Journal of EVOLUTION & Technology

A peer-reviewed electronic journal published by the Institute for Ethics and Emerging Technologies

ISSN 1541-0099

19(1) – September 2008

Cognitive Expansion Technologies

William Sims Bainbridge
http://mysite.verizon.net/wsbainbridge/

Journal of Evolution and Technology - Vol. 19 Issue 1 – September 2008 - pgs 8-16
http://jetpress.org/v19/bainbridge.htm

Abstract

In ancient times, at great effort over the span of many years, people learned to do arithmetic, to read, to write, and to measure reality with rulers and eventually clocks. A person who cannot handle any of these cognitive tools is a very different creature from somebody who can. The changes happening now will be at least as significant, and will occur much faster, probably within a single human lifetime. This article will consider cutting-edge research being done today, and extrapolate its implications some distance into the future. One theme of this survey is that very humble information technology applications could have a cumulative effect that is extremely dramatic, all the more stunning because we ourselves will be involved every step of the way.

A major transformation of the human mind has begun, a revolution as profound as any that has occurred in human history, yet it is occurring peacefully in our own personal lives. It involves nothing less than the integration of human thought with electronic systems for perceiving, processing, and presenting information. Search engines on the World Wide Web are the most obvious part of the first wave of this revolution, but researchers are preparing additional waves that will wash over us from all directions, changing what it means to be human. Perhaps the most visible next stage is the proliferation of semi-autonomous artificial intelligence agents, which are beginning to take over cognitive tasks from the humans who own them, and in so doing to transfer aspects of human personalities into machines and across the wider net.

Personal Artificial Intelligence Agents

The fact that computers are learning to adapt to their users' unique characteristics was impressed upon me when I finished editing the *Encyclopedia of Human-Computer Interaction* in 2004 (Bainbridge 2004). Peter Brusilovsky of the University of Pittsburgh contributed an essay on adaptive help systems, and Alfred Kobsa of the University of California at Irvine contributed one on adaptive interfaces, distinct but related ways in which recent software personalizes itself to the style of perception and decision-making of the user. Attentive user interfaces, as described in another essay by Ted Selker of MIT, actually watch the user and infer what he or she is paying attention to. Humans and computers are learning to talk with each

other, as explained in essays on dialog systems by Susan McRoy of the University of Wisconsin at Milwaukee, speech recognition by Mary Harper of the University of Maryland and V. Paul Harper of the US Patent Office, speech synthesis by Jan van Santen of the Oregon Health and Science University, and natural language processing by James Martin of the University of Colorado.

Erika Rogers of California Polytechnic State University describes the increasing subtlety of human-robot interaction. The ways that smart homes can serve the intimate needs of their inhabitants are described by Diane Cook and Michael Youngblood of the University of Texas at Arlington. How computers, robots and software agents can acquire their own minds is described in essays on artificial intelligence (AI) by Robert St. Amant and multiagent systems by Gal Kaminka of Bar Ilan University. An essay on affective computing by Ira Cohen (Hewlett-Packard Research Labs), Thomas Huang (University of Illinois, Urbana-Champaign), and Lawrence Chen (Eastman Kodak Research Labs) outlines current efforts to give robots and computers emotions, or at least to let them deal appropriately with human feelings. My own encyclopedia essay on personality capture explores the future possibility of transferring significant aspects of individual human personalities to dynamic information systems.

Cognitive assistive information technologies will often be applied first, and in frankly somewhat primitive form, to the needs of disabled people whose problems are especially acute. For example, research is rapidly developing the technologies to orient each person in space and time, through location-aware mobile computing. Project ACCESS (Assisted Cognition in Community, Employment and Support Settings) at the University of Washington has developed an AI advisor that learns the personal routines of cognitively impaired people and guides them whenever they seem to get lost while traveling around the city of Seattle (Liao et al 2004, Patterson et al 2004). With funding from the National Science Foundation, Edmund Durfee and Martha Pollack at the University of Michigan have launched a project

> ...to solve technical problems that need to be overcome to build socio-cognitive orthotic systems, which will augment human cognitive capabilities to promote social interactions. Information technology can help a person with cognitive impairment in managing his or her everyday life, by modeling the activities the person wants or needs to do, monitoring the person's activities as they unfold, and guiding the person to ensure that the most important activities occur. Thus, information technology can provide a cognitive orthotic that augments reduced cognitive abilities and helps the person live independently. (National Science Foundation Awards.)

In order to help their users, these artificial intelligence systems learn some of the goals, habits, and limitations of their human owners, thus coming to resemble their users in some ways. In the near future, we may expect systems that learn alongside their owners, over the years of a rich life, serving nor only as advisors but also as personal archivists. At Carnegie-Mellon University, Howard Wactlar's Experience-on-Demand Project developed "tools, techniques and systems allowing people to capture a record of their experiences unobtrusively, and share them in collaborative settings spanning both time and space." His more recent CareMedia Project monitors elderly people in a skilled nursing facility to capture "a continuous, voluminous audio and video record... that enables more complete and accurate assessment, diagnosis, treatment, and evaluation of behavioral problems for the elderly" (Informedia site) Steve Mann's EyeTap group at the University of Toronto is developing a system that not only archives everything a person sees, but also provides information via augmented reality to the user in realtime (Mann 2004).

One really intriguing idea is the Virtual Lifetime Tutor imagined by Jean Scholtz of the Visualization and Usability Group of the US National Institute of Standards and Technology. This was one of sixteen "Grand Challenges" for computing identified by a task force of the US government, so it is technically realistic and there is a very real chance that it will be developed in the coming years (Strawn et al. 2003). The tutor would consist of an information system communicating through an artificial intelligence agent,

assigned to an individual person. A person would acquire his or her personal AI tutor in childhood, and it would mature as the person did. Importantly, it would adapt to its owner's personal learning style, strengths, weaknesses, and changing knowledge levels across many fields. It would know what the person needed to learn next, manage the best educational plan, and offer one-on-one tutoring for "just-in-time" skill acquisition. Whether the tutoring involves a Spanish language refresher course, quick training with new software on the job, or college science classes, the tutor would be a real partner in the learning experience.

As I envision the tutor, it would come to know its user very well, even to reflect that individual's unique personality in great detail. It would have access to everything the person learns, be able to prompt the user with specific information when needed, but also share the same knowledge as the person. What would some of the consequences be? Perhaps all-purpose AI agents could be the death of the advertising business. Advertising pushes products upon us, whereas an AI agent that knows its owner's tastes could travel the Web looking for things the owner would want, such as novels and music, ignoring advertising but evaluating through the machine equivalent of reading and listening.

Field work in sciences like anthropology and sociology would be transformed, because the agent could transcribe interviews and collect observations into an easily analyzed database. The family album would be replaced by the ability to replay any experience the person enjoyed. Internet-based class reunions could be grand affairs as classmates' agents offered up a phantasmagoria of images that blend both past and present, from multiple interlocking perspectives, and deceased classmates could be represented by their agents, if living classmates wanted, acting as if they were the deceased human through a realistic avatar and artificial intelligence personality emulation.

Games and Drama

The fact that the economic value of video games seems to have surpassed that of movies raises profound questions about the long-term viability of some traditional art forms, notably literature and drama. Why should the proscenium arch separate the spectator from the actor? Why should a play have only one possible conclusion? Why should so many great works focus primarily on one character – Hamlet instead of Ophelia, Oedipus instead of Jocasta – rather than focusing at will on any of the characters? Why should drama and literature primarily be written in the third person in the past tense – he or she did this or that - rather than on the second person in the present tense – you are doing what you want in the context devised by the playwright? The increasing complexity of AI characters in video games suggests that we may want new answers to such questions.

The helicopter disappears into the blinding Antarctica blizzard, leaving you and your three AI assistants to discover what terrible thing happened at this remote outpost, that caused radio contact to be broken a month ago. You cannot do this job alone, because you need the skills of your medic, engineer, and commando, but you are not sure you trust any of them. Nor do they have very good reason to trust you. Indeed, as you find the horribly mutilated bodies of the outpost crew, you must work very hard to build trust among your increasing distraught assistants, while at the same time testing their reliability. This is the situation in the PlayStation 2 videogame *The Thing*. The AI agents in this game are really not very smart, but they provide a good foretaste of what the future will bring. Already researchers like Kathryn Merrick and Mary Loy Maher at the University of Sydney, are beginning to design AI game agents that have their own interests, responding to novelty and learning new skills, rather than merely carrying out pre-programmed behaviours (Merrick and Maher 2006).

Pervasive gaming is a potentially important new computer application that is being developed, especially in Europe. A pervasive game is played both online and in real world environments. Traditionally, most games were played on table tops or athletic fields, but a few were played in the natural environment, notably scavenger hunts, Easter eggs hunts, or Capture the Flag. Today, location aware wearable

computers – or cell phones in communication with a computer system – make it possible to overlay the real world with virtual objects and processes. The simplest example is an online scavenger hunt in which players must go to particular real locations to pick up items that exist only in cyberspace. More complex are Live Action Role Playing games ("larps") in which the players interact partly online.

Feeding Yoshi involves getting seeds from online creatures called Yoshis who can be accessed only at certain locations in town, planting the seeds and growing fruit at other real locations, then feeding the fruit to the Yoshis to earn points (Bell et al. 2006). The goal of *Uncle Roy All Around You* was finding the office where Uncle Roy was waiting, through a long series of steps, such as going to a designated public telephone for the next instruction, interacting with actors who play game-related roles or with passersby who are not aware a game is being played, and even violating norms such as "stealing" postcards or getting into a stranger's car (Benford et al. 2006). In June 2005, twelve people played the pervasive larp *Prosopopeia,* each taking the identity of a real deceased person to track down a ghost in an abandoned Stockholm mental hospital Jonsson et al. 2006). At present, these games require a huge amount of human labor, done following the "Wizard of Oz scenario" by a "man behind a curtain," but in the future AI agents can be designed to manage all the special effects and serve as facilitating characters.

We can easily imagine a time when pervasive gaming is a vast industry, connected to tourism, education, and politics. Every city will offer distinctive games, operating continuously, which relate somehow to the actual local environment. Consider a game I will call *Unterwelt*, played in contemporary Washington but pretending that the date is 1944. A team of twenty Nazi spies competes with a team of twenty American counterspies. The goal of the Nazis is to locate and photograph a complete list of targets that will be destroyed by a barrage of submarine-launched rockets. The counter-spies try to mislead the Nazis and lure them into ambushes. In the virtual world, only buildings that existed in 1944 can be found. A hundred versions of Unterwelt may be running at any given time, involving perhaps 4,000 players. Other Washington-based games may reenact the Civil Rights and anti-war demonstrations of the 1960s, or a science fiction plot set in the year 2100 when hacker heroes attempt to collect secrets that could defeat the dictatorial government. How the real government reacts to such augmented reality activities could become part of the game, for a few of the more adventurous players.

Clearly, pervasive gaming can help the players learn history, geography, and other subjects as part of the fun, if they are designed correctly. But they can also be designed to change the player's personalities. Timid people may act brave, when playing a dramatic role, and their simulated courage may gradually transfer to the real world as increased confidence. I imagine a future therapy, the Displacement Service, unlike a job placement service in that its aim is to displace you from your current unsuccessful social location. The Displacement Service analyzes a person's weaknesses then gives him or her pervasive game experiences that go just beyond what the person would find acceptable in ordinary life. Gradually, the person's abilities to handle feelings and situations would improve, and this learning may transfer to real-world experiences because the game takes place in both the real and virtual worlds.

Personality Transfer

One theme of this paper has been the ways in which AI agents and information systems can come to learn about the user through serving his or her needs and sharing his or her experiences. This implicitly captures aspects of the individual's personality, in the service of a variety of smaller practical tasks. Now I shall consider more direct methods to capture a personality so that it can be transferred to an information system where the person can be preserved, emulated, and perhaps eventually transferred to another, more durable form. A secondary benefit of personality capture can be analysis and thus potential improvement of the individual person.

While impressed by the progress going on in novel fields like artificial intelligence, in my own personality capture research I have tried to build upon what has already been accomplished in the

traditional social and behavioral sciences. For example, a classic experiment done by Saul Sternberg in 1966 (Sternberg 1966) measures the short-term memory of an individual research subject, and my 1986 textbook/software *Experiments in Psychology* includes a derived program anyone may use to determine the capacity of his or her own short-term memory (Bainbridge 1986).

At the same time, I developed software systems to administer and analyze questionnaires of any length, beginning with *Experiments in Psychology* and my 1989 textbook/software *Survey Research: A Computer-Assisted Introduction* (Bainbridge 1989). In 1997, I launched a website called the Question Factory to generate the material for thousands of new questionnaire items that could chart an individual's beliefs, preferences, attitudes, and character. Often, the most creative act involves taking something that exists and simply reversing one or two of its assumptions. Consider this: Traditional questionnaire research has one person write a hundred questions to be answered by a thousand people. Why not have thousands of people write thousands of questions to be answered by one person? This places the individual in his or her unique, precise position with respect to the surrounding culture represented by the thousands.

My software module, *The Year 2100*, is a good example. Based on progress with the Question Factory, I was invited to contribute items to online questionnaires sponsored by the National Geographic Society. One open-ended item asked, "Imagine the future and try to predict how the world will change over the next century. Think about everyday life as well as major changes in society, culture, and technology." About 20,000 people wrote their thoughts about the year 2100, and I culled fully 2,000 distinct predictions from the ocean of text they wrote. Each prediction became a pair of fixed-choice questionnaire items, first asking the respondent to rate each one in terms of how good it would be if it came about, and second asking how likely that would be to happen. Several publications have already been based on this project (Bainbridge 2003 and 2004), and Figure 1 shows highly simplified results of one person who responded to all 4,000 questions.

For sake of clarity, Figure 1 groups the 2,000 predictions in 20 groups of 100, roughly described by names, including predictions about labor, human knowledge, and domestic life. Because they are so closely bunched together, eight categories are not named in Figure 1: family, business, population, conflict, government, nature, justice, and the quality of life. Each dot reflects 200 measurements, how good and how likely on average the respondent judges the 100 predictions in the category to be. Strikingly, for this respondent the 100 predictions about the human future in outer space are both much better and much less likely than the other 19 categories. Thus, this respondent is very pessimistic with respect to space development.

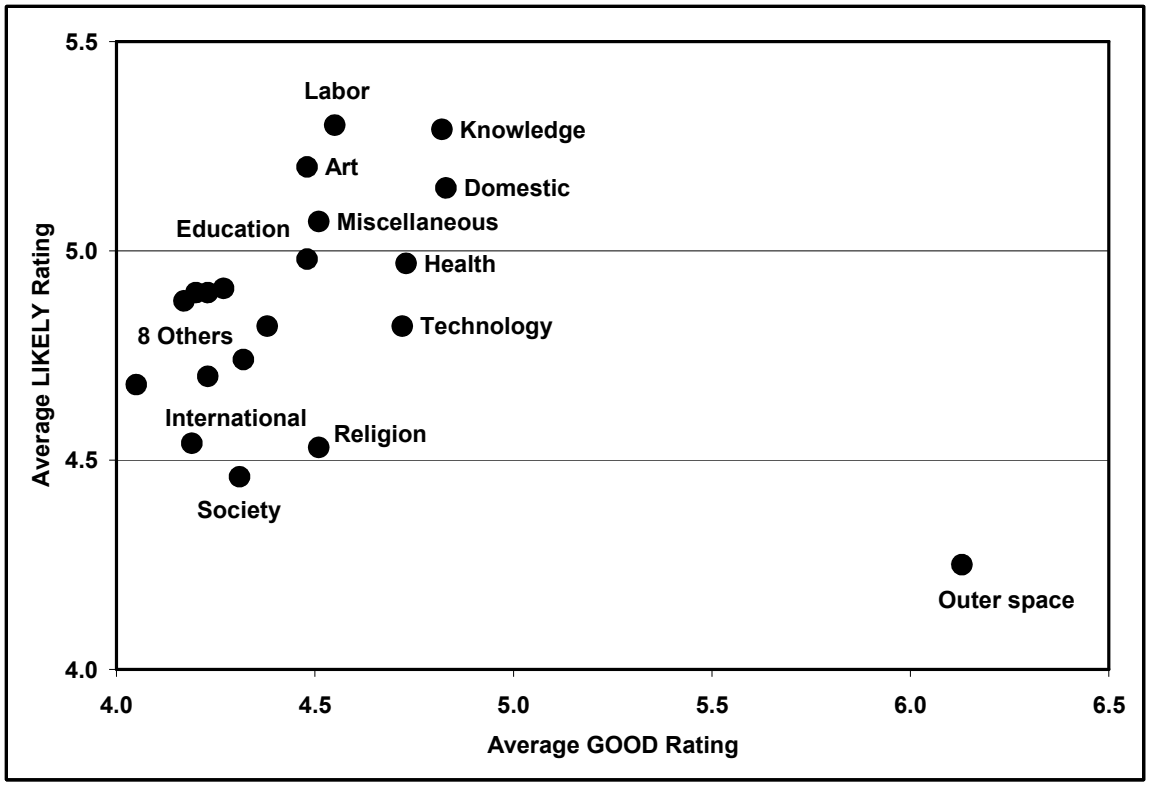

Figure 1: One Respondent's Rating of how Good and Likely 2,000 Predictions Are

A person's views of the future of the world are a major part of his or her worldview. Byproducts of this software include measures of how optimistic or pessimistic the person is in twenty areas of life, a list of the person's utopian ideas that are both very good and very unlikely, and clues about good and likely areas where the person might want to invest his or her own effort. Ten other software modules, in addition to *The Year 2100*, have developed means of measuring other aspects of personality with a total of 44,000 questions.

The *Self* module offers a second example, which we can use to see how personality capture might become the basis for personality emulation. It consists of 1,600 adjectives, consisting of 800 pairs of opposites, that could describe a person. Respondents rate each adjective in terms of how good or bad it is for a person to have that quality, and in terms of how much they themselves feel the adjective describes them personally. This software offers a general analysis of the individual's self image, which is a big part of personality, including measures of self-esteem in twenty areas of life. The software also generates lists of good qualities the person feels he or she lacks, and bad qualities he or she possesses, that could be used as guides for self-improvement.

Data from *The Year 2100* or *Self* could be incorporated in an AI agent, to make the agent behave more like the person. For a very primitive demonstration, a respondent whose *Self* data had already been reported in the scientific literature (Bainbridge 2003) was asked to name some books he had enjoyed reading in childhood. The respondent said that around age twelve he had been an avid fan of the novels of Edgar Rice Burroughs, especially his path-breaking fiction about the planet Mars, but also his better-known Tarzan novels. As it happens, the respondent has also written an unpublished book about his

family's history. I created a fresh computer program that scanned six Burroughs novels plus the family history for the 1,600 adjectives. It turned out that 1,079 of the words appeared in at least one of the books, and the average book included 3,200 instances of the adjectives. Figure 2 shows a slightly more subtle analysis.

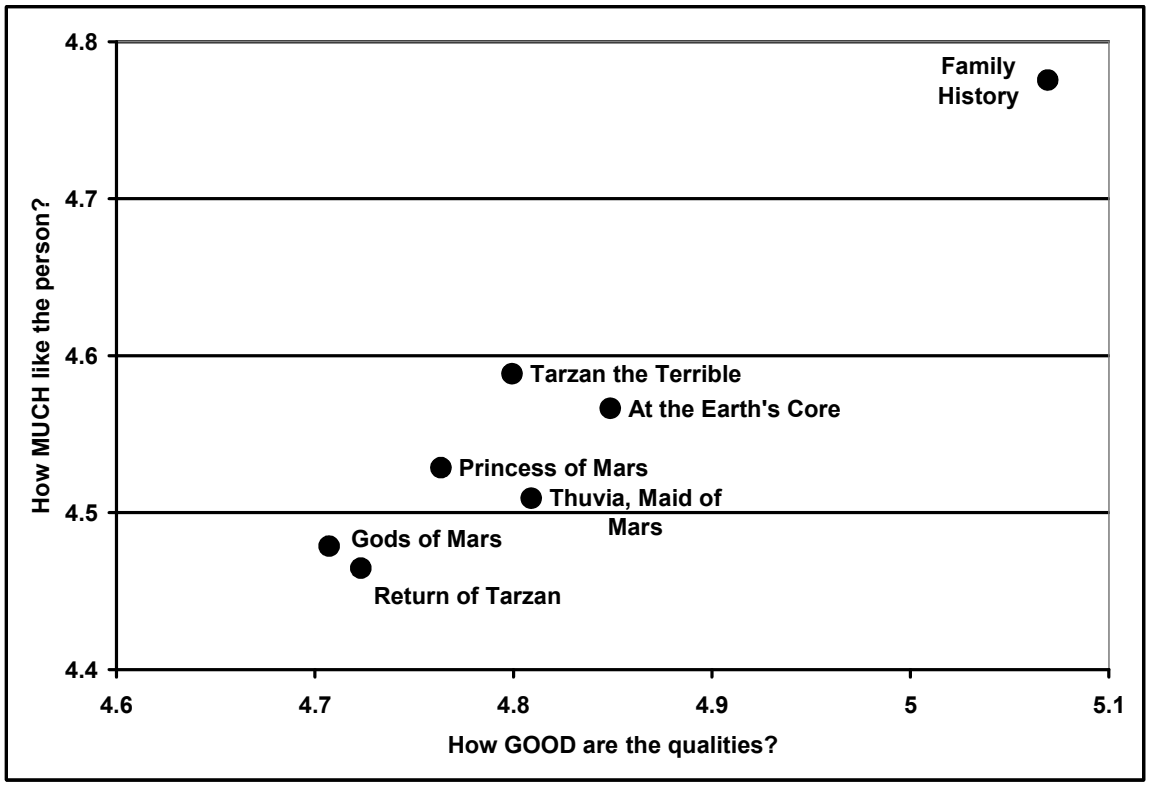

Figure 2: A Computer "Reading" Seven Books with a Person's Values

To make Figure 2, for each time an adjective appeared in a book, the respondent's two ratings of that adjective were tallied. For example, adjectives from the list of 1,600 appear 4,548 times in the respondent's family history manuscript. On average, across the 4,548 cases, they rate 5.07 on the "good" scale and 4.78 on the "much" scale that measures how much the respondent believes the adjective describes himself. The averages are much lower for the six novels, which appear to be arranged away from the family history in three pairs.

The distribution partly reflects the fact that the novels contain villains, described by very negative adjectives. *Tarzan the Terrible* and *At the Earth's Core* are weak in villains, chiefly describing marvelous lands containing dinosaurs. *Gods of Mars* and *Return of Tarzan* are the second books in their respective series, and in both some very wicked villains prevent the hero and heroine from uniting. Once computerized natural language processing has progressed to the point that it can recognize which character each adjective applies to, the software could provide an analysis of how the respondent relates to each of the major characters in a novel, even a novel he has not yet read. An AI surrogate reader could advise its owner about what things to read, and could do an improved job finding desired information, even as it emulates the behavior of the person.

Conclusion

Ray Kurzweil has argued that computers will soon not only surpass human intelligence, but also adopt human personality (Kurzweil 1999 and 2005). I think he is right, except for that four-letter word "soon." However, the process of transferring human personality into information systems has begun, with software agents, assistive technologies, immersive and pervasive games, and natural language processing. My own work is based on the premise that some of the technologies that will expand our minds will be based on traditional social and behavioral science. Combined, these methodologies will create AI agents that extend our own personalities (Bainbridge 2006).

Just as Spirit and Opportunity took human consciousness to Mars, the descendants of these robot explorers could transport aspects of specific human personalities across the galaxy. Whether we consider these interstellar AI agents to be ourselves will be a matter of perspective (Bainbridge 2002).

References

Bainbridge, W. S. 2003. Massive Questionnaires for Personality Capture, *Social Science Computer Review*, 2003, 21(3), pp. 267-280.

Bainbridge, W. S. 2004. The Future of the Internet: Cultural and Individual Conceptions, pp. 307-324 in *Society Online: The Internet in Context*, edited by Philip N. Howard and Steve Jones. Thousand Oaks, California: Sage.

Bainbridge, W.S. 2006. Cognitive Technologies. *Managing Nano-Bio-Info-Cogno Innovations: Converging Technologies in Society*, pp. 203-226. Berlin: Springer.

Bainbridge, W.S. 2002. The Spaceflight Revolution Revisited. S. J. Garber (Ed.) *Looking Backward, Looking Forward* pp. 39-64. (Washington, D.C.: National Aeronautics and Space Administration.

Bainbridge, W.S. (Ed.). 2004. *Encyclopedia of Human-Computer Interaction* Great Barrington, Massachusetts: Berkshire.

Bainbridge, W.S. 1986. *Experiments in Psychology.* Belmont, California: Wadsworth.

Bainbridge, W.S. 1989. *Survey Research: A Computer-Assisted Introduction.* Belmont, California: Wadsworth.

Benford, S., Crabtree, A., Reeves, S., Flintham, M., Drozd, A., Sheridan, J. and Dix, A. 2006. The Frame of the Game: Blurring the Boundary between Fiction and Reality in Mobile Experiences. *Proceedings of CHI*, pp. 429-436. Montréal, Québec, Canada.

Jonsson, S., Montola, M., Waern, A., and Ericsson, M. 2006. Prosopopeia: Experiences from a Pervasive Larp. *Proceedings of ACE 06*, June 14-16, Hollywood, California.

Kurzweil, R. 2005. *The Age of Spiritual Machines: When Computers Exceed Human Intelligence.* New York: Viking.

Kurzweil, R. 2005. *The Singularity Is Near: When Humans Transcend Biology*. New York: Viking.

Liao, L. Fox, D. and Kautz, H. 2004. Learning and Inferring Transportation Routines. *Proceedings of the Nineteenth National Conference on Artificial Intelligence*, San Jose, California.

Mann, M. 2004. Continuous Lifelong Capture of Personal Experience with EyeTap. *Proceedings of CARPE'04*, October 15, New York. Available:
www.eyetap.org/papers/docs/p1-mann.pdf

Marek Bell, M., Chalmers, M., Barkhuus, L., Hall, M., Sherwood, S., Tennent, P., Brown, B., Rowland, D., Benford, S., Capra, M. and Hampshire, A. 2006. Interweaving Mobile Games With Everyday Life. *Proceedings of CHI*, pp. 417-426. Montréal, Québec, Canada.

Merrick, K. and Maher, M. L. 2006. Motivated Reinforcement Learning for Non-Player Characters in Persistent Computer Game Worlds. *Proceedings of ACE '06*, June 14-16, Hollywood, California.

National Science Foundation Awards
www.nsf.gov/awardsearch/showAward.do?AwardNumber=0534280

Patterson, D. J., Liao, L. Gajos, K., Collier, M., Livic, N., Olson, K.,Wang, S. Fox, D.and Kautz, H, 2004. Opportunity Knocks: A System to Provide Cognitive Assistance with Transportation Services. *Proceedings of the Sixth International Conference on Ubiquitous Computing*, Nottingham, England,.

www.informedia.cs.cmu.edu/eod/index.html,

www.informedia.cs.cmu.edu/caremedia/index.html

Saul S. High-Speed Scanning in Human Memory. 1966. *Science*, 153, pp. 652-654.

Strawn, G., Howe, S. E. and King, F. D. (Eds.) 2003, *Grand Challenges: Science Engineering and Societal Advances Requiring Networking and Information Technology Research and Development*. Arlington, Virginia: National Coordination Office for Information Technology Research and Development.