MITSUBISHI ELECTRIC RESEARCH LABORATORIES
http://www.merl.com

Worse is Better for Ambient Sensing

Reynolds, C.J.; Wren, C.R.

TR2006-005 May 2006

Abstract

The intrinsic value of information coupled with the dramatically falling costs of networked sensors suggest that ambient intelligence and ubiquitous computing are inevitable. However, before society resigns itself to a world of constant observation and tracking, a process of moralization and ethical deliberation should occur. In this paper we examine the ethical implications of choosing camera networks or infrared motion detector networks. We employ the Dimensional Metaethics approach to help us structure examination of the complex issues involved. The analysis indicates that choice of sensor technology can powerfully affect the ethical landscape surrounding the final system. This paper also analyzes emprical results from questionnaires that asked participants to rate and choose between scenarios involving pan-tilt-zoom cameras and infrared sensors. In testing against a hypothetical even split in opinion, we find instead a significant preference for the scenario involving infrared sensors (p=0.007). The results show that significant proportion (73%, p=0.05) preferred a scenario with infrared sensors when compared to pan-tilt-zoom cameras. Participants also report that the scenario with infrared sensors was significantly less invasive and expressed a significantly weaker preference toward situations "Without Sensors that Collect Information About Location" (when compared with the scenario involving pan-tilt-zoom cameras). In short, we find that both dimensional metaethics and questionnaire results suggest that infrared sensors are better.

Workshop on Privacy, Trust and Identity Issues for Ambient Intelligence

February 8, 2006: Submitted to the Pervasive 2006 Workshop on Privacy, Trust and Identity Issues for Ambient Intelligence.
March 7, 2006: Accepted to the PTIIAI Workshop.
March 24, 2006: Camera-ready version submittied to publisher.
May 7, 2006: scheduled date of PTIIAI publication.

1 Introduction

There are several pressures pushing toward ubiquitous sensing: concerns about safety, security, and efficiency, as well as the desire to live in a rich, computationally alive world. The current trends point to a solution where cameras and microphones are everywhere, backed by perceptual intelligence. The cost, size and power consumption of cameras and microprocessors are falling, and soon the cost of such devices will become low enough that they will be justifiable by the desires for more contextual information. It has been suggested that much of the functionality that people hope to create in smart environments can be attained with wireless sensor networks at a lower economic cost than can be attained with networks of cameras. We hope to show that there are also compelling ethical reasons to chose sensor networks over camera-based systems using the Dimensional Metaethics framework [8].

When we strive to create an ambient intelligence by deploying sensing technology throughout a building, then there are some issues that arise independent of the sensing technology used. Even if the system is designed with the intention of, for example, making the elevators more efficient, that does not mean that the data cannot be subverted to less agreeable uses: employers spying on their workers, governments spying on their citizens, or deviants seeking potential victims. When we build these systems there is a certain amount of trust that the information will be controlled, used only for the stated purpose, and then destroyed. However, once the systems are in place, there will always be a temptation toward abuse that may injure the privacy, autonomy, or safety of those involved. As these systems become more ubiquitous this contract of trust becomes very diffuse, since it's hard to avoid interacting with such systems once they are everywhere.

It has been argued that the ill-effects can be overcome by creating transparent systems [5]. If everyone has access to the information from the system, the theory goes, then it will be possible to avoid abuse by watching the watchers. This seems impractical. One group in New York City employs a distributed network of volunteers organized over the Internet to track just the locations of security cameras in the city. They currently track over 2,000 cameras. However, tracking the guards who monitor the video content, or worse, the video itself would be an herculean undertaking. The same problem arises when you attempt to offer to people the ability to assess the data that has been collected about them. How can we find all the potentially recognizable images of a person across all the databases separated by representational, institutional, and governmental boundaries? Once found, how to deliver terabytes of data to an average citizens who may not have the same resources to cope with that magnitude of data as a company or government might possess?

Others have argued that the ill-effects can be avoided by placing the sensors on the body of the person, so that all the information is physically controlled by the observed, rather than scattered into the infrastructure [9]. For the technology savvy elite, this might be true, given open source solutions that are available to inspection and complete transparency not only in the manufacture but also

and daily operation of the devices. However, most of the world population does not have the knowledge (or the time) to personally verify the ethical operation of such a device. As a result, carrying around devices does not offer a significant privacy guarantee to the average consumer. At the same time it makes many smart building applications nearly impossible to implement due to a telephone effect: it doesn't make sense to install the infrastructure elements until there are enough people with wearable devices who are willing to share information, and conversely it doesn't make sense to carry the devices until there is infrastructure to realize the benefits. On the other hand, a single building can decide to install a ubiquitous sensing infrastructure to realize, immediately, some benefit within those walls.

One central benefit of sensor networks is that they do not capture very much information in the first place. A lay person can look at a camera and realize that it has a lens. All cameras, no matter how small and well hidden, must present optical glass to the observed. The general public can immediately understand that motion detectors, leaf spring switches, pressure pads, break beam sensors, and other similar devices do not have the ability to post photos to the Internet of them picking their noses. There is a fundamental difference between a system being aware that there are one or more humans in a space, and a system recording high-fidelity video data of those people. We claim that the latter is potentially open to much more serious abuse. If the building system simply needs to know where people are and where they are going, then it should be designed *not only* so that it does not record that those people are male or female, the clothes that they are wearing, their race, if they are conversing with one another, or any other information that could be abused; but further, it should be designed to not even *sense* those attributes.

There are still dangers even with sensor networks composed of mere motion detectors. If an office is outfitted with a motion detector, then the employer might complain that an employee was not putting in appropriate hours, based on the fact that there was no motion in the office before or after particular times. However, there would be no way for the employer to distinguish an employee from a robot that wakes up at 9AM and dances around the office. Nor could the employers utilize their telephoto zoom camera mounted on a pan-tilt base to discover that their employee is taking a little break from that dull standards document by reading an exciting bit of Orwell. Video, when properly documented and handled, is even admissible as evidence in a court of law. While simple sensors can provide useful information, they retain an inherent level of plausible deniability that places fundamental limits on the scope of possible abuses.

In the subsequent sections we carefully examine this claim: that the choice of sensor modality can powerfully affect the ethical context of the system. The next section explores prior work linking ethical considerations to the design of ubiquitous systems. Section 3 carefully illustrates the implications of sensor modality on a hypothetical system using the Dimensional Metaethics approach. Section 4 presents the methodology used to survey a population about their reactions to some hypothetical ubiquitous computing systems. And finally, Sec-

tion 5 discusses the results of the survey.

2 Related Work

Many previous researchers have considered the relationship among privacy, ethics, ubiquitous computing, and ambient intelligence. We review a selection of these in chronological order to provide a context for some of the later argumentation in this paper.

Chaum[6] argued that pervasive "card computers" could help protect privacy in "large-scale automated transaction systems." His paper was concerned that "pattern recognition techniques" could be employed for a variety of "mass surveillance." To help circumvent this he proposed using public key cryptography to protect individual identity from traceability.

In discussing "invasive technologies," Anderson[2] poses ethical questions regarding "organizational and governmental monitoring of individual activities." His approach was to use ethical philosophy in "an analytic or critical" manner to help deduce what sort of problems might exist concerning the morality of invasive technologies.

Want et al.[10] discuss privacy issues that arose with the use of location-tracking "Active Badges." They note that individuals having their location tracked may at some times prefer not to be locatable. Furthermore they acknowledge potentially dystopian abuses such as "secret logging" of employees by a particularly unseemly company. Ultimately, legislation is suggested as a method to help ensure that privacy is protected.

In "Design for Privacy in Ubiquitous Computing Environments," Bellotti and Sellen[3] argue that "feedback and control" over information in a ubiquitous computing environment can help preserve privacy. They provide a "framework for designing for feedback and control in ubiquitous computing environments" which encourages designers to ask questions such as "what is appropriate feedback?" They represent an approach that argues that, when computers are pervasive, special care needs to be taken in analyzing how information is made public.

Palen and Dourish[7] reconsider the concept of privacy given the existence of networked, interactive technology. Privacy is defined as a "dynamic boundary regulation process" following Atman's view[1]. Palen and Dourish describe networked individuals as participating in an active process of "privacy management."

More to the point, Bohn et al.[4] directly consider the effects that ambient intelligence and ubiquitous computing may have upon everyday life. They identify "reliability, accessibility, and transparency" as concerns of ethical import. They go further to decompose reliability, delegation of control, social compatibility, and acceptance into sub-concerns that may be relevant for designers of ubiquitous computing systems.

3 Dimensional Metaethics Analysis

In this section we apply Dimensional Metaethics to evaluate the ethical landscape surrounding a hypothetical ubiquitous computing system. We imagine a system designed to improve elevator efficiency by sensing building occupants and predicting demand for service. In Table 1 we show a summary of the system situated in a multi-dimensional metaethics space.

The left block is an evaluation of a sensor network based on motion detectors. The right block examines the ethical landscape associated with a camera network performing the same task.

Within each block, the center column lists the names of the 23 dimensions we considered. The left column represents the value of the system along that dimension in an utopian context. That is, the ideal context that the system designer intended. The right column presents a dystopian view of the system. Exploring the differences between the dystopian views of the two systems illuminates the potential abuses that might be exacerbated by the design choices.

Motion Sensor Network		
Elevator	Whom	Police State
Motion	**What**	**Inferred ID**
Efficiency	Goal	Monitoring
None	Power	Police State
Efficiency	Stake	Privacy
Events	**Genre**	**Deniable**
Neutral	Valence	Neutral
Workplace	Demeanor	Workplace
Equal	Gender	Unequal
Equal	Ethnicity	Unequal
Equal	Age	Unequal
General	Culture	General
Awareness	**Risk**	**Awareness**
Symmetric	Symmetry	Asymmetric
Cooperative	Trust	Police State
Contractor	Designer	Government
Workday	Time	Ubiquitous
Targeted	Informed	Ubiquitous
Secure	**Security**	**Obfuscated**
Infrastructure	Control	Infrastructure
Easy	**Feedback**	**Easy**
Sensor	**Transparency**	**Sensor**
Building	Proximity	Internet

Camera Network		
Elevator	Whom	Police State
Pixels	**What**	**Incriminating**
Efficiency	Goal	Monitoring
None	Power	Police State
Efficiency	Stake	Privacy
Video	**Genre**	**Evidentiary**
Neutral	Valence	Neutral
Workplace	Demeanor	Workplace
Equal	Gender	Unequal
Equal	Ethnicity	Unequal
Equal	Age	Unequal
General	Culture	General
Imagery	**Risk**	**Pixels**
Symmetric	Symmetry	Asymmetric
Cooperative	Trust	Police State
Contractor	Designer	Government
Workday	Time	Ubiquitous
Targeted	Informed	Ubiquitous
Secure	**Security**	**Webcam**
Infrastructure	Control	Infrastructure
Hard	**Feedback**	**Hard**
Camera	**Transparency**	**Camera**
Building	Proximity	Internet

Table 1: Worksheet for comparing motion sensor networks (left) to camera networks (right). Within each column, left is the utopian case, while the right is the dystopian case. Bold indicates divergence.

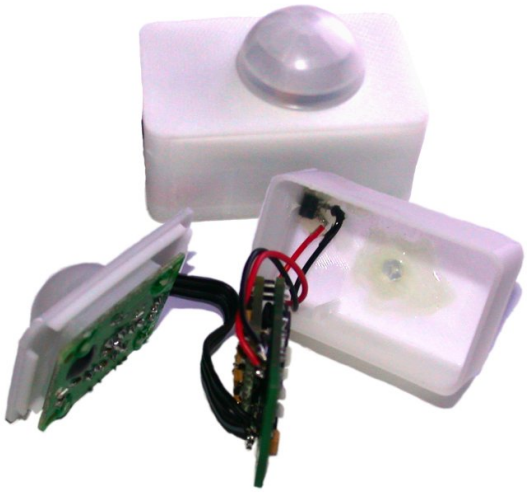

Figure 1: A wireless, passive, infrared motion detector. Note the non-optical lens.

3.1 Utopia

Along many of the dimensions the values are shared between the sensor case and the camera case. These dimensions are not affected by the particular choices of sensor modality but depend instead on the nature of the system.

In the utopian case the world matches the probable expectations of the designer: the elevator control system is the consumer of the information (*Whom* dimension), it accepts presence or video data (*What* dimension) to predict elevator demand and improve scheduling efficiency (*Goal* dimension). There's no meaningful Power relationship between the elevator—a mere machine—and its users. The system is intended to operate in a public workplace environment (*Valence* and *Demeanor*) with no inherent inequities or *Asymmetries* with respect to *Age*, *Gender*, *Ethnicity*, or *Culture*. The system would probably be designed by an elevator installation contractor (*Designer*) who has no particular interest in the occupants of any particular building other than to give them the best possible experience with the product, a goal that they share (*Trust*). The users only interact with the system during the workday (*Time*). The designer and users might share a level of expectation that the building plant, including the elevator system, is guarded against tampering (*Security*), that the data is only being used for the stated purpose (*Informed*), and that it never leaves the building (*Proximity*). Since this is an infrastructure system, there is unlikely to be any control over the system, much as the occupants may be unable to control the lights or the air conditioning without the help of a maintenance worker.

3.2 Different Modalities

There are just a few dimensions where the two networks differ. In one case the system is collecting motion data and in the other it is collecting raw pixels from a camera. These two types of data are very different (*Genre*). The motion data consists of discrete, symbolic events that carry no other information. The video, on the other hand, carries a wealth of collateral information about the age, gender, race, disposition, and possibly even identity of the individual. The *Risk* to the individual is that the imagery from those cameras might somehow be compromised. In the sensor case, the only thing that can leak from the system is that someone or something was in a particular place at a particular time.

Presumably the video is processed in such a way that ignores the extraneous information and extracts only the information necessary for the elevator's efficient operation. However, there is no way, short of a technical audit of the system, to determine if that is actually the case. *Transparency* is an inherent problem with camera-based systems. There is no way for the observed to know for sure that the video data is not being stored or used in some other way. The motion detectors provide a level of *Transparency* by providing up-front guarantees in the from of obvious physical limitations in the sensors themselves.

Bellotti and Sellen[3] attempt to address this by providing *Feedback* in the form of video monitors near the cameras, so that users could see how they appeared in the video stream and be aware of the cameras. This seems impractical when we think about camera networks that would cover an entire building with possibly hundreds of cameras. Since the motion detectors are very low bandwidth, they can provide feedback directly in the form of an LED that blinks when the sensor is activated. Many motion detectors being used in the market today already have this feature.

3.3 Dystopia

These differences become more pronounced when we shift to the dystopian world view. These dimensions are in bold typeface on Table 1. If we imagine that the system could be co-opted by a totalitarian government for the purposes of monitoring its citizens, then it becomes very important what type of information the sensors are actually collecting. Video data has the potential to provide hard, incriminating evidence about any activities that might happen within the building. Video data also has inherent meaning even when it is taken out of the context of the building.

Even if the motion sensor data is accessed nefariously, it is inherently obfuscated. It only indicates the presence of motion in the space. While it might be possible, depending on the context, to infer the identify of someone from patterns of data, there would be no direct evidence to prove that association.

So even in a nightmare scenario of a police state, we see that the design choice to use motion detectors makes the ethical stance of the system more stable. Even enormous shifts in the assumptions only provide thin opportunities for abuse. If we think about smaller attacks, say a building employee merely attempting

Figure 2: Left: camera designed for flush mounting on surface. Right: a pan-tilt-zoom camera system.

to invade the privacy of a building occupant in some way. the threat is even smaller, since the threat of inferring identity from pattern analysis is likely to be out of reach of an adversary with limited resources.

4 Questionnaire

Dimensional Metaethics helped us anticipate and analyze ethical ramifications, however empirical observation provides an excellent supplement for such speculative activities. Thus, to complement the application of Dimensional Metaethics to sensor networks used for ambient intelligence we undertook a questionnaire survey.

4.1 Methods

Drawing upon our Dimensional Metaethics analysis, we hypothesized that participants would prefer a scenario with infrared motion sensors over similar scenario involving pan-tilt-zoom cameras. We also hypothesized that a battery of questions dealing with ethics would show a scenario involving infrared scenarios is viewed more favorably. To test this hypothesis we recruited participants to complete a web-based survey.

4.1.1 Participants

We had a total of 26 participants in our questionnaire survey. Participants recruitment took place over three days with postings to community websites in San Francisco, New York, and Boston. The text of the posting was:

> Your opinion is needed regarding the acceptability of ambient intelligence and scenarios motivated by recent research developments.

> Please fill out this quick questionnaire, which is part of joint research conducted by the University of Tokyo and the Mitsubishi Electric Research Laboratory:
>
> http://www.researchquestionnaire.info/

There were a total of 45 responses to the questionnaire posting. Of these, we made use of data of $N = 26$ participants who filled in the questionnaire completely. The mean age reported by participants was 38.08 with a standard deviation of 16.01. There were 20 female participants and 6 male participants. Most of the participants (80%) reported United States as their nationality, with the remaining distributed uniformly among China, India, Portugal, the United Kingdom, and Vietnam. In terms of education, 15% reported undergraduate, 31% reported post-graduate, and 15% secondary level.

4.1.2 Design

Participants were randomly assigned to one of two different task orderings. Eighteen participants experienced a scenario involving infrared motion-sensor technology first. Eight participants experienced a scenario involving pan-tilt-zoom cameras first.

The independent variable was the sensing technology used in the two hypothetical scenarios that each participant experienced. The two levels for this variable were "infrared motion-sensor" and "pan-tilt-zoom cameras."

The dependent variables were a set of eight-point Likert-scale questions. A first page of these questions dealt with ethically relevant adjectives. A second page had only a single question that asked participants to compare the two hypothetical scenarios they encountered.

4.1.3 Procedure

The questionnaire consisted of six web pages. The first page informed potential participants that the purpose of the questionnaire was "to collect data to evaluate the acceptability of ambient intelligence and related research." Furthermore, this page stated:

- We are seeking participants ages 18 and over.
- You will be asked to fill out a brief questionnaire.
- The information you provide will be anonymous.

No compensation was paid to participants for completion of the questionnaire. Instead, as motivation, the page provided the following:

> The questionnaire gives participants an opportunity to provide opinions regarding new technologies. The results of this research will help us better understand how to design systems. Your opinions may help shape future ambient intelligence systems.

After a page collecting initial demographics (age, gender, nationality, and education) participants encountered a page about one of two randomly chosen scenarios. Participants were first told "Each of the following questions asks about the following scenario." In the case that the participant was assigned to see the pan-tilt-zoom scenario they saw the following:

> Suppose you work in an office where a pervasive network is used. The network uses *pan-tilt-zoom cameras* to track the location of employees. The purpose of this network is to provide an office census in case of emergencies or disasters.

If on the other hand, the participant was randomly selected to view the infrared sensor scenario they saw the following text first:

> Suppose you work in an office where a pervasive network is used. The network uses *infrared sensors* to track the location of employees. The purpose of this network is to provide an office census in case of emergencies or disasters.

The differences between the infrared sensors and pan-tilt-zoom cameras were not explained. Furthermore, participants were not shown any imagery of the sensors. We instead assumed that participants would rely on prior knowledge of these sorts of sensors.

Participants next saw a series of eight-point Likert scale questions. The questions opposed two ethically relevant poles and asked participants to rate which the scenario reflected. Alternatively, participants were allowed to select a no opinion check-box. The questions (and antipodes) were:

- Do you think the scenario is: (Unethical...Ethical)
- Do you think the scenario is: (Invasive...Respectful)
- Does the scenario make you feel: (Comfortable...Uncomfortable)
- Would the scenario be a: (Help...Hindrance)
- Do you think the scenario is: (Moral...Immoral)
- Which of the following does the scenario make you feel: (Suspicious...Trustful)
- Do you feel the scenario is: (Fair...Unfair)
- Given the choice between two scenarios you'd prefer the scenario: (Without Sensors that Collect Information About Location...With Sensors that Collect Information About Location)

Once participants answered these questions, they next saw an identical page with the other scenario. For instance, if the participant first randomly saw the pan-tilt-zoom scenario afterward they encountered the infrared sensor scenario (and vice-versa). Participants then answered the same battery of questions that

are shown above about this second scenario. Form validation was used in the questionnaire, so that some sort of answer had to be selected before it was possible to continue.

After answering questions about both the pan-tilt-zoom and infrared sensor scenarios, participants next compared the two scenarios. A page appeared which stated: "The following question asks you to compare the two previous scenarios." The two randomly-ordered scenarios were shown again in the order the participant encountered them labeled as A and B. The participant was prompted "I prefer:" and given a choice between Scenario A and Scenario B on an eight point Likert scale.

Finally the survey informed participants "Thank you very much for taking the time to complete this questionnaire. Your participation is extremely appreciated, and we hope that the process has been brief and interesting."

5 Results

Two-sample paired Wilcoxon tests compared the within-subject data regarding the pan-tilt-zoom and infrared sensor scenarios. A subset of the data in which "No Response" did not occur was analyzed. In analysis of this data two responses were significant.

When posed the question "Do you think the scenario is" invasive or respectful, a significant difference was observed (p=0.02) with the infrared sensor scenario being less invasive. The effect size for this difference is medium (d=0.57, r=0.26). On the eight point Likert scale, 1 was associated with "invasive" and 8 was associated with "respectful." The mean of the infrared sensor scenario was 3.04 and the standard deviation was 1.93 while the mean of the pan-tilt-zoom camera scenario was 2.08 and the standard deviation was 1.57. Both responses were on the invasive side of the scale, but the infrared scenario was moderately more neutral.

When prompted "Given the choice between two scenarios you'd prefer the scenario" and selecting between "Without Sensors that Collect Information About Location" or "With Sensors that Collect Information About Location" a significant difference also occurred (p=0.02). In this case, the effect size was a medium increase (d=0.56, r=0.27). On the eight point Likert scale, 1 indicated "Without Sensors..." while 8 indicated "With Sensors... ." In the infrared sensor scenario the mean was 2.72 and the standard deviation 2.07 while in pan-tilt-zoom camera scenario the mean was 1.72 and the standard deviation 1.4. Again, participants in both cases expressed a preference toward scenarios "Without Sensors..." but those responding to the infrared sensor situation had a weaker preference.

A much stronger and more straight-forward result occurred in the questionnaire data asking participants to directly compare the two scenarios. On an eight-point Likert scale with 1 associated with the pan-tilt-zoom scenario and 8 associated with the infrared sensors scenario, a mean of 6.09, median of 7, and standard deviation of 2.5 were observed. Testing against the hypothetical

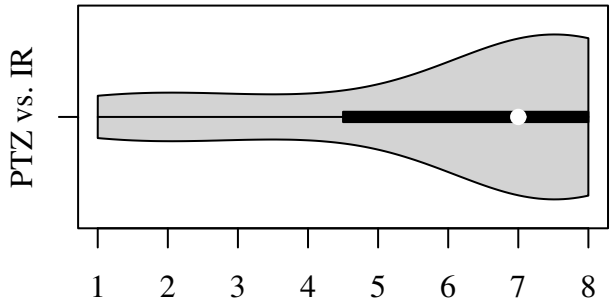

Figure 3: Violin plot of Likert-scale comparison ranging from 1 "pan-tilt-zoom" to 8 "infrared." Note that the majority of responses are on the right (infrared) side, indicating a preference for the location tracking scenario making use of infrared motion detectors.

uniformly split mean of 4.5, a t-test shows that there was a significant preference for infrared sensors (p=0.007). The effect size of this difference was very large (d=1.3, r=0.54). Among the participants who did not select "No Opinion" 16 of 22 or 73% expressed a preference towards the infrared sensor location tracking scenario. This was a significant proportion (p=0.05).

6 Concluding Remarks

We have seen that both Dimensional Metaethics and questionnaire methods suggest that infrared motion detector systems are preferable to their pan-tilt-zoom camera counterparts. In retrospect, this conclusion seems obvious and intuitive. However, it should be said that any sort of speculative studies lack an element of reality that exists with the experience of real systems. Additionally, the questionnaire had a relatively small number of participants and thus the results should be viewed as preliminary until more comprehensive studies are conducted.

At a more social level, it is hoped that these results will influence designers interested in ambient or ubiquitous sensor networks to see that sometimes the most informative sensors are not perhaps the best. Clearly, this work won't convince the community to banish cameras from their sensing milieu, but it may help to highlight the cost of negative reactions that such systems elicit.

References

[1] I. Altman. *The environment and social behavior: privacy, personal space, territory, crowding.* Brooks/Cole Pub. Co., Monterey, Calif., 1975.

[2] R. Anderson. The ethics of research into invasive technologies. Technical Report EPC91-107, Rank Xerox EuroPARC, Cambridge, UK, 1991.

[3] V. Bellotti and A. Sellen. Design for privacy in ubiquitous computing environments. In *Proceedings 3rd European Conference on Computer Supported Cooperative Work*, pages 77–93, Milan, Italy, 1993. Kluwer.

[4] J. Bohn, V. Cororama, M. Langheinrich, F. Mattern, and M. Rohs. Social, economic, and ethical implications of ambient intelligence and ubiquitous computing. http://www.vs.inf.ethz.ch/publ/papers/socialambient.pdf, 2004. Institute for Pervasive Computing. ETH Zurich, Switzerland.

[5] David Brin. *The Transparent Society.* Perseus Books, Reading MA, 1998.

[6] D. Chaum. Security without identification: Transactions systems to make big brother obsolete. *Communications of the ACM*, 24(2), 1985.

[7] L. Palen and P. Dourish. Unpacking "privacy" for a networked world. In *Proceedings of 2003 Conference on Human Factors and Computing Systems*, pages 129–136, Fort Lauderdale, Florida, 2003. CHI, ACM Press.

[8] C. Reynolds and R. Picard. Evaluation of affective computing systems from a dimensional metaethical position. In *First Augmented Cognition International Conference*, Las Vegas, NV, July 2005.

[9] T. Starner. The challenges of wearable computing: Part 2. *IEEE Micro*, 21(4):54–67, July 2001.

[10] R. Want, A. Hopper, V. Falcao, and J. Gibbons. The active badge location system. *ACM Transactions on Information Systems*, 10(1):91–102, January 1992.