CRITICAL MASS: COLLECTIVITY AND COLLABORATION IN THE HISTORY OF MULTI-AGENT INTELLIGENT SYSTEMS

By

Rachel Bergmann

Department of Art History and Communication Studies
McGill University, Montreal

November 2019

A thesis submitted to McGill University in
partial fulfillment of the requirements of the degree of

Master of Arts in Communication Studies

Table of Contents

Figures

Abstract:

This thesis examines the intellectual history of collaboration and collective action in multi-agent systems AI research. Drawing on archival work, including computer science research papers, conference proceedings, technical reports, and journalism, it considers the cultural, institutional, and intellectual forces that shaped this approach in artificial intelligence research and the teams that worked on them. The first chapter examines the cultural history of the Stanford AI Laboratory and Stanford Research Institute and considers how countercultural movements in the Bay Area from 1969-1973 influenced their institutional cultures. The second chapter outlines a network of women in AI research who, together, expanded the range of methodologies and disciplines usually included in AI and MAS research. Borrowing from Michelle Murphy's concept of protocol feminism, I examine their "feminist AI protocol" and outline the sets of practices and techniques they used in their research objectives, scientific method, areas of specialization, and academic service. The third chapter focuses on the work of one researcher, Barbara Grosz, and places her ideas in conversation with academic developments at the same time from philosophy of mind, STS, and feminist critiques of AI. The thesis concludes by considering the limits of this feminist AI protocol, multi-agent or not, without a deeper commitment to feminist epistemologies.

Résumé:

Ce mémoire examine l'histoire intellectuelle de la collaboration et de l'action collective dans la recherche sur l'IA et les systèmes multi-agents. S'appuyant sur des documents d'archives, des travaux de recherche en informatique, des comptes rendus de conférences, des rapports techniques et des publications journalistiques, il prend en compte les forces culturelles, institutionnelles et intellectuelles qui ont façonné les systèmes multi-agents et les équipes qui y ont travaillé. Le premier chapitre se penche sur l'histoire culturelle de SAIL et du SRI et considère la manière dont les mouvements de la contre-culture dans la région de la baie de San Francisco entre 1969 et 1973 ont influencé leurs cultures institutionnelles. Le deuxième chapitre décrit un réseau de femmes dans la recherche sur l'IA qui a élargi l'éventail des méthodologies et des disciplines communément incluses dans la recherche sur l'IA et le MAS. Empruntant à Michelle Murphy son concept de féminisme protocolaire, j'examine leur « protocole d'IA féministe » et décrit l'ensemble des pratiques et des techniques qu'elles ont utilisées dans leurs objectifs de recherche, leur méthode scientifique, leurs domaines de spécialisation et leur service académique. Le troisième chapitre se concentre sur les travaux de la chercheuse Barbara Grosz et met ses idées en dialogue avec les développements académiques issus à la fois de la philosophie de l'esprit, de la STS et des critiques féministes de l'IA. Le mémoire conclue en considérant les limites de ce protocole d'IA féministe, multi-agents ou non, dans l'absence d'un engagement plus approfondi avec les épistémologies féministes.

Acknowledgements

I would like to thank Jonathan Sterne, an advisor more generous and caring than I could have imagined. I am grateful for his mentorship, advice, editing, and intellectual guidance through this new world of the humanities.

The faculty, staff, and students of the Department of Art History and Communication Studies have created a supportive and intellectually rich environment. Thank you especially to Jenny Burman, Matthew Hunter, Becky Lentz, Ara Osterweil, Carrie Rentschler, Jonathan Sterne, and Will Straw, whose courses have shaped my thinking immensely. I am grateful for my fellow graduate students—Burç, Chahinez, Claire, Hannah, Mirra, Rach, Sadie, and everyone in CATDAWG and CPAIS—who inspired me, challenged me intellectually, and made graduate school so much more enjoyable. Thank you to Media@McGill and the AHCS Graduate Student Association for providing a workspace for us to think, read, write, and eat together.

I am thankful for the financial support I received through the McKinnon E-Research Award, the Graduate Excellence Fellowship, and two Media@McGill Awards. Thank you also to Jonathan Sterne, who provided funding through research assistantships. I would like to thank the Social Media Collective at Microsoft Research New England for accommodating me in the final phases of this project.

Lastly, I would like to thank my family, friends, mentors, and places in Montreal that have supported me and without whom this project would not have been possible. Thank you to CKUT 90.3 FM, Urban Element Zone, and all the cafés and public libraries that enabled me to read, write, and create while working on this project. Thank you to my family—Shelli, Bill, Willie, Louie—for their endless love, trust, and support. Thank you to my chosen family— Ammar, Alyssa, Dörte, Hayley, Julia, Pascale, Tyrel, and others— who have made Montreal such a loving home for the last seven years.

In loving memory of Dana

Introduction

> You know that everything you think and do is thought and done by you. But what's a "you"? What kinds of smaller entities cooperate inside your mind to do your work? To start to see how minds are like societies, try this: *pick up a cup of tea*!
>
> *Your GRASPING agents want to keep hold of the cup.*
> *Your BALANCING agents want to keep the tea from spilling out.*
> *Your THIRST agents want you to drink the tea.*
> *Your MOVING agents want to get the cup to your lips.*
>
> Yet none of these consume your mind as you roam about the room talking to your friends. You scarcely think at all about *Balance; Balance* has no concern with *Grasp*; *Grasp* has no interest in *Thirst*; and *Thirst* is not involved with your social problems. Why not? Because they can depend on one another. If each does its own little job, the really big job will get done by all of them together: drinking tea.
>
> Marvin Minsky, *Society of Mind* (1985)[1]

On June 9, 1980, a group of twenty computer scientists gathered at the Endicott House, 10 miles south of Boston, Massachusetts, for the first Distributed Artificial Intelligence (DAI) Workshop. Reports from the workshop make the purpose of the meeting clear: to define the scope and meaning of "distributed artificial intelligence."[2] The researchers cloistered themselves in the house for three days, explaining to each other their labs' DAI projects and hosting group discussions about special topics, like models for collaborative problem-solving and cooperation among AI systems.

Some of the researchers at that first DAI workshop were concerned with how their systems could represent the processes of human intelligence. Nils Nilsson, an artificial intelligence

---

[1] Minsky, Marvin. *The Society of Mind*. 2nd ed. New York: Simon and Schuster, 1986, 20.
[2] Davis, Randall. "Report on the Workshop on Distributed AI." Working Paper. AI Lab: MIT, September 1980. https://dspace.mit.edu/bitstream/handle/1721.1/41155/AI_WP_204.pdf?sequence=4.

researcher at Stanford Research Institute (SRI), gave a presentation at the workshop called "Two Heads Are Better Than One," in which he explained the importance of distributed AI approaches to AI research more broadly.[3] Distributed AI, he argued, could represent how AI systems form beliefs, reason about, and communicate with other AI systems and other dynamic processes in their environments. Understanding how AI systems reason and communicate with other systems would enable researchers to understand how an AI system can reason about itself and help issues in another area of AI research, natural language communication. Most importantly, Nilsson argued, DAI might even be a prerequisite for ordinary artificial intelligence. "To be sufficiently 'intelligent,' a system may have to be so complex and may have to contain so much knowledge that it will be able to function efficiently only if it is partitioned into many loosely coupled subsystems." Models like Minsky's society of mind, Nilsson pointed out, argue this must be true: "No AI without DAI."[4]

In the decades that followed that first workshop in 1980, DAI has grown substantially and been integrated with other fields and technologies in engineering, economics, cognitive science, and computer science. It grew so big that subfields grew within it. The DAI workshop continued for another decade before it renamed itself the International Conference on Multi-Agent Systems, focusing on one approach within distributed AI called multi-agent systems (MAS).[5] MAS borrows concepts and techniques from game theory, economics, human-computer interaction, linguistics, sociology, and political philosophy to model how agents interact in a shared environment to

[3] Randall, 4.
[4] Randall, 4.
[5] Lesser, Victor. "Preface." Menlo Park, California: AAAI, 1995. https://www.aaai.org/Library/ICMAS/icmas95contents.php.

accomplish tasks. In particular, it models how agents with different goals, beliefs, and abilities work together to collaborate, coordinate, and negotiate with others in a shared environment.

The Need for a Longer History

This thesis examines the intellectual history of the concepts of collaboration and collective action in multi-agent systems AI research. Drawing on interviews, archival work, computer science research papers, conference proceedings, technical reports, and journalism, it considers the cultural, institutional, and intellectual forces that shaped multi-agent systems and the teams that worked on them. At the centre of this history is a network of women in AI research who, together, expanded the range of methodologies and disciplines usually included in AI and MAS research. MAS as a field and these researchers in particular offer an entry point to consider the social, cultural, and political forces that shaped the research, development, and use of AI technologies in the twentieth century. The histories of these technologies depend on complex interactions between intellectual trends, government policies, social norms, and funding opportunities, in addition to the particular cultures of the labs, companies, and engineers that influence the development of certain technologies instead of others.

This history of MAS contributes to the growing body of literature on how computing became "male." MAS research in the 1970s and 1980s offered a way for researchers to model the way intelligent agents interacted with each other in a shared environment. System architectures that could model multiple intelligent agents destabilized previous approaches in AI research that assumed firm boundaries between individual intelligent machines and the stimuli from their environments. By basing its theories of agency and collaboration primarily from economics and game theory, however, the field of multi-agent systems fortified the idea that AI was the study of rational agents. Implicit in multi-agent models are narrow imaginations of intelligence, agency, and collaboration based primarily on cybernetics and game theory.

Critical and feminist histories of computing have begun to explore the way race, class, gender, culture, and other social factors have shaped the development of computing technologies, and how these technologies in turn help construct and enforce cultural and social norms— but there remains more work to be done. Joy Lisi Rankin's *A People's History of Computing in the United States* is one recent work that tells the story of computing citizens— teachers, schoolchildren, and university students— using time-sharing networks in the US Northeast and Midwest.[6] Rankin describes her motivation to move beyond the stories of Steve Jobs, IBM, and other figures in the "Silicon Valley mythology," and instead to "develop a history of the digital age that emphasizes creativity, collaboration, and community."[7] Rankin offers an alternative lineage of personal computing— one based on BASIC and the time-sharing networks developed by professors at MIT and Dartmouth College in the early 1960s. In *Programmed Inequality* Mar Hicks explores how the field of computing in the United Kingdom shifted from being seen as "women's work" in the 1940s to, as early as the 1960s, acquiring a "distinctly masculine image."[8] Hicks traces how the tangled web of technological advancements, gendered labour organization, and cultural norms about marriage and sexuality shaped Britain's halting efforts to computerize in the late twentieth century. Rankin and Hicks both offer histories of computing that run counter to individualistic, male-dominated, teleological narratives of invention, innovation, and technological progress. Both frame the history of computing as deeply embedded in a web of

---

[6] Rankin, Joy Lisi. *A People's History of Computing in the United States*. Cambridge, Massachusetts: Harvard University Press, 2018.
[7] Rankin, 4.
[8] Hicks, Mar. *Programmed Inequality: How Britain Discarded Women Technologists and Lost Its Edge in Computing.* Cambridge, MA: MIT Press, 2017.

social, political, and cultural forces which have shaped the labour forces, forms of access, and imaginaries surrounding computing technologies.

This thesis aims to add to these critical histories of computing that primarily centre men. More scholarship to uncover the ways computing became "male" in the twentieth century is needed—and, indeed, some aspects of this history contribute to this project—but there is nevertheless much work needed to recover the many contributions women have made to the development of artificial intelligence as a field of research in the twentieth century. This thesis outlines one network of women computer scientists in the mid- to late-20th century and their contributions to the field of artificial intelligence. A central character in this thesis is Barbara Grosz, an AI researcher who has collaborated and published with dozens of other women computer scientists in the last five decades. In 1977, when Barbara Grosz competed her PhD in computer science, women earned 13.2% of all doctoral degrees in math and computer science in the United States.[9] Women continued to earn 12-20% of PhDs in math and computer science throughout the 1980s and 1990s.[10] Despite the many factors dissuading women from staying in computer science departments— including social isolation, harassment, and lack of university support— women continued to produce research and build technologies.

In addition to this historical recovery work, this thesis analyzes these women in the history of MAS research as complex, politically imperfect characters. It explores how gendered paradigms shape research and how research influenced gendered paradigms in artificial intelligence. In *Seizing the Means of Reproduction* Michelle Murphy explores the protocols of radical feminists in

[9] Fiegener, Mark K. "Science and Engineering Degrees: 1966–2010." Detailed Statistical Tables. NSF: National Center for Science and Engineering Statistics, June 2013. https://nsf.gov/statistics/nsf13327/content.cfm?pub_id=4266&id=2, Table 25.
[10] Fiegener, Table 25.

California in the 1970s as part of the emerging women's health movement.[11] Citing W. E. B. Du Bois, Murphy describes her attempt to work another kind of double vision in her analysis, "ruthlessly historicizing these past feminist efforts as one might any other scientific endeavour, while doing so from a point of deep investment in feminist technoscience studies as a critical epistemological and material project that values entanglement and sits in a genealogic relation to the practices examined."[12] I hope to work a similar kind of double vision. Through their feminist protocol these researchers opened up AI research in major ways, creating computer science departments more welcoming for women as well as doing research that countered dominant assumptions about language and human-computer interaction. Their social and political contexts, however, limited the extent to which their protocols offered more radical imaginings of a feminist approach to AI. These scientists enacted their protocols before feminist technoscience and the vocabulary of intersectionality was widespread to computer science and engineering cultures. These figures and their careers are deeply entangled with the histories and implications of US imperialism, extensive military funding of scientific research, and the shadow of the Cold War. In this thesis I hope to draw attention to the important feminist contributions made by this network while also critically historicizing their work like any other scientific project.

This thesis draws methodological inspiration from what Donna Haraway describes in *Primate Visions* as her four theoretical "temptations."[13] Haraway describes four frameworks that inform her analytic approach— each important, but insufficient by themselves. First is the kind of

---

[11] Murphy, Michelle. *Seizing the Means of Reproduction: Entanglements of Feminism, Health, and Technoscience.* Duke University Press, 2012.
[12] Murphy, 23.
[13] Haraway, Donna J. *Primate Visions: Gender, Race, and Nature in the World of Modern Science.* New York: Routledge, 1990.

social constructivism, articulated by scholars like Bruno Latour, which rejects positivistic notions of reality.[14] Haraway's second temptation comes from Marxist feminism and in particular Marxist analyses of institutions like wage labour, sexual and reproductive appropriation, and racial hegemony. The third temptation, Haraway says, is the "siren call of the scientists themselves," or the epistemologies of working scientists.[15] Science is not just about power and control, the scientists argue: there *are* organisms in the world, and they behave in certain ways and not others. The final temptation is to consider the profound ways race and gender shape the very lenses through which we see the world. These four methodologies each provide tools that clash and fit together to demonstrate the complex, tightly knit entanglements of the constructions of nature, culture, science, and objectivity in this work.

Approaching MAS research historically offers a way to consider algorithms and AI technologies beyond just their technical specifications, instead analyzing them as complex sociotechnical assemblages shaped by material, geographic, aesthetic, cultural, and social forces. Scholarship in critical algorithm studies and feminist STS has provided techniques and frameworks to uncover and challenge the assumptions and values which have been foundational to scientific and technological research. For instance, in *Technofeminism*, Judy Wacjman describes her technofeminist framework, which theories and methods from science and technology studies (STS) and feminist studies of technoscience.[16] A technofeminist framework, she argues, emphasizes how gender and technology mutually shape each other: in one direction, gender relations are materialized in technology, and on the other, notions of masculinity and femininity

[14] Haraway, 7.
[15] Haraway, 7.
[16] Wajcman, Judy. *TechnoFeminism*. Cambridge ; Malden, MA: Polity, 2004.

acquire their meanings and characters through the enrolment and embeddedness in working machines.[17] I use such a technofeminist approach to consider the ways gender and technology shape each other in the history of MAS approaches.

This history of MAS seeks to recenter conversations around bias in AI to the wider history of the field. Machine learning and image processing are not the only techniques and applications of AI research. MAS approaches can certainly be used for AI research in visual perception, but part of the reason researchers turned to MAS approaches was to deal with speech, sound, and other environmental information. The development of MAS runs alongside or counter to the story of AI that centres visual perception. In one telling, "It all started with a frog," writes N. Katherine Hayles in a chapter on cybernetics from her book *How we Became Posthuman.*[18] Hayles's opening line refers to the 1959 paper "What the Frog's Eye Tells the Frog's Brain," written by Jerome Lettvin, Humberto Maturana, Warren McCullock, and Walter Pitts.[19] This paper, in which the authors argue the frog's eye has specific "bug detector" neurons which organize and interpret sensor information before sending the signal to the frog's brain, inspired decades of research in cognitive science and AI about visual perception and image processing. Since then, a vast portion of attention and humanistic scholarship about AI has focused on the histories and developments of visual pattern recognition. In recent years, much attention has focused on how researchers use AI techniques to recognize faces, objects, license plates, and handwriting in images. Part of this focus is related to the relative success of these machine learning techniques— due in large part to the

---

[17] Wacjman, 107.

[18] Hayles, N. Katherine. *How We Became Posthuman: Virtual Bodies in Cybernetics, Literature, and Informatics.* University of Chicago Press, 2008.

[19] Lettvin, J., H. Maturana, W. McCulloch, and W. Pitts. "What the Frog's Eye Tells the Frog's Brain." *Proceedings of the IRE* 47, no. 11 (November 1959): 1940–51. https://doi.org/10.1109/JRPROC.1959.287207.

availability of large, high-quality, labeled datasets of images (ImageNet), objects (Microsoft COCO), and handwritten digits (MNIST). But these developments, and the fantasies that accompany them, point to a kind of ocularcentrism present in many of the myths surrounding artificial intelligence.

In this history of multi-agent systems, speech, sound, and language are central. In the first chapter, I describe the cultures of two AI labs in Palo Alto: the Stanford Artificial Intelligence Laboratory (SAIL) and the Stanford Research Institute's (SRI) AI Center. Barbara Grosz was hired at SRI to work on a speech recognition project, funded by the United States Department of Defence's Advanced Research Projects Agency's (DARPA) five-year research and development program called Speech Understanding Research (SUR).[20] From 1971-1976, DARPA funded a small group of robotics and AI laboratories across the United States to research and develop a human-computer communication system that could recognize and understand continuous speech (e.g., sentence-length utterances instead of individual words spoken with pauses in between). This research project funded Grosz's early work on computational models of discourse and her emphasis on including contextual information in dialogue models.

In addition to funding Grosz's early work SRI, the DARPA Speech Understanding Research project funded Stanford AI alumni Raj Reddy and Victor Lesser. As a graduate student at Stanford, Lesser had spent the summer of 1967 working with Reddy at the "AI Project"

[20] ARPA and DARPA are the same agency and I use the two names interchangeably. The agency was founded as ARPA, renamed DARPA in 1972; ARPA again in 1993; and finally DARPA again in 1996. For more, see Norberg, A. L. "Changing Computing: The Computing Community and DARPA." *IEEE Annals of the History of Computing* 18, no. 2 (Summer 1996): 40–53. https://doi.org/10.1109/85.489723.
Klatt, Dennis H. "Review of the ARPA Speech Understanding Project." *Journal of the Acoustical Society of America* 62, no. 6 (December 1977): 1345–66. https://doi.org/10.1121/1.381666.

(renamed the AI Laboratory shortly after). In 1969 Reddy joined the faculty at Carnegie Mellon University (CMU) and Lesser joined him there after he finished his doctoral work in 1972. Reddy's team at CMU created two speech understanding systems for the DARPA project, the Harpy and the Hearsay-II, which won first and second place, respectively. The Harpy system performed best, satisfying or exceeding all of DARPA's design goals specified at the beginning of the program.[21] The Hearsay-II system did not perform quite as well as Harpy, but it used a novel approach: Hearsay-II's speech understanding system was divided into semi-independent processors that operated asynchronously and in parallel, each providing a piece of the information about the utterance and putting it on a shared "blackboard" before the system as a whole finally decoded the utterance.[22] A reviewer from *The Acoustical Society of America* describes the approach of Hearsay-II as "of interest" and its final performance "encouraging."[23] He elaborated: "Of interest is the fact that only 77% of the sentences were correctly *recognized* (all words correct) while 91% were *understood* correctly. It appears that the CMU grammar contains some desirable characteristics for the realization of computer understanding."[24] It was Hearsay-II that Lesser worked on as a researcher in Reddy's lab from 1972-1977.

From their first projects with DARPA to their later contributions that helped establish MAS as an approach worth studying, sound and speech have been central to the research careers of

---

[21] Klatt, 1345.

Huang, Xuedong, James Baker, and Raj Reddy. "A Historical Perspective of Speech Recognition." *Communications of the ACM* 57, no. 1 (January 1, 2014): 94–103. https://doi.org/10.1145/2500887.

[22] Erman, Lee D., Frederick Hayes-Roth, Victor R. Lesser, and D. Raj Reddy. "The Hearsay-II Speech-Understanding System: Integrating Knowledge to Resolve Uncertainty." *ACM Comput. Surv.* 12, no. 2 (June 1980): 213–253. https://doi.org/10.1145/356810.356816.

[23] Klatt, 1363.

[24] Klatt, 1363.

Barbara Grosz, Victor Lesser, and their many students. After Grosz's work at SRI, Grosz has become an expert in computational methods for studying language. As I describe in the second chapter, Grosz spent much of her career developing these early ideas about computers, language, and context into decades of work on multi-agent approaches in chatbots, natural language processing, human-computer interfaces, and speech-enabled AI systems. Lesser, after he was hired as a professor at the University of Massachusetts, Amherst in 1978, worked on ways to build distributed, cooperative speech recognition systems like the Hearsay-II he had helped design at CMU. In 1980 Lesser attended the first workshop on Distributed AI and in 1981 he published his first article about "cooperative distributed systems." In the 1980s and 1990s Lesser and his students worked on a Distributed Vehicle Monitoring Testbed (DVMT), where they developed sensor networks and sound understanding systems for a military-funded ocean surveillance project. It was through working on these projects that Lesser's lab started designing multi-agent systems.

In the decades of MAS research since, multi-agent systems continue to be used in applications that require sound and speech recognition. Many virtual reality environments use multi-agent architectures for dialogue and other forms of interaction among VR users and virtual humans. [25] So-called smart homes contain many heterogeneous agents interacting with one another—especially if the smart speakers have multiple user profiles saved and the voice assistants have been linked with Internet-enabled appliances. Internet- and voice assistant-connected home appliances can communicate with other agents to, for example, direct a robot vacuum to clean the house or order more groceries, based on the knowledge and skills of the smart vacuum and smart refrigerator. Voice assistants use AI techniques like natural language processing (NLP) and natural

[25] Traum, David, and Jeff Rickel. "Embodied Agents for Multi-Party Dialogue in Immersive Virtual Worlds." *Proceedings of the International Conference on Autonomous Agents*, August 10, 2002. https://doi.org/10.1145/544862.544922.

language generation (NLG) to communicate with users. Very quickly, these interactions form complex webs of different agents with heterogeneous skills and dynamically changing tasks, goals, and environments— likely candidates for MAS architecture. Although proprietary rules protect the exact mechanisms behind these technologies, as "smart" objects become the building blocks for Internet of Things (IoT) systems, it is possible voice assistant-enabled home systems like Google Home and Apple's HomePod will (or already do) use some form of MAS in their systems.[26]

As MAS approaches continue to grow, the kinds of "agents" in the systems vary more and more: some MAS agents are mechanical robots, like the players on a robot soccer team;[27] others MASs exist entirely as software agents, like those used in trading stocks.[28] Agents can be geographically distributed, like in the MASs used to manage IoT, e-commerce, and security and surveillance systems.[29] Some MAS agents navigate dynamic environments, like the robot rescuers

---

[26] Savaglio, Claudio, Giancarlo Fortino, Maria Ganzha, Marcin Paprzycki, Costin Bădică, and Mirjana Ivanović. "Agent-Based Computing in the Internet of Things: A Survey." In *Intelligent Distributed Computing XI*, edited by Mirjana Ivanović, Costin Bădică, Jürgen Dix, Zoran Jovanović, Michele Malgeri, and Miloš Savić, 737:307–20. Springer International Publishing, 2018. https://doi.org/10.1007/978-3-319-66379-1_27.

Fortino, Giancarlo, Antonio Guerrieri, Wilma Russo, and Claudio Savaglio. "Integration of Agent-Based and Cloud Computing for the Smart Objects-Oriented IoT." In *Proceedings of the 2014 IEEE 18th International Conference on Computer Supported Cooperative Work in Design (CSCWD)*, 493–98. Hsinchu, Taiwan: IEEE, 2014. https://doi.org/10.1109/CSCWD.2014.6846894.

[27] Barrett, Samuel, and Peter Stone. "Cooperating with Unknown Teammates in Complex Domains: A Robot Soccer Case Study of Ad Hoc Teamwork." In *Proceedings of the Twenty-Ninth AAAI Conference on Artificial Intelligence*, 7. Austin, Texas, 2015.

[28] Yuan Luo, Kecheng Liu, and D. N. Davis. "A Multi-Agent Decision Support System for Stock Trading." *IEEE Network* 16, no. 1 (January 2002): 20–27. https://doi.org/10.1109/65.980541.

[29] Luong, Peter, and Liming Lu. "Multiagent Web for the Internet of Things." Seoul, South Korea: IEEE, 2014. https://doi.org/10.1109/ICISA.2014.6847432.

that find and retrieve people from avalanches and buildings on fire.[30] Some multi-agent architectures, such as those used for "smart" parking and traffic flow simulations, model pedestrians and human drivers as agents in their systems. Multi-agent systems are prevalent in logistics and manufacturing, where it is often necessary to coordinate tasks, production, quality control, and transport among many warehouses and geographically distributed parties. Many (if not most) MASs include *intelligent* agents—agents that use artificial intelligence techniques like natural language processing (NLP) or machine learning (ML)—but these agents often interact with changing environments and other agents which may or may not use AI. Recently, multi-agent approaches to reinforcement learning (MARL) has become a fertile research area with hundreds of journal articles published on the topic in the last ten years.

By the end of that first DAI workshop, the group of researchers came up with a working definition of distributed AI. A report written for the MIT A.I. Laboratory explains, "Distributed AI is concerned with those problems for which a single problem solver, single machine, or locus of computation seems inappropriate. Instead we turn to the use of multiple, distinct problem solvers, each embodied in its own system."[31] Recent technological developments had enabled other distributed approaches in computing: this workshop brought together computer scientists applying these distributed approaches to problems in artificial intelligence. The researchers distinguished DAI projects from distributed processing— a widely studied topic in computer science at that time— in several ways. Unlike distributed processing systems, which centrally

[30] Takahashi, Tomoichi, Satoshi Tadokoro, Masayuki Ohta, and Nobuhiro Ito. "Agent Based Approach in Disaster Rescue Simulation - From Test-Bed of Multiagent System to Practical Application." In *RoboCup 2001: Robot Soccer World Cup V*, edited by Andreas Birk, Silvia Coradeschi, and Satoshi Tadokoro, 2377:102–11. Berlin, Heidelberg: Springer Berlin Heidelberg, 2002. https://doi.org/10.1007/3-540-45603-1_11.
[31] Davis, "Report on the Workshop on Distributed AI," 4.

controls many machines doing disparate tasks, the processors in DAI systems operate cooperatively and control is distributed across the system. The methods and motivations of the DAI researchers also differed: their research, they concluded, concerned "developing frameworks for *cooperative behaviour between willing entities,* rather than frameworks for enforcing cooperation as a form of compromise between potentially incompatible entities."[32] This conceptualization of multiple, independently acting agents within a shared environment forms the basis of MAS.[33]

Notes on Terminology

As a final note I want to clarify some of the terms which appear in the body of the thesis. Artificial intelligence may have as many definitions as there are academic disciplines that study it: in this thesis, I refer to AI as a field of study within computer science that seeks to mimic "cognitive" functions— usually with subfields to approach different aspects of human cognition, like computer vision, speech recognition, robotics, machine learning, reasoning, planning, decision-making, and natural language processing. Artificial intelligence and computer science overlap, but they are not mutually exclusive categories: other computer science research areas that are *not* artificial intelligence include designing algorithms and data structures; managing databases; designing programming languages and methods; computer security and cryptography; distributed computing; and computer architecture, among many others. Artificial intelligence research in computer science almost always includes computational methods, but AI researchers

[32] Davis, 3.

[33] MAS are related to agent-based models (ABM), another approach in computer science, in that they both consider a system and its environment in terms of agents acting semi-autonomously. ABMs, however, are usually used to study the collective behaviours that emerge from agents following simple rules, like modelling the flocking patterns of a school of fish. MASs, in contrast, are often task-oriented, like programming a group of agents to try to win a robot soccer game. Collaboration, communication, and achievement are central concepts to MAS.

often use concepts or research about human intelligence—usually from fields like cognitive science, neuroscience, cognitive psychology, philosophy of mind, and linguistics—to inform their research methods.

DAI and MAS emerged at the intersection of distributed computing and AI research. Because the development of MAS is entangled with many overlapping subfields in computer science and artificial intelligence, the definitions of terms are not always consistent. The motivations, technologies, and development of MAS overlaps with other subfields in computer science and AI, like human-computer interaction, machine learning, and other forms of distributed AI. I use chatbots to refer to AI systems which communicate with people through dialogue—written or spoken. The category of chatbots includes intelligent personal assistants and voice assistants. Speech recognition and natural language processing (NLP) are related but distinct fields of research in computer science and AI. In general, speech recognition relies on acoustic information to find patterns in audio files. NLP requires both recognition and generation of natural language, but NLP can refer to both text-based and speech-based language processing. In computer science, "natural language" refers to languages used by humans, as opposed to programming languages or machine-readable code. Computational linguistics (CL) shares many techniques and applications with NLP, but in general, computational linguists use computational methods to pursue research questions in linguistics, while NLP researches and develops computer systems which can process natural language.

Chapter outline

The thesis begins 12 years before the first DAI workshop, at Stanford University in 1968, where many MAS/DAI researchers spent time as graduate students and early career researchers. The first chapter offers a pre-history of multi-agent systems by examining the cultures of two

Stanford AI labs in the late 1960s and 1970s. I explore how SRI was the target for student movements like the April Third Movement, who succeeded in pushing Stanford to divest from SRI and their research supporting the US's occupation of Vietnam. I compare SRI with SAIL, Stanford's AI lab in a remote building "up in the hills" south of Palo Alto. Similar to the way the first DAI workshop cloistered itself at the Endicott House, SAIL also occupied a building far from Stanford's main campus. Publicly, SAIL positioned itself as a transcendent haven, physically and intellectually distant from the political unrest happening on campus. Documents from SAIL's disk backups tell a different story, however. The lab's archives include images and texts describing a woman named "Zoe" whom SAIL researchers paid to pose nude with their computers, supposedly for a class project, in 1971. The story of Zoe at SAIL— and the fondness with which it has been remembered decades later— suggests, at minimum, a mistaken conflation of sexual liberation and women's liberation and the inseparability of SAIL's lab culture from the social and political currents of the era. What SAIL's disk backups seem to illustrate is a culture of misogyny and deep indifference to women's issues in AI research. It suggests a particular, narrow, and masculine imagination of collaboration, collectivity, and inclusion in AI research.

It was in this chilly climate that Barbara Grosz and her colleagues created networks of support, feminist protocols, and new research methods to change how AI research was done. In the second chapter, I map a group of researchers who emerged from this climate in Palo Alto and created their own distributed network of researchers in AI in the 1980s and 1990s. Borrowing from Michelle Murphy's concept of protocol feminism, I examine their "feminist AI protocol" and outline the sets of practices and techniques they used in their research objectives, scientific method, areas of specialization, and involvement in academic service. These researchers, most of whom were white women, created spaces and support networks for themselves in the often hostile climates of computer science departments in the 1970s to 1990s. They published papers together, provided mentorship for other rising women scholars, and chaired committees and authored reports on the status of women in computer science at their universities. Many of these researchers worked at the intersection of MAS and computational linguistics, emphasizing linguistic theories like speech act theory and pragmatics in their philosophies of computing and human-computer interaction. Their AI protocol meant language could not be separated from context, action, and collaboration, and AI systems could not be separated from everyday natural language. Together they helped shift multi-agent systems research to include ideas from speech act theory and collaborative approaches to accomplishing shared plans, both among AI agents and between humans and AI systems.

In chapter three I focus on the work of one of the network members, Barbara Grosz, and place her ideas in conversation with academic developments at the same time from feminist philosophy and STS. I consider Grosz's rejection of the master-slave analogy in computer science— a critique that has gained traction in recent years,[34] but which in computing culture in the early 1980s was not nearly so widespread. Instead, Grosz advocated for the metaphor of a team or partnership to describe the relationship humans and AI systems ought to have— and the kind of systems computer scientists ought to design. Despite the radical potentials of this proposition, however, I argue Grosz maintains some of the same key objectives and assumptions prevalent in AI and multi-agent research. She does not question assumptions of rational agency, nor does she problematize symbolic AI's fundamental goal to represent "the truth" about "the world" in a

[34] Eglash, Ron. "Broken Metaphor: The Master-Slave Analogy in Technical Literature." *Technology and Culture* 48, no. 2 (May 21, 2007): 360–69. https://doi.org/10.1353/tech.2007.0066.

computer system. I consider the limits of this feminist AI protocol, multi-agent or not, without a deeper commitment to feminist epistemologies.

Artificial intelligence technologies capture the imaginations of many around the power and potential of scientific and technological innovation. I hope to root changes in these technologies in their social and historical moments by focusing on the development of one approach in recent AI research. While the particular cultural moments that shaped early MAS research have passed, these decisions reverberate in the values and assumptions inherent in current AI technologies. In the development of MAS approaches, researchers constructed, ignored, and mobilized certain notions and theories about the user, researcher, and AI systems. These particular theories of collectivity and human-AI collaboration enable particular forms of collective action and limit others. In many ways, these metaphors, ideas, and institutional entanglements have set the terms of conversation among computer scientists about what human-AI relations, systems, and societies can and ought to look like.

Chapter 1. "The Vietnamese Don't Live on the Quarter System:" Countercultural Politics at SAIL and SRI

Many of the researchers who would go on to establish multiagent systems as a field of research spent time in Palo Alto in the late 1960s and early 1970s. Many of them did their graduate work at Stanford University; others were hired from MIT to teach in Stanford's newly founded Computer Science department. Most of them did research at SAIL or at the other AI lab in Palo Alto— the AI Centre at the Stanford Research Institute (SRI). At the second DAI conference in 1982, one third of the workshop attendees were current or past researchers at SAIL or SRI.[35] As one of the few institutions with AI research centres in the 1960s and 1970s, Stanford played an important role in training and funding many of the AI researchers in the following decades.

In order to investigate the cultural history of multiagent systems, it is therefore useful to consider a pre-history of the field by examining the research cultures of SAIL and SRI at Stanford. In this chapter I will explore the cultural history of SAIL and SRI and examine how countercultural movements in the Bay Area from 1969-1973 influenced the later emergence of multiagent systems research. Both labs were located off-campus, but their histories, research methods, and values were tightly wound up with events happening on campus. SAIL researchers viewed their lab as an oasis from political causes like the civil rights movement and feminist issues— even when women and civil rights activists were employed inside SAIL's walls. SAIL's culture promoted a vision of collaboration through techno-optimism, either ignoring contemporary political issues or actively coopting their language for their own interests. SRI, the other AI lab at Stanford, was tightly entangled with student politics as the primary target of the April Third Movement in 1969-1971. Student leaders, envisioning SRI as the embodiment of their university's complicity in the Vietnam

---

[35] "Report on the Second Workshop on Distributed AI." Working Paper. AI Lab: MIT, January 1982. https://dspace.mit.edu/bitstream/handle/1721.1/41171/AI_WP_228.pdf?sequence=4.

War and the military-industrial-academic complex, used demonstrations and peaceful protest to remove SRI from Stanford's investments; in 1970 Stanford University sold SRI to itself. Although both SAIL and SRI were funded by DARPA, it was SRI's military research that became the target for some of the largest student demonstrations in Stanford's history. The AI research cultures at SRI and SAIL in the late 1960s and early 1970s were not isolated from political movements in the Bay Area, but tightly and messily entangled with them.

To study the culture of the Stanford AI Lab, I examine the traces of the lab's material culture that remain on SAILDART.org — a site maintained by a former SAIL graduate student, Bruce Baumgart. SAILDART describes itself as "an archive of the first Stanford Artificial Intelligence Laboratory derived from its final backup tapes."[36] It includes past memos, PhD thesis work, technical manuals of SAIL equipment, source code, snapshots of the SAIL building and researchers, and early digital images. Baumgart captions the photos and provides brief descriptions on each page and has included several manuscripts of a history of the lab. On the "Album" page, Baumgart includes links to photos of lab mates playing volleyball, thumbnail portraits of the lab researchers, and a 9-minute slideshow of the SAIL building and its driveway.

---

[36] "SAILDART Stanford Artificial Intelligence Lab DART Archive." Digital archive. SAILDART, May 2012. https://www.saildart.org/.

SAILDART.ORG
Highlights

link	*description*
354 AIMS 75 SAILONS	SAIL A. I. Memos SAIL Operating Notes
68 Theses	1970s PhD Thesis work
10+ Books	that were born digital
UTF-8	The SAIL character set
Arm	robot arm VDS LOU
Collage	99 early digital images
SYSTEM 1974	W.A.I.T.S. Time Sharing System 1974 Reenactment
100+ DMP	executables with source
My book	about SAILDART archive
GEOMED	viewable 3-D models
Music	samples by Andy Moorer
Films	digitized 16mm films
Manuals	both analog and digital
Album	snapshots from the 1970s
1972	**Spacewar** by Stewart Brand
Reunion	Events 2009 and on

SAILDART.ORG
Programmer codes

Top 200 PRG codes w/names

code	*name*
LES	Les Earnest
JMC CLT	John McCarthy Carolyn Talcott
DEK	Don Knuth
JC	John Chowning
LCS	Leland Smith
JAM	Andy Moorer
REG	Ralph Gorin
ME	Marty Frost
BH	Brian Harvey
WD	Whit Diffie
RWG	Bill Gosper
TVR	Tovar
HPM	Hans Moravec
TES	Larry Tesler
PMP	Phil Petit
BGB	Bruce Baumgart

Visitor 1976
Walkabout 2009

SAILDART.ORG
Project areas

code	*description*
DOC	Documentation
CSR	Computer Science
SYS	Systems [* , 3]
HE	Hand Eye
LISP SAIL	Programming Languages
NET	Network Guest
MUS	Music
TeX PUB	Typography & Publication
SUDS	Drawing System
Cart	Autonomous Vehicle

0:00

This is an archive of the ***first***
Stanford Artificial Intelligence Laboratory
derived from its final backup tapes.

This site is maintained by **Bruce Guenther Baumgart**.
My email address is **BgBaumgart at MAC dot COM**.
The most recent full rebuild was May 2012.

Figure 1. Home page of SAILDART.org. Accessed 25 May 2019. http://saildart.org.

Building a lab's culture

Accounts from within the field of multiagent systems call Victor Lesser the "godfather of Multiagent Systems."[37] Lesser organized the first workshop on Distributed Artificial Intelligence (DAI) in 1980 while a professor at the University of Massachusetts. After 22 years of various formats and organizing bodies, the workshop changed its name in 2002 to the International Conference on Autonomous Agents and Multi Agent Systems (AAMAS).[38] Lesser chaired several of the early conferences and AAMAS named their annual dissertation award the IFAAMAS Victor Lesser Distinguished Dissertation Award. After Lesser's retirement, many of his former graduate students have continued to organize and chair AAMAS each year.

Lesser had been thinking about distributed approaches to AI before that first DAI workshop, however. In an article describing the history of multiagent systems at Amherst, Lesser and his former student Dan Corkill describe the influence of the Hearsay-II speech understanding system at Carnegie Mellon University (CMU) on their own work.[39] Lesser had worked on the Hearsay-II project from 1972-1978 as a postdoctoral researcher under Professor Raj Reddy. Lesser and Reddy had already worked together at Stanford University, where Lesser had been a PhD student in Computer Science and Reddy had been a professor of Computer Science until he moved to CMU in 1970. Although Lesser did his doctoral research in a physics lab at Stanford, he had spent the summer of 1967 working with Dr. Reddy at the Stanford AI Project— later renamed the

---

[37] IFAAMAS. "Victor Lesser Distinguished Dissertation Award." Accessed November 5, 2018. http://www.aamas-conference.org/award-victorlesser.html.
[38] "IFAAMAS Home Page." Accessed August 27, 2018. http://www.aamas-conference.org/.
[39] Lesser, Victor, and Daniel Corkill. "Comprehensive History of the Multi-Agent Systems Lab at the University of Massachusetts Amherst 1978-2014," 2016. http://mas.cs.umass.edu/Documents/LabHistory_Web-Article.pdf.

Stanford Artificial Intelligence Lab, or SAIL. Reddy's doctoral advisor was SAIL's founder John McCarthy, popularly known as a "father of artificial intelligence."[40]

From 1965-1991, the Stanford AI Lab was housed five miles away from Stanford's main campus, in the D.C. Power Building. Researchers at SAIL during this period seem to remember the building extremely fondly: they describe the "scenic location" of the building, nestled in the hills south of Palo Alto and near Felt Lake and seem nostalgic about its remoteness.[41] In interviews with researchers who visited and worked at SAIL in the 1970s and 1980s, they describe the feeling that SAIL was its own world. Its lack of nearby restaurants meant that researchers often stayed at the lab for meals and through the night. SAIL's longtime lab manager Lester Earnest takes credit for inventing the first "smart" venting machine, the Prancing Pony: lab members used their three-letter programmer codes to purchase beer and snacks and the bill would be sent to their desks.[42] People described the semi-spherical building, which had been built by G.T.E. in the 1960s and donated to the university, as looking like a spaceship.[43] Many SAIL researchers described the

[40] Childs, Martin. "John McCarthy: Computer Scientist Known as the Father of AI." *The Independent*. November 1, 2011, sec. Obituaries. http://www.independent.co.uk/news/obituaries/john-mccarthy-computer-scientist-known-as-the-father-of-ai-6255307.html.
Metz, Cade. "John McCarthy -- Father of AI and Lisp -- Dies at 84." *Wired*, October 25, 2011. https://www.wired.com/2011/10/john-mccarthy-father-of-ai-and-lisp-dies-at-84/.
Woo, Elaine. "John McCarthy Dies at 84; the Father of Artificial Intelligence." *Latimes.Com*. October 27, 2011. https://www.latimes.com/local/obituaries/la-me-john-mccarthy-20111027-story.html.

[41] Lyman, Richard W. "The Martin Luther King Jr. Crisis." In *Stanford in Turmoil: Campus Unrest, 1966–1972*. Stanford University Press, 2009, 74.. https://doi.org/10.11126/stanford/9780804760799.001.0001.

[42] Baumgart, Bruce. "Stanford Artificial Intelligence Laboratory Prolegomenon," January 2017. Stanford University Libraries Department of Special Collections and University Archives. purl.stanford.edu/dr245hh6464.

[43] Baumgart.

unusual shape of the building to support their reminisces that they were on the cutting edge of scientific research— just like the astronauts and NASA scientists on Project Apollo.

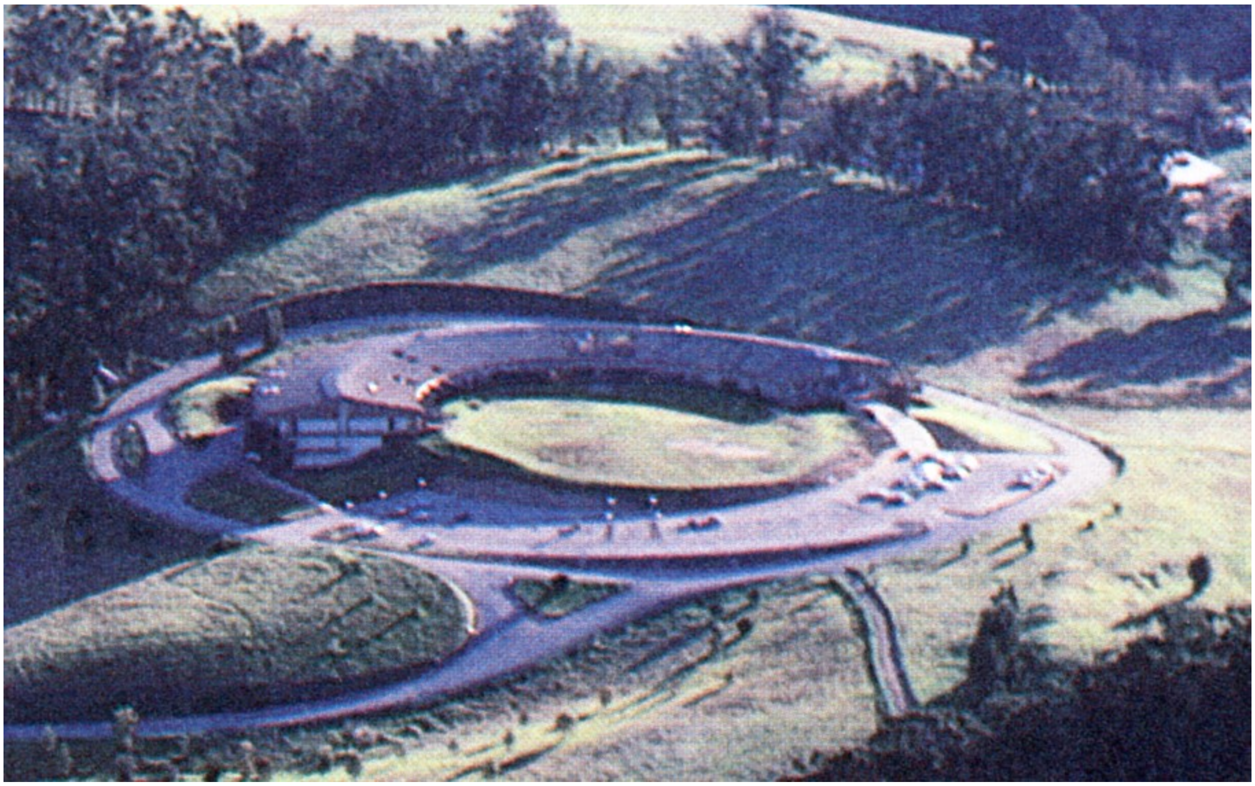


Figure 2. Photo of the D.C. Power Building. Accessed 25 May 2019. https://www.saildart.org/simple/index-book-simple.html.

In 2009 Lester Earnest made a web page called "SAIL Sagas" where he gathered stories lab members had emailed to him for their recent SAIL reunion.[44] Multiple people wrote in about their memories of the Building. Many of the interviews and emails he gathered from the mid-2000s expressed resentment toward Stanford for supposedly not having realized in the 1970s how special SAIL and its building were. In 1980, when Stanford moved SAIL onto the main campus, many researchers saw the move as the end of an era. Many emails from the 2009 web page tell stories about a grass fire that nearly burned down the building in 1986. One email writes:

> Former SAIL people may recall that rows of Blue Gum Eucalyptus trees used to line both sides of Arastradero Road, making it a beautiful country lane beneath the overhanging foliage. People who attend the Walkabout will notice that the trees have now largely disappeared. That was a result of the runaway fire that was started just downhill from the DC Power Lab on the afternoon of July 1, 1985, evidently by a well-known firebug who was seen enjoying the show shortly after it started. However nobody saw him light it. The Lab was occupied at that time by the Center for Computer Research in Music and Acoustics (CCRMA) which had been abandoned there in 1979 as a result of a disgusting political maneuver when SAIL moved back to campus.
>
> It was a very dry summer and the eucalyptus trees by the road turned into torches. With help from a strong west wind the crown fire swept down Arastradero, traveling a half mile to Liddicoat Drive, near the freeway, in a few minutes and burning down a dozen homes. CCRMA was also threatened and when the people there promptly broke out the fabric fire hoses they were found to be full of holes, causing major water leakage, with nothing much coming out the end. This was a typical result of Stanford's failure to maintain the building. Fortunately there were enough people around to put their hands tightly over the holes, reducing the leakage to the point where they could fight the fire. They successfully defended the building but it was a close call.[45]

The author describes Stanford's "failure to maintain the building" and their "disgusting political maneuvers" which had led to SAIL being moved to a new building. The author implies Stanford took the building—and thus the research that happened there—for granted. At this point

---

[44] Earnest, Les. "SAIL Sagas." Les Earnest, December 13, 2009. https://web.stanford.edu/~learnest/spin/sagas.htm.
[45] Earnest, "SAIL Sagas."

in time, however, much of SAIL's research had already been moved to the Stanford campus: only the Center for Computer Research in Music and Acoustics (CCRMA) still used the DC Power Lab. Many of the SAIL lab members seem to have been in the DC Power building, though, to remember this story and defend the building from being destroyed. It was not the actual research at risk of being burned, then, but something else: an object that symbolized the culture of SAIL as an institution.

In that same collection of emails, another author wrote about Spacewar, an early video game created at MIT but played frequently by SAIL's lab members. The author writes:

> Meanwhile a company called Atari was formed independently to convert Spacewar into a commercial video game but Bill Pitts, with help from SAIL colleagues Ted Panofsky and Phil Petit, beat them to it, putting it into the Stanford coffee shop and a local bowling alley. Bill called it "Galaxy Game" because the term "war" was a very unpopular on campus, which was deep in an anti-war movement regarding Vietnam. In fact, the DC Power Lab, where SAIL was located, got firebombed in an unoccupied room in the 1970s, probably by protesters who had learned that most of our funding came from the Defense Department. This resulted only in water damage as the sprinkler system went off and doused the Molotov Cocktail. While the Galaxy Game was quite popular, Atari observed that their version of Spacewar was expensive to reproduce and somewhat hard for people to learn, so they instead introduced the game of Pong which was cheap to make, easy to understand, and a great commercial success. Thus they ate Bill's lunch, alas.[46]

The author seems most excited about the fact that SAIL members managed to distribute their version of the game before a company could turn it into a commercial game and profit from it. This ethos resonates with a computing hobbyist ethos, gaining traction in the Bay Area in the 1960s and 1970s. In the author's excitement about telling *this* story, however, they gloss over a political

---

[46] Earnest, "SAIL Sagas." SAIL and Spacewar are famously discussed in Brand, Russell. "SPACEWAR - Fanatic Life and Symbolic Death Among the Computer Bums." *Rolling Stone*, December 7, 1972. For in-depth discussion of Russell Brand, SAIL, and Silicon Valley, see Turner, Fred. *From Counterculture to Cyberculture: Stewart Brand, the Whole Earth Network, and the Rise of Digital Utopianism.* Chicago, Ill.: University of Chicago Press, 2008.

context that seems to have been much more physically felt—the anti-Vietnam movement on campus. The author brushes aside students' discomfort with playing games with "war" in the title, even when protestors had tried to set SAIL's building on fire. This passage suggests a lab atmosphere relatively unconcerned with popular social movements that surrounded the lab and its research.

"Zoe SAIL 1971"

This research at SAIL was happening at the same time as the second-wave feminist movement in the 1960s and 1970s. Estelle Freedman, a Stanford historian of feminist movements in the US, describes feminist consciousness raising groups that formed at Stanford and a socialist collective called the Women's Union, created around 1969.[47] The collective taught Stanford's first Women's Studies courses at the women's centre they created at the YWCA. A Stanford Daily article from 1969 features an article written by Jane Franklin, "one of the sisters of the Women's Liberation movement;" it features a photo of WITCH, the Women's International Terrorist Conspiracy from Hell, and invites readers to attend the Women's Liberation movement meeting that night.[48] A front page article from 1973 describes various initiatives to enrol women in Stanford's engineering department, including "Consider the Possibility," a pamphlet for women considering engineering as incoming students, and WISE, the newly formed campus group for women in science and engineering.[49] The article's photo caption reads, "Pam Cosby exemplifies the 'new breed' of engineering students. Pam, a sophomore, says the biggest problem is convincing

---

[47] Freedman, Estelle. "Women at Stanford: Inclusion, Exclusion, and Activism from the 1890s to the 1990s." presented at the Stanford Historical Society, Stanford University, March 7, 2019. https://www.youtube.com/watch?v=CPH3ZDUrL7A.
[48] Franklin, Jane. "Call for Equality: Join the Women's Liberation." *The Stanford Daily*, October 7, 1969. The Stanford Daily Archives.
[49] Jacobs, Joanne. "Engineering School Seeks to Enrol More Women." *The Stanford Daily*, February 5, 1973. The Stanford Daily Archives.

people that she can do the work that is required of her." It seems the "new breed" of engineer the article refers to is women with long blonde hair.

—Daily photo by Helen Boyer

WOMEN IN ENGINEERING? — Pam Cosby exemplifies the "new breed" of engineering students. Pam, a sophomore, says the biggest problem is convincing people that she can do the work that is required of her.

Figure 3. Photo from Jacobs, Joanne. "Engineering School Seeks to Enrol More Women." ***The Stanford Daily*****, February 5, 1973. The Stanford Daily Archives.**

Call For Equality

Join The Women's Liberation

By Jane Franklin

(This article was submitted by Jane Franklin, one of the sisters of the Women's Liberation movement. They will be holding a meeting tonight at 7:30, room 270 Tresidder, to meet with anyone interested in working with the Women's Liberation movement.)

"Flashlight in hand, wearing high heels and a sedate but stylish black dress, Amelia J. Gordon stepped briskly into the dingy dimness of a bar. Just behind her were two unsmiling young men with carbines. Another half dozen paces back were two more Filipinos with rapid-firing grease-guns. A pair of U.S. military policemen moved in with them. Four other armed guards leaped from a Jeep and posted themselves around the bar's entrance."

With ten armed guards, the mayor of "sin city" entered one of the 300 night spots in "an acre of frenzy" next to the U.S. Naval Base at Subic Bay in the Philippines. She had come to measure the skirts of the "hostesses." "We don't allow extremely short miniskirts here," Mrs. Gordon explained. "It is obscene the way some girls dress. We must not forget our traditional Filipino modesty." (S.F. Examiner & Chronicle, Aug. 17, 1969.)

Able to get a "man-sized" job? Yes. Paid more rather than less than most men? Yes. Looked upon as an equal by most men? No; feared by most men around her because she has the power to throw them in jail—or worse. Liberated? No; oppressed by her own narrow view of humanity, she has equality in oppression with Madame Nhu, Madame Chiang Kai-shek, Pat Nixon, Happy Rockefeller, and other sisters of Marie Antoinette.

cell with two prostitutes whose unemployed husbands were at home looking after their children and with no money coming in since the bread-winners were "safely" behind bars. Those women love their children as I love mine, and neither my husband nor theirs—male-supremacist as they may be—is the cause of that threatening starvation. But a large number of the men and women who do perpetuate starvation—International Imperialist Industrialists—were banqueting that evening at the Fairmont.

My mind spun around with the desire for women's liberation—liberating those women from behind those bars. But when I was leaving, all I said was, "I hate to leave you here," while they said, "It's all right, Baby." It isn't allright, Baby, and it won't be allright until our whole thing is right.

Join People's Revolt

Although we have just begun the struggle, more and more women are realizing that the only way we are going to be liberated is by changing the profit motive system to the human motive system. We will organize working women, unemployed women, welfare women, housewives, women students to rise up angry alongside our oppressed brothers, linking our liberation struggle to that of all the people.

A Vietnamese woman, arrested during a secret mission, refused to answer her torturers. "Who are your comrades? Who is your husband? Where are your children?" they asked. They killed her for her silence, but on the prison wall they must have seen her answer:

> **A rosy-cheeked woman, here I am fighting side by side with you, men.**
> **On my shoulders weighs that hatred which is common to us.**
> **The prison is my school, its mates my friends,**
> **The sword is my child, the gun my husband.**

Figure 4. Franklin, Jane. "Call for Equality: Join the Women's Liberation." ***The Stanford Daily*****, October 7, 1969. The Stanford Daily Archives.**

The feminist organizing on Stanford's campus was seemingly far from the minds of SAIL researchers in 1971. On the SAILDART.org home page, a hyperlink at the bottom links to something called "Zoe pictures 1971." The link leads to a web page with 26 greyscale pictures of a woman, naked and smiling for the camera in different poses.[50] Many of the images are zoomed or cropped to the woman's waist as she poses against a blank wall. Others include her whole body as she flips her hair and looks coyly at someone behind the camera, leaning against what looks like a piece of computer equipment. There is no information about who this woman is, why she is in the lab, why she is naked, and whether Zoe is even her real name, or why these photos are in the SAIL archive. The only clue is in the title of the page: "Zoe at SAIL - March 1971."

Baumgart does, however, provide a brief description of the technical aspects of the photos on the page. His caption reads, "Raw television images, digitized by a PDP-10 computer with four bits of grey scale, (so missed data words slew by nine pixels) and recorded to disk files around 2am on Monday 8 March 1971." These images are placed without any social context or explanation, featured on in the SAIL archive to demonstrate a technical feat or a method of digitization. The sanitized language and the number of almost identical images on the page make the technical details of the image the clear intended centre of focus.

Elsewhere in the SAILDART, clues provide more context for Zoe and her pictures. In the last message sent on SAIL's timesharing system on June 7, 1991, someone sent an email with the subject line "Subject: life as a computer for a quarter of a century."[51] The text of the email is an essay titled "TAKE ME, I'M YOURS: The autobiography of SAIL." The essay, written from the

[50] I will not include the pictures in this text.
[51] For the full text of the essay, see "SAIL Farewell." Accessed May 24, 2019. http://infolab.stanford.edu/pub/voy/museum/pictures/AIlab/SailFarewell.html.

perspective of the lab itself, is only signed "SAIL." According to one of Baumgart's footnotes, however, it seems likely the essay was ghost written by longtime lab manager Les Earnest.[52] The SAIL message says the essay was sent to 875 email addresses. Today, the essay's unabridged text is archived on Baumgart's website SAILDART and available in the online holdings of the Computer History Museum.[53]

The essay, written for the timesharing system's 25th birthday, describes the beginnings of SAIL and reminisces about the people, adventures, technological systems, and inventions over the years within the walls of the DC building. The essay is written tongue-in-cheek, with many teasing jokes about various lab members. One of the sections, titled "Sex," tells the story of a few (unnamed) SAIL researchers taking a class in abnormal psychology who supposedly needed a term project. According to the essay, they had the idea to make a film about a woman who was sexually attracted to computers. They posted an advertisement in the Stanford Daily newspaper asking for an "uninhibited female" and invited her to the SAIL building after midnight, after telling the lab manager Les Earnest the computers needed to go down for maintenance. They asked the woman—described as "their budding starlet" in the essay— to take off her clothes and photograph her posing with the lab's tape drives and other equipment. From the essay, it sounds like many SAIL researchers were involved: "Other students who were in on this conspiracy remained in other parts of my building, but I catered to their voyeristic [sic] interests by turning one of my television cameras on the action so that they could see it all on their display terminals." The essay describes

---

52 Baumgart, Bruce. "Stanford Artificial Intelligence Laboratory Prolegomenon," January 2017. Stanford University Libraries Department of Special Collections and University Archives. purl.stanford.edu/dr245hh6464.

53 Wiederhold, Voy. "Computer History Exhibits." Accessed May 28, 2019. http://infolab.stanford.edu/pub/voy/museum.html.

one student who took off all his clothes when he entered the room with the woman— "in order to avoid disrupting the mood."[54]

Between February and April 1971, 10 advertisements were posted in the Stanford Daily requesting "uninhibited" women. Four of the ads were posted in April by a man named Steve offering $3,000—equivalent to nearly $20,000 in 2019—for a "female, attractive, uninhibited." The other six posts were posted earlier—first on Feb 26, 1971 and then every day from March 1-5—requesting an "uninhibited girl for experimental film" to call Chuck, Rick, or Mike and offering $15/hour and $50 minimum (approximately $315 in 2019). The SAIL researchers who recruited Zoe could have posted either of these ads, but given the date the Zoe images were supposedly uploaded, and the likely budget of students, it is likely this earlier set of advertisements.[55] Even so, both ads offer what seems to be large sums of money for students to be able to offer for a class project. According to the essay, after they posted the ad for an "uninhibited" woman they received two responses—"that was in the liberated early 70s"—but after interviewing one of the women, "They decided that one of them was too inhibited."[56] This line in the essay suggests two things. First, it is precisely this culture of sexual humour— about women being willing or unwilling to engage in sex acts— which was such a major target in the reports women wrote in the 1980s and 1990s about harassment, alienation, and exclusion in their computer science and engineering

[54] Baumgart, "Stanford Artificial Intelligence Laboratory Prolegomenon," 240.
[55] The earliest available SAIL directory on SAILDART, from 1975, lists several researchers and graduate students named Steve, Chuck, Rick, and Mike.
"Make $3,000 by Jply If You Are Female, Attractive, Uninhibited." Newspaper advertisement. *The Stanford Daily*, April 1, 1971. The Stanford Daily Archives.
"Wanted: Uninhibited Girl for Experimental Film." Newspaper advertisement. *The Stanford Daily*, February 26, 1971. The Stanford Daily Archives.
[56] SAIL, "SAIL Farewell."

departments.[57] Second, this anecdote suggests a conflation of women's liberation and sexual liberation of the 1970s. In this story, "liberated" women seem to be women willing to pose naked for them in their lab. The woman unwilling to pose was declared "not liberated enough." This points to a fundamental misunderstanding of women's agency, consent, and the broader goals of feminist movements. Zoe may have happily posed for the photos as a liberated woman, comfortable with her body— but *that* is what makes it feminist, not the consumption and management of the pictures by men.

$3.50 TO INTERNATIONAL
FISHERMAN OPPORTUNITIES
BOX 12822. SEATTLE, 98111

Make $3,000 by Jply if you are female, attractive, uninhibited. Call Steve 8 to 5. 964-8678

OLSTEN TEMPORARY

Work 8 to 40 hours a week and spplement your income. Clerical. All skills.

Figure 5. "Make $3,000 by Jply If You Are Female, Attractive, Uninhibited." ***The Stanford Daily*****, April 1, 1971. The Stanford Daily Archives.**

[57] Spertus, Ellen. "Why Are There so Few Female Computer Scientists?" MIT Artificial Intelligence Laboratory Technical Report. Cambridge, Massachusetts: MIT, 1991. https://dspace.mit.edu/bitstream/handle/1721.1/7040/AITR-1315.pdf?sequence=2.

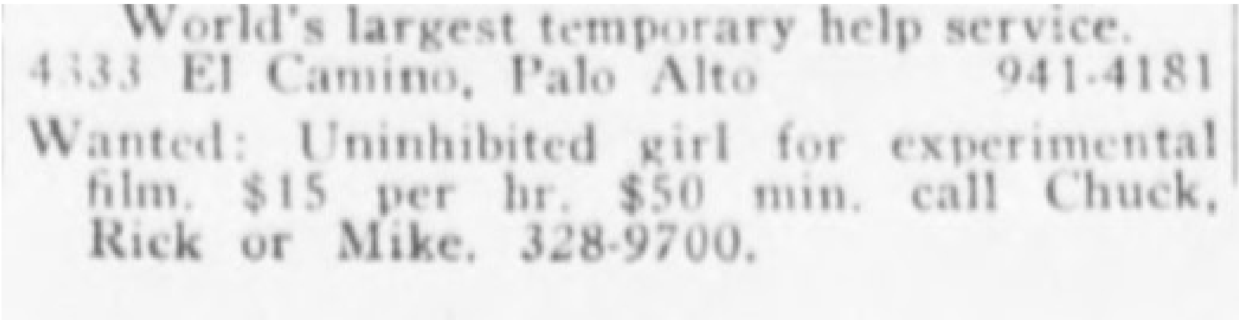

World's largest temporary help service.
4333 El Camino, Palo Alto 941-4181
Wanted: Uninhibited girl for experimental film. $15 per hr. $50 min. call Chuck, Rick or Mike. 328-9700.

Figure 6. "Wanted: Uninhibited Girl for Experimental Film." ***The Stanford Daily*****, February 26, 1971. The Stanford Daily Archives.**

After taking the first round of pictures, it seems the SAIL researchers became bored with the woman statically posing with their gadgets. "After a number of boring shots of this young lady hanging on to me while I rotated, the filmmakers set up another shot using one of my experimental fingers. It consisted of an inflatable rubber widget that had the peculiar property that it curled when it was pressurized. I leave to your imagination how this implement was used in the film. Incidentally, the students reportedly received an 'A' for their work."[58] If this story is true, it suggests these SAIL researchers asked a woman to come to their workplace, take off all her clothes, and engage in sex acts with their equipment while others were doing research—and then submit the images from the event to a class, where the professor approved. Even if not fully true, the fantasy enough is telling—and damning.

Regardless of whether the incident actually happened the way it is memorialized in the newsletter, nearly 50 years later the same version of the story continues to circulate and be remembered fondly. One could argue—hypothetically—that in 1971, it was still the early days of computer science, with fewer women at SAIL and less explicit discussion about the negative

[58] SAIL, "SAIL Farewell."

experiences of women in science and engineering departments. Organized efforts to diversify and improve computer science, like Anita Borg's "Systers" mailing list, were not until the late 1980s.[59] Maybe the researchers at SAIL later realized how the "Zoe" images did not create a lab environment that encouraged women to feel welcome and participate in research. But the essay that describes the context of Zoe's visit in such tactless detail was sent out in 1991, not 1971. It wasn't shared privately, either—it was sent to 875 email addresses on the SAIL mailing list. Even as recent as 2018, Baumgart has included the story as a chapter in his forthcoming coffee table book on the history of SAIL.[60] He features a nine-page excerpt of the essay in the prolegomenon of the book. In the paragraph introducing the "SAIL Farewell" essay, he writes in italics: "*Or you may fast forward to figure 1.5 on page 11 to study early digital pictures of a nude female*."[31] Although the images are still publicly available on SAILDART, the original essay emailed to the network does not include photos of the woman. Baumgart has made it even easier for his readers to see the pictures. In his book he includes four photos of Zoe, running her hands through her hair, squeezing her breasts, and posing with SAIL's equipment. At least in some respects, there seems to be little remorse about "Zoe Pictures 1971."

These "Zoe" images from 1971 foreshadow the circulation and controversy around the "Lena image," a centrefold from *Playboy* cropped, digitized, and transformed into popular test image in image processing. In 1972 a Swedish woman named Lena Söderberg posed in the centrefold for Playboy's November issue. As the legend goes, in 1973 a group of engineers at the University of Southern California were looking for a new test image for their conference paper.[61] They found

[59] "Systers - AnitaB.Org." Accessed June 28, 2019. https://anitab.org/systers/.
[60] Baumgart, "Stanford Artificial Intelligence Laboratory Prolegomenon."
[61] Munson, Jr., David C. "A Note on Lena." *IEEE Transactions on Image Processing* 5, no. 1 (January 1996): 1.

the image, cropped it to just below her shoulders, digitized it, and used it as an image in for their upcoming paper. Other attendees at the conference liked it so much, they started using it as a test image for their own papers. The image became immensely popular, turning into an industry standard for image processing . It remains one of the most popular digital test images of all time.[62] Unlike the Lena image, the "Zoe" images do not seem to be commonly used in computer science publications (although many of the SAIL reports from the 1970s are not publicly available, so it is possible these images were used in various ways SAIL). The two images share similar stories, however: they were both images cropped and removed from their social context; digitized and shared to demonstrate an engineer's skill; shared in the same masculine computing cultures; and left to circulate on computers, online, and in the world for decades to come.

The story of Zoe at SAIL offers a glimpse into the institutional culture of one of AI's most famous labs during what many consider its "golden years." Stanford has had twenty-seven affiliates awarded the ACM Turing Award, more than any other university—sixteen of those winners have been affiliates of SAIL.[63] Many so-called pioneers of robotics, artificial intelligence, personal computing, and internet technology were working at SAIL in 1971 when Zoe visited the lab. 1971 was the year the lab's director, John McCarthy, won the ACM Turing Award.[64] Marvin Minsky and Donald Knuth, both SAIL affiliates, won their Turing awards in 1969 and 1974,

---

[62] Hutchinson, Jamie, "Culture, Communication, and an Information Age Madonna," *IEEE Professional Communication Society Newsletter* 45, no. 3 (May/June 2001): 1.
Mulvin, Dylan, and Jonathan Sterne. "Scenes from an Imaginary Country: Test Images and the American Color Television Standard." *Television & New Media* 17, no. 1 (2016): 21–43. https://doi.org/10.1177/1527476415577211.
[63] For list of all ACM Turing Award winners, see "A.M. Turing Award Winners by Year." Accessed November 6, 2018. https://amturing.acm.org/byyear.cfm.
[64] "John McCarthy - A.M. Turing Award Laureate." Accessed April 20, 2019. https://amturing.acm.org/award_winners/mccarthy_1118322.cfm.

respectively. Minsky, who later co-founded both the MIT AI Lab and MIT Media Lab, was a known associate of convicted sex offender and billionaire Jeffrey Epstein.[65] According to a deposition unsealed in 2019, one of Epstein's victims was forced to have sex with Marvin Minsky as a teenager on Epstein's island in the Virgin Islands.[66] This testimony, as well as the continued circulation and celebration of Zoe's story, demonstrates that, in certain spaces, the "boy's club" of AI in the 1960s and 1970s continues.

The Other AI Lab: SRI and the Anti-War Movement

On May 16, 1969, students from Stanford University's April Third Movement (A3M) protested SRI's Counterinsurgency Centre. Protestors blocked entrances to the SRI parking lot and all four lanes of Page Mill Road for four hours the morning of May 16, 1969.[67] Protestors carried signs reading "STOP SRI," "RESEARCH LIFE NOT DEATH," and "SRI KILLS." They handed out flyers depicting SRI as synonymous with American imperialism, and pamphlets explaining the relationship between Stanford, SRI, and the US Department of Defense. In May 1969 they staged several demonstrations to shut down research at SRI and pressure the university to cut its ties with the institute.

---

[65] Mervis, Jeffrey. "What Kind of Researcher Did Sex Offender Jeffrey Epstein like to Fund? He Told Science before He Died." *Science | AAAS*, September 19, 2019. https://www.sciencemag.org/news/2019/09/what-kind-researcher-did-sex-offender-jeffrey-epstein-fund-he-told-science-he-died.

[66] Brandom, Russell. "AI Pioneer Accused of Having Sex with Trafficking Victim on Jeffrey Epstein's Island." *The Verge*, August 9, 2019. https://www.theverge.com/2019/8/9/20798900/marvin-minsky-jeffrey-epstein-sex-trafficking-island-court-records-unsealed.

[67] "Notes from the Protest Outside SRI," May 16, 1969. April Third Movement Archives. http://www.a3mreunion.org/archive/1968-1969/68-69_may16_sri/files_68-69_may_16/A3M-5-16_Notes_1.pdf.

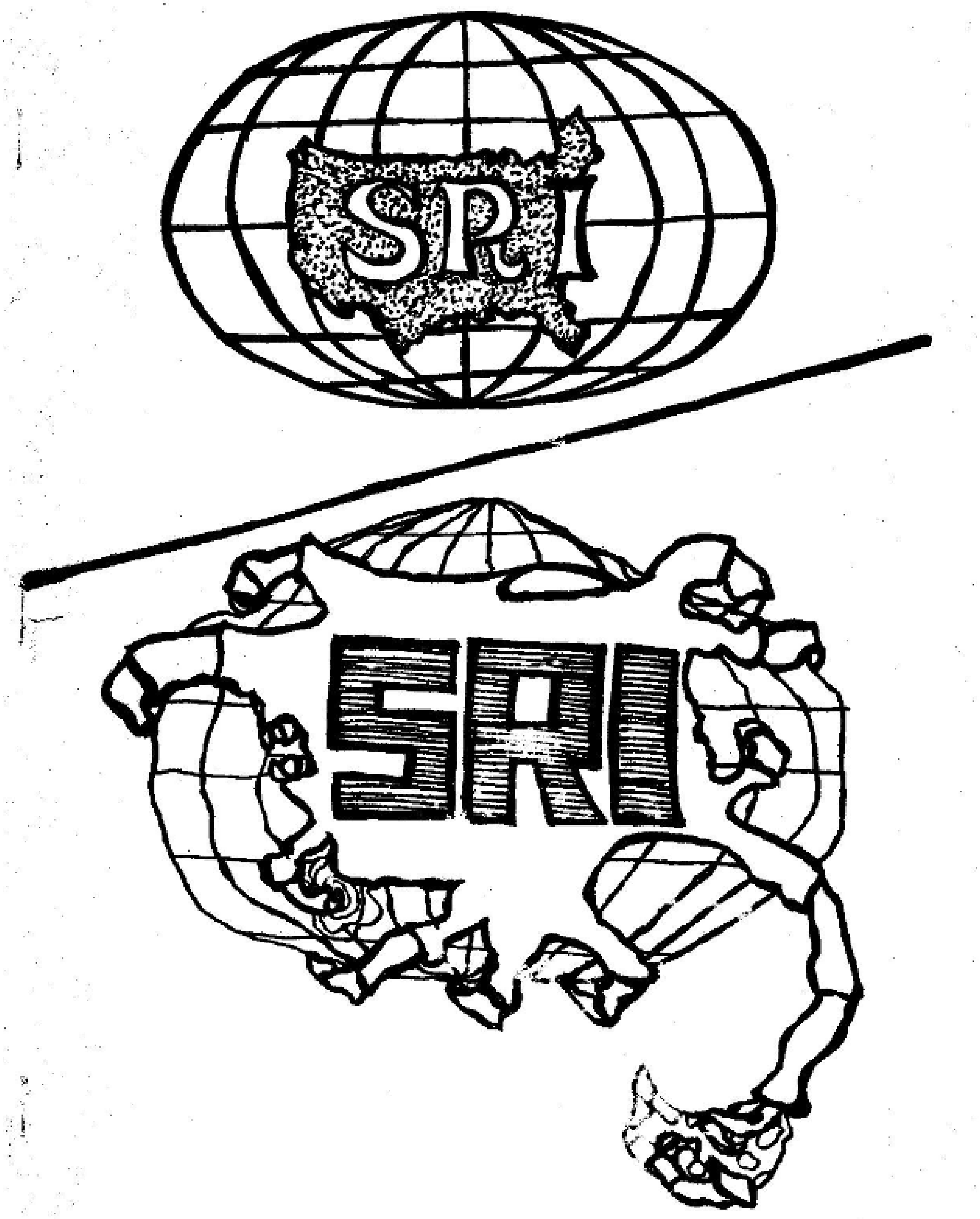


Figure 7. Front page of "SRI/SRI" pamphlet. April Third Movement. April 12, 1969. April Third Movement Archives.

The Stanford Research Institute was not always depicted in student flyers as the clawed fingers of the US military, sucking the earth dry and belching fumes into the atmosphere. In *The Cold War and American Science* Stuart W. Leslie provides a history of the military-industrial-academic complex at MIT and Stanford University.[68] This "golden triangle," as Leslie describes it, of academic research departments, high-tech industry, and military agencies usually funded research in engineering and the physical sciences in the mid-twentieth century.[69] He describes the military's contracts with physics departments and how it created a new kind of research programme, "a new kind of physics that easily blended theory and experiment, science and engineering, understanding and application, unclassified and classified, research and policy advocacy."[70] A similar argument can be made about the computer science programmed funded by military contracts. Indeed, the histories of military-funded research in physics and computer science at SRI are shared: before the AI centre at SRI was officially founded in 1966, SRI's activities in AI grew out of their Applied Physics Laboratory in 1957, where they researched learning machines and self-organizing systems.[71] SRI struggled to make profits in its first decade as a nonprofit research institute from 1946-1955; under new management, however, SRI began to secure lucrative R&D contracts with the Navy, Air Force, the Army, and the Atomic Energy Commission (AEC).[72] By 1965, 82 percent of SRI's revenues came from government contracts,

---

[68] Leslie, Stuart W. *The Cold War and American Science: The Military-Industrial-Academic Complex at MIT and Stanford.* New York: Columbia University Press, 1993.
[69] Leslie, 2.
[70] Leslie, 159.
[71] "AI Center :: AIC Timeline." Accessed May 27, 2019. http://www.ai.sri.com/timeline/.
[72] Leslie, 243.

with military contracts accounting for 78 percent of that; by 1968, SRI had 1,500 staff members and annual contract revenues of $64 million— half of which came from military contracts.[73]

In 1967 a radical group at Stanford called The Experiment published a detailed account of war research on campus. In Leslie's telling, at the top of their list was SRI—"which, for many student radicals, symbolized the military's presence on campus."[74] Some of SRI's more controversial research included "land reform in Vietnam, counterinsurgency surveillance in Thailand, and chemical weapons."[75] On April 3, 1969, a group of hundreds of Stanford students formed the April Third Movement. Their demonstrations included staging a nine-day sit in at Stanford's Applied Electronics Laboratory; an occupation of the administration building; and creating human blockades outside SRI.[76] Ultimately, the students succeeded: after a year of sit-ins, protests, informational sessions, and being tear gassed by the police, Stanford sold SRI to itself for one percent of its gross operating revenues and SRI was renamed SRI International.[77] Stanford's two AI research labs were bifurcated in part because of the work these student organizers did to convince Stanford to cut ties with SRI.

---

[73] Leslie, 242.
[74] Leslie, 242.
[75] Leslie, 243.
[76] "About the Stanford Movement | April Third Movement." Accessed June 30, 2019. http://www.a3mreunion.org/about-1.html.
[77] Leslie, 247.

Figure 8. Flyer for Stanford SRI demonstrations. April Third Movement. May 15-16 1969. April Third Movement Archives.


FREEDOM
not
LICENSE
SRI
KILLS

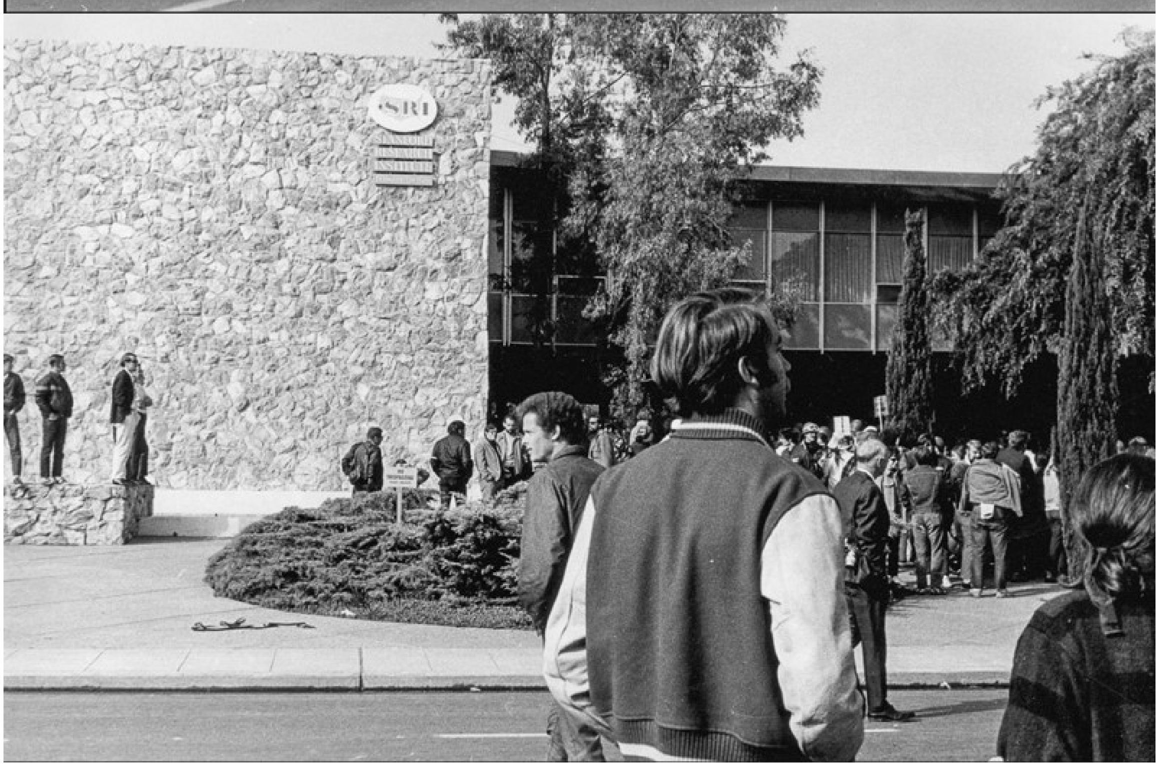

SRI

Figure 9. Photos from A3M's Demonstration at the SRI Counterinsurgency Center, May 16 1969. Accessed May 1 2019. http://www.a3mreunion.org/archive/photos/1968-1969_photos/sri/sri_hanover/index.html.

When considering this history in relation to the development of multi-agent systems research, it is worth noting the extent to which the military-industrial-academic complex at Stanford shaped Victor Lesser's experience there. SAIL and SRI were both funded extensively by DARPA in the first decades of their existence. Victor Lesser spent a summer working at SAIL with Raj Reddy, but he did his doctoral research in Stanford's Linear Accelerator Lab (SLAC). In Leslie's history, he describes Stanford's linear accelerator, a full 2 miles long and taking $114 million to create, as the largest, most powerful and most expensive scientific instrument of its day when it was built in 1966.[78] Leslie argues it "made Stanford a world centre of high-energy physics" and describes it as "perhaps the most visible symbol of Stanford's push to postwar prominence."[79] Additionally, Victor Lesser's advisor in the linear accelerator lab was William Miller, who served as Stanford's vice president and provost in the late 1960s and 1970s while he led SLAC's Computation Group.[80] Richard Lyman, the provost of Stanford from 1967-1970 and Stanford's president from 1970-1980, wrote a memoir called *Stanford in Turmoil: Campus Unrest, 1966–1972,* with chapters about SRI, anti-Vietnam protests, and the civil rights demonstrations during his tenure.[81] Lyman thanks Miller several times throughout his book. Miller, Lyman writes, "served as acting president during my absences, notably when the hospital sit-in took place and then during my Danforth leave of absence. He promised when I started that leave to preserve radio silence while I was gone, and he did so."[82] It seems Lyman and Miller made a great team in silencing student movements at Stanford— until Miller left to become president of SRI International in 1978.[48] Lesser's politics do not necessarily match those of his advisor, but their collaboration suggests complicity at the very least.

Conclusion

A 1974 issue of the *Stanford Daily* includes an announcement about a weekly "Computer Freak Potluck Dinner," hosted at the Stanford AI Lab.[83] The post is sandwiched between notices for the Black Prelaw Society, the Chicano Prelaw Society meeting, and Consciousness Raising

---

[78] Leslie, 160.
[79] Leslie, 160.
[80] "Department Timeline | Stanford Computer Science." Accessed May 23, 2019. https://cs.stanford.edu/about/department-timeline.
[81] Lyman, Richard W. *Stanford in Turmoil: Campus Unrest, 1966–1972*. Stanford University Press, 2009. https://doi.org/10.11126/stanford/9780804760799.001.0001.
[82] Lyman, 204.
[83] "Announcements." *Stanford Daily*. November 6, 1974. The Stanford Daily Archives.

Groups at the Stanford Women's Centre. This post—two lines on the back of a Stanford student newspaper—illustrate the way the history of AI research is deeply embedded in the geographical and social histories of the Bay Area in the 1960s and 1970s. This chapter has tried to disentangle some of the threads that knit the story of multiagent systems AI research together.

When SAIL was in the DC Power building and in the years after they moved, SAIL's members imagined themselves as cloistered away in a spaceship, far from Stanford's main campus and the messy political issues happening on campus in the late 1960s and 1970s. These labs, however, were far from the apolitical oasis they described. While SRI certainly received the brunt of the protests, Stanford students demonstrated at both SAIL and SRI for their DARPA funding and involvement in Vietnam. Despite growing numbers of women scientists and engineers in the 1970s, the institutional cultures were slow to change— especially the more prestigious labs like SAIL. As I will explore in the next chapter, women doing AI research in the 1970s felt the need to develop their own social practices and approaches to research to create a space for themselves.

Chapter 2. Critical Mass: Protocol Feminism in AI Research

In 2002, Martha Pollack co-authored a report called "Becoming a Computer Scientist" with five other women on the Association of Computing Machinery (ACM) Committee on the Status of Women in Computer Science. The authors describe the difficulties women faced in entering the field of computer science, like cultural barriers to the "hacker elite" system; safety concerns about staying in the department at night; diminished self-esteem; lack of mentors and role models; gender discrimination; and difficulties balancing career and family responsibilities.[84] They concluded with eleven recommendations: these included efforts to provide role models, professional experience and research opportunities for young women; safe, 24-hour access to public terminal areas for students and faculty; and grants so that women may purchase terminals, workstations, printers and modems for home use.

From a perspective of contemporary intersectional feminism, Pollack's report leaves much to be desired: it assumes a white, straight, cisgendered, child-bearing woman as its subject and leaves class, institutional privilege, and immigrant status wholly unmentioned. The committee at least acknowledges some of the perspectives they exclude: "We have not, for example, addressed problems unique to women in industrial computer science, nor have we considered how the problems we have described are exacerbated for women of colour or disabled women. We believe it is important that these issues be examined, but we leave them for other articles."[85] They also critique the assumed academic model they call the "helpmate-in-the-background," which they argue works against women without children and men with employed wives in addition to working

[84] Pearl, Amy, Martha E. Pollack, E.A. Riskin, Becky Thomas, Elizabeth Wolf, and Alice Wu. "Becoming a Computer Scientist." *SIGCSE Bulletin* 34, no. 2 (June 1, 2002): 135–43. https://doi.org/10.1145/543812.543847.
[85] Pearl et al, 142.

against women with children: “few of these people have such helpmates.”[86] It is important to remember this report, nearly 20 years old, came before the many widespread critiques of the leaky pipeline metaphor and liberal feminism-via-Sheryl Sandberg— and it is a testament to that fact that, unfortunately, feminist issues have taken a long time to be acknowledged and prioritized in computer science.

Ten years before Pollack’s work, Barbara Grosz led a committee to report on the status of women who had already *become* computer scientists. The 1991 report— referred to as the Grosz Report, according to the *Harvard Crimson*— described sexual harassment, inadequate childcare, and other problems affecting graduate students, junior faculty, and other women in science at Harvard.[87] “Harvard University must live with its times,” the first line of the report commanded.[88] Grosz and the fourteen members of Harvard’s Standing Committee on the Status of Women interviewed women graduate students and junior faculty across Harvard’s eleven science departments, discussing hostile department environments, lack of mentorship from senior (male) faculty, social isolation, sexual harassment, and massively inadequate resources for child care, parental leave, and department scheduling for scientists with children. Grosz and her committee made recommendations for the university to implement at the dean, department, and individual faculty level. The suggestions included steps for clearer policies about maternity and parental leave; recruiting more women graduate students and junior faculty; and including women in informal ways like including them in departmental lunches, introducing them to visiting speakers,

---

[86] Pearl et al, 141.
[87] Wilde, Anna D. “TFs Will Be Tested On Spoken English | News | The Harvard Crimson.” *Harvard Crimson*. May 21, 1993. https://www.thecrimson.com/article/1993/5/21/tfs-will-be-tested-on-spoken/.
[88] Grosz, Barbara J. “Report on Women in the Sciences at Harvard. Part I: Junior Faculty and Graduate Students.” Faculty of Arts and Science Committee on the Status of Women. Cambridge, Massachusetts: Harvard University, February 13, 1991, 2.

and nominating them for leadership positions. The committee was unequivocal: "The Faculty of Arts and Sciences must make a commitment to the recruitment, retainment, and professional development of women graduate students and junior faculty in the sciences, or the pool of women for senior science faculty positions will not be significantly greater in the twenty-first century than it is now."[89]

Department	Students	Female Pop.	Female WD	% WD/P	Male Pop.	Male WD	% WD/P
DAS	286	46	4	8.70%	240	17	7.08%
Astronomy	40	5	1	20.00%	35	2	5.71%
Biochemistry	96	31	5	16.13%	65	6	9.23%
Biophysics	57	11	0	0.00%	46	4	8.70%
Biology/CDB	98	46	1	2.17%	52	2	3.85%
Biology/OEB	105	33	1	3.03%	72	2	2.78%
Chemistry	276	48	7	14.58%	228	23	10.09%
EPS	69	13	2	15.38%	56	3	5.36%
Mathematics	104	8	4	50.00%	96	6	6.25%
Physics	235	22	0	0.00%	213	12	5.63%
Statistics	38	7	1	14.29%	31	1	3.23%
TOTAL	1404	270	26	9.63%	1134	78	6.88%
Anthropology	181	79	7	8.86%	102	8	7.84%
Psychology	105	58	15	25.86%	47	12	25.53%

Table 3: GSAS: Withdrawal by Gender, Natural Sciences, 1985-86 through 1989-90

Figure 10. Table from Grosz's 1991 Report. "Report on Women in the Sciences at Harvard. Part I: Junior Faculty and Graduate Students." Faculty of Arts and Science Committee on the Status of Women. Cambridge, Massachusetts: Harvard University, February 13, 1991.

[89] Grosz, "Report," 3.

In *Seizing the Means of Reproduction* Michelle Murphy examines the "technologies, practices, protocols, and processes—the 'means'— of technoscience as crafted by feminist health activists in the 1970s and beyond."[90] Martha Pollack and Barbara Grosz are two nodes, I argue, of a broader network of AI researchers who enacted a similar kind of protocol feminism within their computer science departments in the 1980s and 1990s. This network began to form in the labs of SRI International in the 1970s and spread to computer science departments at universities across the United States, the United Kingdom, and Israel in the following decades. Members of this network advocated for the status of women in computer science in many forms, whether through gathering data from other women in AI, publishing essays about their own experiences, or writing reports for their universities, disciplinary associations, and the wider computer science community. They created extensive interview guides with dozens of questions about the viewpoints and needs of women in AI, in computer sciences, and in scientific careers. They collected data from other sources and compiled it in charts, tables, and percentages. They supported each other and other junior colleagues through an assemblage of co-publications, citational practices, hosting each other's students as visiting scholars. Many women in this network studied humanities subjects at the graduate level before retraining as computer scientists. This variety of backgrounds, paired with a shared commitment to the users of the AI systems they were building, engendered an approach to AI research that was markedly different from their computer science peers. Their research incorporated concepts from discourse analysis and speech act theory to create NLP models of discourse that include conversational context and collaboration. This assemblage of techniques, values, methods, and practices illustrates a feminist AI protocol rooted in community,

[90] Murphy, Michelle. *Seizing the Means of Reproduction: Entanglements of Feminism, Health, and Technoscience.* Duke University Press, 2012, 21.

interdisciplinarity, and care— in their research, for each other, and for the users of the systems they were trying to create.

While Martha Pollack and Barbara Grosz authored reports on the status of women in the sciences, other women in this network advocated for feminist issues in computer science by discussing their personal experiences of discrimination in computer science departments. Karen Spärck Jones, who attended her first computing conference in 1958,[91] described the isolation and lack of career opportunities for women in computing researching the 1950s: "At that stage there were no opportunities for women. You have no conception of how narrow the career options were."[92] Jones published several papers in the 1960s and 1970s that computer scientists now consider foundational to the fields of NLP and information retrieval. Despite these major contributions to her field, the University of Cambridge continued to employ Jones as a contract employee for several decades— until she was in her mid-fifties.[93] Rather than being explicitly excluded as a woman in computer science, she explained, "I think the discrimination is more that it didn't seem surprising that I should be living on soft money for so long."[94] Jones's experiences of workplace isolation, lack of career options, and being under-recognized and under-paid were some of the foundational issues for working women and second-wave feminists alike.

---

[91] *Finding the Information Wood in Natural Language Trees, Lecture by Karen Jones Spärck*. YouTube video. 1994 Grace Hopper Celebration of Women in Computing Conference, 1994. https://www.youtube.com/watch?v=5fYeKiebpuo.

[92] Jones, Karen Spärck. Computing's too important to be left to men. Interview by Brian Runciman. Online, March 2007. https://www.bcs.org/content/ConWebDoc/10791.

[93] Jones, Interview by Brian Runciman.

[94] Jones, Karen Spärck. Oral History: Karen Spärck Jones - Engineering and Technology History Wiki. Interview by Janet Abbate. Transcription, April 10, 2001. IEEE History Centre. https://ethw.org/Oral-History:Karen_Sp%C3%A4rck_Jones.

Candace Sidner was one of the first of these women to write publicly about the difficulties of being a woman in computer science. In 1980, the year after she finished her dissertation, Sidner published an essay entitled "On Being a Woman at MIT: or, How to Miss the Stumbling Blocks in Graduate Education."[95] In her essay, Sidner describes how receiving a degree from MIT is "rather like being admitted to a fraternity," with certain rituals and performances of "competence and confidence" one must perform to be accepted.[96] These "strutting behaviours," as Sidner calls them, are especially difficult for women. "Women in the everyday world are not supposed to appear very confident and competent…. As a result, women must not only build and show confidence and competence, just as their male counterparts do, but unlike the men, they must *decide* first to unlearn their normal behaviour patterns."[43] Sidner describes how, often, the women who succeed in learning this strutting behaviour feel less feminine and find that friends and partners find her less attractive. "Eventually a woman can learn to find personal friends who value her confident image, but the time in between is frightening."[44] Several paragraphs of Sidner's essay are included in Ellen Spertus's influential report "Why are There so Few Female Computer Scientists?," written as a technical report for the MIT AI Lab in 1991.[97] Sidner's description paints a gloomy portrait of the experience of being a woman at MIT and they resonate with the misogynistic cultures at Stanford. Sidner finished her PhD at MIT CSAIL the year before the "Zoe Pictures 1971" event was documented at Stanford's AI Lab.

Many of the researchers in this feminist AI network worked in labs and departments peripheral to the most famous AI labs like CSAIL at MIT and SAIL at Stanford. Many of the

---

[95] Sidner, Candace. "On Being a Woman at MIT: Or, How to Miss the Stumbling Blocks in Graduate Education." *Association for Women in Mathematics Newsletter*, February 1982, 13–17.
[96] Sidner, quoted in Spertus, "Why Are There so Few Female Computer Scientists?" 17-18.
[97] Spertus, "Why Are There so Few Female Computer Scientists?"

"giants" canonized in the history of artificial intelligence—John McCarthy, Marvin Minsky, Allen Newell, Nils Nilsson, Seymour Papert, and Joseph Weizenbaum, among many others—were based at MIT, CMU or in the Bay Area. These female researchers, in contrast, were concentrated at universities like Harvard and the University of Pennsylvania. These institutions, while both Ivy League universities, were not centres for AI or computer science research in the 1980s and 1990s—especially not compared to nearby MIT and CMU. When Barbara Grosz was hired at Harvard in 1986, there was no AI research and their computer science department was still small. Perhaps by working in these more peripheral labs and departments, these researchers had more leverage to imagine and create alternative institutional cultures. Further, these new, smaller departments might have left open more possibilities for incorporating ideas from speech act theory and approaching the major goals of AI differently than others in more well-known histories of AI.

Forming the network

In 1969, Barbara Grosz began her graduate work in "computing science" at the University of California at Berkeley. She chose a thesis topic in natural language processing (NLP) after Alan Kay, a computer scientist at Xerox PARC, told her to do "something ambitious" with her thesis.[98] In Martin Ford's oral history *Architects of Intelligence,* Grosz says she was more interested in the questions of theoretical computer science than the answers. Kay suggested a project like writing a computer program that would read a children's story and tell it back from one of the character's points of view. Her graduate years overlapped with the five-year period in which DARPA funded millions of dollars of speech recognition and NLU research, described in the introduction above. Grosz began working at SRI International's AI Centre in 1973 as a Research Mathematician and

[98] Ford, Martin. "Barbara Grosz." In *Architects of Intelligence: The Truth about AI from the People Building It*, 333–57. Packt Publishing, 2018.

after completing her PhD was hired as a full-time Computer Scientist. In 1982 she became the program director for SRI International's Natural Language and Representation project and in 1983 she co-founded Stanford University's Centre for the Study of Language and Information (CSLI) with others from SRI International and Xerox PARC.[99]

As a graduate student at SRI, Grosz met Jane Robinson, a historian by training who had discovered computational linguistics late in her career and found herself recruited to the speech group at SRI International's AI Centre in 1973. Robinson held a PhD in History from UCLA, but, according to her obituary, had not been able to become a faculty member when she received her doctorate in the 1950s because she was a woman.[100] Instead, she taught English and writing classes at UCLA and California State College, Los Angeles. Robinson's obituary, written by Grosz, says she "became a computational linguist accidentally."[101] Robinson reportedly attended a talk on Chomsky's transformational grammar which "marked a turning point" in her career: afterwards, she became interested in grammars for computational linguistics and in the late 1950s began working for RAND corporation as a consultant for their machine translation projects. In the late 1960s, Robinson moved to New York to work with IBM's Automata Theory and Computability Group. Robinson and Grosz both arrived at SRI International in 1973. Grosz was one of many young computer scientists working with Robinson on SRI's speech understanding project.[102] Although Robinson never returned to the academy, she served on committees for the Association

[99] *Groszfest Session 1: 1980's*. Video recording. Vol. 1. 4 vols. GroszFest. Maxwell Dworkin Laboratory, Harvard University, Cambridge, Massachusetts, 2018. https://matterhorn.dce.harvard.edu/engage/player/watch.html?id=21b373c3-5cac-4271-88ae-32aea140fd3e.
[100] Grosz, Barbara J, Eva Hajicova, and Aravind Joshi. "Obituary Jane J. Robinson." *Computational Linguistics* 41, no. 4 (2015): 723–26. https://doi.org/10.1162/COLI_a_00235.
[101] Grosz et al. 723. Despite her long career as a lecturer, consultant, and researcher, it is difficult to find many sources on Robinson's biographical information.
[102] Grosz et al, "Obituary Jane J. Robinson," 723.

for Computational Linguistics throughout the 1970s and served terms as ACL President and Vice President in the 1980s.

Figure 11. Photo of Jane Robinson, far left, and Barbara Grosz, bottom centre, at SRI International. No date. Source: *Groszfest Session 1: 1980's*. Video recording. Vol. 1. 4 vols. GroszFest. Maxwell Dworkin Laboratory, Harvard University, Cambridge, Massachusetts, 2018. https://matterhorn.dce.harvard.edu/engage/player/watch.html?id=21b373c3-5cac-4271-88ae-32aea140fd3e.

Grosz describes how Robinson served as a mentor— "before that word was widely used in academia"— for a large group of students involved in building NLP systems at SRI. As the lab grammarian, Robinson edited her colleagues's papers, gave thoughtful critiques of the researchers' paper drafts, and "debugged their ideas."[103] In addition to the day-to-day mentorship, Grosz describes how Robinson helped her and her cohort build professional networks in computational linguistics, introducing them to "the most senior people in linguistics and computational linguistics."[104] Grosz writes that Jane Robinson's contributions at SRI International went "far beyond" her official responsibilities, regularly taking colleagues and students for guided hikes and backpacking trips in Yosemite and other national parks, well into her seventies[105]

Grosz and Candace Sidner met at a graduate conference when they were both PhD students.[106] In 1979, Sidner asked Grosz to serve on her dissertation committee, even though Grosz had only finished her own dissertation in 1977. In Sidner's dissertation acknowledgements, she thanks Grosz, Jane Robinson, and the "Natural Language Group" at the AI Centre at SRI International. The connection between Grosz and Sidner has been deep and long-lasting: throughout their careers they regularly cite each other's work. They published their first publication together in 1985 and wrote seven more papers together through the 1980s and 1990s.

Many early connections in this feminist AI network were formed in the Computer Science department at Penn. Former graduate students describe the environment in the 1980s as very welcoming to women and they give thanks, in large part, to the support of a professor named

[103] Grosz et al, 725.
[104] Grosz et al, 725.
[105] Grosz et al, 726.
[106] University of Pennsylvania School of Engineering and Applied Science. *JoshiFest: "Centring Recollections" by Barbara Grosz*. YouTube video, 2012. https://www.youtube.com/watch?v=d67nNtSAirk.

Aravind Joshi. Grosz first met Joshi at an ACL conference in the mid-1970s—probably introduced to each other by Robinson—when Grosz was still a graduate student and working at SRI International. It was Joshi who invited Grosz to Penn's computer science department in 1982 as a visiting scholar. After she visited in 1982, she returned as visiting faculty from 1984-1986 until she began teaching at Harvard. Grosz described Penn's computer science department environment as having "more women in the same room than I remembered in all the rest of my career."[107]

Indeed, the sheer number of women in computer science at Penn seems to have informed the suggestions Grosz and others made when they wrote their reports and suggestions for improving institutional cultures. In Grosz's report at Harvard, the document's first suggestion is called "Critical Mass." Grosz and the other committee members write:

> Achieving critical mass by hiring more women faculty in the sciences should be a high priority for the university. We explicitly emphasize the need for the university to set critical mass, not the hiring of a few "role models," as its goal. This goal will only be satisfied when the number of women in a department is sufficient for students to perceive it as quite normal for women to be in the field and when the idiosyncrasies of individual women faculty matter no more than those of individual male faculty. In many cases to achieve this goal the climate in some departments will need to change to overcome problems discussed elsewhere in this report. Women students and faculty will not have equal opportunities for participation in the sciences until such a critical mass is achieved.[108]

Contemporary discussions about diversity, equity, and inclusion (DEI) in computer science departments and technology companies continue to make this point. Even the recent shift in terminology from "diversity in tech" to "diversity, equity, and inclusion in tech" reflects this: merely hiring a few members from underrepresented groups risks tokenizing or alienating the members of these groups and is not sufficient, on it own, to actually change an institution's culture.

---

[107] *JoshiFest: "Centring Recollections" by Barbara Grosz*.
[108] Grosz, "Report."

The culture at Penn demonstrated to these women that there was power in numbers— an idea that these women used in their political activism and their later multi-agent approaches to AI research.

When Grosz was at Penn, she co-taught a class with Aravind Joshi and Bonnie Webber, another junior member in the department. In addition to women faculty like Bonnie Webber, many of Penn's graduate students were women. At JoshiFest, a celebration to honour Joshi's retirement, one of Joshi's former graduate students Julia Hirschberg describes her experience in the department:

> We all thought— some us, mostly of the female contingent— we thought at Penn it was actually an *advantage* to be female. And that's probably the first time I had ever felt that to be the case. Because we thought we were the coolest… The girls ruled. We just felt like we were the leaders, and the best. And that is such an unusual and powerful experience, particularly as it goes on over the years. That was amazing. Aravind collected role models, he collected colleagues, we had Bonnie [Webber], we had Martha Palmer, it was just amazing.[109]

Even in other fields of computer science, women in the department mention Joshi's supportiveness in a discipline otherwise indifferent to women.[110] Of course, it is never only one person responsible for the entire culture of a university department. These women likely discount the work *they* did to make other women feel welcome in the department—both in their actions and their very presence, forming the "critical mass" Grosz describes in her Harvard report.

---

[109] University of Pennsylvania School of Engineering and Applied Science. *JoshiFest:"Did You Feed the Animals?" By Julia Hirschberg*. YouTube video, 2012. https://www.youtube.com/watch?v=DgzhIkSFAQk.

[110] It is clear that both Grosz and Joshi knew Robinson: Grosz and Joshi are two of the three authors who penned Robinson's obituary in *Computational Linguistics*. In the obituary, Joshi wrote he met Robinson at an NLP workshop in 1975. Joshi and Robinson were both ACL presidents, Joshi in 1975 and Robinson in 1982. Grosz says that as her mentor, Robinson introduced her to "the most senior people in linguistics and computational linguistics." As past ACL president and a professor with a well-established career by the late 1970s, Aravind Joshi may have been one of those people Grosz was introduced to.

Through this department at Penn, women working at the intersection of linguistics and computer science met, were able to work together, and went on to be frequent collaborators. Martha Pollack and Julia Hirschberg, two graduate students in the seminar with Grosz, Joshi, and Webber, turned their final project into their first conference paper, which they co-presented with Bonnie Webber at the 2nd annual AAAI conference in 1982.[111] Pollack and Hirschberg both describe this conference presentation as an important step in their professional development, launching both of their careers in the computer science world.[112] Other students in this seminar who went on to collaborate with Grosz, Webber, and Joshi include Kathy McCoy, Kathy McKeown, and Ethel Schuster. Martha Pollack started working at SRI International shortly after meeting Grosz and continued to work on SRI projects until 1992. Grosz served on Pollack's dissertation committee in 1986.

After graduating, the students at Penn were hired at industry jobs and the network continued to grow. Julia Hirschberg finished her PhD in computer science at Penn and was hired as a member of Technical Staff in the Linguistics Research Department at AT&T Bell Laboratories.[113] There, Hirschberg began collaborating with Diane Litman, a woman who had finished her PhD in computer science at the University of Rochester in 1986 and worked in the Artificial Intelligence Principles Research Group at AT&T Bell Labs. Hirschberg and Litman went on to co-author 18 publications together, starting in 1987 and continuing long after they both left

---

[111] Pollack, Martha E, Julia Hirschberg, and Bonnie L Webber. "1982 - User Participation in the Reasoning Processes of Expert Systems." *AI Magazine* AAAI-82 Proceedings (1982): 358-361.
[112] *Groszfest Session 1: 1980's*. Video recording. Vol. 1. 4 vols. GroszFest. Maxwell Dworkin Laboratory, Harvard University, Cambridge, Massachusetts, 2018. https://matterhorn.dce.harvard.edu/engage/player/watch.html?id=21b373c3-5cac-4271-88ae-32aea140fd3e.
[113] Hirschberg, Julia. "Julia Hirschberg." Columbia University Computer Science Department, April 19, 2019. http://www.cs.columbia.edu/~julia/.

AT&T Bell Labs. In the 2000s, after both becoming professors, Hirschberg and Litman collaborated in the academic context: they co-organized workshops in computational linguistics, won large NSF grants for collaborative research projects, published journal articles together, and won awards for their conference papers they co-presented. Litman has collaborated with Kathy McKeown, Barbara Grosz, and Aravind Joshi.

The members created networks of support across academic departments and industry research labs like SRI International and Bell/AT&T Labs. Some of the support manifested in formal mentorships, like serving on dissertation committees or as PhD advisors. Many connections were formed through informal academic mentorship and support, like hosting each other for talks, facilitating reading groups, and providing feedback on each other's articles. Many co-authored book chapters, edited books, organized symposia, and gave conference presentations together. The more senior computer scientists regularly cited former students in their papers and included their publications in course syllabi. Scholars like Bonnie Webber, Candace Sidner, and Barbara Grosz are considered experts and leaders of their fields: for the last several decades, the papers they cite get read. These citational practices introduce the work of their more junior colleagues to others in the AI community and the wider world of computer science. As the more junior members became more established in their career, they were able to hire each other's students as postdoctoral fellows and recommend each other for professorships. In their own labs, these women have described their explicit intentions to make their departments more equitable spaces for women and members of other underrepresented groups.

When Barbara Grosz visited the University of Cambridge on sabbatical, she spent time with Karen Spärck Jones— a computer scientist the British Computing Society has described as "one of the most remarkable women in computer science."[114] Although Jones's main area of interest is in information retrieval and Grosz worked in computational models of discourse, they still found a way to collaborate: in 1986 they co-edited a book, *Readings in Natural Language Processing*, along with Bonnie Webber. In the book's acknowledgments, they thank Martha Pollack for helping them make the book's index and advising them which papers to include in the volume.[115] The members of this network incorporate new members into existing relationships and they continued to use each other's work in their own projects: in a Spring 2000 syllabus for Grosz's course on computational models of discourse at Harvard, she included their textbook as a required text alongside papers and book chapters by Candace Sidner, former Penn students Julia Hirschberg and Martha Pollack, and her former PhD students Douglas Appelt and Christine Nakatani.[116] Grosz also visited Israel frequently and made it a priority to collaborate with scholars there.[117] During her 1992 sabbatical at Hebrew University in Jerusalem, she met Sarit Kraus, a graduate student at the time. They presented a paper at IJCAI in 1993 and have since written ten papers together

---

[114] The BCS has created a Karen Spärck Jones Award and an annual Karen Spärck Jones Lecture Series. While Karen Spärck Jones is indeed a node in this network, it is outside the scope of this paper to fully describe all her contributions to computer science. For more, see:
Bowles, Nellie. "Overlooked No More: Karen Spärck Jones, Who Established the Basis for Search Engines." *The New York Times*, January 2, 2019, sec. Obituaries. https://www.nytimes.com/2019/01/02/obituaries/karen-sparck-jones-overlooked.html.
"Karen Spärck Jones Award." British Computing Society Information Retrieval Specialist Group, 2009. https://irsg.bcs.org/ksjaward.php.
"Karen Spärck Jones Lecture." BCS - The Chartered Institute for IT. Accessed June 23, 2019. https://www.bcs.org/events-home/karen-spaerck-jones-lecture.

[115] Grosz, B.J., K.S. Jones, and B.L. Webber. *Readings in Natural Language Processing*. Los Altos, California: Morgan Kaufmann Publishers, Inc., 1986.

[116] Grosz, Barbara J. "CS288: Computational Models of Discourse Reference List." Harvard University, 2000. https://www.seas.harvard.edu/courses/cs288/handouts/references.pdf.

[117] According to her cousin's presentation at GroszFest, Grosz's family members are descendants of Hungarian Jews and Auschwitz survivors and they have family in Israel.

developing computational models of collaboration and complex group action.[118] It seems it is largely through her papers with Kraus that Grosz become involved with the AAMAS community and the field of multiagent systems.

Many of these researchers spent their academic careers advocating for feminist issues in the computer science community. Bonnie Webber has been an advisor for the Society of Women Engineers since 1992. Karen Spärck Jones and Barbara Grosz both spoke at the inaugural Grace Hopper Celebration of Women in Computing conference in 1994. Since joining Harvard in 1986, Grosz has led or contributed to committees and university reports on the status of women in academic science and engineering; Harvard faculty diversity; equity and Title IX for NCAA certification; and sexual harassment and other issues for junior faculty and graduate students at Harvard.[119] She worked as Harvard's Interim Associate Dean for Affirmative Action in 1993-1994. Outside Harvard, Grosz served on the ACM's Women's Council Executive Board; National Academy of Science's Committee on Women in Academic Science and Engineering; and CRA-W (Computing Research Association Committee on the Status of Women in Computing Research). Grosz spoke at the first annual Grace Hopper Celebration of Women in Computing and gave the CRA-W Distinguished Lecture in 2005. In 2005 the Berkeley computer science department invited Grosz to lead their Task Force on Women in Science.[120]

People and language itself: feminist AI research protocols

[118] Ford, Martin. "Barbara Grosz." In *Architects of Intelligence: The Truth about AI from the People Building It*, 333–57. Packt Publishing, 2018.
[119] Grosz, Barbara. "Barbara Grosz CV," May 2017. https://grosz.seas.harvard.edu/files/grosz/files/barbara_grosz_cv_may_2017.pdf.
[120] Crowley, Magdalene L. "Diversity Milestones." Text. EECS at UC Berkeley, February 22, 2018. https://eecs.berkeley.edu/about/diversity/milestones.

This network shared similarities in the methods goals, and topics of their AI research. Most researchers in this network published research that combined computational linguistics, NLP, AI, and MAS. Many of these researchers have graduate degrees in disciplines other than computer science or linguistics: two have PhDs in history and one has a PhD in philosophy. Perhaps as a result, many researchers in this network use interdisciplinary and unconventional methods in their AI research. Some apply computational methods to topics in linguistics, like how prosody and intonation influence dialogues. Those who study multiagent systems primarily research language-based collaborations and plan-making in multiagent systems. As Grosz mentioned in her speech, her research has long been user-focused, and the same is true for her collaborators: in general, these researchers study how humans and AI systems can communicate with each other through to coordinate actions, tasks, and goals. Many of the researchers approach their research philosophically, citing theoretical work in linguistics and computer science in addition to philosophies of language, action, and computation.

In 2017 Barbara Grosz accepted the Lifetime Achievement Award at the Annual Meeting of the Association for Computational Linguistics (ACL). In her acceptance speech, entitled "Smart Enough to Talk With Us? Foundations and Challenges for Dialogue Capable AI Systems," Grosz described some of her contributions over her nearly fifty year career in computer science and reflected on past and current challenges in the field.[121] Grosz's speech is interesting for several reasons, and I will return to some of them in the third chapter, but for the moment I want to focus on two things. The first is how she describes her approach to computational linguistics and AI systems. "Two themes have guided my dialogue research from its start: people matter and language

---

[121] Grosz, Barbara J. "ACL Lifetime Achievement Award: Smart Enough to Talk With Us? Foundations and Challenges for Dialogue Capable AI Systems." *Computational Linguistics* 44, no. 1 (March 2018): 1–15. https://doi.org/10.1162/COLI_a_00313.

itself matters," She said. "To build systems capable of conversing sensibly with people requires close observation and analysis of ways people talk with one another, collecting data in real situations, and celebrating the full glory of language use rather than building systems that require people to control their language use."[122] Since she began her career in the 1970s, Grosz has contributed to AI research that prioritizes users over systems; the glory of everyday language, not dialogue that follows user scripts; and listening to the ways people *actually talk* with each other to inform the design of AI systems.

Second, Grosz cites linguists and computer scientists, but she also cites philosophers—namely, the philosophers of language H. P. Grice and J. L. Austin. Grice's theories of implicature and Austin's theory of speech acts are canonical texts for pragmaticians and contextualist approaches to language, but surprising citations for a computer scientist. Grosz mentions AI giants John McCarthy and Noam Chomsky, but irreverently— they were naysayers whose skepticism she overcame. Noam Chomsky remarked that her thesis was an interesting topic, she said, "but advised I would never succeed because dialogue could not be formalized."[123] John McCarthy, who founded the Stanford AI Lab, told her her research was irrelevant because "an understanding of people's cognitive processes was irrelevant to AI."[124] These examples are heartwarming encouragements to persist despite negative feedback, but they say more about the environment for young women entering the fields of AI and computational linguistics in the 1970s.

One example of this network's cross-discipline research is the book *Intentions in Communication,* co-edited by Martha Pollack and featuring chapters by of the women discussed

---

122 Grosz, "ACL Lifetime Achievement Award."
123 Grosz, "ACL Lifetime Achievement Award," 12.
124 Grosz, "ACL Lifetime Achievement Award," 12.

in this chapter. In March 1987, the System Development Foundation and AAAI funded an interdisciplinary workshop on "Intentions and Plans in Communication and Discourse," held in Monterrey, California.[125] This volume includes edited versions of the papers and commentaries by other workshop contributors. The contributors included several researchers at SRI International's AI Centre and the CSLI, as well as linguists, philosophers, psychologists, and computer scientists working at universities and research laboratories across the US and Canada.[126] The book features chapters authored and co-authored by Grosz, Hirschberg, Pollack, and Sidner, as well as chapters by philosophers Michael Bratman and John Searle. At the time of the workshop, Grosz had just finished teaching the Penn seminar with Joshi and Webber and was in her second year as a professor at Harvard. Candace Sidner was in Cambridge with Grosz, working in the research lab for Digital Equipment Corporation (DEC). The Penn graduate students were all in the first years of their post-PhD jobs: Julia Hirschberg and Diane Litman were at AT&T Bell Laboratories in New Jersey (their colleagues Janet Pierrehumbert and Henry Kautz also contributed to the book). Pollack, working at SRI International at their AI Centre and the CSLI, helped edit the volume with Philip Cohen, a colleague from SRI, and Jerry Morgan, a linguist from the University of Illinois.

The introduction by Pollack, Cohen, and Morgan situates the book's contributions firmly within contextualist approaches to linguistics, citing concepts from Gricean pragmatics and speech act theory inspired by J.L. Austin's *How to Do Things with Words.* The book begins and ends with chapters by analytic philosophers considering the nature of intention. The book is more than just a collection of papers from different disciplines: many chapters are co-authored by a computer scientist and a scholar from another field, like linguistics, psychology, or philosophy. The research

[125] Cohen, Philip, Jerry Morgan, and Martha Pollack, eds. *Intentions in Communication*. System Development Foundation Benchmark Series. Cambridge, Massachusetts: MIT Press, 1990.
[126] The CSLI is perhaps most famous for its project the Stanford Encyclopedia of Philosophy.

features scholars who write with, analyze, or directly respond to the other workshop contributors and to ideas from fields adjacent to cognitive science and computational linguistics.

These women have helped each other gain prestige and recognition in AI research. The first issue of *AAAI AI Magazine* in 1980 featured two articles, one of which was written by Barbara Grosz. Bonnie Webber presented a paper in the first annual AAAI Conference in 1980, thanking Candace Sidner in the paper's acknowledgements.[127] Later in their careers, Grosz and Sarit Kraus won the 2007 IFAAMAS Influential Paper Award for their 1996 paper — an award for papers that continue to contribute in the multiagent systems research community for years after publication (Kraus went on to win the award a second time in 2014).[128] Grosz, Hirschberg, McKeown, Pollack, Sidner, and Webber have all served as past officers and councillors for the Association for the Advancement of Artificial Intelligence (AAAI); Grosz and Pollack were AAAI Presidents in 1993-1995 and 2009-2010, respectively.[29] Many of them are department heads and distinguished chairs in Ivy League computer science departments. Martha Pollack is now the president of Cornell University. Aravind Joshi, Karen Spärck Jones, and Barbara Grosz have all won the ACL Lifetime Achievement Award (in 2002, 2004, and 2017, respectively).[129]

Conclusions and contradictions

This group of researchers created a feminist AI protocol in the last decades of the twentieth century to change the scientific cultures of their universities, to improve working conditions of

---

[127] Bobrow, Robert, and Bonnie Webber. "Knowledge Representation for Syntactic / Semantic Processing." In *AAAI-80 Proceedings*. Stanford University, CA: AAAI Press/MIT Press, 1980. https://www.aaai.org/Library/AAAI/1980/aaai80-090.php.
[128] "IFAAMAS: Awards: Influential Paper." International Foundation for Autonomous Agents and Multiagent Systems. Accessed June 23, 2019. http://www.ifaamas.org/award-influential.html.
[129] ACL Wiki. "ACL Lifetime Achievement Award Recipients." Accessed June 6, 2019. https://aclweb.org/aclwiki/ACL_Lifetime_Achievement_Award_Recipients.

women scientists, and to expand the kinds of research included in the umbrella of AI research. The protocols were based on values of collaboration, interdisciplinarity, and designing AI systems with users' needs in mind. These women entered AI research just after the "golden years" of the 1960s and found themselves in an institutional culture built for men. Through collecting data about the experiences and rates of success of women in compeer science fields, this network of women brought attention to issues they faced in their own universities and departments. This protocol enabled them to cite each other's scholarship, support each other in their careers, and make spaces for their work to reach the broader AI audience. They worked within their institutions to outline the social, cultural, theoretical, and material barriers that they felt limited AI research— both the scope of AI research itself and the material conditions that prevented them from conducting research as women in computer science. In many ways, they succeeded— most of these women became tenured professors and administrators at Ivy League universities.

This feminist AI protocol was not without its contradictions, however. Only certain disciplines were included in the interdisciplinarity of their approaches: primarily computer science, linguistics, and other fields related to cognitive science like analytic philosophy and cognitive psychology. As I will elaborate in the next chapter, concepts from these fields can expand what is included in AI research, but only in certain directions. Just like there were only certain disciplines this network includes in their interdisciplinary approaches, there are only certain political issues these researchers have spoken publicly about in their political involvement. Like most computer science researchers in the last 60 years, many of the members in this network have received grants from DARPA and other military funds that total millions of dollars over their careers.[130] Recently,

---

[130] Even critical mass, the phrase Grosz and her collaborators often repeated, refers to the "critical mass" needed in development of nuclear weapons.

a few of these researchers have spoken against using their research for military defence applications, like when Grosz, Webber, and Pollack signed their names to the IJCAI 2015 open letter against the development of autonomous weapons.[131] Beyond this example, however, I have found little public discussion by these researchers about the role of military funding in computer science research, past or present. Some of the researchers in this network have collaborated extensively with Israeli universities and institutions; while it is outside the scope of this paper, but I have not found any public discussion by these members about how US-Israel collaborations in computer science relate to human rights issues in Israel and Palestine. My critiques are certainly true of many—if not most—American computer scientists, and to adequately address them would go far beyond the scope of this paper. At risk of appearing to present a presentist argument to these historical events, I think these critiques are still worth noting— especially now that intersectionality has (rightfully) become the new standard by which we consider whether something is feminist.

---

[131] Future of Life Institute. “Open Letter on Autonomous Weapons,” July 28, 2015. https://futureoflife.org/open-letter-autonomous-weapons/.

Chapter 3. Whose *SharedPlans*? Scripts, Collaboration, and the Possibilities of Feminist AI Research

In the previous chapter, I discussed a network of computer scientists who shared common methods, philosophies of language, and research goals for interdisciplinarity and incorporating human needs into their AI systems. In this chapter I will investigate more deeply how Barbara Grosz and her colleagues encoded concepts like agency, collaboration, and human-computer relations into the AI systems they designed and built. I will focus on Barbara Grosz's writings from the 1980s, 1990s and early 2000s about user scripts, master-slave analogies, and her broader philosophy of human-AI relations. I use as a point of entry her concept of SharedPlans, developed with Candace Sidner in 1990 to enable computational models of collaboration in discourse. In some ways, Grosz's approaches to AI align with critical and feminist work on standpoint theory and gender scripts. I argue, however, that Grosz's vision of human-AI partnerships— despite the potentially radical implications of a partner-like relationship between humans and AI systems— does not question the claims of universal truths, computational models of the mind, and the myth of the *homo economicus* dominant in AI research in the twentieth century.

In June 1994 Barbara Grosz gave a lecture at the first annual Grace Hopper Celebration of Women in Computing (GHC). In her lecture, entitled "Collaborative Plans and Dialogue Participation," Grosz extends her feminist protocol of collaboration to the future of AI as a field. She focuses on the importance of collaboration for AI systems at every step of the design process. Grosz uses several examples to illustrate her vision of human-AI interactions that go beyond what she calls the "master-slave" relationship— an analogy, common in computer science at the time, to refer to communication between two devices when one device has control over the other. In one example, Grosz describes a dialogue between a network manager (human) and a computer system about how to perform a maintenance task. The network manager speaks to the computer in clear but colloquial language, and the computer responds by volunteering useful information about an aspect of the task the human had not previously mentioned. The network manager is not the master of the computer, Grosz argues, and neither is the computer the manager's slave: they each provide different skills and sets of knowledge to complete the task together. Grosz explains, "The network manager in this case isn't a master just directing the network's presentation system as a slave. One of the things that I think we want to get around in the systems that we build is this notion that there's a master in charge and a slave just trying to do his or her work." She continues, "What I'd like to see us able to do is to provide systems that can be more supportive of their users and become problem solving partners with them."[132]

Grosz's suggestion that AI systems ought to provide a supportive partnership to their human users points to the focus of this third chapter. One of Barbara Grosz's contributions to the field of natural language understanding (NLU) was to provide models that could incorporate contextual information into the computer system's dialogue responses. In her research on how to incorporate information beyond the utterances of the speakers, Grosz explored more abstract philosophical questions about how agents— human and computer— collaborated in their plans, goals, and actions. It appears it was these questions that led Grosz into the field of multi-agent systems.

In 1990 Grosz began presenting and publishing about collaborative planning for discourse with Candace Sidner. She began collaborating with Sarit Kraus in 1992, where they began to publish work on goal-sharing and models for how AI systems and human users could relate. By

[132] Grosz, Barbara. "Women in Computing: Collaborative Plans and Dialogue Participation, Lecture by Barbara Grosz." presented at the 1994 Grace Hopper Conference on Women in Computing, June 9, 1994. https://www.youtube.com/watch?v=2VrrH-hHVmQ&list=PLDEiwPwOpXnaI416H7xik1omEZq0nEdtU&index=7&t=0s.

1999, Grosz was presenting at multi-agent systems conferences and on the board of directors for the International Foundation for Multi-Agent Systems. In interviews from the early 2000s, Grosz describes her approach of thinking of AI systems as helpers or teammates— quite an unusual framework for thinking about human-AI relations at the time.[133] There appears to be a path in Grosz's research, then, from computational linguistics to multi-agent systems— motivated by her research goal of building AI systems that could incorporate the goals and intentions of the user and other agents.

The issues surrounding Barbara Grosz's approaches to AI, agency, and collaboration points to the complexity of envisioning what a feminist AI would actually look like. In many ways, the work of Grosz and her collaborators promoted a notion of human-computer collaboration that was less hierarchical than the analogy of master and slave. As I outline later in the chapter, their research challenged the individualist orientation of AI research which had been rooted in cognitive science. It emphasized the way agents were continuously creating interpersonal plans and provided space to incorporate ideas from social theory and sociology into computer models of language and collaboration. In these respects, Grosz's theoretical approach in this period resonates with contemporaneous work produced in feminist theory and science and technology studies. A deeper consideration, however, reveals that in many ways Grosz's approach to human-AI relations reinforces many of the hierarchical relations dominant in AI research. It is still humans making the goals, and AI systems still work for them. The computers are simply more helpful because they have been programmed with more accurate models. Grosz's approaches to AI research still aimed

---

[133] In recent years Grosz has collaborated with health care professionals to build AI systems that help doctors.

to make computational models of human behaviour, explained by economic principles of rationalism and optimization.

Scripts, Plans, and *SharedPlans*

In the 1980s and 1990s, Barbara Grosz's vision for the future included a world in which humans and AI systems collaborate as partners on shared goals. In many ways, Grosz's approach to AI research was more critical than many of her colleagues in computer science at the time. In 1977 Roger Schank and Robert Abelson published their influential book *Scripts, Plans, Goals, and Understanding: An Inquiry Into Human Knowledge Structures.*[134] Schank, an AI researcher, and Abelson, a social psychologist, had met at a workshop on the emerging field of cognitive science in 1971.[135] The two began collaborating, publishing several papers together in the six years before the publication of *Scripts, Plans, Goals, and Understanding*. Schank moved to join Abelson at Yale in 1974 and start a research group about knowledge systems in humans and AI. At an AI conference in 1975, Schank and Abelson introduced their concept of "scripts" to "account for knowledge about mundane situations" in natural language processing.[136] "A script, as we use it, is a structure that describes an appropriate sequence of events in a particular context. A script is made up of slots and requirement about what can fill these slots."[137] In the conference paper they use examples of a customer ordering food at a restaurant to show "predetermined, stereotyped

[134] Schank, Roger C., and Robert P. Abelson. *Scripts, Plans, Goals, and Understanding: An Inquiry Into Human Knowledge Structures.* New York: Psychology Press, 1977.
While it is outside the scope of this chapter, it is worth noting that in their work on scripts, Schank and Abelson engage with the theories of psychologist Silvan Tomkins— an important figure in the development of affect theory.
[135] Schank & Abelson, *Scripts, Plans, Goals,* 1.
[136] Schank, Roger, and Robert Abelson. "Scripts, Plans, and Knowledge." *IJCAI'75 Proceedings of the 4th International Joint Conference on Artificial Intelligence* 1 (September 3, 1975): 151-57.
[137] Schank & Abelson, "Scripts, Plans, and Knowledge," 151.

sequence of actions that define a well-known situation. A script is, in effect, a very boring little story."[6] They describe the following scenario:

1. John went into the restaurant.

2. He ordered a hamburger and a coke.

3. He asked the waitress for the check and left.

Did John eat the hamburger? Schank and Abelson argue that, although it is never written, most people know the answer is yes. Most people know that when someone goes to a restaurant, orders food, pays, and leaves, most of the time they eat what they ordered. People use these scripts to infer what is implicit in these everyday situations. Scripts are useful, Schank and Abelson argue, because they contain information that is *good enough* for *many or most* circumstances. They are "boring little stories" that most people follow, most of the time, Schank and Abelson argue. In the early days of cognitive science, where the dominant paradigm was to consider the mind as a computer with much of its functionality contained in the "hardware" of the brain, the human mind must have to do an immense amount of optimization to get results that are good enough given a limited amount of resources. Scripts fit well within this paradigm because they provide the mind with structures for frequent or well-known situations. AI researchers found Schank and Abelson's theory of scripts useful because they could be operationalized: a computer scientist could theoretically program a computer with these same scripts to guide the actions and responses of the computer.

From a humanistic perspective, Schank and Abelson's scripts are far from "very boring little stories." Even in the toy scenario above, Schank and Abelson rely on numerous norms that they assume people to follow. I want to take a moment to describe some of the many assumptions built into even their most basic examples. The third statement includes gendered assumptions about people who work in restaurants. Schank and Abelson operate with the assumption that John must have eaten the hamburger, or else he would not have paid for the bill—but this is simply not true. For one, they assume John is alone. What if John was at the restaurant with his children and had ordered a meal for them? Maybe John ordered a hamburger and coke for his friend in the car. Even if John was alone at the restaurant, there are so many reasons for John to violate the norms of entering a restaurant, ordering food, eating it, paying, and leaving. John could have gone into the restaurant to hide out from the rain and ordered something to be polite, even though he was not hungry. Maybe John was treated poorly by the other patrons of the restaurant and left out of frustration. John may have wanted to leave quickly because he saw someone he did not want to run into. Perhaps John is ill and the smell of the burger made him nauseous. What if John left because he had received an urgent call? People behave unpredictably, even in commonplace scenarios like ordering food at a restaurant. This is seen as a difficult problem in computer science and AI research, where many goals involve designing systems which "predict" a scenario based on very little information and respond adequately.

To test their theory in a computational setting, Schank and Abelson created the Script Applier Mechanism (SAM) computer program, an AI system that used scripts like the example above to answer questions about scenarios in a human-like way. Schank and Abelson built SAM to take as inputs a series of "conceptual dependency structures"— scripts for a particular domain, in this case the context of a restaurant— and programmed SAM to make inferences about causes and effects in these domains. In their IJCAI presentation from 1975, they provide SAM with a story about a customer at a restaurant; they demonstrate SAM's ability to paraphrase and summarize the story and, when SAM is questioned, provide answers only implicitly stated in the input story's text. As a result, it appears as if SAM "understands" the scripts people follow— at least limited to the strictly confined context of people ordering food at a restaurant. Many saw SAM's performance as a major step in understanding how humans store knowledge about the world. For many AI researchers in the 1970s and 1980s, the SAM program and Schank and Abelson's script theory seemed to provide a computational model for the cognitive process of understanding a story.

Input:

John went to a restaurant. The hostess seated John. The hostess gave John a menu. The waiter came to the table. John ordered lobster. John was served quickly. John left a large tip. John left the restaurant.

Questions and Answers:

Q: What did John eat?
A: LOBSTER.
Q: Who gave John the menu?
A: THE HOSTESS.
Q: Who gave John the lobster?
A: PROBABLY THE WAITER.
Q: Who paid the check?
A: PROBABLY JOHN.

Figure 12. An example of SAM's input and output. Taken from Schank, Roger, and Robert Abelson. "Scripts, Plans, and Knowledge." ***IJCAI'75 Proceedings of the 4th International Joint Conference on Artificial Intelligence*** **1 (September 3, 1975): 153.**

It is precisely these claims—that SAM's question-answering skills demonstrated understanding—against which John Searle argues in the so-called Chinese Room Argument from his 1980 paper "Minds, Brains, and Programs."[138] In the paper, Searle distinguishes between weak AI and strong AI. According to weak AI, computers provide value in the study of the mind because they are powerful tools. Strong AI, on the other hand, suggests "the appropriately programmed computer really *is* a mind, in the sense that computers given the right programs can be literally said to understand and have other cognitive states."[139] Searle uses as his primary example Schank and Abelson's SAM program, which answers questions about a series of restaurant scenarios. "Partisans of strong AI claim that in this question and answer sequence the machine is not only simulating a human ability, but also 1. That the machine can literally be said to *understand* the story and provide the answers to questions, and 2. That what the machine and its program do *explains* the human ability to understand the story and answer questions about it."[140] Searle rejects both of these claims and describes a thought experiment to demonstrate his critique, known colloquially as Searle's "Chinese room argument."

In the thought experiment, Searle supposes he is locked in a room and given a large batch of Chinese writing to translate. Searle speaks no Chinese and cannot distinguish Chinese characters from "meaningless squiggles."[141] He is given a second document: a set of instructions, written in English, which instruct him to correlate one set of Chinese symbols with another set of Chinese symbols. Finally, he is given a third document, this time with Chinese symbols and sentences written in English. The third document contain rules that correlate elements of the third batch with

[138] Searle, John R. "Minds, Brains, and Programs." *Behavioral and Brain Sciences* 3, no. 3 (September 1980): 417–24. https://doi.org/10.1017/S0140525X00005756.
[139] Searle, "Minds, Brains, and Programs," 417.
[140] Searle, 417.
[141] Searle, 418.

parts of the first and second batch; they also instruct him how to respond with a certain set of Chinese symbols in response to certain other Chinese symbols. Searle explains that, unbeknownst to him, the Chinese researchers call the first of documents a "script," the second set a "story," and the third batch "questions." Searle's responses to the third batch are called "answers to the questions." In this scenario, even though Searle is answering questions about things written in Chinese and even answering them in Chinese, Searle argues he does not *understand* Chinese. "As regards the first claim, it seems to me quite obvious in the example that I do not understand a word of the Chinese stories. I have inputs and outputs that are indistinguishable from those of the native Chinese speaker, and I can have any formal program you like, but I still understand nothing."[142]

Searle argues that, just as the person in the room in the thought experiment does not understand Chinese, Schank and Abelson's AI system does not understand the scenarios about which it answers questions. "For the same reasons, Schank's computer understands nothing of any stories, whether in Chinese, English, or whatever, since in the Chinese case the computer is me, and in cases where the computer is not me, the computer has nothing more than I have in the case where I understand nothing."[143] Searle argues intentionality is a product of the causal features of the brain; regarding the question of whether a machine can think, he argues "*only* a machine could think, and only very special kinds of machines, namely brains and machines with internal causal powers equivalent to those of brains."[144] Searle concludes machines can tell us little about thinking because machines run on computer programs and, as his thought experiment demonstrates, running a computer program is not the same as thinking and understanding.

[142] Searle, 418.
[143] Searle, 418.
[144] Searle, 417.

Barbara Grosz's work in natural language processing from the late 1970s and early 1980s engages with the concepts put forward by Schank, Abelson, and Searle and their debate about the use of scripts in AI and cognitive science. In her article in the 1980 inaugural issue of *AI Magazine*, Grosz reflects on ways theorists, designers, and builders of NLP systems must "consider language as part of a larger situation."[145] "The participants in a conversation and their states of mind are as important to the interpretation of an utterance as the linguistics expressions from which it is formed," she argues.[146] Citing ideas from speech act theory, including Searle's 1969 essay on the subject, Grosz considers the way language both uses and helps build shared models of the world.[147] She argues for the importance of considering utterances, what is intended, and the relation between the two. Grosz argues that an NLP system must use a combination of mechanisms— some about language specifically and others about common-sense reasoning. Language systems must represent the beliefs and knowledge of multiple agents, she argues, and they must be able to operate on incomplete and inconsistent information. Grosz cites Schank and Abelson's 1977 article as an effort that demonstrates "the feasibility of incorporating planning and plan recognition into the common sense reasoning component of a natural language processing system."[148] Grosz describes the serious limitations of theories like theirs, however, in their ability to handle situations that involve actions having multiple effects. She says that for these kinds of theories, "their limitations highlight the need for more robust capabilities in order to achieve the

[145] Grosz, Barbara. "Utterance and Objective: Issues in Natural Language Communication." *AI Magazine*, 1980, 11.
[146] Grosz, "Utterance and Objective," 11.
[147] Searle, J. 1969. *Speech Acts: An Essay in the Philosophy of Language.* Cambridge University Press, Cambridge, England.
[148] Grosz, "Utterance and Objective," 17.

integration of language-specific and general common-sense reasoning capabilities required for fluent communication in natural language."[149]

Grosz's concept of SharedPlans, co-created with Candace Sidner and introduced in a chapter in Martha Pollack's edited volume, provided an important intervention in AI research that moved beyond the script-based theories proposed by Schank and Abelson. In Grosz's and Sidner's chapter in *Intentions in Communication*, they articulate their critique of Schank and Abelson's script theory in much more detail.[150] In their chapter, Grosz and Sidner reflect on their 1986 publication about their theory of discourse structure.[151] They describe how a computational theory of discourse structure must have underlying theories of intention, action, and plans. They describe how existing work on planning failed to account for collaboration:

> Previous work on planning and plan recognition for natural language might seem to provide the basis for such theories. However, as we examined that work, we realized that various assumptions it made about plans, actions, and agents were inappropriate for the general discourse situation and precluded any simple type of generalization. In particular, it did not provide the right basis for explaining collaborative behaviour. Discourses are fundamentally examples of collaborative behaviour. The participants in a discourse work together to satisfy various of their individual and joint needs. Thus, to be sufficient to underlie discourse theory, a theory of actions, plans, and plan recognition must deal adequately and appropriately with collaboration.[152]

Grosz and Sidner introduce their concept of SharedPlans, which they define as an approach that provides a way to account for collaborative behaviour in discourse and formalize the "mutual belief among all participants about one another's intentions and about the way in which those

---

[149] Grosz, "Utterance and Objective," 17.
[150] Cohen, Philip, Jerry Morgan, and Martha Pollack, eds. *Intentions in Communication*. System Development Foundation Benchmark Series. Cambridge, Massachusetts: MIT Press, 1990.
[151] Grosz, Barbara J. and Candace L. Sidner. "Attention, intentions, and the structure of discourse." *Computational Linguistics* 12, no 3 (1986): 175-204.
[152] Cohen et al, 418.

intentions support the achievement of the overall goal."[153] Grosz and Sidner write that Searle's chapter in the same book addresses "similar issues concerning theories for explaining how two (or more) people work together to accomplish goals."[154] "Although his detailed proposals are different," Grosz and Sidner write, "they appear similar in spirit." Like Schank and Abelson's script theory, Grosz and Sidner's concept of SharedPlans incorporates planning among multiple agents at the level of discourse. It addresses some of script theory's limitations, however, because it explains the way plans are collaborative and offers flexibility for actions and plans with multiple possible outcomes. As they explain, the recognition process for SharedPlans "differs significantly from prior work on recognition in that it does not presume a fixed plan on the part of one participant, the form and content of which must be inferred by the other(s). Instead, collaborative planning entails a negotiation in which information about actions, action relationships, desires, and intentions is made sufficiently clear for all participants to know how actions will be used to satisfy desires."[155] Unlike scripts, which are predetermined, simplistic, and stereotyped sequences of actions, SharedPlans requires mutual comprehension and continual, active negotiation among dialogue participants. Instead of an additive approach to many individual agents acting concurrently, SharedPlans requires interpersonal collaboration to co-create new plans as they are happening. In SharedPlans, both computers and humans are agents working together in a multi-agent system.

In a 2018 interview, Grosz spoke at length about the issues she sees surrounding script-based AI systems. Although she has seen significant developments in speech processing, Grosz describes areas that still have room to improve: "If you consider any of the systems that purport to carry on

---

[153] Cohen et al, 429.
[154] Cohen et al, 418.
[155] Cohen et al, 442.

dialogues, however, the bottom line is they essentially don't work. They seem to do well if the dialogue system constrains the person to following a script, but people aren't very good at following a script. There are claims that these systems can carry on a dialogue with a person, but in truth, they really can't."[156] Current AI systems, she suggests, are designed with particular social scripts in mind. When users go off the scripts assumed in the AI system's design, the systems fail. These concerns Grosz raises fit with her broader approach to user-centred design in AI and NLP systems, which I described in the second chapter. In the same interview, Grosz describes the ethical issues that arise when users learn to rely on AI systems in one scenario, but the AI systems cannot provide trustworthy information in a seemingly similar scenario. Grosz uses the example of asking a phone assistant where the nearest hospital is versus asking it where the user can go for a heart attack. "People would assume a system that can answer one of those questions you can answer the other."[157]

Going off-script: resonances in STS and feminist theory

Grosz's work on building systems that can go "off-script," the way she grounds her approach in the diverse goals and desires of real users, and her philosophy of "language as action" mirror ideas circulating in feminist and critical science and technology studies (STS) in the 1990s. In 1990 Steve Woolgar published "Configuring the User: the case of usability trials," in which Woolgar spent eighteen months at a microcomputer company doing ethnographic work on their "DNS" project team.[158] Woolgar describes the way usability trials, technical manuals, and stories

---

[156] Ford, Martin. "Barbara Grosz." In *Architects of Intelligence: The Truth about AI from the People Building It*, 333–57. Packt Publishing, 2018.
[157] Ford, "Barbara Grosz."
[158] Woolgar, Steve. "Configuring The User: The Case of Usability Trials." *The Sociological Review* 38 (1990): 58–99. https://doi.org/10.1111/j.1467-954X.1990.tb03349.x.

from the company's Tech Support department helped construct and configure particular ideas about users within the company. "Indeed," Woolgar argues, "the whole history of the DNS project can be construed as a struggled to configure (that is, to define, enable, and constrain) the user."[159] In 1992, Madeline Akrich published her essay "The De-Scription of Technical Objects."[160] In it she describes the work innovators do of "inscribing" visions about the world in the technical aspects of an object: "Designers this define actors with specific tastes, competences, motives, aspirations, political prejudices, and the rest, and they assume that morality, technology, science, and economy will evolve in particular ways."[161] She calls the end product of this inscription a "script" or a "scenario." Akrich goes on to explore the negotiations between the imagination of the designer and the adjustments, or lack of adjustments, of the users. Grosz's emphasis on continual collaboration and negotiation among "agents"—I will return to the ambiguity of this term later—resonates with Akrich's analysis of subjects and objects "in the making." Woolgar's and Akrich's essays represent a larger trend in STS in the 1980s and 1990s to consider the way certain forms of use and access are built into technologies. In some ways, Grosz's approach of building AI systems that incorporate the diverse and continually developing goals of users resonates with these trends in STS.

Grosz's work on off-script AI systems was also happening at the same time as Alison Adam's work on how gender and feminism related to artificial intelligence. In 1994, the same year Grosz gave her lecture on collaborative AI systems at the inaugural Grace Hopper

---

[159] Woolgar, 69.
[160] Akrich, Madeleine. "The De-Scription of Technical Objects." In *Shaping Technology/Building Society. Studies in Sociotechnical Change*, edited by Wiebe E. Bijker and John Law, 205–24. Cambridge, Massachusetts: MIT Press, 1992. https://halshs.archives-ouvertes.fr/halshs-00081744.
[161] Akrich, 208.

Conference, Allison Adams published her essay "Gendered knowledge — Epistemology and artificial intelligence" in *AI and Society*.[162] In Adam's essay, she critiques the epistemology of strong AI, what she calls the "cognitive science programme" (computationalist understandings of the mind), and specifically the AI subdomain of knowledge-based or expert systems. Adam highlights strong AI's emphasis on the individual cognizer: "The cognitive science view as part of a strong AI which after all has been a dominant school of thought in AI in recent years, places an excessive emphasis on the individual as perceiver and acquirer of knowledge."[163] In this respect, Grosz's concept of SharedPlans responds to this individualist approach in AI research. Instead of pre-set plans of individuals, SharedPlans provides a way for computer systems to model the way information and plans are co-created by multiple people.

Although Grosz's intervention with SharedPlans aligns with Adam's critiques of AI's epistemological individualism, it does not engage with her deeper critiques. Adam uses an approach rooted in feminist epistemology and standpoint theory to question the very project of knowledge-based AI systems. Adam, like Grosz, rejects Schank and Abelson's script theory, but, citing feminist scholars like Donna Haraway and Jane Flax, questions the ideas of universal truth and knowledge that underpin the fundamental goals of natural language understanding. She points out that Schank and Abelson's restaurant scenarios point to the "essentially social, cultural, and conventional nature of our knowledge."[164] Adam argues Searle's critique of Schank and Abelson focuses on the wrong level: "As a critique of AI, intentionality is important, but Searle's arguments in failing to take account of cultural questions, do not constitute a critique at the appropriate level

[162] Adam, Alison. "Gendered Knowledge — Epistemology and Artificial Intelligence." *AI & SOCIETY* 7, no. 4 (December 1, 1993): 311–22. https://doi.org/10.1007/BF01891414.
[163] Adam, "Gendered Knowledge," 319.
[164] Adam, "Gendered Knowledge," 320.

and are tarred with the same epistemological brush as research in strong AI/ the cognitive science school."[165] Searle offers a critique of strong AI, but leaves weak AI untouched. Strong or weak, Adam argues, AI systems assume an individualist account of knowledge and a single account of truth. "AI, especially strong AI, is in epistemological terms essentially conservative, preserving socially legitimated knowledge and offering limited scope for considering of alternative forms of knowledge— in particular for this discussion— seeing knowledge in terms of gender."[166] Grosz's work on SharedPlans questions individualist approaches to knowledge, but she seems to avoid this constructivist critique.

Grosz's hesitancy to make broader critiques of AI's epistemological grounding suggests a remaining commitment to computer science as an epistemic culture. Fields like AI and NLP aim to make machines that can process information, sound, speech, and language, either *imitating* or *inspired* by the way humans do. In either case, the idea of a single, stable human ideal processing information a specific way guides this aim. This assumes one universal human experience, leaving little room for critical exploration of the way cultures, histories, and social forces shape the very categories of what is considered "human." These different histories, bodies, and cultures shape the categories of language, speech, sound, and information— and the way people perceive, believe, and act on these categories. As fields that primarily exist as subfields within computer science, both AI and NLP use computer scientific methods: they approach problems and solutions in terms of formal models, simulations, and testing models using benchmarks like scalability, stability, reliability, and safety.[167] Benchmark values like scalability reward systems built for a universal

---

[165] Adam, 317.
[166] Adam, 321.
[167] Feil-Seifer, David, Kristine Skinner, and Maja J Matarić. "Benchmarks for Evaluating Socially Assistive Robotics." *Interaction Studies: Psychological Benchmarks of Human-Robot Interaction* 8, no. 3 (October 2007): 423–29.

user, easily scalable to work for as many people as possible. It seems difficult to problematize the idea of a single account of human-ness and truth while working in a field that prioritizes values like scalability. Grosz, as a computer scientist working at the intersection of AI and NLP, seems to be caught in this bind.

Destabilizing the master-slave analogy in AI

Grosz and Sidner use their concept of SharedPlans to challenge the assumption of a master-slave relationship between dialogue participants in computer science. They write:

> Serious consideration of dialogue makes it clear that the master-slave assumption is the wrong basis on which to build a theory of discourse. This assumption encourages theories that are unduly oriented toward there being one controlling agent and one reactive agent. Only one agent has any control over the formation of the plan; the reactive agent is involved only in execution of the plan (though to do so he must first figure out what that plan is). We conjecture that the focus of the speech act and plan recognition work on single exchanges underlies its (implicit) adoption of the master-slave assumption. To account for extended sequences of utterances, it is necessary to realize that two agents may develop a plan together rather than merely execute the existing plan of one of them. That is, language use is more accurately characterized as a collaborative behaviour of multiple active participants.[168]

Grosz and Sidner's move away from master-slave metaphor in computing precedes more recent efforts in critical STS to examine the history and implications of using such an analogy about technology. For example, in his 2007 article "Broken Metaphor: The Master-Slave Analogy in Technical Literature," Ron Eglash provides a sketch of the history of the analogy.[169] He traces its origins to the relatively recent past, when a clockmaker in South Africa used it describe an "innovative control relationship between two autonomous devices" in the early 20th century. It

---

[168] Cohen et al, *Intentions in Communication,* 427.

[169] Eglash, Ron. "Broken Metaphor: The Master-Slave Analogy in Technical Literature." *Technology and Culture* 48, no. 2 (May 21, 2007): 360–69. https://doi.org/10.1353/tech.2007.0066.

seems the technical literature for Dartmouth's timesharing system was the first instance of the metaphor in the computing context. By the late twentieth century, it was used widely across technical fields as diverse as "automotive clutch and brake systems, clocks, flip-flop circuits, computer drives, and radio transmitters."[170] In more recent years, researchers in STS and tech workers themselves have discussed whether, and how, to remove master-slave analogies in technical contexts. For example, in one viral HASTAC blog post from 2013, Ari Schlesinger discussed the possibility of a feminist computer language and how feminist critiques of logic might provide a framework for creating such a language.[171] Starting in 2014, multiple technology and open source communities, such as Django, Python, Drupal, and Redis, have started to phase out "slave" and "master" from their technical terminologies. Grosz and Sidner's critique of the analogy predates these initiatives by decades.

After she wrote her critique of the master-slave analogy in that 1990 chapter, Grosz incorporated alternative metaphors for human-AI relationships, like that of a teammate, assistant, or partner, into her work. In a 2002 article in the *Harvard Gazette* Barbara Grosz described projects that contributed to her goal of designing collaborative AI systems.[172] She uses an example of building a system to help users find files on a computer. A user might not know or care exactly where the file is located— their priority is inserting the file. Their approach involved designing other ways for the computer to respond to the user and getting computer systems to try multiple approaches to the solve the user's problem. Grosz explains, "We're aiming to have computer

---

[170] Eglash, 360.
[171] Schlesinger, Ari. "Feminism and Programming Languages." HASTAC. Accessed July 27, 2019. https://www.hastac.org/blogs/ari-schlesinger/2013/11/26/feminism-and-programming-languages.
[172] Powell, Alvin. "AI Evolution: From Tool to Partner." *Harvard Gazette*. January 31, 2002. https://news.harvard.edu/gazette/story/2002/01/ai-evolution-from-tool-to-partner/.

systems be team players, acting collaboratively to help us accomplish our goals."[173] This desire sounds straightforward, but getting a computer to act in accordance with broader goals is much more difficult than the straightforward query of searching for file X in subfolder Y. It includes designing systems to make explicit judgements about how to approach a task, like where the computer should search for a file, how to search for the file, and how long the system should search for the file before giving up. Grosz acknowledges the human judgements required in this approach, but sees this as a positive step toward making AI systems that act like a human teammate would. Like her work in linguistics to look beyond utterances, the solutions of Grosz and her colleagues includes looking beyond specific user queries and instead making broader models of tasks, subtasks, possible user goals, and ways computers systems can contribute to these goals.

In the article, Grosz describes another project she and her colleagues were working on called "Writer's Aid"— a citation assistant that sounds like an early version of websites like EasyBib. Her goal at the time of the interview was to make a writing assistant that was helpful like a research assistant. "Being able to model collaborative behaviour and design collaborative software systems will cause a fundamental change in the systems that are available. We're still asking people to adapt to computer systems (rather than the reverse)," Grosz said.[174] "Writing with a computer is a lot easier than writing with a pen, but it's still not as good as writing with a research assistant. I'd like to move toward the computer acting more like a capable assistant than a pen." I will return to this discussion of the difference between a pen and a research assistant, which strikes me as an odd comparison to make. For now, it seems Grosz's aspirations for AI to serve as an assistant or teammate fall somewhere in between weak and strong AI. Throughout her career Grosz

[173] Powell, "AI Evolution."
[174] Powell, "AI Evolution"

cites Searle, an outspoken critic of strong AI. When Grosz describes her goals for AI, she does not use many of the terms that proponents of strong AI use, like thinking, understanding, or artificial consciousness. She does, however, use verbs like participate, achieve, and comprehend when she talks about computers. In the passage where she describes her Writer's Aid project, Grosz makes clear her goal to make AI systems that serve as more than tools.

To design assistant-like AI systems to help users with their goals, Barbara Grosz and her team turn to research in economics and philosophy about how rational agents make decisions. In the 1990s Grosz and colleagues built a simulation system called SPIRE— SharedPlans Intention-Reconciliation Experiments— to study the ways social and environmental factors influence individual and team decision-making strategies.[175] To conceptualize the social aspects of team behaviour, they coin the term *social commitment policies*, which refers to the implicit rules that "govern the rewards and penalties of an individual agent's behaviour in the context of group activities."[176] Unlike social laws, which are domain-specific and provide constraints an agent's actions, social-commitment policies affect decision-making across domains and tasks and "concern *rational choice* and the ways a society can influence an individual's decision making."[177] They cite literature in economics about rational choice and intention reconciliation. As these make clear, these social commitment policies concern rational agents making rational choices. When they discuss their research goals, the epistemological differences between Grosz's research and feminist theorists like Alison Adam becomes clear. "Our longer-term goal," Grosz et al. write, "is

---

175 Sullivan, David G., Alyssa Glass, Barbara J. Grosz, and Sarit Kraus. "Intention Reconciliation in the Context of Teamwork: An Initial Empirical Investigation." In *Cooperative Information Agents III*, edited by Matthias Klusch, Onn M. Shehory, and Gerhard Weiss, 1652:149–62. Berlin, Heidelberg: Springer Berlin Heidelberg, 1999. https://doi.org/10.1007/3-540-48414-0_10
176 Sullivan et al, 152.
177 Sullivan et al, 152.

to derive principles that system designers can use in constructing computer-based agents that participate in teams."[178] The authors make little distinction between human and computer agents. In one section of the chapter, they describe how some agents have different skills and abilities: as their example they describe how some human agents might be better at checking for security breaches while only software agents might be able to run the backup program for a software. The differences between agents are determined by what the agent can provide to the team.[179]

Multi-Agent Systems and the limits of feminist AI research

SharedPlans and Grosz's broader AI approach helped deconstruct certain notions of individualism and research methods that focused too narrowly on specific scripts, tasks, and utterances. She offers critiques of strong AI that do not seem to be caught up with the philosophical questions of consciousness and intentionality like Searle does in the Chinese Room Argument. Neither does she seem to be concerned with "GOFAI" (good old fashioned AI), which centres human-readable intelligence as the benchmark for the success of AI systems. In fact, in her approach to AI research, she seems to avoid using frameworks of thinking, cognition, and intelligence altogether. Instead, she emphasizes AI's *use*: AI researchers should be making technologies that help users solve their problems. Her approach, grounded in speech act theory and other concepts from linguistics, required making AI systems whose models could account for conversational context and collaboration. Grosz and her collaborators were able to incorporate

---

[178] Sullivan et al, 157.
[179] Although the motivations and philosophical underpinnings differ, this flattening of human and non-human relations resonates with approaches in cultural theory like actor network theory (ANT), object-oriented ontology, and conceptualizations of identity, difference, and assemblage put forward by Gilles Deleuze and Félix Guattari (e.g., Latour, Bruno. *Reassembling the Social: An Introduction to Actor-Network-Theory.* 1st edition. Oxford: Oxford University Press, 2007 and Deleuze, Gilles, and Felix Guattari. *A Thousand Plateaus: Capitalism and Schizophrenia*. Translated by Brian Massumi. 2 edition. Minneapolis: University of Minnesota Press, 1987).

their values of collaboration and contextual understanding into mainstream AI research with concepts like SharedPlans. SharedPlans built collaborative capability into existing work on intelligent agents and helped create multi-agent systems that were better at helping their users.

Grosz and her colleagues were able to incorporate more collaborative models of AI systems within particular cultural and epistemological parameters. Grosz aspired to make AI systems that were more than just tools— as she describes in her Writer's Aid project, "more like a research assistant than a pen"— and could collaborate with their users. Multi-agent systems was one of the primary fields of computer science working on models with multiple actors. Grosz's goals require a theory of AI somewhere between weak and strong AI: the *AI as rational agent* paradigm dominant in intelligent agent research and gaining traction in AI more widely in the 1990s seemed to fit the bill. Multi-agent models of human-AI collaboration enable a certain kind of flattening between human and computer, object and subject, human and nonhuman. On the surface, these visions of human-AI partnership sound similar to cyberfeminist ideas from the 1990s and posthuman feminist theories and from the last few decades. Making AI systems in this context, however, meant that Grosz needed to contain both humans and computers both within the parameters of rational agency and self-interest.

These parameters limit how radical the epistemological interventions of Grosz and her collaborators could have been. A strong alignment with the feminist and constructivist work being done at the same time may have entailed a disavowal of the very aspirations of NLP and AI. A constructivist approach to knowledge emphasizes the way claims to truth and knowledge are historically contingent, socially and culturally mediated, and shaped by prior experiences. It would be impossible to build an expert system with all the tacit knowledge of a human expert. What's more, as Alison Adam argues, it is difficult to imagine an AI system that could account for the social construction of knowledge, feminist postmodernism and constructivist approaches to science. An AI system that incorporated the plurality of standpoints and knowledges might not be a very useful assistant for accomplishing a user's goals. Grosz's AI systems, which model agents motivated by rational self-interest, leave little room to incorporate alternative forms of knowledge. SharedPlans and other NLP systems use a universal notion of truth, a shared reality within which intelligent agents can successfully collaborate, comprehend one another, and share joint goals. But whose goals are they, really? In the Writer's Aid project, the answer is clear: the human is the researcher, the computer the research assistant. They might have a shared goal, but there is a clear power imbalance.

In Linda Martín Alcoff's essay "The Problem of Speaking for Others," she writes, "The practice of speaking for others is often born of a desire for mastery, to privilege oneself as the one who more correctly understands the truth about another's situation or as one who can champion a just cause and thus achieve glory and praise."[180] In the essay, Alcoff argues that, whether speaking for oneself or for another, a speaker never escapes the "crisis of representation" and the way a speaker's position shape their interpretation of people and events. Thinking through the design of these systems with the issues discussed in Alcoff's essay points to difficulty of making a feminist NLP or AI technology. Do these systems enact a kind of speaking for others by claiming mastery over their models of human experiences? Assuming rational agency for all agents in a MAS—and building systems to help people based on MAS technologies—risks erasing marginal and non-normative perspectives, just like the script-based models of Schank and Abelson. How could MAS incorporate the variety of knowledge and experiences of generational trauma, of living in

[180] Alcoff, Linda M. "The Problem of Speaking for Others." *Cultural Critique*, no. 20 (1991): 5. https://doi.org/10.2307/1354221.

gendered, sexed, and racialized bodies? Much more work remains to be done about how—if even possible—to build feminist AI technologies that flatten hierarchies and enable collaboration that can also privilege alternative forms of knowledge and leave room for the multiplicity of life experiences.

Conclusion

The year I was conducting research for this thesis, the UK-based organization *Feminist Internet* hosted a series of workshops at the University of Arts London (UAL) for students and staff to imagine, design, and code a "Feminist Alexa." The project, created as part of the inaugural fellowship at UAL's Creative Computing Institute, included a public seminar series and multiple three-day workshops for people from academia and industry to imagine and create prototypes of what a feminist intelligent personal assistant (IPA) might look like. The project culminated in a feminist chatbot prototype named F'xa— pronounced Effects-a— created by *Feminist Internet* to teach people about bias in AI systems from a feminist perspective.

In many ways, *Feminist Internet*'s Feminist Alexa project would not be possible without the contributions Barbara Grosz and her colleagues made to AI and NLP in the mid- to late-twentieth century. *Feminist Internet*'s project cites research done by women computer scientists from the last 40 years throughout their materials. Grosz, her mentor Jane Robinson, and Candace Sidner were pursuing computer science careers right as computer science departments were being established at universities across the United States. While the history of women doing computer science certainly precedes this network of researchers, the network I describe created spaces for themselves and other women in computer science departments when computer science had become a male-dominated field of study. At Harvard, Penn, Stanford, and beyond, this network established that women studied computer science, they had ideas worth studying, and deserved to be supported at all academic levels.

This thesis has examined some of the social, cultural, and intellectual currents which shaped the development of multi-agent AI systems in the second half of the twentieth century. From its early history in the labs of SAIL and SRI in the 1960s and 1970s, MAS approaches have employed imaginaries of collaboration and collective action limited to depoliticized individual actors, governed by rational choice theory and accomplishing discrete, straightforward tasks. Barbara Grosz and her colleagues at SRI and across computer science departments expanded MAS approaches by putting language and context at the centre of their AI systems. Their feminist AI protocol, in addition to creating a more welcoming and supportive space for women in computer science, included a novel approach to human-AI collaboration. They prioritized building systems to help everyday users accomplish everyday tasks, and this emphasis led this network of researchers to use NLP and MAS approaches in their AI research. Their protocol has expanded AI research in many ways beyond single tasks and narrowly defined environments. Although their interventions challenged and expanded MAS and AI research in several ways, the feminist AI protocol of these researchers stayed within certain norms of liberal feminism and the epistemic culture of computer science research. In some ways it leaves us with even more questions about how to actually build AI systems that include anti-imperialism and intersectionality in their feminisms.

Feminist Internet describes their interest in creating a personal assistant that used speech recognition and would satisfy a meaningful human need. "As the voice 'revolution' unfolds," they write, "we see a fantastic opportunity to build conversational interfaces that have drivers beyond commerce - drivers around positive— social change and wellbeing."[181] These priorities to make AI systems centred on the user echo the research contributions of Grosz, Sidner, and their colleagues for the last 40 years. The work of Barbara Grosz and her fellow researchers was grounded in the ideas that to make AI systems useful for people, the systems should incorporate

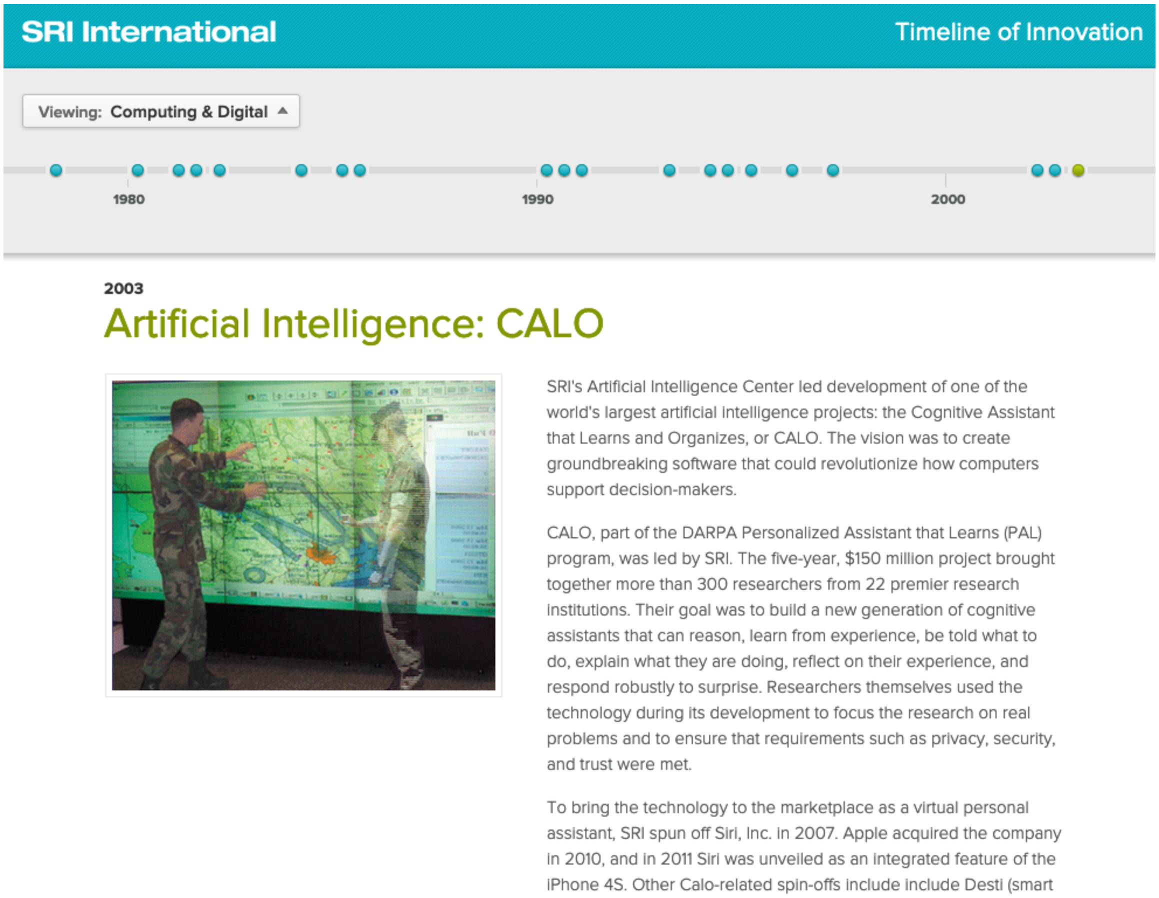
[181] Feminist Internet. "FEMINIST INTERNET." Accessed March 15, 2019. https://feministinternet.com/.

in their models the ideas that people are social; they belong to diverse communities; they use language to do things; and AI systems are none of those things. Any attempts at computer "intelligence" should try to complement these features, not replace them.

Even the "voice revolution" *Feminist Internet* describes would not look the same without the AI research I described in this thesis. Before Amazon's Alexa and the IPAs of Alibaba, Google, Microsoft, and Samsung, Apple's Siri was the first voice assistant to enter the mainstream market in 2011. Before Apple acquired Siri, Inc. in 2010, the company was a spinoff from the CALO project at SRI International. From 2003-2008, DARPA funded several projects as part of their Personal Assistant that Learns (PAL) program. SRI International's project, CALO, stands for Cognitive Assistant that Learns and Organizes, but the name seems to take inspiration from both definitions of *calo* in Latin: as a verb, *calō* comes from *calāre*, which Lewis & Short define as "to call out" or "to summon."[182] As a noun, *cālō* means a servant in the army, a soldier's servant, or any "low servant."[183] CALO, a Department of Defense-funded voice assistant created to follow the orders of military "decision-makers," fits both of these meanings.

[182] Lewis, Charlton T, and Charles Short. "Călo." In *A Latin Dictionary*, 1879. http://www.perseus.tufts.edu/hopper/text?doc=Perseus%3Atext%3A1999.04.0059%3Aentry%3Dcalo1.

[183] Lewis, Charlton T, and Charles Short. "Cālo , -Ōnis, m." In *A Latin Dictionary*, 1879. http://www.perseus.tufts.edu/hopper/morph?l=calo&la=la&can=calo1&prior=calo#Perseus:text:1999.04.0059:entry=calo2-contents.



SRI International

Timeline of Innovation

Viewing: Computing & Digital

1980 1990 2000

2003

Artificial Intelligence: CALO

SRI's Artificial Intelligence Center led development of one of the world's largest artificial intelligence projects: the Cognitive Assistant that Learns and Organizes, or CALO. The vision was to create groundbreaking software that could revolutionize how computers support decision-makers.

CALO, part of the DARPA Personalized Assistant that Learns (PAL) program, was led by SRI. The five-year, $150 million project brought together more than 300 researchers from 22 premier research institutions. Their goal was to build a new generation of cognitive assistants that can reason, learn from experience, be told what to do, explain what they are doing, reflect on their experience, and respond robustly to surprise. Researchers themselves used the technology during its development to focus the research on real problems and to ensure that requirements such as privacy, security, and trust were met.

To bring the technology to the marketplace as a virtual personal assistant, SRI spun off Siri, Inc. in 2007. Apple acquired the company in 2010, and in 2011 Siri was unveiled as an integrated feature of the iPhone 4S. Other Calo-related spin-offs include include Desti (smart

Figure 13. Screen shot taken from "Artificial Intelligence: CALO | SRI International Timeline of Innovation." Accessed May 4, 2019. https://www.sri.com/sites/default/timeline/timeline.php?timeline=computing-digital#!&innovation=artificial-intelligence-calo.

SRI International's AI Center began much of its research on personal assistants as part of the DARPA Speech Understanding Research project in the 1970s— the project for which they hired Grosz— and continued to work on intelligent personal assistants for the rest of the twentieth century. Although Grosz officially left SRI International for Harvard in 1986, her name appears throughout the list of publications associated with CALO and the PAL program from 2003-2011.[184] So, too, does Grosz's longtime collaborator Martha Pollack appear frequently in the publications list: Pollack co-authored nine CALO-funded research publications in 2005 alone.[4] [185] The network of researchers I describe are central figures in the history of SRI International's AI Center, the development of Siri, and consequently other voice assistants entering the US market after Siri. The feminist AI protocol of these researchers foundationally shaped the current the "voice revolution" of the current moment.

In other ways, however, *Feminist Internet*'s Feminist Alexa project reveals some of the limits and contradictions in the Feminist AI Protocol of Grosz and her colleagues. Throughout the Feminist Alexa project, *Feminist Internet* describes their commitment to using an explicitly feminist framework to conceptualize, design, and implement the chatbot workshops and prototypes. In their first series of workshops, called "Designing a Feminist Alexa," *Feminist Internet* introduced participants to their Personal Intelligent Assistant standards— a set of questions they created for the conceptual design of a chatbot or IPA. Their set of standards is a slight adaptation of Josie Swords Young's Feminist Chatbot Design Process— a "series of reflective questions incorporating feminist interaction design characteristics, ethical AI principles

---

[184] "Publications." PAL Project: SRI International, December 8, 2014. https://pal.sri.com/publications/.
[185] PAL Project, "Publications."

and research on de-biasing data"— intended for chatbot design and development teams at tech companies.[186] Young, a feminist AI researcher, designed these questions using scholarship from both feminist theory and computer science. She uses Judy Wacjman's *Technofeminism* as a starting point and incorporates principles from recent papers in feminist HCI and and IEEE's 2016 report *Ethically Aligned Design.*[187] Young's questions invite team members to reflect on the design and implementation of their chatbot at the level of the user, the team, and the overall strategy for creating the chatbot and how it will be represented to the user. It includes questions that encourage team members to reflect on the purposes and ecosystems of the chatbot and articulate assumptions potential users and their goals. Young asks team members to imagine a "marginal user" who might benefit from the chatbot— and how to design for those particular needs, barriers, and viewpoints— instead of designing the system for a single, universal user. Team members are asked to discuss their own values and backgrounds and how they will ensure their own values and expectations are not imposed on the user through their design decisions.

---

[186] Feminist Internet. "Feminist PIA Standards," 2018.
https://drive.google.com/file/d/1J6mMeZxwlOLhxFirIphoJBrOFyZiV6m_.
Swords Young, Josie. "Designing Feminist Chatbots - Research Summary (Sept 2017)." London: Goldsmiths, University of London, September 2017.
https://drive.google.com/file/d/0B036SlUSi-z4UkkzYUVGTGdocXc/view?usp=sharing&usp=embed_facebook.

[187] This report is a result of IEEE's recent Global Initiative on Ethics of Autonomous and Intelligent Systems.

Design & Representation

How are you planning to depict or represent your chatbot to your users?

Have you considered:

1. The ways your user might form social attachments to the PIA, or socially and emotionally connect with it as if it was human?
2. How will your PIA remind the user it's not human?
3. Will you assign a gender to your PIA? Why? In what ways might this reinforce or challenge gender stereotypes? In what ways might this prompt your user to behave unethically or in a prejudiced way?
4. Have you considered a genderless PIA? What possibilities does a genderless chatbot open up for your design?

5. **Representation of the chatbot**
   - How are you planning to depict or represent your chatbot to your users?
     - Have you thought about the ways in which your users will likely form social attachments to your chatbot, and will tend to socially and emotionally connect with it as if it was human?
     - Are you planning on assigning a gender to your chatbot? Why? In what ways might this reinforce or challenge gender stereotypes? In what ways might this prompt your user to behave unethically or in a prejudiced way?
     - Have you considered a genderless chatbot? What possibilities does a genderless chatbot open up for your design?

Figure 14. Some of the questions included in *Feminist Internet*'s PIA Standards and in Josie Swords Young's Feminist Chatbot Design Process, respectively. Screen shots from Feminist Internet. "Feminist PIA Standards," 2018 and Swords Young, Josie. "Designing Feminist Chatbots - Research Summary (Sept 2017)." London: Goldsmiths, University of London, September 2017.

Feminist Internet's feminist protocol also engages explicitly with intersectionality and includes a broad commitment to diversity and inclusion. According to their published materials about the project, the Feminist Alexa workshop leaders, facilitators, and speakers represented different genders, gender identities, ethnicities, nationalities, ages, and values. *Feminist Internet* recruited participants for the Designing a Feminist Alexa workshop with particular emphasis on recruiting people with non-technical backgrounds, first-generation university students, and a balance of UAL students from the UK, EU, and the rest of the world.[188] On the first day of each iteration of the workshop, workshop facilitators introduced the philosophy of the program by referencing feminist theories of Patricia Hill Collins and Chimamanda Ngozi Adichie. In all their written materials, they emphasize the importance of intersectionality and representation of different cultures in any feminist work, and how gender oppression often intersects with colonialism and capitalism. *Feminist Internet*'s feminist AI protocol is one firmly rooted in intersectional, anti-colonial, anti-capitalist perspectives.

An analysis of F'xa, the chatbot prototype borne from *Feminist Internet*'s Feminist Alexa project, suggests an alternate future for intelligent personal assistants in which AI systems are built from a feminist starting point. Unlike other chatbots, whose development might be motivated by commercial success or military applications, F'xa was designed with the specific purpose to teach users about bias in AI systems from a feminist standpoint. *Feminist Internet*'s design decisions for F'xa reflect their commitment to create feminist technologies. For example, when the user asks F'xa a question, instead of providing an "answer from above" F'xa provides possible answers with

[188] Feminist Internet. "Designing a Feminist Alexa: An Experiment in Feminist Conversation Design." UAL creative computing institute. Accessed April 17, 2019. https://drive.google.com/file/d/1vIrIT8dIA9muhvd-XfCCCCUQCujRhMOO/view?usp=drive_open&usp=embed_facebook

information about each perspective. If the user asks F'xa the definition of artificial intelligence, F'xa responds, "There are hundreds of definitions, but here is what some of F'xa's designers think… Whose definition would you like to read?" and offers answers from three of F'xa's creators and a definition from Margaret Boden, a cognitive science and AI researcher who wrote *Mind as Machine*— an extensive, 1,700-page internalist history of the discipline of cognitive science.[189] Providing multiple perspectives on a given inquiry creates room for the user to consider the way the experiences and social positions of the researchers might influence how they conceptualize and experience the world. This design intervention brings chatbots much closer to the kind of AI system Alison Adams argues for in her work, which I discussed in the third chapter: AI designed from a consciously feminist position, based on feminist empiricism and feminist standpoint theory. This decision is one of several design decisions *Feminist Internet* made to incorporate feminist theories and epistemologies into their technology.

Considering the design choices of *Feminist Internet* in juxtaposition with the protocol described in this thesis reveals some of the limits of the feminist AI protocol enacted by Grosz and her fellow researchers. One major difference between the two feminist protocols is *Feminist Internet*'s commitment to creating feminist technologies in solidarity with other movements against imperialism and colonialism. Although feminist critiques of militarism certainly existed in the 1970s and 1980s, in the last 30 years many strands of feminist writing and activism have critiqued of the entanglements of state violence, patriarchal capitalism, and the high-tech military establishment. In her canonical essay "A Cyborg Manifesto," Haraway writes, "The main trouble with cyborgs, of course, is that they are the illegitimate offspring of militarism and patriarchal

---

[189] Boden, Margaret. *Mind As Machine: A History of Cognitive Science Two-Volume Set.* 1 edition. Oxford : New York: Clarendon Press, 2006.

capitalism, not to mention state socialism."[190] In recent years scholars in science and technology studies continue to explore the intersection and potential futures for feminist STS and anti-colonial, decolonial, and postcolonial perspectives.[191] Well-known feminist scholars like Judith Butler and Lila Abu-Lughod have offered numerous critiques of American and Israeli militarism and argued for deep consideration of the way perpetual war contributes to global injustices.[192]

Despite the growing number of these feminist critiques, the actors in this thesis, to my knowledge, have not publicly engaged with these ideas in their AI research or leadership. Far from taking an anti-imperial stance, many of these researchers continue to work with United States military institutions like DARPA, the Department of Defense, and the Office of Naval Research. As the CATO project demonstrates, for many decades now Barbara Grosz and Martha Pollack continue to work with SRI International and receive large research grants from military funders. Throughout the 2000s and 2010s there have been student protests against US military involvement in Iraq and Afghanistan, including demonstrations on the campuses where Barbara Grosz and Martha Pollack were teaching and global movements to demilitarize university research.[193] With

---

[190] Haraway, Donna J. "A Cyborg Manifesto: Science, Technology, and Socialist-Feminism in the Late Twentieth Century." In *Manifestly Haraway*, 5–90. University of Minnesota Press, 2016.

[191] Lyons, Kristina, Juno Parreñas, and Noah Tamarkin. "Engagements with Decolonization and Decoloniality in and at the Interfaces of STS." *Catalyst: Feminism, Theory, Technoscience* 3, no. 1 (October 18, 2017): 1–47.

[192] Butler, Judith. *Precarious Life: The Powers of Mourning and Violence*. Reprint edition. London ; New York: Verso, 2006.

Butler, Judith. *Parting Ways: Jewishness and the Critique of Zionism*. New York: Columbia University Press, 2012.

Abu-Lughod, Lila. "Do Muslim Women Really Need Saving? Anthropological Reflections on Cultural Relativism and Its Others." *American Anthropologist* 104, no. 3 (September 2002): 783–90. https://doi.org/10.1525/aa.2002.104.3.783.

[193] Assad, Bita M. "Crimson Sits Down with Harvard Anti-War Coalition | News | The Harvard Crimson." *Harvard Crimson*. March 19, 2008. https://www.thecrimson.com/article/2008/3/19/crimson-sits-down-with-harvard-anti-war/.

the exception of the few who signed the IJCAI open letter against developing autonomous weapons, this network of AI researchers has not publicly participated in these growing movements to limit military involvement in computer science departments and universities in general. The feminist AI protocol of Grosz, Sidner, Pollack, and their colleagues is based on a kind of liberal feminism, focused on recruiting women into STEM fields, rather than one that challenges norms and institutions in in solidarity with other social movements.

In some ways, the continued entanglements of these researchers with the US military is not abnormal for those working in computer science departments at American universities. The United States Department of Defense continues to provide billions of dollars each year to fund academic research, awarding $2.2 billion in grants in 2017 and $2.3 billion in 2018.[194] According to the Global Campaign on Military Spending, the Pentagon spends $4 billion each year to support university research in the United States.[15] Much of this funding goes to projects in engineering, computer science, and the physical sciences. Given how much money the US spends on its military, in the last 70 years many—if not most—American computer scientists have received military funding to do their research.

As so many feminist scholars have shown in more recent scholarship, however, any commitment to feminism— and specifically intersectional feminism— requires consideration of the way the US military, colonial projects, and the military industrial academic complex contribute

---

Geng, Julie. "Cornell Students Travel to D.C. to Protest Iraq War." *The Cornell Daily Sun*. September 28, 2005. https://cornellsun.com/2005/09/28/cornell-students-travel-to-d-c-to-protest-iraq-war/.
Galef, Daniel. "Demilitarize McGill: The Unexpected Applications of Military Research." *The McGill Tribune*. October 14, 2015. https://www.mcgilltribune.com/sci-tech/demilitarize-mcgill-the-unexpected-applications-of-military-research-151014/.
[194] "USAspending.Gov." USAspending.gov. Accessed September 3, 2019. https://www.usaspending.gov/.

to forms of oppression and injustice worldwide. A feminist practice that does not engage with issues of race, class, sexuality, nationality, ethnicity, and capitalism runs the risk of excluding voices already marginalized and prioritizing the experiences of straight, cisgendered, upper- and middle-class white American women. The need for intersectionality is especially relevant for feminist studies of AI and other computer technologies. Military and surveillance applications of AI systems disproportionately affect women living in zones occupied by the US military, poor women, Indigenous women, and women of colour.[195] There remains an urgent need for computer scientists to use a feminist AI protocol that engages with critiques of colonialism, imperialism, and the military industrial academic complex.

Many of the figures in this thesis are the faces of contemporary institutes and initiatives to promote fairness and ethics in AI. In the last five years, Stanford University has launched several projects to study the social impact of AI technologies. In 2015, they launched their One Hundred Year Study on Artificial Intelligence (AI100), whose inaugural committee members included Julia Hirschberg, Sarit Kraus, and Barbara Grosz's recent graduate student Ece Kamar.[196] From 2015-2019, Barbara Grosz served as the Inaugural Chair of project's Standing Committee. Grosz is the first author of their 2018 report for the Association of Computing Machinery (ACM).[197] In March 2019, Stanford launched their centre for Human-Centred Artificial Intelligence (HAI) and included

---

[195] Browne, Simone. *Dark Matters: On the Surveillance of Blackness*. Durham: Duke University Press Books, 2015.
O'Neil, Cathy. *Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy*. 1 edition. New York: Crown, 2016.
[196] "One Hundred Year Study on Artificial Intelligence (AI100)." Accessed August 21, 2018. https://ai100.stanford.edu/.
[197] Grosz, Barbara J., and Peter Stone. "A Century-Long Commitment to Assessing Artificial Intelligence and Its Impact on Society." *Communications of the ACM* 61, no. 12 (November 20, 2018): 68–73. https://doi.org/10.1145/3198470.

Grosz in their inaugural group of Distinguished Fellows. In April 2019, Barbara Grosz won a $150,000 grant from the Responsible Computer Science Challenge, an initiative of the Mozilla Foundation.[198] The award will support Grosz's Embedded EthiCS program, an initiative to integrate ethics and "ethical reasoning" into computer science courses by collaborating with philosophers to teach modules throughout the standard computer science curriculum at Harvard.[199] These initiatives at Stanford and Harvard show promise, but as I have hoped to show, their lineages come from computer science cultures that to date have engaged very little with feminist epistemologies or critical perspectives on imperialism and colonialism. Given the prestige and influence of these institutions, it is all the more important to pay attention to what—and whom—they exclude and to think deeply about a future of feminist AI studies that makes room for and includes many different standpoints.

The amount of money, power, and prestige in technology research—and AI specifically—means that a social justice-informed approach to AI will be difficult and likely full of contradictions. SharedPlans and the other AI research these women did expanded the kinds of interpersonal interactions and human-machine collaborations that could be possible in AI research. Much work remains, however, to critically examine what kinds of collaboration current AI systems enable and the kinds of hierarchies they continue to maintain. Linda Alcoff argues in "The Problem of Speaking for Others" that, in addition to critical examination of the position of the speaker, feminist scholarship must examine the probable and actual effects of what is said on the broader

---

[198] "Responsible Computer Science Challenge." Mozilla Foundation. Accessed September 6, 2019. https://foundation.mozilla.org/en/initiatives/responsible-cs/.
[199] "Embedded EthiCS @ Harvard." Accessed September 6, 2019. https://embeddedethics.seas.harvard.edu/index.html.

discursive and material context.[200] Analogously, the future of critical, feminist, and socialist AI/NLP research should consider the positions of those who make the technology *and* the effects of these technologies. As feminists, social justice activists, and critical scholars, we must hold each other accountable to the ways technologies can maintain or deepen gendered, sexual, national, racial, and other kinds of hierarchies.

The United States military's ongoing interest in AI complicates the potentials of feminist AI research in computer science departments and other institutions where most sources of funding have some connection to US military interests. The researchers I discuss in this thesis were able to enact significant changes in their departments and the broader AI research community, but to do so required many compromises about other aspects of feminist values. In the mid-twentieth century, the computing power needed to do AI limited AI research to only a handful of institutions, but this is changing. For future waves of feminist and leftist AI research, perhaps futures lie in spaces on the periphery, funded by grassroots organizing and led by community interests. Even in community-led AI research initiatives, however, the current technological landscape requires people doing AI to rely on computing resources from tech companies like Google and Amazon. In Alison Adam's feminist critique of AI epistemology she emphasizes the need for more concrete examples of a "successor science" based on feminist epistemology and social constructivism.[201] In addition, I see a need for "successor technologies" that are built on feminist principles of situated knowledges and standpoint theory. Future critical, feminist, and leftist AI research will likely always involve some amount of complicity, error, compromise, and contradiction, but an

---

[200] Alcoff, Linda. "The Problem of Speaking for Others." *Cultural Critique*, no. 20 (1991): 5. https://doi.org/10.2307/1354221.
[201] Adam, Alison. "A Feminist Critique of Artificial Intelligence." *European Journal of Women's Studies* 2, no. 3 (August 1995): 355–77. https://doi.org/10.1177/135050689500200305.

ecosystem with more feminist technologies would offer possibilities for building AI in new, more equitable directions.

Bibliography

"About the Stanford Movement | April Third Movement." Accessed June 30, 2019. http://www.a3mreunion.org/about-1.html.

Abu-Lughod, Lila. "Do Muslim Women Really Need Saving? Anthropological Reflections on Cultural Relativism and Its Others." *American Anthropologist* 104, no. 3 (September 2002): 783–90. https://doi.org/10.1525/aa.2002.104.3.783.

ACL Wiki. "ACL Lifetime Achievement Award Recipients." Accessed June 6, 2019. https://aclweb.org/aclwiki/ACL_Lifetime_Achievement_Award_Recipients.

Adam, Alison. "A Feminist Critique of Artificial Intelligence." *European Journal of Women's Studies* 2, no. 3 (August 1995): 355–77. https://doi.org/10.1177/135050689500200305.

_______. "Gendered Knowledge — Epistemology and Artificial Intelligence." *AI & SOCIETY* 7, no. 4 (December 1, 1993): 311–22. https://doi.org/10.1007/BF01891414.

"AI Center :: AIC Timeline." Accessed May 27, 2019. http://www.ai.sri.com/timeline/.

Alcoff, Linda M. "The Problem of Speaking for Others." *Cultural Critique*, no. 20 (1991): 5-32. https://doi.org/10.2307/1354221.

Akrich, Madeleine. "The De-Scription of Technical Objects." In *Shaping Technology/Building Society. Studies in Sociotechnical Change*, edited by Wiebe E. Bijker and John Law, 205–24. Cambridge, Massachusetts: MIT Press, 1992. https://halshs.archives-ouvertes.fr/halshs-00081744.

"A.M. Turing Award Winners by Year." Accessed November 6, 2018. https://amturing.acm.org/byyear.cfm.

"Announcements." *Stanford Daily*. November 6, 1974. The Stanford Daily Archives.

Assad, Bita M. "Crimson Sits Down with Harvard Anti-War Coalition | News | The Harvard Crimson." *Harvard Crimson*. March 19, 2008. https://www.thecrimson.com/article/2008/3/19/crimson-sits-down-with-harvard-anti-war/.

Barrett, Samuel, and Peter Stone. "Cooperating with Unknown Teammates in Complex Domains: A Robot Soccer Case Study of Ad Hoc Teamwork." In *Proceedings of the Twenty-Ninth AAAI Conference on Artificial Intelligence*, 7. Austin, Texas, 2015.

Baumgart, Bruce. "Stanford Artificial Intelligence Laboratory Prolegomenon," January 2017. Stanford University Libraries Department of Special Collections and University Archives. purl.stanford.edu/dr245hh6464.

Bobrow, Robert, and Bonnie Webber. "Knowledge Representation for Syntactic / Semantic Processing." In *AAAI-80 Proceedings*. Stanford University, CA: AAAI Press/MIT Press, 1980. https://www.aaai.org/Library/AAAI/1980/aaai80-090.php.

Boden, Margaret. *Mind As Machine: A History of Cognitive Science Two-Volume Set.* 1 edition. Oxford : New York: Clarendon Press, 2006.

Bowles, Nellie. "Overlooked No More: Karen Spärck Jones, Who Established the Basis for Search Engines." *The New York Times*, January 2, 2019, sec. Obituaries. https://www.nytimes.com/2019/01/02/obituaries/karen-sparck-jones-overlooked.html.

Brand, Russell. "SPACEWAR - Fanatic Life and Symbolic Death Among the Computer Bums." *Rolling Stone*, December 7, 1972.

Brandom, Russell. "AI Pioneer Accused of Having Sex with Trafficking Victim on Jeffrey Epstein's Island." *The Verge*, August 9, 2019. https://www.theverge.com/

2019/8/9/20798900/marvin-minsky-jeffrey-epstein-sex-trafficking-island-court-records-unsealed.

Browne, Simone. *Dark Matters: On the Surveillance of Blackness*. Durham: Duke University Press Books, 2015.

Butler, Judith. *Parting Ways: Jewishness and the Critique of Zionism*. New York: Columbia University Press, 2012.

———. *Precarious Life: The Powers of Mourning and Violence*. Reprint edition. London ; New York: Verso, 2006.

Childs, Martin. "John McCarthy: Computer Scientist Known as the Father of AI." *The Independent*. November 1, 2011, sec. Obituaries. http://www.independent.co.uk/news/obituaries/john-mccarthy-computer-scientist-known-as-the-father-of-ai-6255307.html.

Cohen, Philip, Jerry Morgan, and Martha Pollack, eds. *Intentions in Communication*. System Development Foundation Benchmark Series. Cambridge, Massachusetts: MIT Press, 1990.

Crowley, Magdalene L. "Diversity Milestones." Text. EECS at UC Berkeley, February 22, 2018. https://eecs.berkeley.edu/about/diversity/milestones.

Davis, Randall. "Report on the Workshop on Distributed AI." Working Paper. AI Lab: MIT, September 1980. https://dspace.mit.edu/bitstream/handle/1721.1/41155/AI_WP_204.pdf?sequence=4.

Deleuze, Gilles, and Felix Guattari. *A Thousand Plateaus: Capitalism and Schizophrenia*. Translated by Brian Massumi. 2 edition. Minneapolis: University of Minnesota Press, 1987.

"Department Timeline | Stanford Computer Science." Accessed May 23, 2019. https://cs.stanford.edu/about/department-timeline.

Earnest, Les. "SAIL Sagas." Les Earnest, December 13, 2009. https://web.stanford.edu/~learnest/spin/sagas.htm.

Eglash, Ron. "Broken Metaphor: The Master-Slave Analogy in Technical Literature." *Technology and Culture* 48, no. 2 (May 21, 2007): 360–69. https://doi.org/10.1353/tech.2007.0066.

"Embedded EthiCS @ Harvard." Accessed September 6, 2019. https://embeddedethics.seas.harvard.edu/index.html.

Erman, Lee D., Frederick Hayes-Roth, Victor R. Lesser, and D. Raj Reddy. "The Hearsay-II Speech-Understanding System: Integrating Knowledge to Resolve Uncertainty." *ACM Comput. Surv.* 12, no. 2 (June 1980): 213–253. https://doi.org/10.1145/356810.356816.

Feminist Internet. "Designing a Feminist Alexa: An Experiment in Feminist Conversation Design." UAL creative computing institute. Accessed April 17, 2019. https://drive.google.com/file/d/1vIrIT8dIA9muhvd-XfCCCCUQCujRhMOO/view?usp=drive_open&usp=embed_facebook

———. "FEMINIST INTERNET." Accessed March 15, 2019. https://feministinternet.com/.

———. "Feminist PIA Standards," 2018. https://drive.google.com/file/d/1J6mMeZxwlOLhxFirIphoJBrOFyZiV6m_.

Fiegener, Mark K. "Science and Engineering Degrees: 1966–2010." Detailed Statistical Tables. NSF: National Center for Science and Engineering Statistics, June 2013. https://nsf.gov/statistics/nsf13327/content.cfm?pub_id=4266&id=2,

Feil-Seifer, David, Kristine Skinner, and Maja J Matarić. "Benchmarks for Evaluating Socially Assistive Robotics." *Interaction Studies: Psychological Benchmarks of Human-Robot Inteaction* 8, no. 3 (October 2007): 423–29.

Finding the Information Wood in Natural Language Trees, Lecture by Karen Jones Spärck. YouTube video. 1994 Grace Hopper Celebration of Women in Computing Conference, 1994. https://www.youtube.com/watch?v=5fYeKiebpuo.

Ford, Martin. "Barbara Grosz." In *Architects of Intelligence: The Truth about AI from the People Building It*, 333–57. Packt Publishing, 2018.

Franklin, Jane. "Call for Equality: Join the Women's Liberation." *The Stanford Daily*, October 7, 1969. The Stanford Daily Archives.

Freedman, Estelle. "Women at Stanford: Inclusion, Exclusion, and Activism from the 1890s to the 1990s." presented at the Stanford Historical Society, Stanford University, March 7, 2019. https://www.youtube.com/watch?v=CPH3ZDUrL7A.

Fortino, Giancarlo, Antonio Guerrieri, Wilma Russo, and Claudio Savaglio. "Integration of Agent-Based and Cloud Computing for the Smart Objects-Oriented IoT." In *Proceedings of the 2014 IEEE 18th International Conference on Computer Supported Cooperative Work in Design (CSCWD)*, 493–98. Hsinchu, Taiwan: IEEE, 2014. https://doi.org/10.1109/CSCWD.2014.6846894.

Future of Life Institute. "Open Letter on Autonomous Weapons," July 28, 2015. https://futureoflife.org/open-letter-autonomous-weapons/.

Galef, Daniel. "Demilitarize McGill: The Unexpected Applications of Military Research." *The McGill Tribune*. October 14, 2015. https://www.mcgilltribune.com/sci-tech/demilitarize-mcgill-the-unexpected-applications-of-military-research-151014/.

Geng, Julie. "Cornell Students Travel to D.C. to Protest Iraq War." *The Cornell Daily Sun*. September 28, 2005. https://cornellsun.com/2005/09/28/cornell-students-travel-to-d-c-to-protest-iraq-war/.

Grosz, Barbara J. "ACL Lifetime Achievement Award: Smart Enough to Talk With Us? Foundations and Challenges for Dialogue Capable AI Systems." *Computational Linguistics* 44, no. 1 (March 2018): 1–15. https://doi.org/10.1162/COLI_a_00313.

_______. "Barbara Grosz CV," May 2017. https://grosz.seas.harvard.edu/files/grosz/files/barbara_grosz_cv_may_2017.pdf.

_______. "CS288: Computational Models of Discourse Reference List." Harvard University, 2000. https://www.seas.harvard.edu/courses/cs288/handouts/references.pdf.

_______. "Report on Women in the Sciences at Harvard. Part I: Junior Faculty and Graduate Students." Faculty of Arts and Science Committee on the Status of Women. Cambridge, Massachusetts: Harvard University, February 13, 1991, 1-36.

_______. "Utterance and Objective: Issues in Natural Language Communication." *AI Magazine*, 1980, 11-20.

_______. "Women in Computing: Collaborative Plans and Dialogue Participation, Lecture by Barbara Grosz." presented at the 1994 Grace Hopper Conference on Women in Computing, June 9, 1994. https://www.youtube.com/watch?v=2VrrH-hHVmQ&list=PLDEiwPwOpXnaI416H7xik1omEZq0nEdtU&index=7&t=0s.

Grosz, Barbara J. and Candace L. Sidner. "Attention, intentions, and the structure of discourse." *Computational Linguistics* 12, no 3 (1986): 175-204.

_______, Eva Hajicova, and Aravind Joshi. "Obituary Jane J. Robinson." *Computational Linguistics* 41, no. 4 (2015): 723–26. https://doi.org/10.1162/COLI_a_00235.

_______, Karen .S. Jones, and Bonnie L. Webber. *Readings in Natural Language Processing*. Los Altos, California: Morgan Kaufmann Publishers, Inc., 1986.

_______, and Peter Stone. “A Century-Long Commitment to Assessing Artificial Intelligence and Its Impact on Society.” *Communications of the ACM* 61, no. 12 (November 20, 2018): 68–73. https://doi.org/10.1145/3198470.

Groszfest Session 1: 1980’s. Video recording. Vol. 1. 4 vols. GroszFest. Maxwell Dworkin Laboratory, Harvard University, Cambridge, Massachusetts, 2018. https://matterhorn.dce.harvard.edu/engage/player/watch.html?id=21b373c3-5cac-4271-88ae-32aea140fd3e.

Haraway, Donna J. “A Cyborg Manifesto: Science, Technology, and Socialist-Feminism in the Late Twentieth Century.” In *Manifestly Haraway*, 5–90. University of Minnesota Press, 2016.

_______. *Primate Visions: Gender, Race, and Nature in the World of Modern Science.* New York: Routledge, 1990.

Hayles, N. Katherine. *How We Became Posthuman: Virtual Bodies in Cybernetics, Literature, and Informatics.* University of Chicago Press, 2008.

Hicks, Mar. *Programmed Inequality: How Britain Discarded Women Technologists and Lost Its Edge in Computing.* Cambridge, MA: MIT Press, 2017.

Hirschberg, Julia. “Julia Hirschberg.” Columbia University Computer Science Department, April 19, 2019. http://www.cs.columbia.edu/~julia/.

Huang, Xuedong, James Baker, and Raj Reddy. “A Historical Perspective of Speech Recognition.” *Communications of the ACM* 57, no. 1 (January 1, 2014): 94–103. https://doi.org/10.1145/2500887.

Hutchinson, Jamie, "Culture, Communication, and an Information Age Madonna," *IEEE Profesional Communication Society Newsletter* 45, no. 3 (May/June 2001): 1.

“IFAAMAS: Awards: Influential Paper.” International Foundation for Autonomous Agents and Multiagent Systems. Accessed June 23, 2019. http://www.ifaamas.org/award-influential.html.

“IFAAMAS Home Page.” Accessed August 27, 2018. http://www.aamas-conference.org/.

IFAAMAS. “Victor Lesser Distinguished Dissertation Award.” Accessed November 5, 2018. http://www.aamas-conference.org/award-victorlesser.html.

Jacobs, Joanne. “Engineering School Seeks to Enrol More Women.” *The Stanford Daily*, February 5, 1973. The Stanford Daily Archives.

“John McCarthy - A.M. Turing Award Laureate.” Accessed April 20, 2019. https://amturing.acm.org/award_winners/mccarthy_1118322.cfm.

Jones, Karen Spärck. Computing’s too important to be left to men. Interview by Brian Runciman. Online, March 2007. https://www.bcs.org/content/ConWebDoc/10791.

_______. Oral History: Karen Spärck Jones - Engineering and Technology History Wiki. Interview by Janet Abbate. Transcription, April 10, 2001. IEEE History Centre. https://ethw.org/Oral-History:Karen_Sp%C3%A4rck_Jones.

“Karen Spärck Jones Award.” British Computing Society Information Retrieval Specialist Group, 2009. https://irsg.bcs.org/ksjaward.php.

“Karen Spärck Jones Lecture.” BCS - The Chartered Institute for IT. Accessed June 23, 2019. https://www.bcs.org/events-home/karen-spaerck-jones-lecture.

Klatt, Dennis H. "Review of the ARPA Speech Understanding Project." *Journal of the Acoustical Society of America* 62, no. 6 (December 1977): 1345–66. https://doi.org/10.1121/1.381666.

Latour, Bruno. *Reassembling the Social: An Introduction to Actor-Network-Theory.* 1st edition. Oxford: Oxford University Press, 2007.

Leslie, Stuart W. *The Cold War and American Science: The Military-Industrial-Academic Complex at MIT and Stanford.* New York: Columbia University Press, 1993.

Lesser, Victor, and Daniel Corkill. "Comprehensive History of the Multi-Agent Systems Lab at the University of Massachusetts Amherst 1978-2014," 2016. http://mas.cs.umass.edu/Documents/LabHistory_Web-Article.pdf.

_______. "Preface." Menlo Park, California: AAAI, 1995. https://www.aaai.org/Library/ICMAS/icmas95contents.php.

Lettvin, J., H. Maturana, W. McCulloch, and W. Pitts. "What the Frog's Eye Tells the Frog's Brain." *Proceedings of the IRE* 47, no. 11 (November 1959): 1940–51. https://doi.org/10.1109/JRPROC.1959.287207.

Lewis, Charlton T, and Charles Short. "Călo." In *A Latin Dictionary*, 1879. http://www.perseus.tufts.edu/hopper/text?doc=Perseus%3Atext%3A1999.04.0059%3Aentry%3Dcalo1.

_______. "Cālo , -Ōnis, m." In *A Latin Dictionary*, 1879. http://www.perseus.tufts.edu/hopper/morph?l=calo&la=la&can=calo1&prior=calo#Perseus:text:1999.04.0059:entry=calo2-contents.

Luong, Peter, and Liming Lu. "Multiagent Web for the Internet of Things." Seoul, South Korea: IEEE, 2014. https://doi.org/10.1109/ICISA.2014.6847432.

Lyman, Richard W. *Stanford in Turmoil: Campus Unrest, 1966–1972*. Stanford University Press, 2009. https://doi.org/10.11126/stanford/9780804760799.001.0001.

Lyons, Kristina, Juno Parreñas, and Noah Tamarkin. "Engagements with Decolonization and Decoloniality in and at the Interfaces of STS." *Catalyst: Feminism, Theory, Technoscience* 3, no. 1 (October 18, 2017): 1–47.

"Make $3,000 by Jply If You Are Female, Attractive, Uninhibited." Newspaper advertisement. *The Stanford Daily*, April 1, 1971. The Stanford Daily Archives.

Mervis, Jeffrey. "What Kind of Researcher Did Sex Offender Jeffrey Epstein like to Fund? He Told Science before He Died." *Science | AAAS*, September 19, 2019. https://www.sciencemag.org/news/2019/09/what-kind-researcher-did-sex-offender-jeffrey-epstein-fund-he-told-science-he-died.

Metz, Cade. "John McCarthy -- Father of AI and Lisp -- Dies at 84." *Wired*, October 25, 2011. https://www.wired.com/2011/10/john-mccarthy-father-of-ai-and-lisp-dies-at-84/.

Minsky, Marvin. *The Society of Mind*. 2nd ed. New York: Simon and Schuster, 1986.

Mulvin, Dylan, and Jonathan Sterne. "Scenes from an Imaginary Country: Test Images and the American Color Television Standard." *Television & New Media* 17, no. 1 (2016): 21–43. https://doi.org/10.1177/1527476415577211.

Munson, Jr., David C. "A Note on Lena." *IEEE Transations on Image Processing* 5, no. 1 (January 1996).

Murphy, Michelle. *Seizing the Means of Reproduction: Entanglements of Feminism, Health, and Technoscience.* Duke University Press, 2012.

Norberg, A. L. "Changing Computing: The Computing Community and DARPA." *IEEE Annals of the History of Computing* 18, no. 2 (Summer 1996): 40–53. https://doi.org/10.1109/85.489723.

"Notes from the Protest Outside SRI," May 16, 1969. April Third Movement Archives. http://www.a3mreunion.org/archive/1968-1969/68-69_may16_sri/files_68-69_may_16/A3M-5-16_Notes_1.pdf.

"One Hundred Year Study on Artificial Intelligence (AI100)." Accessed August 21, 2018. https://ai100.stanford.edu/.

O'Neil, Cathy. *Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy*. 1 edition. New York: Crown, 2016.

Pearl, Amy, Martha E. Pollack, E.A. Riskin, Becky Thomas, Elizabeth Wolf, and Alice Wu. "Becoming a Computer Scientist." *SIGCSE Bulletin* 34, no. 2 (June 1, 2002): 135–43. https://doi.org/10.1145/543812.543847.

Pollack, Martha E, Julia Hirschberg, and Bonnie L Webber. "1982 - User Participation in the Reasoning Processes of Expert Systems." *AI Magazine* AAAI-82 Proceedings (1982): 358-361.

Powell, Alvin. "AI Evolution: From Tool to Partner." *Harvard Gazette*. January 31, 2002. https://news.harvard.edu/gazette/story/2002/01/ai-evolution-from-tool-to-partner/.

"Publications." PAL Project: SRI International, December 8, 2014. https://pal.sri.com/publications/.

Rankin, Joy Lisi. *A People's History of Computing in the United States*. Cambridge, Massachusetts: Harvard University Press, 2018.

"Report on the Second Workshop on Distributed AI." Working Paper. AI Lab: MIT, January 1982. https://dspace.mit.edu/bitstream/handle/1721.1/41171/AI_WP_228.pdf?sequence=4.

"Responsible Computer Science Challenge." Mozilla Foundation. Accessed September 6, 2019. https://foundation.mozilla.org/en/initiatives/responsible-cs/.

"SAIL Farewell." Accessed May 24, 2019. http://infolab.stanford.edu/pub/voy/museum/pictures/AIlab/SailFarewell.html.

"SAILDART Stanford Artificial Intelligence Lab DART Archive." Digital archive. SAILDART, May 2012. https://www.saildart.org/.

Savaglio, Claudio, Giancarlo Fortino, Maria Ganzha, Marcin Paprzycki, Costin Bădică, and Mirjana Ivanović. "Agent-Based Computing in the Internet of Things: A Survey." In *Intelligent Distributed Computing XI*, edited by Mirjana Ivanović, Costin Bădică, Jürgen Dix, Zoran Jovanović, Michele Malgeri, and Miloš Savić, 737:307–20. Springer International Publishing, 2018. https://doi.org/10.1007/978-3-319-66379-1_27.

Schank, Roger C., and Robert P. Abelson. *Scripts, Plans, Goals, and Understanding: An Inquiry Into Human Knowledge Structures.* New York: Psychology Press, 1977.

_______. "Scripts, Plans, and Knowledge." *IJCAI'75 Proceedings of the 4th International Joint Conference on Artificial Intelligence* 1 (September 3, 1975): 151-57.

Schlesinger, Ari. "Feminism and Programming Languages." HASTAC. Accessed July 27, 2019. https://www.hastac.org/blogs/ari-schlesinger/2013/11/26/feminism-and-programming-languages.

Searle, John R. "Minds, Brains, and Programs." *Behavioral and Brain Sciences* 3, no. 3 (September 1980): 417–24. https://doi.org/10.1017/S0140525X00005756.

Searle, John. 1969. *Speech Acts: An Essay in the Philosophy of Language.* Cambridge University Press, Cambridge, England.

Sidner, Candace. "On Being a Woman at MIT: Or, How to Miss the Stumbling Blocks in Graduate Education." *Association for Women in Mathematics Newsletter*, February 1982, 13–17.

Spertus, Ellen. "Why Are There so Few Female Computer Scientists?" MIT Artificial Intelligence Laboratory Technical Report. Cambridge, Massachusetts: MIT, 1991. https://dspace.mit.edu/bitstream/handle/1721.1/7040/AITR-1315.pdf?sequence=2.

Sullivan, David G., Alyssa Glass, Barbara J. Grosz, and Sarit Kraus. "Intention Reconciliation in the Context of Teamwork: An Initial Empirical Investigation." In *Cooperative Information Agents III*, edited by Matthias Klusch, Onn M. Shehory, and Gerhard Weiss, 1652:149–62. Berlin, Heidelberg: Springer Berlin Heidelberg, 1999. https://doi.org/10.1007/3-540-48414-0_10

Swords Young, Josie. "Designing Feminist Chatbots - Research Summary (Sept 2017)." London: Goldsmiths, University of London, September 2017. https://drive.google.com/file/d/0B036SlUSi-z4UkkzYUVGTGdocXc/view?usp=sharing&usp=embed_facebook.

"Systers - AnitaB.Org." Accessed June 28, 2019. https://anitab.org/systers/.

Takahashi, Tomoichi, Satoshi Tadokoro, Masayuki Ohta, and Nobuhiro Ito. "Agent Based Approach in Disaster Rescue Simulation - From Test-Bed of Multiagent System to Practical Application." In *RoboCup 2001: Robot Soccer World Cup V*, edited by Andreas Birk, Silvia Coradeschi, and Satoshi Tadokoro, 2377:102–11. Berlin, Heidelberg: Springer Berlin Heidelberg, 2002. https://doi.org/10.1007/3-540-45603-1_11.

Traum, David, and Jeff Rickel. "Embodied Agents for Multi-Party Dialogue in Immersive Virtual Worlds." *Proceedings of the International Conference on Autonomous Agents*, August 10, 2002. https://doi.org/10.1145/544862.544922.

Turner, Fred. *From Counterculture to Cyberculture: Stewart Brand, the Whole Earth Network, and the Rise of Digital Utopianism.* Chicago, Ill.: University of Chicago Press, 2008.

University of Pennsylvania School of Engineering and Applied Science. *JoshiFest: "Centring Recollections" by Barbara Grosz*. YouTube video, 2012. https://www.youtube.com/watch?v=d67nNtSAirk.

_______. *JoshiFest:"Did You Feed the Animals?" By Julia Hirschberg*. YouTube video, 2012. https://www.youtube.com/watch?v=DgzhIkSFAQk.

"USAspending.Gov." USAspending.gov. Accessed September 3, 2019. https://www.usaspending.gov/.

Wajcman, Judy. *TechnoFeminism*. Cambridge ; Malden, MA: Polity, 2004.

"Wanted: Uninhibited Girl for Experimental Film." Newspaper advertisement. *The Stanford Daily*, February 26, 1971. The Stanford Daily Archives.

Wiederhold, Voy. "Computer History Exhibits." Accessed May 28, 2019. http://infolab.stanford.edu/pub/voy/museum.html.

Wilde, Anna D. "TFs Will Be Tested On Spoken English | News | The Harvard Crimson." *Harvard Crimson*. May 21, 1993. https://www.thecrimson.com/article/1993/5/21/tfs-will-be-tested-on-spoken/.

Woo, Elaine. "John McCarthy Dies at 84; the Father of Artificial Intelligence." *Latimes.Com*. October 27, 2011. https://www.latimes.com/local/obituaries/la-me-john-mccarthy-20111027-story.html.

Woolgar, Steve. “Configuring The User: The Case of Usability Trials.” *The Sociological Review* 38 (1990): 58–99. https://doi.org/10.1111/j.1467-954X.1990.tb03349.x.

Yuan Luo, Kecheng Liu, and D. N. Davis. “A Multi-Agent Decision Support System for Stock Trading.” *IEEE Network* 16, no. 1 (January 2002): 20–27. https://doi.org/10.1109/65.980541.

“Zoe Pictures 1971.” Accessed May 24, 2019. https://www.saildart.org/Zoe/.