, USA: B. Blackwell.

———. 2005. ‘Materiality: An Introduction’. Online Article. *UCL Anthropology*, 2005. https://www.ucl.ac.uk/anthropology/people/academic-and-teaching-staff/daniel-miller/materiality-introduction.

Minsky, Marvin, and Seymour Papert. 2017. *Perceptrons: An Introduction to Computational Geometry*. 2017 edition. Cambridge, MA: The MIT Press.

Minsky, Marvin, and Doug Riecken. 1994. ‘A Conversation with Marvin Minsky About Agents’. *Communication of the ACM* 37 (7): 22–29.

MIRI. 2021. ‘Intelligence.Org/Team/’. Machine Intelligence Research Institute. Accessed November 13, 2021. https://intelligence.org/team/.

Moravec, Hans. 1995. *Mind Children: The Future of Robot and Human Intelligence*. 4. Cambridge: Harvard Univ. Press.

Morgan, Lewis Henry. 2003. *Ancient Society*. Tucson: University of Arizona Press.

Mooney, Gerry. 2019. *Good Companion*. Photograph. *The Independent*. https://www.independent.ie/irish-news/robo-carer-ai-prototype-stevie-2-could-be-futuristic-friend-for-elderly-38116810.html

Mori, Masahiro. 1970. 'The Uncanny Valley'. *Energy* 7: 33–35.

———. 2012. 'The Uncanny Valley: The Original Essay by Masahiro Mori'. Translated by Karl F. MacDorman and Norri Kageki. *IEEE Spectrum*, June 12, 2012. https://spectrum.ieee.org/the-uncanny-valley

Moya, Patricia. 2014. 'Habit and Embodiment in Merleau-Ponty'. *Frontiers in Human Neuroscience*, July 25, 2014. https://www.frontiersin.org/articles/10.3389/fnhum.2014.00542/full

Mozur, Paul. 2017. 'Google's AlphaGo Defeats Chinese Go Master in Win for A.I.' *The New York Times Business,* May 23, 2017. https://www.nytimes.com/2017/05/23/business/google-deepmind-alphago-go-champion-defeat.html.

MTurk. 2021. 'Amazon Mechanical Turk'. Home page. Accessed October 19, 2021. https://www.mturk.com/.

Ng, Andrew. 2016. 'What Artificial Intelligence Can and Can't Do Right Now'. *Harvard Business Review*, November 9, 2016. https://hbr.org/2016/11/what-artificial-intelligence-can-and-cant-do-right-now.

Nichols, Greg. 2021. 'Delivery Robots Are Taking over College Campuses'. *ZDNet*, November 9, 2021. https://www.zdnet.com/article/delivery-robots-are-taking-over-college-campuses/.

Nilsson, Nils J. 2010. *The Quest for Artificial Intelligence: A History of Ideas and Achievements*. Cambridge; New York: Cambridge University Press.

Nomura, Tatsuya, Takayuki Uratani, Takayuki Kanda, Kazutaka Matsumoto, Hiroyuki Kidokoro, Yoshitaka Suehiro, and Sachie Yamada. 2015. 'Why Do Children Abuse Robots?' In *Proceedings of the Tenth Annual ACM/IEEE International Conference on*

Human-Robot Interaction Extended Abstracts, 63–64. Portland Oregon USA: ACM.

Norton-Wise, M. 2007. 'The Gender of Automata in Victorian Britain'. In *Genesis Redux: Essays in the History and Philosophy of Artificial Life*, edited by Jessica Riskin, 163-195. Chicago: University of Chicago Press.

Nussbaum, Martha C. 2001. *The Fragility of Goodness: Luck and Ethics in Greek Tragedy and Philosophy*. Cambridge, UK; New York: Cambridge University Press.

Ocasio-Cortez, Alexandria. 2021. 'Shout out to everyone who fought against community advocates who demanded…' *@AOC*. Twitter, February 25, 2021. https://twitter.com/AOC/status/1365021717144420354.

O'Connell, Sanjida. 2000. 'What the Tortoise Taught Us'. *The Guardian Science*, December 7, 2000. https://www.theguardian.com/science/2000/dec/07/robots.

Ogden, Sam. 1993. *Cog Exercises Its Fingers*. Photograph. Science Source. https://robots.ieee.org/robots/cog/.

olinerd. 2008. 'Boston Dynamics Big Dog (New Video March 2008)'. YouTube video, 3:28, March 17, 2008. https://www.youtube.com/watch?v=W1czBcnX1Ww.

O'Neil, Cathy. 2017. *Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy*. London: Penguin Books.

Ong, Walter J. 2012. *Orality and Literacy: The Technologizing of the Word*. 30th anniversary ed.; 3rd ed. Orality and Literary. London; New York: Routledge.

Robotics openletter.eu. 2018. 'Open letter to the European Commission'. Open Letter, 2018. http://www.robotics-openletter.eu/.

Open Philanthropy. 2016. 'UC Berkeley — Center for Human-Compatible AI (2016)'. Focus Areas. Potential Risks from Advanced Artificial Intelligence. Accessed November 13, 2021. https://www.openphilanthropy.org/focus/global-catastrophic-risks/potential-risks-advanced-artificial-intelligence/uc-berkeley-center-human-compatible-ai.

Ortner, Sherry B. 2006. *Anthropology and Social Theory: Culture, Power, and the Acting Subject*. Durham: Duke University Press.

Papert, Seymour. 1968. ‘The Artificial Intelligence of Hubert L. Dreyfus: A Budget of Fallacies’. MIT Project Mac Artificial Intelligence Memo. No. 154.

Patton, Michael Quinn. 2015. *Qualitative Research & Evaluation Methods: Integrating Theory and Practice*. Fourth edition. Thousand Oaks, California: SAGE Publications, Inc.

Pfaffenberger, Bryan. 1988. ‘Fetishised Objects and Humanised Nature: Towards an Anthropology of Technology’. *Man (New Series)* 23: 236–52.

Pietz, William. 1985. ‘The Problem of the Fetish, I’. *RES: Anthropology and Aesthetics*, no. 9: 5–17.

———. 1987. ‘The Problem of the Fetish, II: The Origin of the Fetish’. *RES: Anthropology and Aesthetics*, no. 13: 23–45.

———. 1988. ‘The Problem of the Fetish, IIIa: Bosman’s Guinea and the Enlightenment Theory of Fetishism’. *RES: Anthropology and Aesthetics*, no. 16: 105–24.

Premack, David, and Guy Woodruff. 1978. ‘Does the Chimpanzee Have a Theory of Mind?’ *Behavioral and Brain Sciences* 1 (4): 515–26.

Purtill, Corinne. 2019. ‘The Robot That Could Change the Senior Care Industry’. *Time Magazine*, October 4, 2019. https://time.com/longform/senior-care-robot/.

Rapp, Rayna. 1988. ‘Moral Pioneers: Women, Men and Fetuses on a Frontier of Reproductive Technology’. *Women & Health* 13 (1–2): 101–17.

Reynolds, Jack. 2002. ‘Merleau-Ponty, Levinas, and the Alterity of the Other’. *Symposium* 6 (1): 63–78.

Richardson, Kathleen. 2010. ‘Disabling as Mimesis and Alterity: Making Humanoid Robots at the Massachusetts Institute of Technology’. *Ethnofoor* 22 (1): 75–90.

———. 2015. *An Anthropology of Robots and AI: Annihilation Anxiety and Machines*. Routledge Studies in Anthropology 20. New York: Routledge, Taylor & Francis Group.

Riskin, Jessica. 2003. 'The Defecating Duck, or, the Ambiguous Origins of Artificial Life'. *Critical Inquiry* 29 (4): 599–633.

———, ed. 2007. *Genesis Redux: Essays in the History and Philosophy of Artificial Life*. Chicago: University of Chicago Press.

Robertson, Jennifer. 2007. 'Robo Sapiens Japanicus: Humanoid Robots and the Posthuman Family'. *Critical Asian Studies* 39 (3): 369–98.

Rose, Deborah Bird. 2011. *Wild Dog Dreaming: Love and Extinction*. Under the Sign of Nature: Explorations in Ecocriticism. Charlottesville, VA: University of Virginia Press.

Rosenblueth, Arturo, Norbert Wiener, and Julian Bigelow. 1943. 'Behaviour, Purpose and Teleology'. *Philosophy of Science* 10 (1): 18–24.

Rosenthal-von der Pütten, Astrid M., Nicole C. Krämer, Laura Hoffmann, Sabrina Sobieraj, and Sabrina C. Eimler. 2013. 'An Experimental Study on Emotional Reactions Towards a Robot'. *International Journal of Social Robotics* 5 (1): 17–34.

Rumelhart, David E., G. E. Hinton, and James L. McClelland. 1987. 'A General Framework for Parallel Distributed Processing'. In *Parallel Distributed Processing: Explorations in the Microstructure of Cognition: Foundations*, 45–76. Cambridge, Mass: MIT Press.

Russell, Stuart J., and Peter Norvig. 1995. *Artificial Intelligence: A Modern Approach*. Prentice Hall Series in Artificial Intelligence. Englewood Cliffs, N.J: Prentice Hall.

Saerbeck, Martin, and Christoph Bartneck. 2010. 'Perception of Affect Elicited by Robot Motion'. In *2010 5th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*, 53–60. Osaka, Japan: IEEE.

Sahlins, Marshall. 1994. 'Evolution: Specific and General'. In *Evolution and Culture*, edited by Thomas Harding and Marshall Sahlins, Nachdr. Ann Arbor Paperbacks. Ann Arbor, Mich.: Univ. of Michigan Press.

———. 2013. *What Kinship Is - And Is Not*. Chicago: University of Chicago Press.

Sample, Ian. 2015. ‘AI: Will the Machines Ever Rise Up?’ *The Guardian*, June 26, 2015. https://www.theguardian.com/science/2015/jun/26/ai-will-the-machines-ever-rise-up.

Sansi Roca, Roger. 2015. ‘Fetishism’. In *International Encyclopedia of the Social & Behavioral Sciences*, edited by James Wright, 105–10. Amsterdam: Elsevier.

Sartre, Jean-Paul. 1968. *Search for a Method*. Translated by Hazel E. Barnes. New York: Random House.

Scassellati, Brian. 2002. ‘Theory of Mind for a Humanoid Robot’. *Autonomous Robots* 12 (1): 13–24.

Schuster, Joshua, and Derek Woods. 2021. *Calamity Theory*. Online Book. Manifold Edition. Minneapolis & London: University of Minnesota Press. https://manifold.umn.edu/read/calamity-theory/section/f7e1ce50-b026-428a-81a1-785be400a378

Sharkey, Noel. 2007. ‘The Programmable Robot of Ancient Greece’. *New Scientist Technology*, July 4, 2007. https://web.archive.org/web/20201231104107/https://www.newscientist.com/article/mg19526111-600-the-programmable-robot-of-ancient-greece/.

Shelley, Mary. 1993. *Frankenstein, or, The Modern Prometheus*. Herts, United Kingdom: Wordsworth Classics.

Simon, Herbert A. 1996. *Models of My Life*. 1st MIT Press ed. Cambridge, Mass: MIT Press.

Simon, Herbert A., and Allen Newell. 1958. ‘Heuristic Problem Solving: The Next Advance in Operations Research’. *Operations Research* 6 (1): 1–10.

Singer, P. W. 2010. *Wired for War: The Robotics Revolution and Conflict in the Twenty-First Century*. New York: Penguin Books.

Smith, Craig S. 2020. ‘Computers Already Learn From Us. But Can They Teach Themselves?’ *The New York Times Technology*, April 8, 2020. https://www.nytimes.com/2020/04/08/technology/ai-computers-learning-supervised-unsupervised.html.

Soja, Edward W. 1996. *Thirdspace: Journeys to Los Angeles and Other Real-and-Imagined Places*. Cambridge, Mass: Blackwell.

Sperber, Dan. 1996. *Explaining Culture: A Naturalistic Approach*. Oxford, UK; Cambridge, Mass: Blackwell.

Spyer, Patricia, ed. 1998. *Border Fetishisms: Material Objects in Unstable Spaces*. Zones of Religion. New York: Routledge.

SRI International. 2017. *Shakey the Robot: The First Robot to Embody Artificial Intelligence*. YouTube video, 5:39, February 24, 2017. https://www.youtube.com/watch?v=7bsEN8mwUB8.

Standage, Tom. 2003. *The Mechanical Turk: The True Story of the Chess-Playing Machine That Fooled the World*. London: Penguin.

Starship. 2019. ‘Starship Technologies - Robot Operator on Campus’. Starship Careers. Accessed September 30, 2019. https://jobs.lever.co/starship/6386c6ed-04fd-4ee4-8fae-d8e55768de12.

Star Wars. 2014. ‘The Cinema Behind Star Wars: Metropolis’. StarWars.Com. Accessed 10 November, 2021. https://www.starwars.com/news/the-cinema-behind-star-wars-metropolis.

Stocking, George W., ed. 1985. *Objects and Others: Essays on Museums and Material Culture*. History of Anthropology, v. 3. Madison, Wis: University of Wisconsin Press.

Strathern, Marilyn. 1990. ‘Artifacts of History: Events and the Interpretation of Images’. In *Culture and History in the Pacific*, edited by Jukka Siikala. Transactions / the Finnish Anthropological Society, no. 27. Helsinki: Finnish Anthropological Society.

———. 1992. *Reproducing the Future: Essays on Anthropology, Kinship and the New Reproductive Technologies*. Manchester: Manchester University Press.

———. 1996. ‘Cutting the Network’. *The Journal of the Royal Anthropological Institute* 2: 517–35.

———. 2001. *The Gender of the Gift: Problems with Women and Problems with Society in Melanesia*. Nachdr. Studies in Melanesian Anthropology 6. Berkeley, Calif.: Univ. of California Press.

Suchman, Lucy. 1985. *Plans and Situated Actions: The Problem of Human-Machine Communication*. Palo Alto, California: Xerox Corporation.

———. 2007. *Human-Machine Reconfigurations: Plans and Situated Actions*. 2nd ed. Cambridge; New York: Cambridge University Press.

———. 2011. ‘Subject Objects’. *Feminist Theory* 12 (2): 119–45.

———. 2016. ‘Situational Awareness and Adherence to the Principle of Distinction as a Necessary Condition for Lawful Autonomy’. Briefing Paper. Panel ‘Towards a Working Definition of LAWS’: Autonomy. Geneva: CCW Informal Meeting of Experts on Lethal Autonomous Weapons.

———. 2019. ‘Demystifying the Intelligent Machine’. In *Cyborg Futures*, edited by Teresa Heffernan, 35–61. Social and Cultural Studies of Robots and AI. Cham: Springer International Publishing.

Sutton, Richard S., and Andrew G. Barto. 2018. *Reinforcement Learning: An Introduction*. Second edition. Adaptive Computation and Machine Learning Series. Cambridge, Massachusetts: The MIT Press.

Suzuki, Yutaka, Lisa Galli, Ayaka Ikeda, Shoji Itakura, and Michiteru Kitazaki. 2015. ‘Measuring Empathy for Human and Robot Hand Pain Using Electroencephalography’. *Scientific Reports* 5 (1): 1-9

Takanishi, Atsuo. 2019. ‘Historical Perspective of Humanoid Robot Research in Asia’. In *Humanoid Robotics: A Reference*, edited by Ambarish Goswami and Prahlad Vadakkepat, 35–52. Dordrecht: Springer Netherlands.

Taussig, Michael. 1993a. ‘Maleficium: State Fetishism’. In *Fetishism as Cultural Discourse*, edited by Emily S. Apter and William Pietz, 217–47. Ithaca, N.Y: Cornell University Press.

———. 1993b. *Mimesis and Alterity: A Particular History of the Senses*. New York: Routledge.

———. 2010. *The Devil and Commodity Fetishism in South America*. 30th anniversary ed. Chapel Hill [N.C.]: University of North Carolina Press.

The Guardian. 2016. *Google's Deep Dream Photograph*. Image. https://www.theguardian.com/artanddesign/2016/mar/28/google-deep-dream-art

The New York Times. 1927. *Front Page of New York Times October 23, 1927*. Image. New York Times Digital Archive. http://timesmachine.nytimes.com/timesmachine/1927/10/23/96676638.html

TIME magazine. 2019. *Stevie the Robot*. Image. https://time.com/longform/senior-care-robot/

Traweek, Sharon. 1992. *Beamtimes and Lifetimes: The World of High Energy Physicists*. 1. ed. Cambridge, Mass: Harvard Univ. Press.

Turing, A. M. 1937. 'On Computable Numbers, with an Application to the Entscheidungsproblem'. *Proceedings of the London Mathematical Society* 2.1: 230–65.

Turkle, Sherry. 2005. *The Second Self: Computers and the Human Spirit*. 20th anniversary ed., 1st MIT Press ed. Cambridge, Mass: MIT Press.

———. 2011. *Alone Together: Why We Expect More from Technology and Less from Each Other*. New York: Basic Books.

Turner, Bryan S. 2008. *The Body & Society: Explorations in Social Theory*. 3rd ed. Theory, Culture & Society. Los Angeles: SAGE.

Tylor, Edward Burnett. 1871. *Primitive Culture: Researches into the Development of Mythology, Philosophy, Religion, Art and Custom*. Vol. 1. London: John Murray, Albermarle Street.

UNDP. 2017. 'UNDP in Asia and the Pacific Appoints World's First Non-Human Innovation Champion'. United Nations lead agency on international development in Asia and the Pacific. Accessed November 8, 2021. https://www.asia-

pacific.undp.org/content/rbap/en/home/presscenter/pressreleases/2017/11/22/rbfsingapore.html.
Veling, L. 2017. *Becoming Real: An Anthropological Approach to Evaluating Robots in the Real World*. In *International Conference on Social Robotics*, 638-646. Springer, Cham, 2017.
Veling, Louise, and Conor McGinn. 2021. ‘Qualitative Research in HRI: A Review and Taxonomy’. *International Journal of Social Robotics*, 13: 1-21.
Verdonck, Kris. 2016. ‘DANCER #1’. atwodogscompany. YouTube video, 4:17, May 30, 2016. https://www.youtube.com/watch?v=xKbBImj3mWM.
Viveiros de Castro, Eduardo. 2012. ‘Cosmologies: Perspectivism’. *Hau* Masterclass: 45–168.
Vuleta, Branka. 2021. ‘How Much Data Is Created Every Day? [27 Powerful Stats]’. *SeedScientific* (blog). 28 January 2021. https://seedscientific.com/how-much-data-is-created-every-day/.
Walter, William Grey. 1963. *The Living Brain*. 5. The Norton Library 153. New York: Norton.
Waseda University. 1973. *WABOT – Waseda robot-*. Image. https://www.humanoid.waseda.ac.jp/booklet/kato_2.html
Weheliye, Alexander G. 2008. ‘After Man’ 20 (1/2): 321–36.
Wehrle, Maren. 2021. ‘“Bodies (That) Matter”: The Role of Habit Formation for Identity’. *Phenomenology and the Cognitive Sciences* 20 (2): 365–86.
Weizenbaum, Joseph. 1976. *Computer Power and Human Reason: From Judgment to Calculation*. San Francisco: W. H. Freeman.
Westfahl, Gary. 1992. ‘“The Jules Verne, H.G. Wells, and Edgar Allan Poe Type of Story”: Hugo Gernsback’s History of Science Fiction’. *Science Fiction Studies* 19 (3): 340–53.
Westling, Louise. 2010. ‘Merleau-Ponty’s Human-Animality Intertwining and the Animal Question’. *Configurations* 18 (1–2): 161–80.
Whelan, Christal. 2011. ‘Robot Revolution: Significant Otherness in Japan’. *The Daily Yomiuri Travel*, August 7th, 2011.

———. 2014. *Kansai Cool: A Journey into the Cultural Heartland of Japan*. North Clarendon, VT: Tuttle Pub.

White, Leslie A. 1943. ‘Energy and The Evolution of Culture’. *American Anthropologist* 45 (3): 335–56.

———. 2007. *The Evolution of Culture: The Development of Civilization to the Fall of Rome*. Walnut Creek, Calif: Left Coast Press.

Wiener, Norbert. 1949. ‘Cybernetics or Control and Communication in the Animal and the Machine’. *Psychological Bulletin* 46 (3): 236–37.

Wikimedia. 2021. *Gakutensoku*. Image. https://upload.wikimedia.org.

Winner, Langdon. 1980. ‘Do Artifacts Have Politics?’ *Daedalus*, 121–36.

World Economic Forum. 2011. ‘Personal Data: The Emergence of a New Asset Class’. World Economic Forum, Bain & Company, Inc.