3 FIGURING SERVICE IN DISCOURSES OF ICT: THE CASE OF SOFTWARE AGENTS

Lucy Suchman
Centre for Science Studies
Department of Sociology
Lancaster University
United Kingdom

The much discussed shift, particularly within the industrialized countries of North America and Europe, to a service economy is underwritten in myriad ways by information and communications technologies (ICT). In this paper, I explore some aspects of how the provision of service is figured, both rhetorically and materially, in contemporary discourses of ICT. I use the idea of *figuration* here in the sense developed by cultural historian of science Donna Haraway (1997, p. 11). Haraway's argument is, first, that all language, including the most technical or mathematical is figural; that is, it is made up of tropes or turns of phrase that at least evoke, if not directly invoke, associations across diverse realms of meaning and practice. Technologies, Haraway argues, are forms of *materialized figuration*; that is, they bring together particular assemblages of stuff and meaning into more and less stable arrangements. These arrangements—commonly termed *configurations* in the practices of technology research and development—imply in turn particular ways of relating humans and machines. One form of intervention into current practices of ICT development, then, is through a critical consideration of how humans and machines are currently figured—and figured together—in those practices, and how they might be configured differently.

Given this general approach, my more particular aim here is to bring together two established critiques of the way that humans, and their relations to machines, are currently configured in ICT research and development. The first of these concerns efforts to develop intelligent, interactive machines—interactive not just in the sense that computational media make possible new, very distinctive dynamics of writing and reading, but in the sense of machines that can engage in conversation with us. My argument, in brief, is that efforts so far to create intelligent machines (whatever our views on the ultimate possibility of that enterprise) are deeply conservative, in their assumption that the model human is the rational, autonomous individual. This assumption, in turn, trivializes the embodied competencies involved in intelligence and interaction as these are

enacted by humans. Nonetheless, the aim of artificial intelligence (AI) is to replicate this particular image of the human in the form of a rational, autonomous, interactive machine.

The second line of critique starts from the observation that discourses of ICT have tended to erase the human labor that continues to be involved in technological production, implementation, maintenance and the like. This erasure is tied to the more general ways in which information has been rhetorically dematerialized—has "lost its body" in Katherine Hayles' apt phrase (1999, p. 2). Through the particular case of software agents, I explore the proposition that contemporary ICT projects restage the very problematic dream of a perfect, invisible infrastructure, in ways that join together the fantasies of AI with the promises of a service economy.

The stage is set well by a figure courtesy of British writer P. G. Wodehouse, circa 1923:

> "Morning, Jeeves," I said.
>
> "Good morning, sir," said Jeeves.
>
> He put the good old cup of tea softly on the table by my bed, and I took a refreshing sip. Just right, as usual. Not too hot, not too sweet, not too weak, not too strong, not too much milk, and not a drop spilled in the saucer. A most amazing cove, Jeeves. So dashed competent in every respect. I've said it before, and I'll say it again (Wodehouse 1999, p. 1).

So opens the first chapter of *The Inimitable Jeeves*, subtitled "Jeeves Exerts the Old Cerebellum." The inimitability (or not) of Jeeves, and the cultural imaginaries within which Jeeves' competencies are attributed to his cerebellum, provide the starting place for my observations. Jeeves is the icon of the consummate service provider, the ever-resourceful "gentleman's personal gentleman." The just-visible-enough worker, he directs his considerable skills to maintaining the comfort and respectability of his employer, the upper class, good-natured, but slightly dim-witted Bertie Wooster.

While created close to a century ago, it is evident that Jeeves prefigures a central object of contemporary projects in computing and interface design; that is, the interactive software agent.

As the robot was to industrial imagery, so the software agent is to the desires and fantasies of the service economy. Rather than machines that can do our heavy lifting for us, the dream now is that every one of us can be a Bertie Wooster, commanding a staff of servants that gets to know us intimately, watches out for us, keeps us informed in just the ways that we need to be

(knowing better what those ways are than we do ourselves), and represents us faithfully in our everyday affairs.

To understand these latest materialized figurations, it is useful to look back briefly at their hardware ancestors in the form of robots and their computational kin. Perhaps most obviously, the history of robotics and AI is the latest episode of the ongoing serial "the separation of body and mind." While industrial robots are designed to be super-strong or special purpose bodies, intelligent machines are designed to think. A progeny of the Cold War, the first thinking machines develop in the context of dreams and anxieties over military command and control, code-breaking, and the like (see Edwards 1996; Hayles 1999). Perhaps one of the most consequential moves in the early days of AI was the importation of language used to describe human behavior—search, recognition, learning, problem-solving—to describe computational processes. At the same time, researchers in the emerging field of cognitive psychology adopted more and more of the language of mathematics and information processing to describe human thought.

My interest in the projects of AI and interactivity at the interface dates back to the early 1980s at the Xerox Palo Alto Research Center (PARC), where I became intrigued by an effort to design an interactive interface to a particular product. The effort was initiated in response to a delegation of Xerox customer service managers who traveled to PARC to report on a problem with the machine and to enlist research advice in the problem's solution. The machine was a relatively large, feature-rich photocopier that had just been launched, mainly as a placeholder to establish the company's presence in a particular market niche that was under threat from other, competitor companies. The machine was advertised with a figure dressed in the white lab coat of the scientist or engineer, but reassuring the viewer that all that was required to activate the machine's extensive functionality was to "push the green [start] button." It seemed that customers were refuting this message, however, complaining instead that the machine was, as the customer service managers reported it to us, "too complicated."

My interest turned to investigating just what specific experiences were glossed over by that general complaint, a project that I followed up among other things by convincing my colleagues that we should install one of the machines at PARC and invite our coworkers to try to use it. My analyses of the troubles evident in these video-taped encounters by actual scientists and engineers with the machine led me to the conclusion that its obscurity was less a function of any lack of general technological sophistication on the part of its users, than of their lack of familiarity with this particular machine. I argued that the machine's complexity was tied less to its esoteric technical characteristics than to mundane difficulties of interpretation characteristic of any unfamiliar artifact. My point was that making sense of a new artifact is an inherently problematic activity.

Moreover, I wanted to suggest that however improved the machine interface or instruction set might be, this would never eliminate the need for active sense-making on the part of prospective users. This called into question, then, the very viability of marketing the machine as self-explanatory, or self-evidently easy to use.

My colleagues, meanwhile, had set out on their own project, to design an intelligent, interactive computer-based interface to the machine that would serve as a kind of coach or expert advisor in its proper use. Working within the context of research in artificial intelligence at the time, the phrase self-explanatory machine for them was coming to take on new meaning. Along with the more traditional notion that an artifact is self-explanatory just to the extent that a prospective user is able to reconstruct how its designer intended it to be used, my colleagues were interested in the prospect that an artificially intelligent machine might actually be able to explain itself in something more like the sense that a human being does. In this second sense, the goal was that the artifact should not only be *intelligible* to the user as a tool, but that it should be *intelligent*—that is, able to understand the actions of the user, and to povide for the rationality of its own.

My own research turned to a critical analysis of my colleagues efforts to create this expert advisor, and of the field of AI and human-computer interaction more generally (Suchman 1987). I took as my focus the question of interactivity, and assumptions about human conversation within the field of AI, working those against findings that were emerging in sociological studies of face-to-face human conversation. The main observation of the latter was that human conversation does not follow the kind of message-passing or exchange model that formal, mathematical theories of communication posit. Rather, humans dynamically co-construct the mutual intelligibility of a conversation through an extraordinarily rich array of embodied interactional competencies, strongly situated in the circumstances at hand (which are, in turn, unfolding through that same interaction). I accordingly adopted the strategy of taking the premise of interaction seriously, and applying a similar kind of analysis to peoples' encounters with the machine to those being done in conversation analysis. The result of this analysis was a renewed appreciation for some important differences, and more particularly asymmetries, between humans and machines as interactional partners, and for the profound difficulty of the problem of interactive interface design.

While the dreams of AI have fallen on some difficulties since that time, the project has responded by shape-shifting into ever new (and at the same time, I will argue, ever old and familiar) forms. One of these is the intelligent software agent, revived in contemporary ICT discourse by the growth of the Internet. So we find a renaissance of enthusiasm about "knowbots," online personal assistants, and other computational artifacts attributed with a capacity for intelligent,

interactive behavior. By way of introduction, here is a definition from a recent paper in the *International Journal of Human-Computer Studies*:

> Interface agents are computer programs that aid a user in accomplishing tasks carried out at the computer, such as sorting email, filtering information and scheduling meetings. These agents differ from conventional computer programs in that *they can act autonomously on behalf of the user*, that is, without requiring the user to enter a command or click a button whenever she wants the task to be carried out. In addition to autonomy, a characteristic of intelligent agents is *their ability to perform tasks delegated to them in an intelligent, that is context- and user-dependent way* (Dehn and van Mulken 2000, p. 1, emphasis added).

The first claim made here I think is most interesting for its choice of the word "autonomy," in contrast to the relatively straightforward examples of functionality ("sorting email, filtering information and scheduling meetings") that the authors suggest, while the second claim regarding intelligence, albeit framed now in terms of "context" and "user-dependence," begs the same set of questions regarding machine interactivity that my own, and others', critiques had earlier raised.

One thing that has unquestionnably changed, however, is the rise of graphics and animation as resources for interface design. Dehn and van Mulken report that advances in computer graphics now enable *animated interface agents.* Of the latter, they write:

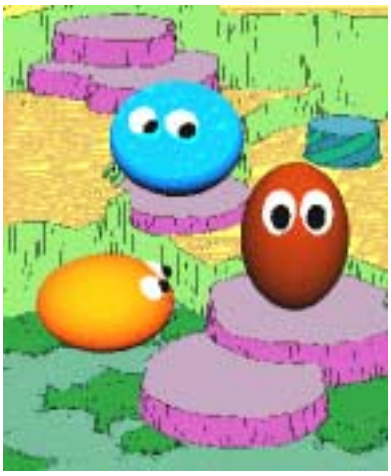
> Such agents appear on the screen as embodied characters and exhibit various types of life-like behaviours, such as speech, emotions, gestures and eye, head and body movements (p. 2).

Setting aside for the moment the sense of the term "embodied" here, I am interested in the question of just how these agents differ from conventional cartoon characters. Here is what the classic reference work on Disney animation has to say about cartooning:

> Disney animation makes audiences really believe in...characters, whose adventures and misfortunes make people laugh— and even cry. There is a special ingredient in our type of animation that produces *drawings that appear to think and make decisions and act of their own volition; it is what creates the illusion of*

> *life* (*Disney Animation: The Illusion of Life*, cited in Bates 1994, p. 122, emphasis added).

This seems quite straightforward, using the language of *appearances* and *illusions*. So what is different about the claims being made for agents at the interface?



© Oz Project, Carnegie Mellon University, used by permission

This quote is taken from an article by Joseph Bates in a special issue of *Communications of the ACM* on Intelligent Agents. The approach taken by Bates is to import techniques developed to portray emotion in cartoon characters into a computer program, called "Edge of Intention," populated by three cartoon creatures named Woggles.

The medium of cartooning is appropriate here in more than a technical sense. What *emotions* become in this system are a series of emotional/behavioral attributions mapped to visual features of the figures. So for example a state labeled "sadness" triggers a "moping behavior," expressed through a "decreased muscle tone," shorter jumps and slower actions (Bates 1994, p. 124). As with cartoon animation, the artful synthesis of cartoonist's design work and viewers' readings results in successful animations. But for Bates, the achievement is more than that. As he puts it, the result of the work is

> creatures with definite emotional reactions to events. A simple example is a Woggle creating an analog of anger when it both experiences an important goal failure and judges that the failure was caused by another Woggle....We took care to design an architecture that provided Woggles with strong internal emotional states (pp. 123-124).

In this single passage, Bates' creatures are simultaneously presented as just illusions of life *and* as important steps along the path to the real thing. Why, if a Woggle has emotional reactions, experience, judgement, and strong internal emotional states does it create only "an analog of anger"? The rhetoric, it seems, s getting very slippery indeed.

Woggles and pets notwithstanding, the most popular role for software agents remains that of personal representative or assistant to the user, which brings us

back to Jeeves. The idea of personal agents was animated perhaps most vividly in the form of Phil, the bow-tied agent in Apple's 1984 video "The Knowledge Navigator," but more modest implementations abound. Jeeves' travels through the interface are exemplified most directly, of course, in the Web search service Ask Jeeves®. And in a feature article in the May 2001 issue of the popular technoscience magazine, *Scientific American*, Tim Berners-Lee and his coauthors present their vision for the successor to today's World Wide Web, named (before its birth, in the manner typical of many software projects) "The Semantic Web." The authors animate their project with a scenario reminiscent of the Knowledge Navigator, although updated to include a hand-held Web device:

> The entertainment system was belting out the Beatles' "We Can Work It Out" when the phone rang. When Pete answered, his phone turned the sound down by sending a message to all the other local devices that had a volume control. His sister, Lucy, was on the line from the doctor's office: "Mom needs to see a specialist and then has to have a series of physical therapy sessions…I'm going to have my agent set up the appointments." Pete immediately agreed to share the chauffering.
>
> At the doctor's office, Lucy instructed her Semantic Web agent through her handheld Web browser. The agent promptly retrieved information about Mom's prescribed treatment from the doctor's agent, looked up several lists of providers, and checked for the ones in-plan for Mom's insurance within a 20-mile radius of her home and with a rating of excellent or very good on trusted rating services. It then began trying to find a match between available appointment times (supplied by the agents of individual providers through their Web sites) and Pete and Lucy's busy schedules (Berners-Lee et al. 2001, p. 36).

From Bertie Wooster's trials as a member of the British leisure class, we move to the dilemmas of the baby boomer engaged in a busy working life, called upon to care for aging parents under a regime of privately insured health care. While Mom apparently still needs to be transported bodily to her physical therapist, the rest of the logistics are adeptly handled by Pete and Lucy's software agents, and with just the right degree of deference (the first agent-generated plan is submitted for approval and sent back for modification). Issues of privacy, trust, and the like are dispatched through the application of appropriate techniques alluded to at relevant moments in the scenario.

As the authors explain, "Pete and Lucy could use their agents to carry out all these tasks thanks not to the World Wide Web of today, but rather the Semantic Web that it will evolve into tomorrow" (p. 36). The article describes how a new

language of machine-readable Web content—a system of "well defined meanings"—will underwrite that evolutionary process (p. 37). They conclude that,

> Properly designed, the Semantic Web can assist the evolution of human knowledge as a whole.…This structure will open up the knowledge and working of humankind to meaningful analysis by software agents, providing a new class of tools by which we can live, work and learn together (p. 43).

The ideal that unites these scenarios is that agents should be enough like us to understand our desires and to figure out on their own how to meet them, but without either their own desires or ambitions, or other human frailties that might get in the way of efficient and effective accomplishment of their assigned tasks. Another example, announced by the online news service Ananova™, is their personified newscaster, imaged as a somewhat Barbie-doll-like female figure, and described as follows:

> The world's first virtual newscaster...Ananova is a super-fast, super-intelligent news and information computer system with a difference—she has a "human" face and personality. She has been programmed to act like a human newscaster, responding with relevant emotions and actions depending on the nature of the information she is imparting. Unlike a flesh-and-blood newsreader she is in action every second of the day and can deliver any number of different bulletins or pieces of information simultaneously (Ananova 2000).

What appears constant across these cases is that the litmus test of a good agent is the agent's capacity to be autonomous, on the one hand, and just what we want, on the other. We want to be surprised by our machine servants, but not displeased. At the same time we live in an age that embraces the ideal of the independent, self-motivated, entrepreneurial worker. As Henry Lieberman asks in his paper "Autonomous Interface Agents":

> Why autonomous agents? An assistant may not be of much practical help if he or she needs very explicit instruction all the time and constant supervision while carrying out actions. Assistants can be time-savers when they are allowed to act independently and concurrently. Allowing an interface agent to run off-line and in parallel, with the user directing attention to other activities, enables the user to truly delegate tasks to the agent (1997, p. 2).

Here then is a classic tension. As management theory has pointed out with respect to the problem of agents and delegation in business administration, the more empowered these others, and the more capable of pursuing their own self-interests rather than ours, the less reliable they are. There is a deep and enduring ambivalence, in other words, inherent in the image of the agent: on the one hand, the agent as faithful representative, on the other, the agent as autonomous, self-directed, and therefore able to pursue its own agenda.

Somewhat paradoxically, one could argue that it is actually the persistence of the modernist human/machine divide rather than its disappearance that makes the prospect of machine autonomy so compelling to those interested in the design of intelligent, interactive artifacts. The modernist assumption is that agency is something contained within singular individuals, and in this respect the project of designing intelligent artifacts—however distributed intelligence is understood to be—remains consistent with a tradition that treats separation and autonomy, rather than relatedness, as the mark of humanity. Having systematically established the division of humans and machines, many technologists now seem worried that once separated from us machines are rendered lifeless, and by implication less. They need to be revitalized, restored to humanness—in other words, to be made like us—in order that we can be reunited with them.

What we see, moreover, in these recent initiatives in software agency is the persistence of what I take to be a central figure across industrial and post-industrial initiatives around new technologies. This figure has been insightfully discussed within science and technology studies under the name of the invisible worker and invisible infrastructures, from Steve Shapin's (1989) observations about the role of technicians in scientific discovery to recent work by Bowker and Star (1999) on systems of classification. Just as the dream of the robot worker was to relieve us of hard labor, or of the contingencies of managing others so engaged, so the dream of animated agents at the interface promises to relieve us from having either to perform the mundane work involved in providing services for ourselves, or to negotiate the moral dilemmas and practical inconveniences of delegating that work to others who might, more and less faithfully, represent us.

In *How We Became Posthuman* (1999), N. Katherine Hayles' offers a critical interrogation of the evacuation of the body and materiality from the field of informatics in the early days of the field's formation. Hayles recounts at length the proceedings of the Macy Conferences on Cybernetics, a series of 10 meetings held between the years 1946 and 1953. As she closes her discussion, she brings a previously absent figure onto the stage: the conference assistant and transcriptionist Janet Freed, pictured in a photograph from the 1953 meeting. The photograph, as Hayles recounts it, shows a U-shaped table around which are seated a large group of men and two women—the anthropologist Margaret Mead and, "with her back to the photographer, her arms extended, hands reaching out to a machine I can't quite see," Janet. As Hayles reflects:

> Thinking of her, I am reminded of Dorothy Smith's suggestion that men of a certain class are prone to decontextualizaton and reification because they are in a position to command the labor of others. "Take a letter, Miss Freed," he says. Miss Freed comes in…the man speaks, and she writes on her stenography pad (or perhaps her stenography typewriter). The man leaves. He has a plane to catch, a meeting to attend. When he returns the letter is on his desk, awaiting his signature. From his point of view, what has happened? He speaks, giving commands or dictating words, and things happen. A woman comes in, marks are inscribed onto paper, letters appear, conferences are arranged, book are published. Taken out of context, his words fly, by themselves, into books. The full burden of the labor that makes these things happen is for him only an abstraction, a resource diverted from other possible uses, because he is not the one performing the labor.
>
> Miss Freed has no such illusions. Embedded in context, she knows that words never make things happen by themselves. They can't put marks on paper. They can't get letters in the mail. They can't bring twenty-five people together at the right time and in the right place, at the Beekman Hotel in New York City, where white table cloths and black chalkboards await them. For that, material and embodied processes must be used—processes that exist never in isolation but always in contexts where the relevant boundaries are permeable, negotiated, instantiated (pp. 82-83).

The software agent is figured within a discourse that makes service the imperative for a global economic infrastructure. The relations of upstairs and downstairs, frontstage and back, that the service economy presupposes are constituted within a closed world that simultaneously presumes and regenerates the needs, desires, identities and inequalities that those relations comprise. Just as the decorum of Bertie Wooster's world is maintained by the supporting activities and discrete interventions of Jeeves, the dream of technology innovators in the service economy is that new sociomaterial agents and infrastructures will make it possible for more and more of *us* to be hailed as persons residing upstairs rather than down. My concern, then, is with the kinds of *we's* that are posited by this future vision, widening the circle of those who employ, manage, and command to include more and more of us, while the others who serve us are refantasized from problematic human workers to the now-quite-imitable in silicon Jeeves. Discourses of software agents at once naturalize the desirability of service provision, while further obscuring the specific sociomaterial infrastructures, including growing numbers of human workers, on which smooth interactions at the interface continue to depend.

REFERENCES

Ananova™ "Virtual Newscaster," http://www.ananova.com/ (accessed June 2000).
Ask Jeeves® http://www.askjeeves.com/.
Bates, J. "The Role of Emotion in Believable Agents," *Communications of the ACM* (37:7), 1994, pp. 122-125.
Berners-Lee, T., Hendler, J., and Lassila, O. "The Semantic Web," *Scientific American*, May 2001, pp. 36-43.
Bowker, G., and Star, S. L. *Sorting Things Out: Classification and Its Consequences.* Cambridge, MA: MIT Press, 1999.
Dehn, D., and van Mulken, S. "The Impact of Animated Interface Agents: A Review of Empirical Research," *International Journal of Human-Computer Studies* (52), 2000, pp. 1-22.
Edwards, P. N. *The Closed World: Computers and the Politics of Discourse in Cold War America.* Cambridge, MA: MIT Press, 1996.
Haraway, D. *Modest_Witness@Second_Millenium.FemaleMan©_Meets_OncoMouse™: Feminism and Technoscience.* New York: Routledge, 1997.
Hayles, N. K. *How We Became Posthuman: Virtual Bodies in Cybernetics, Literature and Informatics.* Chicago: University of Chicago Press, 1999.
Lieberman, H. "Autonomous Interface Agents," in S. Pemberton (ed.), *Proceedings of the Conference on Human Factors in Computing Systems*, 1997, pp. 67-74.
Shapin, S. "The Invisible Technician," *American Scientist* (77), 1989, pp. 554-563.
Suchman, L. *Plans and Situated Actions: The Problem of Human-Machine Communication.* New York: Cambridge University Press, 1987.
Wodehouse, P. G. *The Inimitable Jeeves.* Harmondsworth, UK: Penguin, 1999 (1923).

About the Author

Lucy Suchman is Professor of Anthropology of Science and Technology in the Department of Sociology at Lancaster University. She spent 20 years as a researcher at Xerox's Palo Alto Research Center before taking up her present position. Her research has centered on the sociomaterial practices that make up technical systems, explored through critical studies and through experimental, interdisciplinary, and participatory interventions in new technology design.

She is currently preparing a second revised edition of her book, *Plans and Situated Actons: The Problem of Human-Machine Communication*, which will include an extended new introduction looking at relevant developments since the mid-1980s both in computing and social studies of technology. The focus will be on agent systems and new forms of interactivity at the interface on the one hand, and on recent theorizing regarding humans, machines, and relations between them on the other.

Lucy can be reached by e-mail at l.suchman@lancaster.ac.uk.