To appear in: John Krumm (ed.): "Ubiquitous Computing,"
Chapman & Hall / CRC Press, Sep. 2009. ISBN: 9781420093605

# PRIVACY IN UBIQUITOUS COMPUTING

MARC LANGHEINRICH

## 1 INTRODUCTION

Privacy is by no means a recent addition to the Ubiquitous Computing (ubicomp) research curriculum. In his 1991 Scientific American article, Mark Weiser already identified it as one of its biggest challenges: "Perhaps key among [the social issues that embodied virtuality will engender] is privacy: hundreds of computers in every room, all capable of sensing people near them and linked by high-speed networks, have the potential to make totalitarianism up to now seem like sheerest anarchy." [**1**] It would be nice if by now, almost two decades later, we would have a standard set of solutions that we could easily prescribe for any ubicomp system (or any computer system in general): "in order to protect privacy, implement subroutines A, B, and C."

Unfortunately, privacy is such a complex issue that there is no single solution, no recipe for success, no silver bullet (or set of silver bullets) that will "fix" a system for us so that it is "privacy-safe". What exactly does it mean anyway, for a system to be "privacy-safe"? Whose privacy does it protect, when, and to what extent? None of these questions can be answered in general. Instead of a simple one-size-fits-all recipe, one needs to look into each single system and application in great detail, first understanding what the system does and what the implications of this are, and then working out how (and why) this needs to be changed in order to reach the "right" behavior.

This chapter attempts to provide some guidance for this process. First, by explaining the concept of privacy in more detail, so one understands what it is that should be protected. Second, by giving some examples on how technology can safeguard personal information in ubicomp systems. This chapter is not meant as a cookbook that allows one to quickly find a solution for a particular problem, but rather as a starting point for recognizing and approaching privacy issues in one's own design, development, or use of ubicomp systems and applications.

### 1.1 WHY A PRIVACY CHAPTER IN A UBICOMP BOOK?

Privacy might indeed be an important topic, but one could argue that it would be much better suited for a text on legal or social issues. Why include it in book about ubicomp technology? Privacy and technology are closely intertwined. Shifts in technology require us to rethink our attitude towards privacy, as suddenly our abilities to see, hear, detect, record, find and manipulate others and their lives is greatly enhanced. Ubicomp represents such a technology shift, and its widespread adoption will significantly influence the way we handle personal information.

One could still suggest putting this topic in front of people interested in databases or information retrieval in particular, rather than systems or usability evaluation in general. Is privacy not about the storage and processing of data, which in turn is really the job of database and information retrieval specialists? The answer is: Yes and no. Clearly, the ability to store, process, and analyze information is at the heart of the privacy debate. However, in order to make meaningful choices within any system parameters, one needs to understand the entirety of the system and its applications: What kind of information is collected and in what manner? Who needs to have access

to such information and for what purpose? How long should this information be stored and in what format, with what levels of accuracy and precision? Effective privacy protection can only work if it addresses the entirety of the information life cycle in each individual ubicomp application.

### 1.2 ISN'T PRIVACY THE SAME AS SECURITY?

Security – i.e., the confidentiality, integrity, and authenticity of information – is often a necessary ingredient to privacy, as it facilitates the control of information flows (i.e., who gets to know what when?) and helps to ensure the correctness of data. However, it is possible to have high levels of security but no privacy (think surveillance state), or even some sort of privacy without security (e.g., a private table conversation in a busy restaurant). The important insight is that simply implementing some form of security is not enough to ensure privacy. Ensuring the confidentiality and authenticity of a particular information does not say anything about how and when this particular piece of information will be *used* by its designated recipient.

### 1.3 WHAT IS IN THIS CHAPTER?

The main focus of this chapter is how ubicomp affects privacy and what technical methods can be used to counter or mitigate this influence. This requires a clear understanding of what exactly should be protected, however. A large part of this chapter is therefore dedicated to understanding the concept of privacy first. Only then can one discuss the particular effects of ubicomp on privacy and the required technological countermeasures. Of course, privacy issues of "regular" computer systems such as databases often apply equally to ubicomp system, simply because most ubicomp applications are built with such standard components. Including technical solutions for such computing systems in general would be beyond the scope of this book, so the discussion in this chapter is strictly limited to examples on how particular threats induced by ubicomp technology can be addressed.

**Conclusions**

1. Data collection and processing are core components of ubicomp technology. Privacy issues are thus of utmost importance to ubicomp researchers, designers, service providers, and users.
2. Simply offering strong security does not solve privacy issues. While security is an integral part of any privacy solution, it fails to address questions such as scope, purpose and use, adequacy, lifetime, or access.
3. To build privacy-aware or privacy-compliant ubicomp systems, we need to understand the nature of privacy, its social and legal realities, and the technical tools at our disposal. This chapter attempts to introduce those three aspects in detail.

## 2 UNDERSTANDING PRIVACY

Imagine your kitchen of the future (it might look like the one in Figure 1). All appliances – your fridge, freezer, stove, microwave, but also cabinets, utensils, faucets and lamps – are now "smart", i.e., they are able to sense their environment and communicate: among themselves, with you, and with other things and people via the Internet. The famous *Internet-fridge* monitors its contents and orders milk and other ingredients before they run out. Your microwave-grill combo interrogates the pizza package to make sure it properly heats and bakes your TV-food to perfection. And your faucets, freezer, and stove coordinate the use of warm water and electricity with the local power plant to minimize your energy costs. Ubicomp heaven!

**Figure 1: The Smart Kitchen of the Future: Spy in Your Home or Ultra Convenience?**


Now imagine a few extras: your fridge not only orders milk and other staples if you run low, but also scouts for offers and coupons from the supermarket. Interestingly enough, the offers you receive are often very different from what your friend's fridge receives, who regularly gets discounts for expensive organic products (you don't). For your convenience, you have allowed your home insurer to periodically query your belongings in order to verify that you are sufficiently covered. However, after having had a number of visitors during the last few months, your insurer yesterday suddenly doubled your premium, claiming that the contents and activities in your kitchen indicated that you no longer have a single-person household. And just now a police officer stopped by, asking you to explain the large quantities of hydrogen peroxide stored in your shed (sensed by the chemical sensors that monitor its contents for fire safety). Ever since the government passed the *Preventing Irregularities Through Smart Appliances* (PITSA) Act, all household appliances are required by law to report suspicious items and activities directly to law enforcement agencies. Even though you showed the officer the antique wooden sailboat in your backyard that you are planning to bleach with the hydrogen peroxide, he informed you that your name will be kept on a list used by pharmacies around the country, alerting the authorities of any additional products you buy that could be used for bomb making. Well, at least your appliances haven't been broken into yet: Your chaotic neighbor forgot to renew the firewall update for his fridge and promptly had a hacker monitor its use in order to detect longer absences. As soon as your neighbor had left for a business trip, his apartment address was traded on a local underground bulletin board that, for a small fee, would show "inactive" households on a Google Maps mash-up. Ubicomp heaven?

Whether or not any of the above examples make you question this brave new world of ubiquitous computing depends largely on your personal conception of privacy. Is the automated creation of profiles for the purpose of offering you discounts a good thing, even if it might offer you worse deals than others? Is the detection of inaccurate statements on insurance applications (and thus savings of several millions of dollars) a good thing, even if the system *occasionally* gets it wrong? And what is so bad about smart appliances that report unlawful behavior to the police, if they stay silent otherwise? If you don't do anything wrong, what have you got to hide? And if your data got stolen, maybe you didn't protect it well enough?

In order to build ubicomp systems that are "privacy-aware" or "privacy-respecting", one obviously has to first define what exactly is meant by "privacy". There certainly is no shortage of definitions for privacy, yet no single one seems to work for all cases and disciplines. Section 2.1 will briefly summarize some of them and illustrate the

changes the concept of privacy has undergone over time. Motivating privacy is an important part of such a definition, and section 2.2 will look into current trends and surveys in order to understand what consumers and citizens expect in terms of privacy protection. Section 2.3 then gives a very short summary of the legal status of privacy protection, as any technical solution will most likely need to conform to local laws and regulations. However, not all interactions are governed by laws, in particular if they do not involve companies or government agencies, but friends and families. Section 2.4 introduces some concepts from social sciences and psychology that can help understanding how people "regulate" access to themselves.

## 2.1 DEFINING PRIVACY

Privacy has been on people's minds even before credit cards or the Internet came along. Everybody seems to have some sort of intuitive understanding of what privacy means on a case by case basis, yet it is difficult to *objectively* define what exactly it is. Legal scholars have long since grappled with such a definition, in order to write laws and regulations that describe what kind of privacy protections should be granted, and they provide a good start for analyzing this complex topic.

One of the earliest legal references to privacy is found in the 1361 *Justices of the Peace Act* in England, which laid down sentences for peeping toms and eavesdroppers. The famous saying "My Home is My Castle" dates back from the 18th century, where the British parliamentarian William Pitt said in a 1763 speech in parliament: "The poorest man may in his cottage bid defiance to all the forces of the Crown. It may be frail — its roof may shake — the wind may blow through it — the storm may enter — the rain may enter — but the King of England cannot enter — all his force dares not cross the threshold of the ruined tenement!" However, as old as the concept of privacy is, it stills seems unclear exactly what it means today.

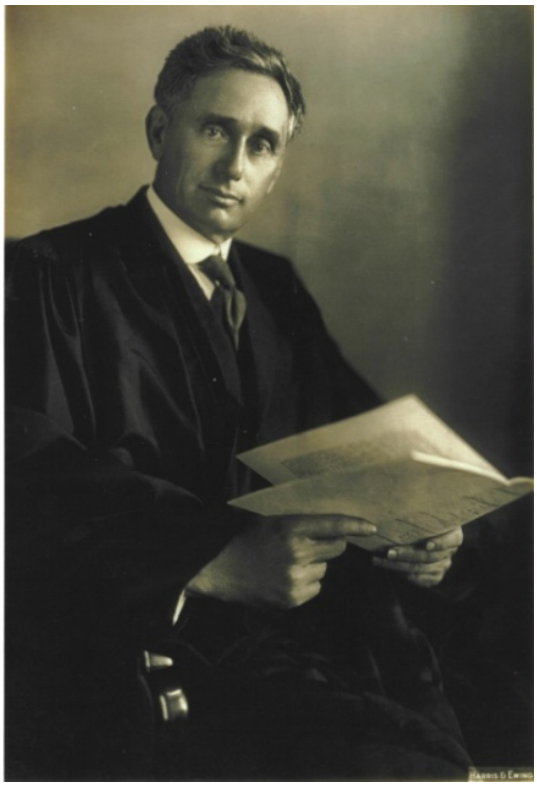


**Figure 2: Louis Brandeis, Co-Author of "The Right to Privacy"**

One of the most popular definitions of privacy comes from two US-American lawyers, Samuel Warren and Louis Brandeis (see Figure 2), who wrote the first legal article that framed privacy as a "tort action", i.e., a civil wrong that one could sue for compensation of injuries. In their 1890 Harvard Law Review paper [**2**], Warren and Brandeis described privacy as "the right to be let alone," a state of solitude and seclusion that would ensure a "general right to the immunity of the person, the right to one's personality."


**Figure 3: The Kodak Camera Suddenly Allowed Anybody To Take An Instant Photograph**


It is interesting to look into the particular circumstances that prompted Warren and Brandeis to write that article. In their introduction, they write "numerous mechanical devices threaten to make good the prediction that 'what is whispered in the closet shall be proclaimed from the housetops'." What may sound like an accurate description of the new possibilities of ubiquitous computing systems is actually a reference to the technical progress in the field of photography at that time. Before 1890, getting one's picture taken usually required visiting a photographer in his studio and sitting still for a considerable amount of time, otherwise the picture would be blurred. But on October 18, 1884, George Eastman, the founder of the Eastman Kodak Company, received U.S.-Patent #306 594 for his invention of the modern photographic film. Instead of having to use the heavy glass plates in the studio, everybody could now take Kodak's "Snap Camera" out on the streets and take a snapshot of just about anybody without their consent (see Figure 3). It was this rise of unsolicited pictures – pictures that started to appear more and more often the ever expanding tabloid newspapers – that prompted Warren and Brandeis to paint this dark picture of a world without privacy.

Today's developments of "Smart Labels", "Memory Amplifiers", and "Smart Dust" seem to mirror the sudden technology shifts experienced by Warren and Brandeis, opening up new forms of social interactions that change one's expectation of privacy. However, even the strong resemblance of technological progress cannot ignore the fact that their "right to be let alone" looks hardly practicable today: With the multitude of interactions in today's

world, consumers find themselves constantly in need of dealing with people that do not know them in person, hence require some form of information from them in order to judge whether such an interaction would be beneficial. From opening bank accounts, applying for credit, obtaining a personal yearly train pass, or buying books on-line – one constantly has to disclose part of one's personal information in order to participate in today's modern society. Preserving privacy through isolation is just not as much an option anymore as it was one hundred years ago.

A more up-to-date definition thus comes from the 1960s, when automated data processing first took place on a national scale. Alan Westin, professor emeritus of public law and government at Columbia University, defined privacy in his groundbreaking book *Privacy and Freedom* [**3**] as follows:

> Privacy is the claim of individuals, groups, or institutions to determine for themselves when, how, and to what extent information about them is communicated to others.

This definition is often described as *information privacy*, contrasting it to Warren and Brandeis definition of privacy as solitude, of being "let alone." While solitude might be an effect of information privacy, Westin stressed the fact that "the individual's desire for privacy is never absolute, since participation in society is an equally powerful desire." However, as Warren and Brandeis' definition suggests, being in control of one's personal data is only one facet of privacy. Back in the 19th century, the protection of the home – or *territorial privacy* – was the most prevalent aspect of privacy protection. Equally important was the idea of *bodily privacy*, the protection from unjustified strip searches or medical tests or experiments (e.g., drug testing). These two facets are also often called local privacy or physical privacy. And with the invention of the telegraph and telephone in the late 19th century, the rise of modern telecommunication required re-evaluation of the well-known concept of *communication privacy*, previously manifested in the secrecy of sealed letters.

Over the last 200 years, the focus of privacy has thus shifted from things that one could perceive directly with one's own eyes and ears (bodily privacy and territorial privacy) to more "remote" forms of privacy such as communication privacy and information privacy, where the privacy violations are undertaken at a distance. It is interesting to note, however, that ubicomp has made those seemingly long-solved issues of bodily and territorial privacy become highly relevant again: Smart appliances, wearable computers, and activity recognition algorithms allow one to invade the bodily and territorial privacy of another person – not with one's own eyes and ears, but from a distance that had previously constituted the realm of communication and information privacy. A smart fridge might disclose the activities in a home to a grocery distributor, while a smart shirt would send a stream of vital signs to a remote health center or insurance provider.

The limitation of both Warren and Brandeis' and Westin's definition of privacy is that they do not specify exactly how one's privacy should be protected. In order to better understand how to build technology that safeguards privacy, one need to look at how one's privacy can be violated. This is important because privacy, just as security, is often not a goal in itself, not a service that people want to subscribe to, but rather an expectation of being in a state of protection without having to actively pursue it. For example, most users would prefer systems without passwords or similar access control mechanisms, as long as they would not suffer any disadvantages from this. Only if any of their data are maliciously deleted or illegally copied, users will regret not having any security precautions in place. So what would be the analogy to a "break-in" from a privacy point of view?

Gary T. Marx, professor emeritus for sociology at MIT, has done extensive research in the areas of privacy and surveillance, identifying personal border crossings as a core concept: "Central to our acceptance or sense of outrage with respect to surveillance ... are the implications for crossing personal borders" [**4**]. Marx differentiates between four such border crossings that are perceived as privacy violations:

- Natural borders: Physical limitations of observations, such as walls and doors, clothing, darkness, but also sealed letters and telephone calls. Even facial expressions can form a natural border against the true feelings of a person.
- Social borders: Expectations about confidentiality for members of certain social roles, such as family members, doctors, or lawyers. This also includes expectations that your colleagues will not read personal fax messages addressed to you, or material that you left lying around the photocopy machine.
- Spatial or temporal borders: The usual expectations of people that parts of their life, both in time and social space, can remain separated from each other. This would include a wild adolescent time that should not interfere with today's life as a father of four, or different social groups, such as your work colleagues and friends in your favorite bar.
- Borders due to ephemeral or transitory effects: This describes what is best known as a "fleeting moment," an unreflected utterance or action that one hopes gets forgotten soon, or old pictures and letters that one puts out in the trash. Seeing audio or video recordings of such events later, or observing someone sifting through our trash, will violate one's expectations of being able to have information simply pass away unnoticed or hopefully forgotten.

Whenever your personal information crosses any of these borders without your knowledge, your potential for possible actions – your decisional privacy – gets affected. When someone at the office suddenly mentions family problems that you have at home, or if circumstances of your youth suddenly are being brought up again even though you assumed that they were long forgotten, you perceive a violation of your zonal, informational, or communication privacy. This violation is by no means an absolute measure, but instead depends greatly on the individual circumstances, such as the kind of information transgressed, or the specific situation under which the information is disclosed. The effects such border crossings have on people's lives, as well as the chances that they actually happen, are therefore a highly individual assertion.

In a similar fashion, Solove [**5**] has attempted to create a *privacy taxonomy*, i.e., an overview of the activities that might lead to privacy problems. He groups such activities into four sets (see Figure 4): information collection, information processing, information dissemination, and invasion. Starting from the affected individual, the data subject, various entities collect information. Most of the time this will be voluntary, but hidden or forced collections lead to *surveillance* or *interrogation* activities that violate the data subject's privacy. Once data holders process the data, i.e., store, combine, search, or otherwise use it, they might engage in a number of activities that directly threaten the subject's privacy: through *aggregation* (multiple information sources might be linked that the subject might prefer to be separated); *identification* (which ties a particular information or activity to a person); acts of *insecurity* (a failure to properly protect the stored information that leads to improper access); *secondary use* activities (using the collected data for a purpose other than what was agreed with the data subject); and *exclusion* (the lack of letting the data subject know what data the holder has on file about her and how it is used). Information dissemination activities then propagate the information outward from the data holder: a *breach of confidentiality* is breaking a promise of keeping information confidential; *disclosure* means the publication of truthful facts about a person that might affect the person's reputation; *exposure* involves revealing private details, e.g., nude pictures or bodily disabilities; *increased accessibility* concerns the publication of, e.g., telephone numbers or email addresses; *blackmail* is the threat of disclosing information; *appropriation* is the use of the data subject's identity to serve someone else's interest; and *distortion* activities concern the dissemination of false or misleading information about the data subject. Solove also lists *intrusion* into one's life and *decisional interference* as privacy problems, even though they do not necessarily involve the use of personal information (but they often do).

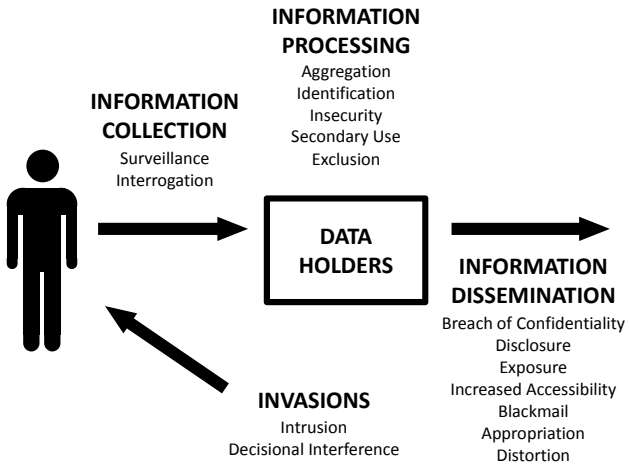



**Figure 4: Solove's Privacy Taxonomy [5]**
**© 2006 D. Solove. Figure reprinted with permission. Also appears in [6]**

While Solove has drawn up the framework primarily to aid in discussing legal protections in the domain of privacy and tort law, his taxonomy is nevertheless also useful for technologists. Given his comprehensive list of privacy-related wrongdoings, one can systematically analyze whether a particular piece of software and/or technology might increase the chances of such problem occurring, and how to mitigate it.

**Conclusions**

1. Privacy has a long history and is by no means a fad of modern times.
2. Technological shifts often open up new ways of how privacy can be affected, thus prompting the need to reassess one's understanding of what privacy is and how it should be protected. With the advent of modern telecommunication and computers, society's focus changed from bodily and territorial privacy to communication and information privacy. Ubicomp's embedded and ubiquitous sensors now reassert the importance of bodily and territorial privacy.
3. A final definition of privacy is difficult. Privacy is related to, but not identical with: secrecy, solitude, liberty, autonomy, freedom, intimacy, and personhood.
4. Privacy violations can be seen as involuntary border crossings, i.e., whenever information permeates barriers without our help (or contrary to our efforts) – barriers like sealed letters, closed doors, the trust of confidentiality with a close friend, or the passage of time.

**Further Reading**

- Daniel Solove: *Understanding Privacy*. Harvard University Press, Cambridge, MA, USA, 2008.
  Solove offers detailed legal analyses of various privacy problems, based on his taxonomy (cf. Figure 4). The book is based on an earlier journal article [**5**], which is freely available for download: http://papers.ssrn.com/sol3/papers.cfm?abstract_id=667622
- Ken Gormley: One Hundred Years of Privacy. *Wisconsin Law Review,* (1335), 1992.
  Examines the evolution of privacy law in the United States, with a particular focus on the varying definitions of privacy over time. Available from: http://cyber.law.harvard.edu/privacy/Gormley--100%20Years%20of%20Privacy.htm
- Paul Sieghart: Privacy and Computers. Latimer, London, UK, 1976.
  One of the earliest books on modern information privacy and still an insightful read about the consequences of computerized data processing.

## 2.2 MOTIVATING PRIVACY: DO PEOPLE CARE ABOUT PRIVACY?

Defining privacy might be a fruitless endeavor if its goals remain unclear. Why is it important to protect one's privacy? What can be gained by providing privacy, what would be at stake if it were lost?

Another way of differentiating the various conceptions of privacy can be found by distinguishing the various effects privacy has on people's lives, grouping them around the three functional concepts of zonal, relational, and decisional privacy. Zonal privacy protects certain spaces, such as one's home, workplace, or car. Relational privacy protects the relationships in an individual's life, such as intimate family relations between husband and wife, or between mother and child. Decisional privacy is what Beate Rössler, professor for philosophy at the University of Amsterdam, calls "securing the interpretational powers over one's life," the freedom to decide for oneself "who do I want to live with; which job to take; but also: what clothes do I want to wear" [**7**].[1]

Privacy is thus needed for the autonomy of the individual, to protect one's independence in making choices central to personhood. Without privacy, such personal choices would inevitably have to be done in the open, with society at large (e.g., friends, family, neighbors, colleagues, superiors, and subordinates) scrutinizing and judging them, thus leaving no "margin for error". Privacy provides us with room to experiment, a space to explore choices and values, in order to find the right balance between our own goals and society's expectations. Westin describes this as follows:

> "Each person is aware of the gap between what he wants to be and what he actually is, between what the world sees of him and what he knows to be his much more complex reality. In addition, there are aspects of himself that the individual does not fully understand but is slowly exploring and shaping as he develops." [**3**]

Westins allegory of playing roles in everyday life has its roots in the interaction theory of sociologist Erving Goffman, who described the fact that people routinely disclose and withhold information about themselves in a very selective fashion in order to maintain different fronts, or *faces*, for different audiences. This also connects with what Westin calls the emotional release functionality of privacy, moments "off stage" where an individual can be himself, finding relief from the various roles he plays on any given day: "stern father, loving husband, car-pool comedian, skilled lathe operator, unions steward, water-cooler flirt, and American Legion committee chairman." Equally important in this respect is the "safety-value" function of privacy, e.g., the "minor non-compliance with social norms" and to "give vent to their anger at 'the system,' 'city hall,' 'the boss':"

> "The firm expectation of having privacy for permissible deviations is a distinguishing characteristic of life in a free society." [**3**]

One important aspect of motivating privacy is of course to ask the data subjects directly. What kind of information about yourself do you consider private? What type of actions would you consider to be privacy invasive? There are and have been many such surveys, the most prominent ones perhaps by Alan Westin, who has conducted over 30 such surveys in the US since 1978. In most of these surveys, Westin provided a summary that classified the respondents into three categories – privacy fundamentalists, privacy pragmatists, and privacy unconcerned – based on their replies. Westin described privacy fundamentalists as "generally distrustful of organizations that ask for their personal information" and "worried about the accuracy of computerized information and additional uses made of it." Pragmatists instead would "weigh the benefits to them of various consumer opportunities and services, protections of public safety or enforcement of personal morality against the degree of intrusiveness of personal information sought and the increase in government power involved," while unconcerned would be

[1] Quote translated from the original German text by the author

"generally trustful of organizations collecting their personal information, comfortable with existing organizational procedures and uses, and ready to forego privacy claims to secure consumer-service benefits or public-order values." Westin measured the distribution of these three types over the years, typically finding about 55-60% pragmatists, 25-30% fundamentalists, and 10-20% unconcerned (see Figure 5).

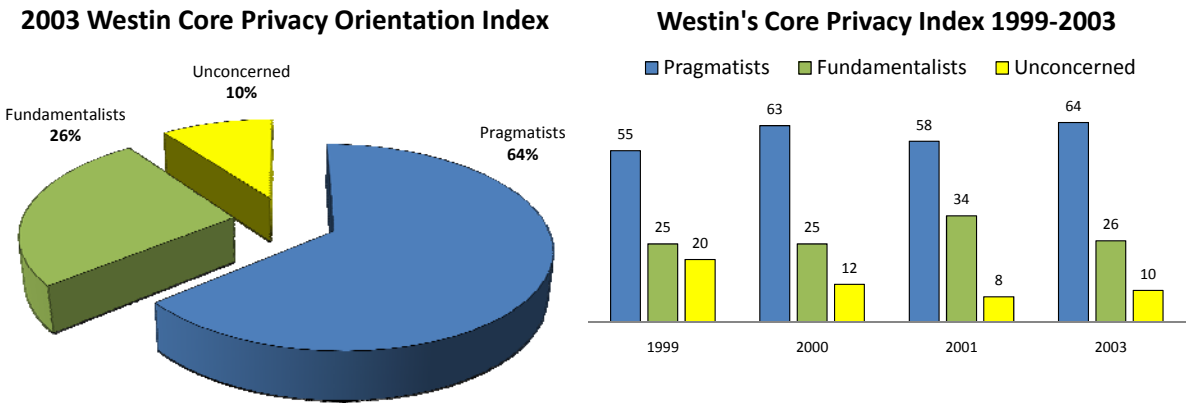



**Figure 5: Westin Core Privacy Orientation Index, differntiating between privacy pragmatists, fundamentalists, and unconcerned. Original Data From Westin, as Cited in [8]**

As to the actual data that is considered private, answers similarly differ. In a November 1998 survey (see Figure 6), administered among 381 Internet users in the US on comfort levels for disclosing personal information online, Cranor, Reagle, and Ackerman [**9**] found that large numbers were always or usually comfortable with disclosing their email address (76%) but only half would be comfortable giving out their full name (54%) or postal address (44%). Few said they would be comfortable giving out their telephone number (11%). A 2007 poll[2] among 1200 US adults also found strong differences in privacy perception between different age groups. For example, only 35.6 percent of 18-24 year-olds consider someone posting a picture of them in a swimsuit to be an invasion of their privacy, compared to 65.5 percent of other respondents, and only 19.6 percent of 18-24 year-olds consider their dating profile to be an invasion of their privacy, compared to 54.6 percent of other respondents.

At the same time, however, there is ample evidence that people value their privacy far less than they indicate in their survey replies. In Germany, recent polls show that some 64% of all citizens carry at least one consumer loyalty card with them at all times, with 34% of respondents using it for almost every purchase [**10**]. Obviously, many of today's consumers are willing to disclose detailed shopping records in return for savings of often less than half a percent. Several surveys found that users of (imaginary or prototypical) location-based services would not mind sharing this information widely with companies or other interested parties, given a small remuneration [**11**]. Some scholars have called for a "privacy as property" approach, where personal data becomes a commodity that people can sell, arguing that market forces might do a better job in protecting privacy than laws and regulations alone.[3] And the recent rise of publicly available personal information on blogs and social networking sites continues to astonish sociologists and legal scholars alike [**12**].[4]

[2] What is Privacy? Poll Exposes Generational Divide on Expectations of Privacy, Zogby/Congressional Internet Caucus Advisory Committee Survey. January 31, 2007 http://www.zogby.com/search/ReadNews.dbm?ID=1244

[3] See [**56**] for an overview and a critique.

[4] See, e.g., http://www.davidhenderson.com/2009/01/21/key-online-influencer/ or http://www.seo-pr-tips.com/2009/01/26/ facebook-new-big-brother/ for recent examples of online self-disclosure of personal information gone wrong.

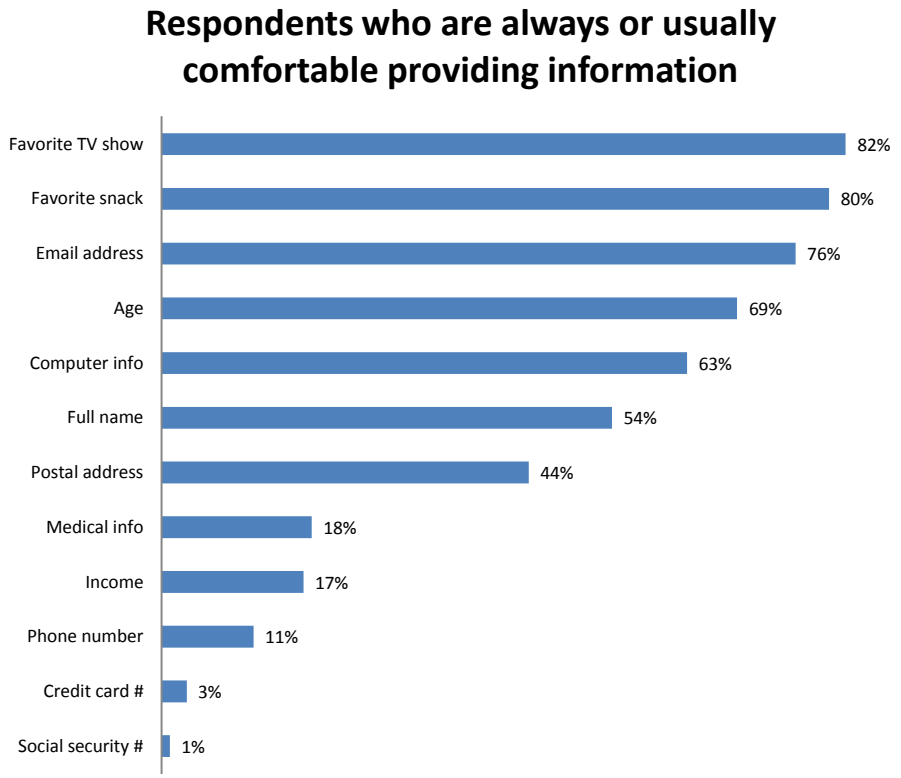



**Figure 6: Results of a 1998 Survey Among 381 US Internet Users Regarding the Disclosure of Personal Information Online [9]**

Should people be allowed to sell large parts of their personal data, maybe even on a lifetime contract? Is there a benefit to having privacy, even if the data subject does not want it? One important alternative motivation for privacy goes beyond the value for the individual, and sees privacy as a social good necessary for the functioning of a democratic society. Representative of this paradigm shift was the so-called "census-verdict" of the German federal constitutional court (Bundesverfassungsgericht) in 1983, which extended the existing right to privacy of the individual (Persönlichkeitsrecht) with the right of *self-determination over personal data* ("informationelle Selbstbestimmung"). The judgment reads as follows:

> Those who cannot with sufficient surety be aware of the personal information about them that is known in certain parts of their social environment, and who cannot judge the amount of information that potential communication partners may know about them, can be seriously inhibited in their freedom to plan and decide in a self-determined manner. A society, in which the individual citizen would not be able to find out who knows what when about them, would not be reconcilable with the right of self-determination over personal data. Those who are unsure if deviant attitudes and actions are ubiquitously noted and permanently stored, processed, or distributed, will try not to stand out with their behavior. [...]This would not only limit the chances for individual development, but also affect public welfare, since self-determination is an essential requirement for a democratic society that is built on the participatory powers of its citizens.[5]

[5] (BVerfGE 65, 154). Translation by the author. See http://www.servat.unibe.ch/law/dfr/bv065001.html for the full text in German. The quote comes from paragraph 154 (as inidicated on the right hand side of the page).

The concept of self-determination over personal data constitutes an important part of modern privacy legislation with respect to ensuring the autonomy of the individual. Firstly, it extended classical data protection principles with a participatory approach, which would allow the individual to decide beyond a "take it or leave it" choice over the collection and use of his or her personal information. Secondly, it frames privacy protection no longer only as an individual right, but emphasizes its positive societal role. Privacy is thus seen not as an individual fancy, but as an obligation of a democratic society, as Julie Cohen notes [**13**][6]:

> Prevailing market-based approaches to data privacy policy ...treat preferences for informational privacy as a matter of individual taste, entitled to no more (and often much less) weight than preferences for black shoes over brown, or red wine over white. But the values of informational privacy are far more fundamental. A degree of freedom from scrutiny and categorization by others promotes important non-instrumental values, and serves vital individual and collective ends.

**Conclusions**

1. Privacy plays an important role in human relationships, enabling us to create intimacy, as well as in personal development, supporting decisional autonomy.
2. Many people wish to control the flow of information about themselves, but they often differ widely about what kinds of information they want to control.
3. While the public interest often conflicts with the privacy of the individual, there is a strong societal benefit in offering its members at least some degree of freedom from scrutiny and categorization.

**Further Reading**

- Daniel Solove: The Future or Reputation. Yale University Press, New Haven, CT, USA. 2008.
  Discusses how recent trends in publishing personal information online – both by ourselves and by our friends or enemies – can seriously affect personal freedom and self-development. The full text is licensed under a creative commons and is freely available at http://docs.law.gwu.edu/facweb/dsolove/Future-of-Reputation/text.htm
- Colin J. Bennett and Charles D. Raab: *The Governance of Privacy*. MIT Press, Cambridge, MA, USA, 2006.
  Gives an excellent overview on international privacy protection policy and examines regulatory instruments, in particular in light of technological change and globalization. Argues that privacy related problems are as much public policy issues as they are legal and technological ones.
- Simson Garfinkel: Database Nation. O'Reilly, Sebastopol, CA, USA, 2000.
  A vivid illustration of the myriads of data collections taking place today, and the danger these might pose to the individual.

## 2.3 LEGAL BACKGROUND

Data protection and privacy laws provide an important aspect for understanding both the "Why?" as well as the "How?" of privacy protection. Laws and regulation mirror a social process that defines, for a particular culture, a set of practices that are acceptable and unacceptable. Looking at privacy laws thus helps to understand what society envisions privacy to be. At the same time, laws are also tools that may complement, support, or in turn rely on technical privacy tools.

The work of Warren and Brandeis in 1890 argued for a "right to privacy" in the realm of *tort law*, a part of the law that provides remedies for civil wrongs. In the 1960s, William L. Prosser described a set of four *privacy torts* that have since become established legal concepts (cf. with Solove's extended taxonomy in Figure 4 in section 2.1):

[6] As reprinted in [**55**]

1. Intrusion upon seclusion or solitude, or into private affairs;
2. Public disclosure of embarrassing private facts;
3. Publicity which places a person in a false light in the public eye; and
4. Appropriation of name or likeness.

While tort law addresses conflicts between two private parties, *privacy law* or *data protection law* regulates how the government and companies can collect and process personal information. The basis for all modern privacy laws are the Fair Information Principles, which were drawn up in the early 1980s by the Organization for Economic Cooperation and Development (OECD) and which describe eight practical measures aimed at harmonizing the processing of personal data in its member countries. By setting out core principles, the organization hoped to "obviate unnecessary restrictions to trans-border data flows, both on- and off-line." The eight principles are as follows:[7]

1. *Collection Limitation Principle*. There should be limits to the collection of personal data and any such data should be obtained by lawful and fair means and, where appropriate, with the knowledge or consent of the data subject.
2. *Data Quality Principle*. Personal data should be relevant to the purposes for which they are to be used, and, to the extent necessary for those purposes, should be accurate, complete and kept up-to-date.
3. *Purpose Specification Principle*. The purposes for which personal data are collected should be specified not later than at the time of data collection and the subsequent use limited to the fulfillment of those purposes or such others as are not incompatible with those purposes and as are specified on each occasion of change of purpose.
4. *Use Limitation Principle*. Personal data should not be disclosed, made available or otherwise used for purposes other than those specified in accordance with the *Purpose Specification* principle except:
    a) with the consent of the data subject; or
    b) by the authority of law.
5. *Security Safeguards Principle*. Personal data should be protected by reasonable security safeguards against such risks as loss or unauthorized access, destruction, use, modification or disclosure of data.
6. *Openness Principle*. There should be a general policy of openness about developments, practices and policies with respect to personal data. Means should be readily available of establishing the existence and nature of personal data, and the main purposes of their use, as well as the identity about usual residence of the data controller.
7. *Individual Participation Principle*. An individual should have the right:
    a) to obtain from a data controller, or otherwise, confirmation of whether or not the data controller has data relating to him;
    b) to have communicated to him, data relating to him
        i. within a reasonable time;
        ii. at a charge, if any, that is not excessive;
        iii. in a reasonable manner; and
        iv. in a form that is readily intelligible to him;
    c) to be given reasons if a request made under subparagraphs (a) and (b) is denied, and to be able to challenge such denial; and
    d) to challenge data relating to him and, if the challenge is successful, to have the data erased; rectified, completed or amended.
8. *Accountability Principle*. A data controller should be accountable for complying with measures which give effect to the principles stated above.

[7] See http://www.oecd.org/document/18/0,2340,en_2649_34255_1815186_1_1_1_1,00.html

Even though the OECD principles carried no legal obligation, they nevertheless constitute an important international consensus that has substantially influenced national privacy legislation in the years since. Consequently, they provide a good starting point to measure any ubicomp application against, in order to assess its privacy compliance: Is the collection limitation principle adequately addressed? Can data quality be ensured? Are reasonable security safeguards in place? And how is data subject participation and data controller accountability supported in the system?

In practice, it depends on national legislation how these principles are incorporated into actual regulations. In today's legal landscape, two main approaches to privacy legislation exist: the sectorial approach, favored in the US, and the omnibus approach in Europe. US legislation features strong, overarching privacy laws only for the federal government, while state governments and private organizations are regulated on an "as-needed" basis with a variety of highly focused laws, such as the "Driver's Privacy Protection Act" of 1994, which safeguards an individual's motor vehicle record, or the "Video Privacy Protection Act" of 1988, which limits access to customers' video rental records. Laws protecting financial (Financial Modernization Act, also known as Gramm-Leach-Bliley Act) and health records (Health Insurance Portability and Accountability Act, HIPAA) are limited and went in effect only very recently (in 2001 and 2003, respectively). Probably most relevant to ubicomp applications are recent efforts in several US states on addressing RFID technology[8] and several rulings involving location privacy,[9] though "traditional" privacy law that addresses privacy at home, at school, or at work, is of course equally relevant (see, e.g., [**14**] for an overview).

On the other side of the Atlantic, Europe has long since favored overarching frameworks that apply to both governments and commercial entities. Probably the most influential pieces of privacy legislation in recent years is the 1995 "Directive 95/46/EC of the European Parliament and of the Council of 24 October 1995 on the protection of individuals with regard to the processing of personal data and on the free movement of such data"[10] (often called the "Data Protection Directive," or "the Directive" for short). It requires all member states of the European Union[11] to enact national law that offers the same level of privacy protections as those set forth in the Directive, in effect harmonizing privacy law across all EU member states and thus ensuring the free flow of information that the OECD had in mind with their fair information principles. The Directive was complemented by "Directive 1997/66/EC"[12] and later "Directive 2002/58/EC"[13] (which replaced 1997/66/EC) to offer specific guidelines on "the processing of personal data and the protection of privacy in the electronic communications sector" (called the "e-Privacy Directive" for short). While directive 1995/46 sets out general guidelines, it is directive 2002/58 that details their implementation in the telecommunications and networking sector, containing, e.g., provisions for "calls," "communications," "traffic data" and "location data." By 2009, the 2002 e-Privacy Directive has again become dated and a "review" of the e-Privacy Directive is currently underway that attempts to take recent technological developments into account, e.g., the rise of semi-public and private networks such as WiFi hotspots, which are not covered by the 2002 directive. In addition, the European Commission is also evaluating whether specific regulations for the "Internet of Things", in particular the rise of RFID technology, will be required.[14]

[8] See for example the Identity Information Protection Act of 2007 (SB30) in California (www.eff.org/issues/rfid/sb30facts), the new Hampshire House Bill 478 (rfidlawblog.mckennalong.com/2009%20NH%20H%20478.pdf) and the recently proposed New York Assembly Bill A275 (rfidlawblog.mckennalong.com/archives/privacy-new-york-assembly-bill-a275.html). At least 12 US states have already introduced RFID legislation, see www.ncsl.org/programs/lis/privacy/rfid05.htm for an overview.
[9] See for example the recent ruling in Pennsylvania on mobile phone tracking (www.eff.org/press/archives/2008/09/11) or GPS tracking in US vs. Jones (www.eff.org/cases/us-v-jones).
[10] See http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=CELEX:31995L0046:EN:HTML
[11] As of 2007, the European Union (EU) has 27 member states: Austria, Belgium, Bulgaria, Cyprus, the Czech Republic, Denmark, Estonia, Finland, France, Germany, Greece, Hungary, Ireland, Italy, Latvia, Lithuania, Luxembourg, Malta, the Netherlands, Poland, Portugal, Romania, Slovakia, Slovenia, Spain, Sweden, and the United Kingdom.
[12] See http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=CELEX:31997L0066:EN:HTML
[13] See http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=CELEX:32002L0058:EN:HTML
[14] See http://ec.europa.eu/information_society/policy/rfid/index_en.htm

**Conclusions**

1. Data collection and processing does not happen in a legal vacuum. Most countries have laws in place that specifically address privacy issues, both for institutions (i.e., government or industry) and private persons. Ubicomp systems must at least be compliant with these national laws.
2. Laws also help to further understand privacy issues as they provide an overview of negative consequences of data collection practices (see, e.g., privacy torts).
3. The fair information principles are at the core of most modern privacy laws. They require that data subjects be informed of a data collection taking place, give their consent, and have access to the collected data. Security safeguards must be in place and data must only be used for the purpose it was collected under.

**Further Reading**

- Daniel J. Solove, Paul M. Schwartz: Information Privacy Law. 3rd Edition, Aspen, New York, NY, USA, 2009.
  A fascinating read that, despite its dry title, manages to vividly illustrate the many challenges and issues of modern (US-American) information privacy law. With the help of many examples, the book provides a comprehensive overview that does not require a law degree nor law school attendance to understand.
- Christopher Kuner: European Data Protection Law: Corporate Regulation and Compliance. 2nd Edition. Oxford University Press, Oxford, UK, 2007
  The definite (corporate) reference on EU data protection law. With a hefty price tag of £140,- and more details on corporate compliance than you ever cared to know, this book is probably not what you want for your home bookshelf. But if you have a law faculty close-by, their library might hold a copy!
- Electronic Privacy Information Center & Privacy International (Eds): *Privacy and Human Rights Report 2006: An International Survey of Privacy Laws and Developments.* Electronic Privacy Information Center, Washington D.C., USA, 2006.
  An annual report that provides an overview of key privacy topics and reviews the corresponding legislation and state of privacy around the world. Full text of latest available edition should be online at the Privacy International website (see "Key PI Resources"): http://www.privacyinternational.org/

## 2.4 INTERPERSONAL PRIVACY

Laws and the Fair Information Principles are important tools to define the roles of institutionalized data collectors such as companies or government agencies. However, they do not help if privacy issues arise between private parties, i.e., between neighbors, friends, or family members. Neither does the law require parents to announce that they are monitoring the time that their children come home, nor do friends exchange privacy policies with each other before allowing continuous access to their instant messenger status. Clearly, the fair information principles are only of limited applicability when it comes to social etiquette and social norms.

While tort law provides remedies for some privacy violations, such as intrusion or disclosure, these seem to be only last resorts when all friendship, neighborliness, or family bonds have already ended. How can ubicomp applications that offer ubiquitous information exchange between peers be designed in order to avoid having to go to court to ensure one's privacy? What aspects of privacy matter for people when they are not dealing with faceless companies or agencies, but with acquaintances, friends, and family?

In the 1970s, psychologist Irwin Altman looked at how people regulate their "environmental" privacy, i.e., being alone vs. joining social interactions [**15**]. Instead of the simple view of privacy as a state of solitude, Altman saw it as a dynamic boundary negotiation process that encompassed the entire spectrum of social interactions, a "selective control of access to the self or to one's group." For Altman, the optimum level of privacy is reached when

the *achieved* level of privacy reaches the *desired* level of privacy (i.e., the level of contact with others). Having more privacy than desired leads to the feeling of loneliness, having less than what is desired leads to annoyance and the feeling of being crowded. Privacy regulation, in Altman's sense, is thus the control of one's openness and closedness to others in response to one's desires and one's environment.

As a psychologist, Altman looked at behavioral mechanisms that support such privacy regulation: verbal interactions with others ("inputs and outputs") as well as spatial interactions ("personal space and territory"). These mechanisms are the tools by which one regulates one's privacy, by listening to others (input), talking to others (output), positioning oneself in relationship to others (personal space, i.e., distance, angle, etc) and choosing one's location (territory).

While Altman developed his theory for real-world interactions, there is much that can be learned from this theory, even in the context of ubicomp privacy:

- Privacy as a non-monotonic function: By conceptualizing privacy not simply as one end of the social interaction spectrum (i.e., being alone), but applying it to the entire range of interactions, Altman shows that more privacy is not always better. Both "crowding" and "isolation" are suggested as "examples of privacy regulation gone wrong." [**16**]
- Privacy as a social process: Humans do not use one-off policies and rules to manage their everyday, interpersonal privacy. Instead, they continually adjust their accessibility along a spectrum of "openness" and "closedness" with a variety of mechanisms, in order to match the achieved privacy state to the desired one.

Dourish and Anderson [**17**] summarize it succinctly: "Privacy is not simply a way that information is managed but how social relationships are managed." Ubicomp systems that facilitate communication and awareness between peers must thus provide similar tools to allow users to dynamically adjust their inputs and outputs, their levels of openness and closedness, their personal space and territory so that they can achieve their desired level of privacy with respect to other users. In ubicomp systems, Altman's privacy regulation mechanisms of verbal, paraverbal (e.g., personal space or territoriality), and non-verbal communication need to be augmented with (or replicated by) explicit and implicit controls that allow "in-situ" adjustments, rather than a simple "configuration panel" that allows one to set the desired privacy level to "high" or "low". Privacy is thus not so much a particular state of being, but a tool for social discourse.

**Conclusions**

1. Interpersonal privacy cannot be approached by requiring purpose specifications and recipients lists. Between individuals, privacy is a tool for social discourse that allows one to manage peer groups.
2. In this context, privacy is an ongoing boundary-negotiation process and not a monotonic function where "more secrecy" translates to "more privacy". Yes, you can have "too much" privacy (resulting in isolation).
3. Privacy tools must support "in-situ" adjustments, rather than separating configuration and action (e.g., in a separate control panel).

**Further Reading**

- Leysia Palen and Paul Dourish: Unpacking "Privacy" for a Networked World. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '03 ), Ftl. Lauderdale, FL, USA, April 5-10, 2003.* ACM Press, New York, NY, USA, pp. 129-136, 2003
  Based on Altman's theory, Palen and Dourish define a framework of privacy in the context of modern information systems and point out tensions that govern interpersonal privacy regulation in everyday life.

- Battya Friedman, Peter H. Kahn, Jr., and Alan Borning: Value Sensitive Design and Information Systems. In: P. Zhang and D. Galletta (Eds.) *Human-Computer Interaction in Management information Systems: Foundations.* M.E. Sharp, New York, NY, USA, pp. 348-372, 2006
  Value Sensitive Design is a process that explicitly prompts designers to think about and incorporate human values in a principled and comprehensive manner. In the context of privacy, this process may help to understand the reasons behind and the mechanisms of existing and planned information sharing practices.
- Paul Dourish and Ken Anderson: Collective Information Practice: Exploring Privacy and Security as Social and Cultural Phenomena. *Human-Computer Interaction*, Vol. 21, Lawrence Erlebaum, pp. 319-342, 2006
  Dourish and Anderson argue that the commonly formulated abstract goals of attaining "privacy" and "security" should be replaced by thinking about "information practices" – common understandings of the ways how information should be shared, managed, and withheld. They offer a range of pointers to ethnographic studies that illustrate how privacy is used as a boundary regulation process between social groups.

# 3 TECHNICAL SOLUTIONS FOR PRIVACY IN UBIQUITOUS COMPUTING

By first understanding what privacy is and what we are trying to solve, we can now set out to design and implement privacy features in ubicomp applications. The previous section illustrated the underlying legal and social aspects of information collection and sharing, thus helping ubicomp application developers to properly formulate their privacy goals. This section attempts to illustrate various technical solutions to common ubicomp privacy challenges. We focus in particular on three examples: smart spaces (section 3.3), RFID (section 3.4) and location-based services (section 3.5). Obviously, these three areas do not fully capture the variety of challenges presented in ubicomp (as discussed in section 3.1 below), but they hopefully serve as useful examples. Also, since many ubicomp privacy issues apply to information systems in general, additional context can be found from the fields of computer networking, databases, and information retrieval (see section 3.2).

## 3.1 NOVEL UBICOMP CHALLENGES TO PRIVACY

Privacy and data protection have always been closely related to what is technically feasible. At the end of the 19th century, it was the invention of modern photography that prompted Warren and Brandeis to rethink the concept of legal privacy protection. At the beginning of the 20th century, laws had to be reinterpreted again to take into account the possibilities of modern telecommunication (again, then supreme court judge Brandeis played a large part in that). And in the 1960s and 1970s, it was the implementation of efficient government through the use of modern databases that required yet another update of privacy laws, resulting in the first of today's modern data protection laws with their focus on data self-determination. In each instance, technology changed what was possible in the everyday and thus prompted – if sometimes with considerable delay – a realignment of our notion of privacy.

After the rise of Internet e-commerce in the 1990s had initiated the last round of updates, the dawn of ubiquitous computing promises the next revolution of "smart things." Even though many ubiquitous computing visions sound like AI-revisited, applications like the "intelligent car," or the "smart home" might not face the same fate as the dreams of intelligent machines that some 20 years ago researchers thought of being just around the corner. Ubiquitous computing often solves a much more mundane yet important problem, namely crossing media boundaries. Using miniature sensors, cheap microchips, and wireless communication, computer technology can penetrate our everyday lives in a completely unobtrusive manner. Similarly, real world facts and phenomena can be mapped on a computer with an unprecedented reliability and efficiency. The boundary between the real and

virtual world seems to disappear – it will soon be possible to comprehensively track real world interactions in real-time on a computer system, making it look like a game simulation (think SimCity®).[15]

Data protection and privacy is all about these mappings: translating facts of the real world into bits of information that can be stored for later retrieval. Ubiquitous computing is about the digitalization of information about our lives in order to allow computer systems to automatically process it. It comes as no surprise that ubiquitous computing has the potential to yet again change our perception of privacy in a significant manner. This qualitative quantum leap can be traced along five aspects of ubiquitous computing systems: the collection scale, manner, and motivation, as well as the data types and the data accessibility.

**Collection Scale**

The conscious surveillance of the actions and habits of our fellow men is probably as old as mankind. In the "good old times," when people lived in small villages and close-knit social circles, this kind of observation was implicit in our daily interactions: Everybody knew everybody, and local news and gossip spread fast. "Non-compatible" people often preferred to move out into the large cities, in which the large number of citizens and their high variance rendered this classical method of direct social monitoring impractical.

With the rise of automated data processing, machines began to take over the role of the curious neighbor. At first only available to governments, automated data processing soon found its way into commerce, in both cases facilitating a much more efficient management by providing detailed population or inventory information. However, while our neighbors would quickly note anything out of the ordinary, machines were now employed to actually determine what was ordinary: Not the deviations of the norm were noticed and tracked, but the average citizen and his or her ordinary everyday.[16]

With ubiquitous computing, real life monitoring – the surveillance of the ordinary – will extend beyond today's credit card transaction, telephone connection records, and Web server logs. Even without assuming a single homogeneous surveillance network like Orwell's Big Brother, the sheer applicability of ubiquitous computing technology in diverse areas such as hospitals and nurseries, kindergartens, schools, universities, offices, restaurants, public places, homes, cars, shopping malls, and elderly care facilities, will create a comprehensive set of data trails that will cover us anywhere we would go.

The "always on" vision of ubiquitous computing – alleviating us from laboriously switching various devices on and off as everything "stands ready" to our attention, right when we need it – will drastically extend this coverage over time. Instead of the spotty trails that can be obtained through our Internet logs when we are on-line, say, after work for an hour or two, smart homes and intelligent environments will not be switched off at night or while we are gone for lunch. In fact, it might not be even possible to turn such devices off, as they would not feature a corresponding on-off-switch, but would sleep most of the time to preserve energy and wake up on their own whenever something of interest happens to them. The digital coverage of our lives will be anywhere and anytime, from sunrise to sunset, from cradle to grave, 24 hours a day, seven days a week. As [**18**] points out, the actual selection of data that is captured and stored will at the same time significantly alter the value of that information: "Anything that is recorded instantly achieves a potential pervasiveness and immortality that it did not have before.... Anything that does not 'make the cut' ... is invisible to someone inspecting the digital record at a different location or time."

**Collection Manner**

[15] See, e.g., Sandy Pentland's SenseNetworks for a taste of what is to come: http://www.sensenetworks.com/
[16] See, e.g., [**57**] for a description of several large-scale activity studies in the US.

When little children play Hide and Seek, they often cover their eyes with their hands assuming that if they cannot see, others will not see them in turn. While they will learn eventually that the principle of reciprocity does not hold in this case, this apparent childish belief is much more difficult to unlearn than we might want to believe. Even years after playing their last game of Hide and Seek, many will assume that if they cannot see anybody else around, their actions will go unnoticed.

In the old days, this principle of reciprocity was actually a reasonable approximation of the collection manner in which people's action were observed. Only when one was out in public, others were able to see and draw their inferences. Once we entered the sanctuary of our own homes or those of others, we were shielded from the prying eyes of the public. This dichotomy of public and private was closely associated with the realities of space – the architecture of walls, windows, and doors, or the natural environment of woods and dense thickets: The presence and quality of a physical boundary provided an immediate indicator of the (potential) quality of privacy. With the rise of electronic transactions, day-to-day actions like talking to a friend (over the phone) or buying groceries (using a credit-card) became noticeable beyond such physical boundaries. The presence or absence of others was not a good approximation of privacy anymore, as the digital trace of a transaction could be observed, stored, and retrieved from potentially anywhere in the world.

The deployment of ubiquitous computing technology will make it even more difficult to differentiate between public and private actions. As ubiquitous computing tries to hide the use of technology, to make computers practically invisible, the level of awareness for such electronic transactions will drop drastically from today's implicit awareness through the use of physical tokens such as credit-cards or mobile phones. In a fully computerized environment, potentially any item could take fingerprints and wire them halfway around the world, take pictures, measure body temperature, or observe one's gait in order to draw far-reaching conclusions about a person's physical and mental state. Neither data collection nor continuous surveillance activities will have recognizable markers that would indicate the publicity of actions – ultimately requiring us to assume that at any point in time, in any location, any of our actions could potentially be recorded electronically and thus made public.

**Data Types**

With ubiquitous computing, also the type of information that is collected will also change. The village gossip was based on the observation of neighbors and fellow citizens and on a person's discussions with others. This information was by definition "soft" information, that is, it was based on an individual's personal reception and more often that not, two different people observing the same fact would retell widely different accounts of it. While this would often result in rather exaggerated claims, it nevertheless retained some level of deniability.

Modern data processing seems far away from the village gossip of old. It concerns itself with "hard" information – with facts, rather than hearsay. Instead of capturing the individual (and error-prone) human perception, it collects factual information such as names, birth dates, addresses, income levels, or lists of purchases. Using statistical models, this information can subsequently be used to draw inferences on a person's life based on his or her residence and shopping preferences.

Ubiquitous computing will extend this selection of hard facts beyond traditional information types: smart shirts and underwear will be able to record health data such as blood pressure, heart rate, perspiration, or glucose levels in real time; smart supermarket shelves will not only know what items a person bought, but also in what order and how long he or she hesitated before reaching out; mobile phones with GPS-locator already allow friends and family to know one's whereabouts at anytime – in the future this will be standard unless one decides to turn the service off and find a good excuse for doing so.

Data mining technology will allow researchers, politicians, and marketers to make sense of this ever increasing stream of minute details, by correlating widely disparate information such as chocolate consummation and shower habits (for example to infer the beginning of a new relationship), and through comparing information from hundreds of similar people in order to discern broad population patterns. This also has significant implications for the anonymization of such data, as perceived information such as one's location over the course of a day, or the particular way of walking as registered by floor pressure sensors, or one's individual breathing pattern, might turn out to be easily identifiable even if collected in a completely anonymous manner.

With a wide array of new sensors and collection mechanisms, ubiquitous computing technology will potentially allow inferring the "soft" gossip of old, based on the "hard" facts of today, thus not only giving it new credibility (by being based on facts, not hearsay) but also eventually incapacitating our own judgments about personal beliefs and feelings based on computerized self-assessments, e.g., inferring our emotional attachment to our partner based on our heart-rate and eye blinking rate.

**Collection Motivations**

As we have seen in the previous section on privacy and its motivations, incentive (i.e., the "Why?") plays an important role when it comes to facilitating or preventing data collection. And just as the reasons for wanting privacy have changed over the years, so have the motivations for collecting this data.

Our neighbor's eyes and ears looked for the unusual, the out-of-ordinary events that would make for attractive gossip. Consequently, people who were adept at "blending in", those who hardly attracted attention due to their ordinary lives and average physical features, would get the least scrutiny. With automated data processing, attention shifted from the unusual to the ordinary: Governments tried to make better policies by having better data on whom they governed, and that meant finding out what the average citizen did, liked, or feared. Companies tried to find out what goods consumers wanted (or did not yet know they wanted). Questionnaires were used (and still are) to solicit the preferences of the masses, in order to better understand what products would work and which would not. With modern data analysis methods, large amounts of statistical information, such as family income, street address, or political preferences, can be statistically correlated in order to segment population groups and predict human behavior (e.g., a family moving into the suburbs might soon decide to buy a lawn mower, as most families living there own one already).

Providing better services and/or better products will still be at the heart of many future ubiquitous computing systems, yet what data is necessary to predict this becomes less and less clear, as more different types of information can be collected. With better data mining capabilities than ever before, virtually anything can be of importance, if only enough statistical data on it can be collected. Context-Awareness is one of the main paradigms in ubiquitous computing, as it is thought to enable otherwise "dumb" systems to predict the user's needs and intents without involving any actual intelligence. Not surprisingly, the more such context information is available, the better these systems are expected to perform. Instead of targeted data collections of specific information for a certain purpose, future ubiquitous computing systems could easily attempt to collect any and all information possibly available, thus maximizing their chances for correctly determining the user's context and intent from it.

**Data Accessibility**

Information is only of worth if one can find it: collecting large amount of data without having efficient retrieval mechanisms in place suggests not collecting it in the first place. In the old days, retrieving gossip was typically limited to a particular village or neighborhood. By moving into a different town or even into a larger, anonymous city, the previously assembled body of "knowledge" would typically be rendered inaccessible for the newly acquired neighbors, requiring them to start out anew.

With modern information networks, information can travel quickly around the globe, and modern database management systems allow for the efficient retrieval of minute details out of huge, federated databases from a wide variety of sources. However, even though standardized interface definitions exist, integrating these sources is far from a trivial problem, as the large number of failed data-integration projects in both government and industry have shown.

In the vision of ubiquitous computing, such kinds of information systems would not be primarily designed with humans in mind (and thus will lead to often non-interoperable systems), but directly target machine-to-machine interactions: Smart things would "talk" to other smart things in order to collaboratively determine the current context, and large networks of autonomous sensor nodes would send sensor readings back and forth in order to arrive at a global state based on hundreds of individual sensor readings. Similarly, improved human-computer interfaces (see the chapter by Aaron Quigley on UUIs in this book) would allow easy access to non-traditional data formats such as video-and audio-streams, e.g., for automated diary applications that would document one's everyday in a continuous multimedia format. Living in a world of smart cooperating objects, the "freedom of movement" for personal information would be greatly increased, both between humans and computers (How well can I search your memory?) and between cooperating artifacts (What is my artifact telling yours?).

## 3.2 THE BASICS: PRIVACY ENHANCING TECHNOLOGIES (PETS)

Using (information) technology to protect privacy is as old as the first databases that appeared in the 1960s. Consequently, a number of techniques exist that form the general toolbox of privacy compliant data processing, often abbreviated as "PETs" – privacy enhancing technologies. Many of them operate on the networking layer or databases, and as such might be readily applicable to ubicomp applications. Others are more reliant, e.g., on traditional Web interfaces and thus require some sort of "ubicompification" in order to be of use. Today's PETs can roughly be grouped into two sets: *transparency* tools and *opacity* tools.

*Opacity tools* describe the more traditional security approaches, namely support for authentication and confidentiality. Traditional communication link encryption, e.g., secure socket layers (SSL) or secure shell (SSH), that prevent an attacker from learning the contents of a communication, can be complemented by *unobservability* tools that attempt to prevent attackers from even learning that a communication took place. This line of research was instigated in the 1980s by David Chaum's Mix-Net protocol [**19**], in which email is not sent directly from sender to recipient, but is routed through a cascade of "mixes" – network nodes that combine (delay) and forward messages in a seemingly random fashion. Using public-key cryptography, the sender embeds the message – like a Russian doll – in several layers of encryption. Each layer is readable to only one particular mix node and reveals only where to send the message next. An external observer will not be able to trace the flow of messages through the system, even if a significant fraction of the mix nodes get compromised. *Online* mixes carry this concept over to Internet traffic in general, allowing participants to route any Internet connection through a "cloud" of mixes, making it impossible for an attacker to associate, e.g., Web site visits to a particular client.[17] A similar application is anonymous or pseudonymous payment, again pioneered by David Chaum in the late 1980s, where completely anonymous online payments can be made [**20**].[18]

In the area of authentication, so-called *identity management tools* attempt to disconnect identity from authority, i.e., allowing a user to prove his or her authority to access a resource without revealing identity. While early commercial systems often equated the term with simple "single-sign-on" solutions where multiple logins would be seamlessly performed by a central login service, today's systems such as Microsoft's CardSpace (which is integrated

[17] For commercial implementations see, e.g., JonDo at https://www.jondos.de/en/ or Anonymizer.com at http://www.anonymizer.com/
[18] Obviously, the majority of today's online payments are not anonymous, which just goes to show that it takes more than good technology to change a multibillion dollar business (market forces and political preferences clearly dominate such developments)

into Microsoft Vista®) use cryptography-based certificates to prove, e.g., that one is over 18, without the need for disclosing any identifying personal information.[19] Figure 7 shows the Identity Selection Screen in Windows Vista®.

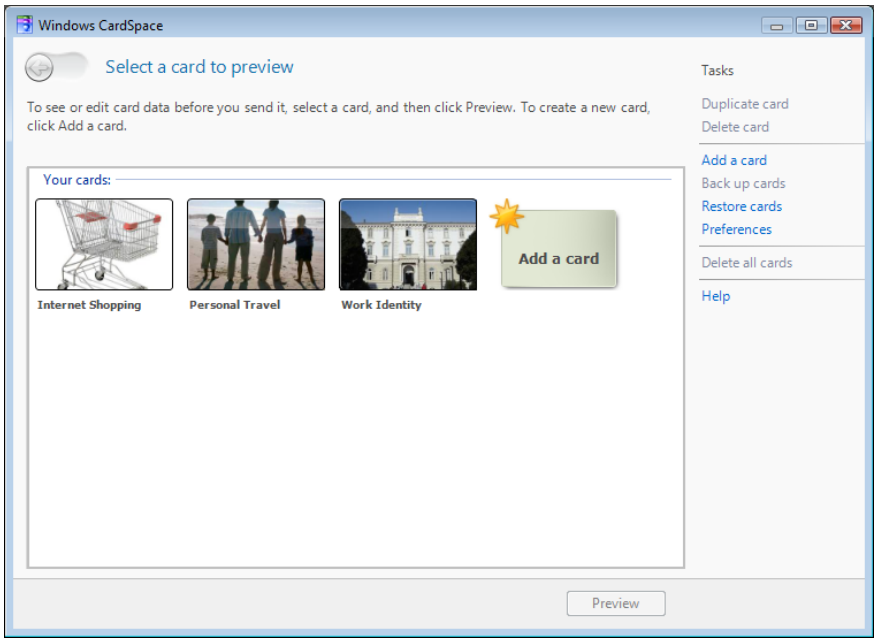



**Figure 7: Microsoft CardSpace (built into Windows Vista®) is an Example of a Modern Identity Management System.**
**© 2009 Microsoft. Microsoft product screen shot reprinted with permission from Microsoft Corporation.**

Maintainers of medical databases have long since identified the need to anonymize their records, without changing important statistical properties that might help researchers, e.g., epidemiologists, to identify important patterns in the collected data. Research in so-called *statistical databases* offers guarantees on the viability of several statistical measures while anonymizing individual datasets. A common measure for the degree of anonymity achieved is *k-anonymity*, indicating that a particular record or piece of information could be from *k* possible people [**21**]. The higher the value of *k*, the higher the level of anonymity.

Opacity tools are complemented by another set of PETs, so-called *transparency tools*, which in turn attempt to improve the data subject's understanding and control of his or her data profile. *Watermarking* systems allow the visible or invisible marking of information in order to trace the origin of a particular piece of data. This is already in widespread use in digital rights management system, e.g., in digital photography, where the rights holder can use such watermarks to prove authorship/ownership of a particular image. While such techniques can readily be used to make personal pictures, the ability to watermark other personal data is limited. People have long since used spelling errors in names and addresses when signing up, e.g., for mail-order catalogs, in order to track the origin of (real-world) spam mail. The same approach can of course be used with one-time email addresses, but its use with other data items (e.g., preferences) is doubtful.

*Policy tools* use a similar approach in storage systems in order to allow for the privacy-compliant data processing of personal information. All data collected is stored together with a set of metadata that describes the allowed recipients, uses, and storage duration. Each database operation requires operators to specify the identity of the

[19] A Canadian company called Zero Knowledge Systems (now called Radialpoint) pioneered this approach in the late 1990s. Their *Freedom Network* product allows paying customers to create several digital identities (called "nyms") that they could use to associate their Internet activity with. While the technology was impressive, the company ultimately failed to get consumers to pay for this (the product was discontinued in 2002). Clearly, the idea of "choosing a nym" in advance of each and every Internet session, and the added overhead of "managing" one's set of nyms (i.e., "which one should I choose for this particular action?") did ultimately not appeal to Internet users.

requestor and the reason for the query (purpose). This can then be used by the policy engine to return only those entries with a compatible privacy policy. If the requester is using a compatible system, the queried data in turn will carry metadata that allows further privacy-compliant processing. This approach is sometimes called the "sticky policy" paradigm [**22**] or "hippocratic databases" [**23**]. With such a system in the background, data subjects can receive detailed information on the use and storage of their personal information, while both companies and data protection agencies can perform automated audits to verify the compliance with data protection law and regulations. Much work has also been undertaken in policy languages, i.e., describing both the data processing guarantees by data collectors and the privacy preferences by data subjects. The Platform for Privacy Preferences (P3P) is probably the most prominent one: it uses XML syntax to describe both data processing policies and subject preferences – the latter in an addendum specification called "APPEL – A Privacy Preferences Exchange Language" [**24**].

Note that all of the above PETs are centered around "traditional" data processing and Web use. While PETs are an important building block to any ubicomp privacy solution, they must be complemented and extended in order to be useful in ubicomp settings, as outlined in section 3.1 above.

## 3.3 EXAMPLE: PROTECTING SMART SPACES

Smart environments are probably the archetypical ubicomp application: A smart room or house that constantly detects the context and activities of its inhabitants and visitors and seamlessly adapts its services and functions to provide the best possible experience. Initial work on ubicomp privacy has thus focused on both infrastructures that collect this data, and on interfaces that allow data subjects to inspect and control the corresponding information flows.

One prominent privacy-aware ubicomp infrastructure is the Confab Toolkit by Hong and Landay [**25**]. Data in Confab is managed in InfoSpaces, network-addressable logical storage units that store context information about a single entity, i.e., a person, a location, a device, or a service. In-and out-filters manage data flows between different InfoSpaces, with *in-filters* only allowing the storage of data from trusted sensors or entities, and *out-filters* enforcing access policies and adding privacy tags to all outgoing data (see Figure 8).

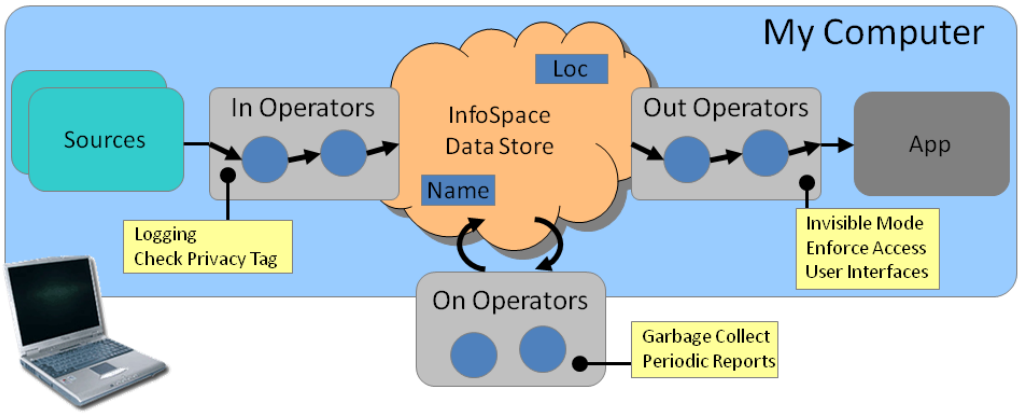



**Figure 8: Confab Architecture. Personal Data is Stored in InfoSpaces That Contain Attribute Tuples. Data Flows are Governed by In-, Out-, and On-Filters** [25]

Privacy tags are similar to the general "sticky policy" concept presented in section 3.2, as they represent metadata that can be used to enforce[20] privacy compliant usage and retention. However, Confab's privacy tags are more custom tailored to the exchange of dynamic context data, featuring elements that declare how many "sightings" the other party may amass of a particular attribute (e.g., only retain the last five locations a person was in) and a "garbage collect" declaration that can contain data-deletion triggers (e.g., deleting information if the data subject leaves a particular area). In order to provide plausible deniability, information that is deemed too sensitive to be released will simply be marked by out-filters as "unknown," making it indistinguishable from technical failures or lack of connectivity.

While Confab provides a framework for disseminating context information that data subjects collect themselves, the PawS system [**26**] focuses on third party data collection instead. PawS addresses smart environments that can communicate and enforce data collection practices for various optional and mandatory data collections. Based on a Hippocratic Database[21] in the back ("*privacy DB*"), it uses *privacy proxies* to control data collection and access for smart devices such as cameras or printers. *Privacy beacons* advertise the URLs of these proxies so that user devices can download information about a smart environment and then directly configure the available services as needed. Figure 9 shows an overview of the architecture, using the example of a smart room equipped with a camera and a publicly accessible printer.

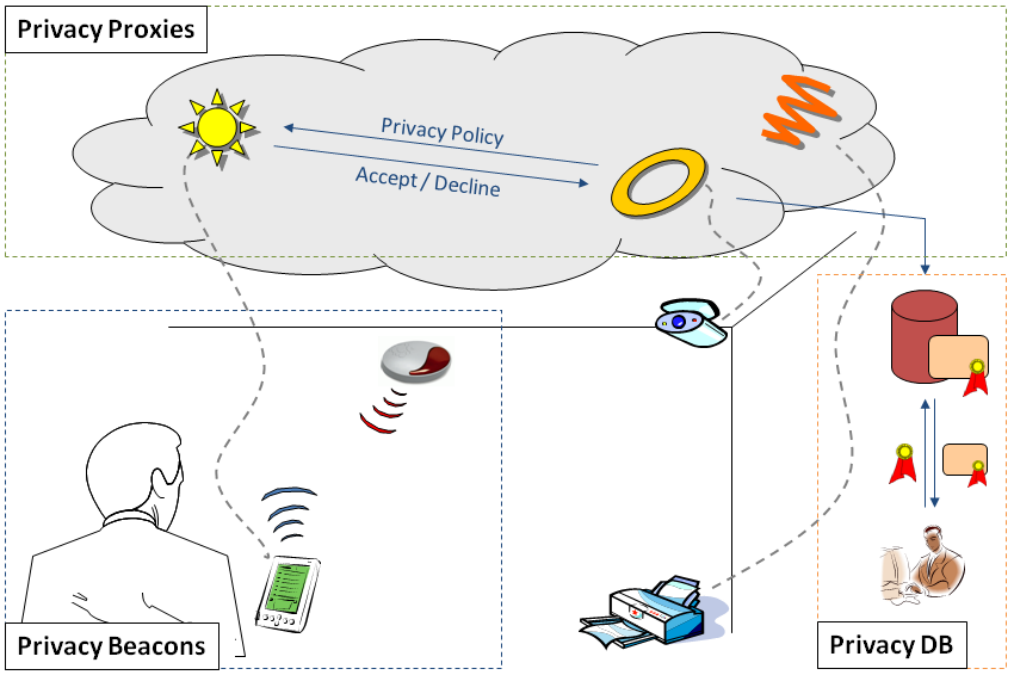



**Figure 9: The PawS System uses Privacy Beacons to Advertise Data Collections Taking Place. Data Subjects can use Privacy Proxies to Configure and Control Available Devices. A PrivacyDB Stores Collected Data Using Sticky Policies [26].**


As the concept of privacy beacons illustrate, it is difficult to convey the act of a data collection taking place in smart environments, which should operate unobtrusively in the background. A handheld user device is needed in PawS to report the current "smartness" of a room to a data subject, and to give the data subject control over the respective data flows and services. However, PawS does not address how to best present such information and controls to the

[20] Enforcement of course relies on the use of InfoSpaces throughout the entire data life cycle.
[21] Cf. transparency tools in section 3.2

user, which is further complicated through the use of small screens on mobile devices. What kind of information is important to users in such smart environments? What aspects of such a user interface are the most critical?

One of the earliest projects that tried to answer such questions was undertaken by Bellotti and Sellen at Xerox's EuroPARC in the early 1990s [**27**]. As part of an audio-video presence and collaboration environment called RAVE, cameras, monitors, microphones, and speakers were deployed in offices in order to allow EuroPARC's staff to glance into other offices (i.e., get a few seconds of video-only transmission), make *v-phone calls* using both audio and video, or install a longer lasting office-share (i.e., a semi-permanent v-phone call). Hong and Landay's Confab toolkit also addresses user interface issues. Confab not only attempts to deliver simple and appropriate controls, as well as clear and timely feedback of the current data collection status, but also incorporates the principle of *plausible deniability* into its design.

A sample application built with Confab was the "Lemming Location-Enhanced Instant Messenger", which allowed users to automatically and semi-automatically publish their current location to their contacts. Figure 10 shows an example prompt. The large "1" illustrates a one-off query and contrasts it with queries that seek to continuously track the user's movements. If the user ignores the request, the Confab toolkit returns his or her location as "unknown" to the requestor, which does not allow one to differentiate between network errors, system errors, and user decisions. Users can thus plausibly deny that they ever got the request.

Common to all these examples is trust users need to have in the smart infrastructure surrounding them. Once personal data leaves the protection that these spaces offer, no guarantees can be made. Equally important is the challenge of properly configuring these spaces, so that the infrastructure does what data subjects want them to do.

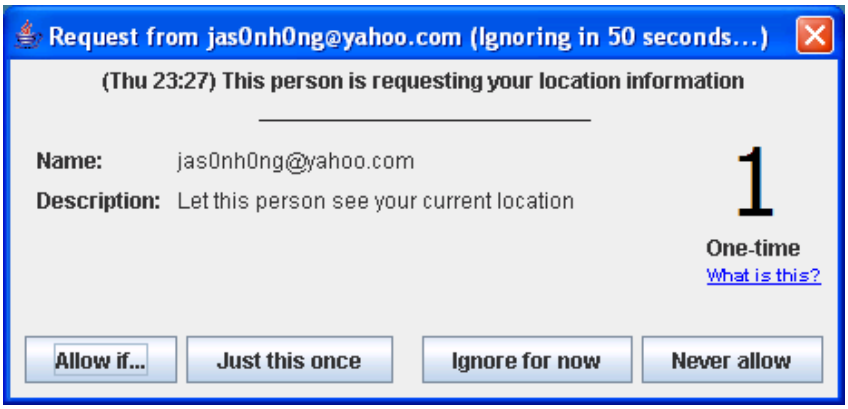



**Figure 10: A Location Prompt in the Lemming Location-Enhanced Instant Messenger. If the User Ignores the Request, Confab Reports Back an "Unknown" Location, Thus Facilitating the Principle of "Plausible Deniability" [25].**


## 3.4 EXAMPLE: PROTECTING RFID TAGS

Radio Frequency Identification (RFID) Tags represent probably the most prominent ubicomp technology, at least when it comes to privacy issues. Their increasing use, especially on consumer products in showcase supermarkets, as implants for entering "in"-clubs, and as a security feature of "electronic passports", have prompted widespread public concern over RFID privacy issues.[22] The privacy challenges of RFID tags are fourfold:

[22] There is a bewildering variety of technologies that are often lumped together under the umbrella term "RFID". This section will focus on *passive* RFID tags, i.e., those that do not come with their own power source (battery) but instead receive the energy to operate from the reader's field. For the definite source on RFID technology, see [**59**]

1. Automation: Reading an RFID tag typically does not require the help of the person carrying the tag, nor any manual intervention on behalf of the reader. Thus, simple reader gates can easily scan large numbers of tags, making data acquisition much easier.
2. Identification: The ability to identify individual items instead of only whole classes of items significantly improves the ability to identify an individual. This would facilitate, e.g., the creation of detailed consumer or citizen profiles.
3. Integration: Not only that the act of reading a tag can be completely hidden from the tag carrier (especially when operating at larger distances), also the fact that a tag is present in a particular product will be hard to ascertain for an individual without special detection equipment.
4. Authentication: The above points become especially critical given the increasing amount of sensitive information, e.g., health information, payment details, or biometric data that are stored on or linked to tags used in authentication systems.

These four attributes of RFID applications threaten two classes of individual privacy: data privacy and location privacy. The location privacy of a person is threatened if a tag ID that is associated with that person is spotted at a particular reader location. For example, by knowing that a person's car has passed a certain toll station, or that a person's shoes have entered a particular building, others might be able to infer (though not prove) the location and ultimately the activity of that person. These IDs do not need to be unique – certain combinations of non-unique tags might still form unique constellations of items that can be used to identify an individual (Weis, 2003).

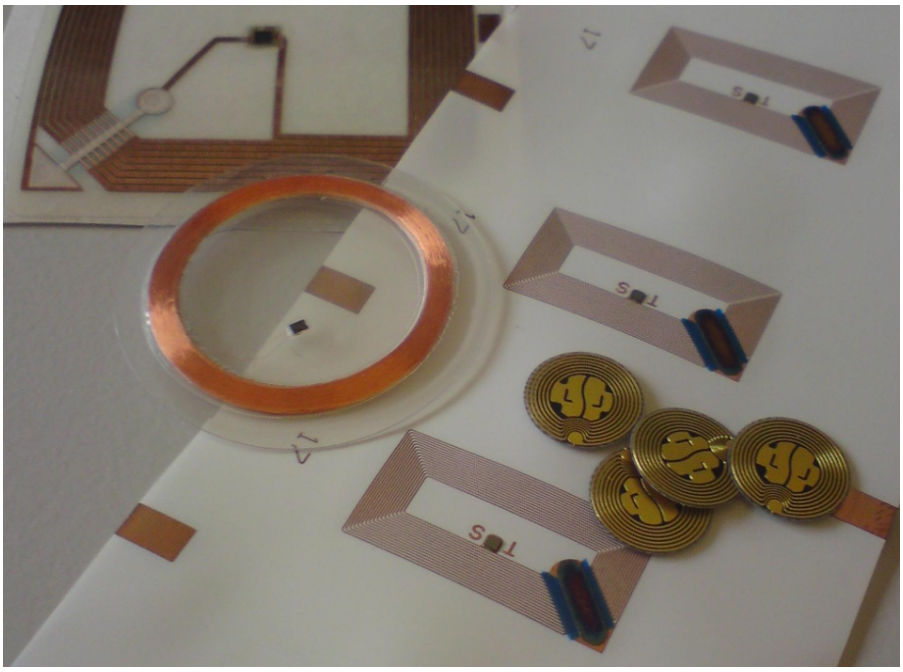


**Figure 11: Set of Example RFID Tags (all Passive)**

Once tags carry more than just an identifier, but also a person's name or account number, data privacy may be violated. This happens if unauthorized readers eavesdrop on a legitimate transaction, or if rogue readers trick a tag into disclosing its personal data. These types of attacks are typically called *skimming*. A special case of data privacy are product IDs that disclose details of the (otherwise not visible) belongings of a person, e.g., the types and brands of clothing one is wearing, the items in one's shopping bag, or even the furniture in a house. Note that in the latter case, the actual identity of the victim might very well remain unknown – it might be enough to know that this person carries a certain item.

There are three principal ways of violating an individual's data and/or location privacy: clandestine scanning, eavesdropping, and data leakage:

- Clandestine Scanning: The tag data is scanned without the tag-carrier's consent. This might disclose personal information (data privacy) either indirectly, e.g., by revealing the contents of bags that one cannot see through otherwise, or directly, e.g., by revealing personal data such as the name of a user or the date that a particular item was bought. If several clandestine scans are pooled, clandestine tracking can reveal a data subject's movements along a tag reading infrastructure (location privacy).
- Eavesdropping: Instead of reading out a tag directly, one can also eavesdrop on the reader-to-tag channel (or even the tag-to-reader channel) and receive the IDs of the tags being read due to the employed anti-collision protocol (cf. section 3.4.1 below).
- Data Leakage: Independent of the actual RFID technology is the threat of having applications read out more information from a tag than necessary, or storing more information than needed. This is of course a threat common to all data gathering applications, though the envisaged ubiquity of RFID-based transactions renders it highly relevant in this context.

Security is obviously the primary issue here, i.e., ensuring the confidentiality of the information stored on such tags, so that only authorized parties are able to detect, read, and potentially write to such tags. Ensuring the confidentiality of information, both through the encryption of transmissions and the use of authentication mechanisms to limit access to it, is an old problem and many solutions exist that make it virtually impossible for an attacker to succeed. However, RFID tags add two novel challenges to the problem that render many existing approaches infeasible:

- Limited resources: Just as other small devices (e.g., sensor nodes) in ubicomp, RFID tags are extremely limited in their computational capabilities. Implementing advanced encryption algorithms such as AES or public key systems has so far been only feasible for high-end chips that are specifically engineered for security applications. Cheap mass-market tags that are envisioned to be embedded in virtually all products or product packages cannot currently be protected this way.
- Key selection: In order to authenticate an authorized party against an RFID tag, some shared secret must exist between the two.[23] This in itself is no different from securing, say, a computer with a password. However, with potentially hundreds or thousands of tags, knowing which particular secret to use becomes a problem. This is difficult because one typically does not know which tag one is interacting with – after all, being able to identify an RFID tag is exactly what one wants to prevent.

The following sections will present a number of approaches that have been developed to enhance the privacy of RFID.

### 3.4.1 COMMUNICATION CONFIDENTIALITY AND ANTI-COLLISION PROTOCOLS

Encrypting the wireless communication channel between an RFID tag and a reader is the obvious solution to eavesdropping attacks. A large body of work in the crypto community specifically targets low-power and low-complexity ciphers, and RFID chips used in ePassports or contactless train passes typically employ some sort of communication encryption. As pointed out above, however, key management is still critical and thus this approach is only usable for selected applications.

[23] A shared key assumes the use of symmetric encryption. While public key cryptography does not require a shared key, the general principle remains, as one still needs to select the right public or private key.

However, even if tag communication is encrypted, the nature of RFID communication might still allow an attacker to track the tag. This is because RFID readers first need to *singularize* a tag before they can read from it, i.e., it must determine which tags are present in its range and then address exactly one of them with a command.[24] To this end, many RFID protocols use a unique ID (UID), similar to the hardware MAC address of a WiFi card or a Bluetooth adapter, to aid tag singularization at a lower protocol level. If this UID is fixed, rogue readers can simply perform a round of singularization and thus be able to track individual tags, even if the tag payload (e.g., the name of a product and its product ID) is encrypted. This process is called an *anti-collision* protocol.

Some RFID anti-collision protocols offer variants where the UID is randomized upon every reader session, i.e., whenever a tag enters the field of a reader, it chooses this UID anew and uses it only until it leaves the field again. This of course drives up tag prices, as the each tag needs to include a pseudo random number generator (PRNG) for this. However, if this PRNG is well implemented, tracking tags by their UID is much harder, if not impossible.

A particular attribute of RFID communication is the power asymmetry of the wireless link: while tag replies can only be read from up to a few meters or even centimeters, reader commands can typically be received from up to 100 meters away. This is because the battery-less, passive tags are receiving their energy supply through the reader radio field, which thus needs to be very powerful. If any of the reader commands contain the ID (or UID) of an identified tag, an attacker could obtain this information from a very large distance, even though the defined read-out range of the tags is only several centimeters.

While this is not much of a problem if random UIDs are used, many high-performance RFID systems use fixed UIDs in their anti-collision protocols in order to lower read-out times. In these so-called *deterministic* anti-collision protocols, readers probe for conflicting UIDs by systematically querying for all possible prefixes, i.e., “all tags beginning with 0”, “all tags beginning with 1”, “all tags beginning with 00”, “all tags beginning with 01”, and so on. If the UIDs of two or more tags begin with queried prefix, all of them will reply at the same time, causing a collision that the reader is able to detect. If this happens, the reader simply increases the length of the prefix by one digit (e.g., by adding a “1” to it) and tries again, until only a single tag replies. After it successfully singularized a single tag this way, it will first instruct this particular tag to be silent in the upcoming rounds, then backtrack and replace the last bit that it added with its inverse (e.g., if it added a “1” before it will now use a “0” instead) and continue. Should more collisions occur, it again increases the length of the prefix until it can singularize a tag, then backtrack. Such protocols are called *binary tree walking* protocols, as this behavior can be seen as traversing a binary tree, in which the individual bit positions are the branches and the tags are the leaves (see Figure 12 for an example).

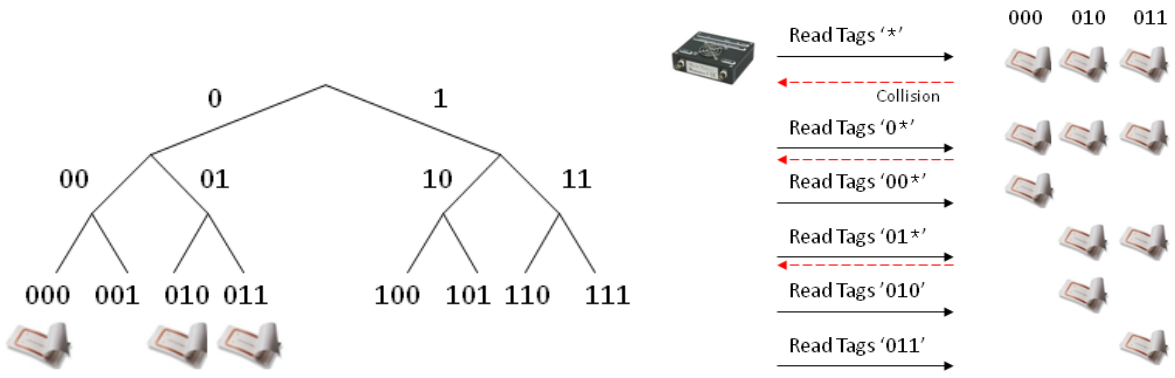



**Figure 12: Binary Tree Walking Protocol With Three Tags Present. Note How the Reader Commands Transmit the Tag IDs in the Clear**

[24] The need for singularization stems from the fact that RFID tags are unable to detect the presence of other tags. Thus it is up to the reader to detect collisions between tags.

An attacker can overhear the prefixes sent from a reader to its tags and thus infer many partial UIDs of the tags present. A simple fix for this is to have the reader traverse the tree one bit at a time, i.e., the reader never explicitly sends any prefixes, but only uses the command "transmit next bit". Tags in turn only reply with the *n*-th bit of their UID, not with their full UID. As long as the corresponding bit positions of all available tags are identical, no collision occurs and the reader is able to clearly receive the common bit prefix, building it up incrementally. Once two or more tags differ at a particular position, the reader detects a collision and now has to branch into one of the two subtrees. Instead of explicitly naming the subtree that it wants to explore (i.e., sending "0" or "1"), the reader XORs its selection with the previous, error-free bit in the prefix. As the value of this bit was only sent from the tags to the reader, an attacker outside the tag's communication range (but inside the reader's forward channel) will not be able to know the true value of the next selected bit. The tags, on the other hand, know their own ID, and accordingly the bit value at the previously queried position, thus sharing a common secret with the reader that can be exploited for every conflicting bit position.

### 3.4.2 ACCESS CONTROL/TAG DEACTIVATION

A simple solution to access control is to obstruct the reader signal by means of a metal mesh or foil that encloses the tag. With the inclusion of RFID tags into passports, a number of vendors now offer coated sleeves for protecting the passport while not in use (Figure 13). Other options would be aluminum-lined shopping bags for groceries. However, tagged clothing or personal items can often not be protected in this extreme way.


**Figure 13: Low-tech privacy solution – an electromagnetically shielded sleeve for ePassports (Paraben '07)**


For tree-based anti-collision protocols, researchers have proposed a specially engineered "blocker-tag" to jam readers that attempt to traverse the UID-tree. Such a blocker-tag uses two antennas to simply send out both a "0" and a "1" for each reader query, thus creating a collision at each branch of the tree. This creates the impression of trillions[25] of tags being present and will in effect fully stall even the fastest reader. The difficulty of this approach lies in controlling the range of tags that are being blocked. Otherwise a single blocker-tag could accidentally disable any (tree-based) RFID system it comes close to.

Alternatively, some form of password could be set on each tag, only allowing readers that use the right password to read out any data from the tag. This could be realized in a very simple manner using so-called *hash-locks.* An RFID

[25] Fully simulating all possibilities of a, say, 64-bit ID would be actually more than just a few trillions. An (implausibly) fast reader able to read 100'000 tags per second would be busy for over four billion years reading all $2^{64}$ tags

tag is by default open and readable, but it can be locked by sending it a lock-command with a password. The tag then creates a hash of this password (which could simply be its true ID number) and hitherto only replies to a query if an unlock-command is sent that, when hashed, matches the stored hash-value.

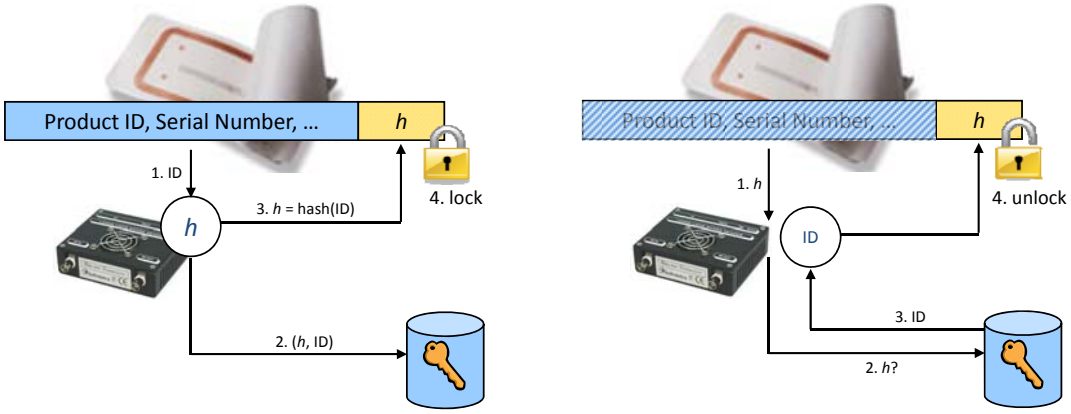



**Figure 14: Example of an ID-based Hash-Lock**

As elegant and simple the solution is, it squarely falls into the *key selection* issue mentioned above: how would a reader (or user) know which password to use with which tag? While individual passwords work well for single accounts or special items like an RFID-based health card, it seems unlikely that one would be able to remember hundreds of passwords for the individual groceries one buys each day. If users have only a few tagged items, fast readers might simply try out all possible passwords until one of them eventually unlocks the tag. Clearly, such an approach does not scale well.

One option might be to have locked tags always reply with the hashed password, but nothing else. This would still protect the tag's contents without disclosing the real password, while allowing authorized readers to lookup the right password in a hash-to-password lookup-table. However, such an approach might still protect the data privacy of the tag owner but fail miserably at protecting the location privacy, as static identifiers can trivially be tracked. A number of researchers are thus investigating the use of *hash-chains*, where a locked tag still replies with its stored hash but continually re-hashes it after each reply [**28**]. This results in constantly changing hashes, which makes tracking much more difficult (though not impossible). Authorized readers that know the right password can similarly follow such a chain and thus can properly predict the password to use. The challenge of hash-chain schemes lies in synchronizing the two, even in the presence of an attacker who might try to desynchronize them. Another important aspect of such schemes is their *forward-security*, i.e., if an attacker learns the current key of a tag, she is nonetheless unable to identify previous outputs of the tag (e.g., in log files).

Whether hash-chains are used or not: the problem of managing various passwords for a plethora of tagged items remains the biggest challenge in any access control scheme for RFID. In addition to the sheer number of items that need protection, it is also the number of ownership changes that challenge traditional approaches: An item is shipped from a supplier to a distributor to a retailer, who sells it to a consumer, who might later give it to a friend or even resell it. At each step along the way, the current password must either be passed along or some means for setting a new password must be given. This problem of *ownership transfer* has been the focus of several approaches, i.e., allowing either the temporary delegation or full transfer of ownership to a product, without allowing the previous owner to know the new key or the new owner to know the old key [**29**].

### 3.4.3 PROXIES

A common response to many of the problems detailed above has been the proposition of a powerful *proxy device* that would locally manage one's tagged items. By incorporating an RFID reader into a commodity consumer device like a mobile phone or a wrist watch, this device could act as a proxy for all tag interaction: individually blocking or allowing access to its owner's tagged items (i.e., acting like a smart blocker-tag), setting individual passwords on items bought in the supermarket or resetting those of items given to others. Other proposed features are logging and alerting functions that would require readers to identify themselves and offer links to machine-readable privacy policies that could be the basis for automatically regulating access to the owner's tags.


**Figure 15: Privacy Guardian Reader Hardware and User Interface**
**© 2006 M. Rieback. Images reprinted with permission from http://www.rfidguardian.org**

While such devices have been proposed several times in the literature, few actual implementations exist. One such example is the *Privacy Guardian* at the Free University of Amsterdam [**30**]. A portable RFID reader in a 10x20cm box is paired via Bluetooth with a control application running on a mobile phone (see Figure 15, left to right). Users can scan for tags in the vicinity, control access to them, and log readout attempts (and successes) of other readers. The privacy guardian works similar to a blocker tag, jamming the channel whenever a third party reader attempts to access a protected tag (see Figure 16). This only works, however, for a particular set of RFID protocols in which tag replies are deterministic, i.e., where their UID determines when they will reply to a reader inquiry.

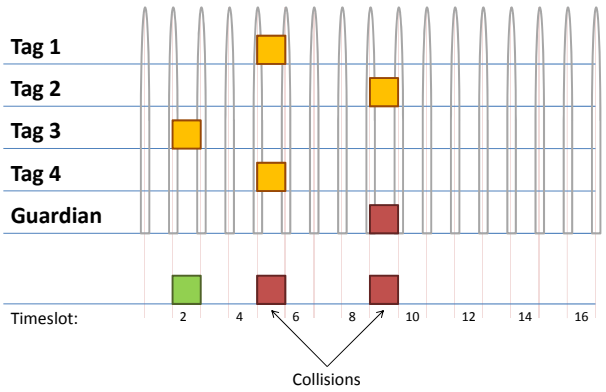



**Figure 16: Privacy Guardian Blocking Access to Tag No. 2 [30]**
**© 2006 M. Rieback. Image reprinted with permission.**

As an alternative to control approaches based on powerful consumer proxy devices (which might imply a significantly management overhead for individuals), several researchers have advocated password-less access

control approaches. One of the earliest ideas was to have tags measure the signal strength of a reader signal and reply with different level of detail, depending on the inferred distance of the reader [**31**]. In order to read out the highest level of detail, an attacker would need to come very close to a target and thus become noticed, making stealth skimming attacks very difficult. The two main problems of this approach are the difficulty of controlling one's tags this way ("how close is enough for a level 3 disclosure, without reaching the too detailed level 2?") and the costs and reliability of an integrated signal strength module on an RFID tag (an attacker could use a very high powered reader to simulate closeness).

An alternative, password-less solution wraps tag data into several encryption layers that require continuous read access for significant amounts of time, thus slowing down an attacker [**32**]. Together with small antennas that restrict tag read ranges to several centimeters, an attacker would not only need to come close but also stay there for, say, a few minutes. Based on Adi Shamir's theory of shared secrets, the tag's real ID is encoded into several pieces ("shares") and these get stored on the tag instead. Due to the properties of Shamir's theory, the original tag ID can only be reconstructed if all of those pieces are known. While all shares are stored on the same tag, readout is complicated by allowing only a random trickle of bits from the tag. Together with a short read range, this requires an attacker to spend a considerable amount of time in close proximity to the "target", making quick unnoticed readouts difficult. At the same time, however, legitimate owners are able to use simple caching strategies to identify their items instantaneously, as an initial burst of disclosed bits is enough to probabilistically identify a tag from a known set. In order to prevent the repeated querying of such a larger initial subset, which would give an attacker faster access to the entire key, tags use random temporary IDs for tag singularization, thus making it more difficult for an attacker to correlate two such bit-strings across consecutive queries.

## 3.5 EXAMPLE: PROTECTING LOCATION INFORMATION

Knowing when a particular person was at a particular point in time is at first no different from knowing other facts about a person, e.g., her age, place of birth, or favorite hobby. However, location information is typically associated with a particular place (e.g., "home" or "pawn shop"), which in turn often implies an activity (e.g., "visiting family" or "selling items") or a certain personal interest (e.g., "betting"). As such, location information gives rise to a large number of implications, with more or less correlations. Moreover, knowing the *trajectory* of a person's movements often allows to corroborate such implications through connections (e.g., "he first went to a gun dealer and then to the victim's home"). Knowing a person's current or favorite location can moreover threaten her personal well-being (e.g, stalker attacks) or make her more susceptible to disturbances (e.g., spam).

Location information is thus in many cases a particularly sensitive piece of personal data, even though people might not realize it (cf. survey results reported in section 2.2 above). In particular, the problem of location privacy can be framed as the problem of separating three distinct pieces of information: Who? Where? And When? Knowing only one or at most two of these pieces typically lowers the value of this data significantly (but not in all cases). For example, knowing that someone entered a supermarket at 2pm might not carry much value, but knowing that someone entered a particular bedroom at 3am might tell parents when their son came home on Friday night.

Many technical systems for providing location privacy thus attempt to disassociate such time-identity-location-tupels. One of the first location systems, the Active Badge System at Xerox's Palo Alto Research Center [**33**], already separated location tracks from users with the help of two separate architectural entities: *User Agents* and *Location Query Services*. Active Badges emitted a *pseudonymous* ID while being tracked by the location infrastructure. The badge sighting for each place is managed by one Location Query Service, which would register all sightings of these pseudonyms without knowing the user's identity. The true identity behind such an ID was only known to the user's personal User Agent, which would explicitly register an interest in a particular pseudonym with each Location

Query Service. In order to query a user's identity, one would need to contact the user's User Agent, which could then handle the appropriate access control verification.

The location privacy of Active Badge users thus rested on the integrity of the location infrastructure: As long as an attacker does not have access to the full log files, the only way to find out about someone's present and past whereabouts was through his or her personal User Agent. However, once the system is compromised, all static pseudonyms can be trivially associated with a user, instantly providing access to all past and future location traces. Even if an attacker gains only access to the pseudonymous logs, without being able to find the matching User Agents to resolve the pseudonyms, it might be trivial to infer the true identity behind each pseudonym, as movement patterns often allow the inference of one's home or office (see Figure 17).

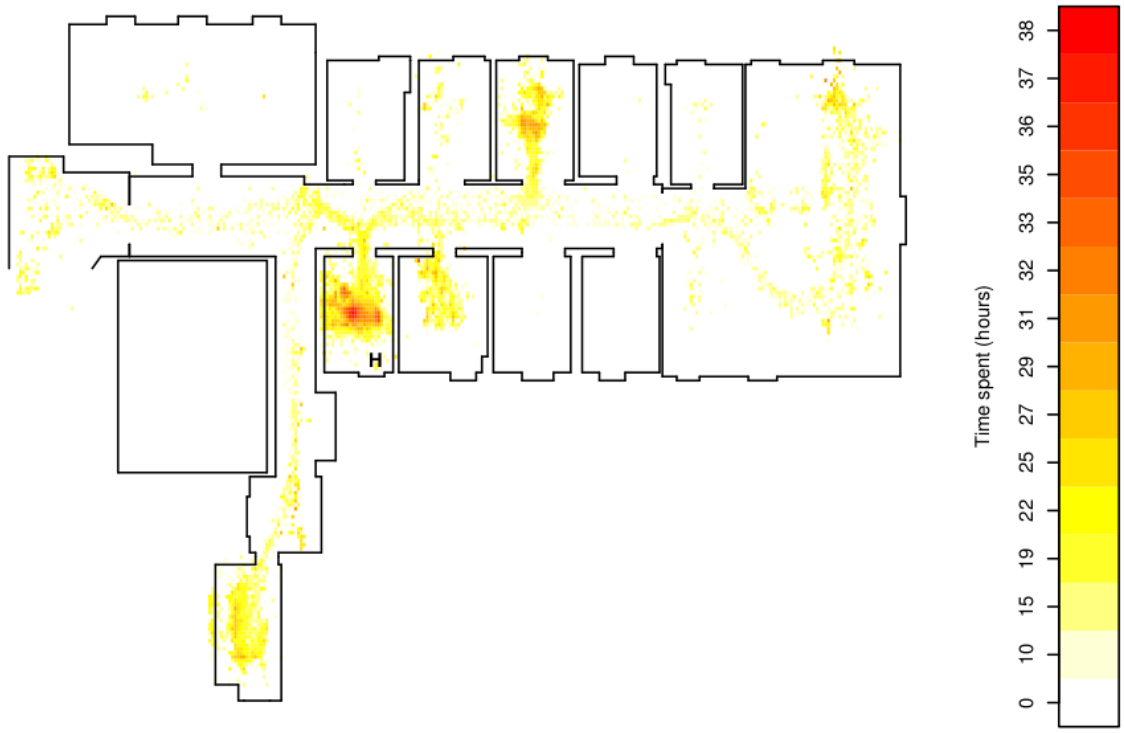
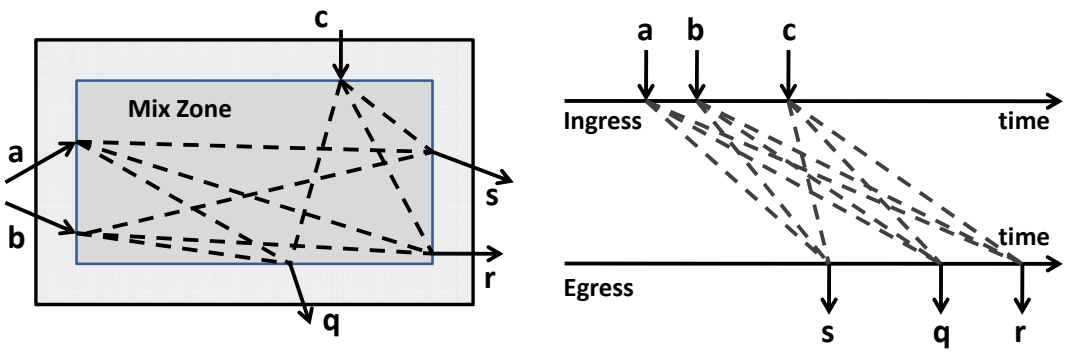



**Figure 17: Static Pseudonyms Can Often be Trivially Resolved if Additional Information is Available. For Example, the Above Location Trace is Most Likely From the Owner of Office "H" [34]**
**© 2004 A. Beresford. Image reprinted with permission.**

Consequently, in order to truly decouple identity from location-time tuples, pseudonyms must frequently change. This change cannot be simply changing one random number into another, as the two location tracks from these pseudonyms would be trivial to join again later. Instead, switching pseudonyms has to be done in a way that prevents the location track of the new pseudonym to be associated with the previous one. Beresford and Stajano [**35**] propose so-called *mix-zones* for this – areas in which no location tracking takes places and which are large enough so that at any point in time, a large enough number of targets are present that can be "mixed". Figure 18 shows an example of three users entering a mix-zone, using pseudonyms *a*, *b*, and *c*. With location tracking disabled inside the mix-zone, it is difficult to infer who *q*, *r*, and *s* are. The challenge of this approach is the proper definition of mix-zones (would people mind those "dead" zones where no tracking is available?), as well as adjusting them dynamically to the actual traffic (late night traffic might require much larger zones in order to enclose enough users to mix). Algorithms that attempt to use the crossing of two or more location tracks in order

to increase the chances that an attacker confuses the path of different users are also called “path-perturbation” algorithms.



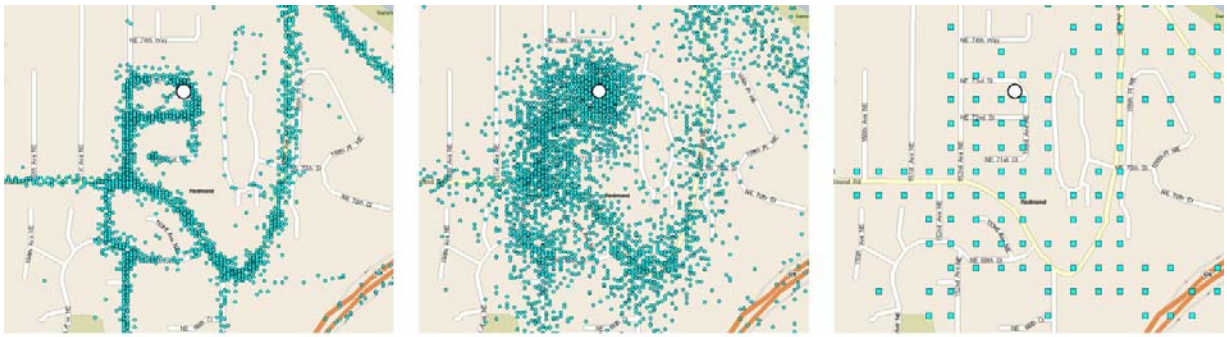

(a) Plan view of the mix zone (b) Timeline of movements

**Figure 18: Three Users Enter a Mix-Zone, Three Different Ones Exit it Some Time Later. Who Went Where?** [36]


One obvious solution to location privacy seems to be the use of *self-positioning systems*, such as GPS, Cricket [**37**], or BlueStar [**38**], where the infrastructure sends out positioning information that clients can use to compute their own location. While this obviously alleviates concerns that arise with a positioning infrastructure that tracks its clients, it might still disclose this information once a location-based service is used. For example, a GPS-enabled mobile phone that uses an online mapping tool to visualize its location obviously discloses its location to the online service.

Instead of trying to separate identity from location and time, one can also try to *obfuscate* them, i.e., lower their precision or accuracy. For location information, this could mean either to widen the area in which the subject is located in (e.g., from exact coordinates to street level to city level) or to offset the reported position by some (seemingly random) value from the true position (see Figure 19). A related technique is to create additional queries from “dummy positions” and to hide the true position in there. Identity information can similarly be obfuscated, meaning that one enlarges the pool of possible identities, e.g., making it impossible for an attacker to distinguish between several potential senders of a location-based request. As even anonymous location information might reveal an identity, given additional information about one’s home or office, this process of identity obfuscation thus necessarily implies location obfuscation as well.

It is important to realize how obfuscated location systems must be used in practice. One obvious option is the use of stored location tracks that are only of statistical or historical significance. This might prevent attackers from using such logs to their advantage, while still allowing, e.g., statistical inference over many such tracks. Their online use is more problematic, as degrading the accuracy or precision of position information when using a location-based service might directly affect the quality of the service, or increase the cost of its use (e.g., sending 19 false queries in order to hide one’s true location might incur 20 times as much service charge).



(a) Original GPS data (b) Additive Gaussian noise (c) Discretized to points on grid

**Figure 19: Obfuscation Can Simply Mean Adding Some Noise to the True Position, or to Discretize it [39]**


An often used measure for the strength of such an approach is using the above-mentioned concept of *k*-anonymity within the context of a location system: The location infrastructure pools location requests from at least *k* users and changes all queries so that the location information given by each user is sufficiently imprecise to make them indistinguishable from the *k-1* other users. This pooling can also be done by the clients themselves, by using ad-hoc communication and routing location requests among several peers first – similar to a mix network (see section 3.2) – until an "exit" node then actually sends off the request (using a properly adjusted location area that also comprises the original node).

[**40**] propose an access control system to stored location information that takes the past location information of the inquirer into account: queries are only executed on data that the inquirer could have seen herself at the time, i.e., the returned data is centered around a small area of the inquirer's own location trace. This can still offer useful services, e.g., for finding lost objects of the owner ("Where did I last see this item?") or to refresh one's own memory ("Who did I meet this morning in the hallway that I wanted to send email to later?"). Obviously, users need to trust the infrastructure to properly enforce this concept of *physical access control*. Thus this solution is much more about how to give users control over who can access their data – in this case a default rule would limit access to people who were co-located.

**Conclusions**

1. Ubicomp challenges privacy due to its novel technical capabilities, large coverage (both in time and space), novel types of data collected, vague collection purposes, and envisioned data interchange. While standard technological fixes exists (PETs), they do not fully cover the many peculiarities of ubicomp applications.
2. Research in privacy for smart environments attempts to uncover and control information flows between data subjects and system operators, and between users of the system. "Sticky policies" allow for the propagation of use policies, but defining and adjusting such policies is a challenge for users. Smart environments require novel user interfaces to allow for the inspection and control of their information flows.
3. RFID applications envision highly automated environments that allow for the unobtrusive detection of tagged artifacts, and implicitly the people carrying those. Due to the high number of items and the lack of user interface, offering users to control who can read what tags is difficult. The low resources available on passive RFID tags additionally challenge the use of traditional security protocols.
4. The data created by location-based services can either be anonymized or obfuscated in order to protect the identity of the data subject, yet with the help of some simple background information, attackers can circumvent much of this. These techniques also affect service quality, as the fidelity of information is often

significantly decreased. Even if the localization service is trusted, there still remains the problem of allowing users to simply control location disclosure to others.

**Further Reading**

- John Krumm: A Survey of Computational Location Privacy. *Personal and Ubiquitous Computing*, Vol. 13, 2009. Special Issue on Privacy in Ubiquitous Computing (forthcoming). A prior workshop submission with the same title is available from www.vs.inf.ethz.ch/events/uc07privacy/program.html.
  Provides a concise overview of the literature and serves as an excellent starting point.
- Matt Duckham and Lars Kulik: A Formal Model of Obfuscation and Negotiation for Location Privacy. In: Pervasive Computing. Third Intl. Conference, PERVASIVE 2005, Munich, Germany, May 8-13, 2005. LNCS 3468, Springer, Berlin Heidelberg New York, pp. 152-170, 2005.
  Duckham and Kulik define a formal model for obfuscation and outline methods for using obfuscation in practice.
- Alastair R. Beresford: Location Privacy in Ubiquitous Computing. PhD dissertation, published as technical report UCAM-CL-TR-612, University of Cambridge, Cambridge, UK, January 2005.
  A good overview of the threats and issues in location privacy, and a detailed discussion of the concept of mix-zones.

# 4 HOW TO ADDRESS PRIVACY IN YOUR UBICOMP WORK

What is there to do if you want to accommodate privacy in your ubicomp system? How do you have to design your ubicomp application to make sure it respects and supports the privacy of its users? As we said in the introduction, there is no simple answer, no set of algorithms or routines that, when "applied", will 'fix' the privacy issue once and for all.

What the preceding sections hopefully illustrated is the *scope* of the privacy problem (technical, legal, social), and that simply having "a good firewall" or implementing "strong 128-bit security" are not enough. Instead, one needs to carefully analyze each and every ubicomp application: What is it supposed to do, and how does it do it? How are users using it in their daily routines? And what technical and organizational tools are available to support privacy and security under these circumstances? The three steps described below try to illustrate the type of question you should be asking, in order to increase the chances of "getting it right".

## 4.1 UNDERSTAND YOUR APPLICATION (CONSIDER USERS AND USE)

If there is no single answer, no single set of "todo's" that will ensure the privacy of our users, then each solution must be more or less unique, depending on the particular ubicomp system and application one wants to build. Consequently, you will first need to understand how this application (or the potential applications that are being built on top of this system) is supposed to work: What problem does it try to solve? How do people currently "solve" or work around this problem? How are people expected to use your application? And maybe also: how are people *actually* using your system?

These questions are by no means specific to privacy, but simply good interaction design practices, as summarized, e.g., by Reinmann [**41**]:

> Interaction Design is a design discipline dedicated to:
> * Defining the behavior of artifacts, environments, and systems (i.e., products)
> …and therefore concerned with:
> * Defining the form of products as they relate to their behavior and use
> * Anticipating how the use of products will mediate human relationships and affect human understanding
> * Exploring the dialogue between products, people, and contexts

Understanding users and their use of your application is critical for assessing the privacy implications [**42**]. As we have seen in section 2 above, privacy issues do not happen in a vacuum: privacy is not a monotonically behaving function (i.e., the more, the better), so one will need to understand when someone needs what access and what kind of control to what kind of data. Simply providing anonymity or access control does not create "privacy".

[**43**] illustrate this point nicely in a field study involving a location-aware mobile phone application called “Reno”. Participants could use Reno to automatically or manually share their current location with their friends or family. Users that shared their location information were mostly spouses or close friends, so a high level of trust existed and thus the willingness to share one’s current location was high. However, users still reported incidents of perceived privacy violations, e.g., when an automatic “trigger” that they configured suddenly notified others of their location, but in a context that they did not anticipate:

> Participant g: “My phone disclosed my location to [participant a] last night when I was out running an errand and was returning to my house ([a] has a trigger for my house). I’m now reconsidering that trigger because I felt weird about that one. I didn’t feel weird when he got notified I was “home from work” but did about the late-night errand running.”

Battya Friedman et al. additionally propose to explicitly incorporate *values* into the design of information system, a process they call *value sensitive design* [**44**] [**44**]. Value sensitive design incorporates conceptual, empirical, and technical investigations in order to identify both direct and indirect stakeholders, uncover their values and beliefs, describe the social context in which technology is used, and uncover the hidden values that are inherent in a particular technological solution. An example on how to apply this methodology in the context of ubicomp can be found in Freier et al. [**45**] [**45**] or in the works of Camp and Connelly, e.g., [**46**].

## 4.2 DEFINE THE PROBLEM (THINK ATTACKER MODEL IN SECURITY)

Once you understand how your application is supposed to work, how you expect your users to interact with it, and what assumptions and values are coded into the system design, you can start defining what needs protection, i.e., what kind of privacy you want to provide to your users. If you do not know what you are trying to protect, and what you want to protect it against, how can you know when you are successful? Defining privacy is a “fractal problem” and if you don't set out clear limits, you'll find yourself endlessly chasing “but, what if...?” situations.

In computer security, these kinds of questions have a long tradition, running under the term of “threat model” or “attacker model”:[26]

> The first rule of security analysis is this: understand your threat model. Experience teaches that if you don't have a clear threat model – a clear idea of what you are trying to prevent and what technical capabilities your adversaries have – then you won't be able to think analytically about how to proceed. The threat model is the starting point of any security analysis. [**47**]

As we described in section 2, privacy is not only about an active “adversary” that is “threatening” your privacy (though this of course also happens). Privacy is also about opportunistic data use (i.e., data that has been given for one purpose is “recycled” for another) or involuntary disclosures (e.g., someone who is entitled to receive information *in general* should not have gotten this information in a *particular*, unexpected situation). Consequently, we will need to cast a slightly broader net and not only think about *attacks* on our data, but on *opportunities* for unwanted disclosure, based on actual and potential information flows (cf. Figure 4 in section 2.1)

[26] There is actually a subtle but important difference between the two terms, but for our purposes this does not really matter. See, e.g., [**58**] for a discussion.

within and between applications. [**48**] consequently prefer to talk about *privacy risk models* instead of privacy *threat* models.

Hong et al. propose to analyze personal information flows in an application from two viewpoints: the *social and organizational* context, and the *technological* context. The social and organizational context looks at application stakeholders and their use of the system (see section 4.1 above), but with a particular focus on the actual information flow, e.g., "Who are the *data sharers*, who are the *data observers*?" or "What kind of data is shared and what is the value proposition for sharing it?" The technological context then examines the mechanisms that collect information (push vs. pull; one-time vs. continuous; granularity, accuracy, precision), the means of solicitation (opt-in or opt-out; automatic or manual), and how storage and access is handled (retention, encryption, access control). This analysis then forms the basis for privacy risk management, where identified privacy risks are prioritized and translated into design guidelines. Prioritization is based on a cost-benefit analysis that factors the likelihood for an unwanted disclosure, the damage resulting from this, and the cost of protection.

It is important to note that similar kinds of exercises have become common in industrial and governmental projects involving the collection and processing of personal data. So-called *privacy impact assessments* (PIAs) are recommended by several data protection agencies to ensure legal compliance to national and international privacy laws, e.g., by the Australian Privacy Commissioner[27] [**49**] or the UK Information Commissioner[28].

## 4.3 KNOW YOUR TOOLS (GET THE TECHNICAL DETAILS RIGHT)

Last but not least, you have to make sure that you know the capabilities and limits of current security and access control technology, so that you neither re-invent the wheel nor prematurely assume a problem solved. This is especially true if multiple technologies converge. Ubicomp applications often face severe technological challenges, both in terms of resources (energy, processing power, storage) and interfaces (e.g., small screens, gesture input). Relevant work includes security for wireless sensor networks [**50**] or RFID [**51**], as well as security and privacy interfaces [**52**].

The data collected in ubicomp applications also poses novel challenges to anonymization methods [**23**][29] and online subject access [**53**]. As we have discussed in section 3 above, even anonymous data collected in a ubicomp application may easily be de-anonymized, e.g., in the context of location privacy. The problem of subject access[30] is exacerbated by the latent identifiabilty of anonymized human activity, physiognomic and movement data. The rapidly growing area of *reality mining* [**54**] might not only provide means to make sense of such sensor data, but also offer clues as to their proper anonymization.

**Conclusions**

1. "Solving" the problem of personal privacy in ubicomp applications requires the careful analysis of application use and users, information flows and privacy risks, and technical options.
2. Privacy risk analysis and privacy impact assessments can help to identify privacy issues in actual and planned deployments of ubicomp applications.
3. Traditional technology tools to provide security, anonymity, and privacy need to take the special challenges of ubicomp applications into account, e.g., resource-constraint operations and large-scale behavioral data sets.

[27] See http://www.privacy.gov.au/publications/pia06/
[28] See http://www.ico.gov.uk/upload/documents/pia_handbook_html/html/foreword.html
[29] See http://en.wikipedia.org/wiki/AOL_search_data_scandal and http://www.aolstalker.com/ for an example of how difficult it is to anonymize search records.
[30] Many privacy and data protection laws guarantee the data subject access to the data that is collected about him or her.

**Further Reading**

- James Waldo, Herbert S. Lin and Lynette I. Millett, *Engaging Privacy and Information Technology in a Digital Age*, Computer Science and Telecommunications Board, National Academies Press, Washington, DC, 2007
  The book is the result of a multi-year study committee on Privacy in the Information Age, sponsored by the Computer Science and Telecommunications Board (CSTB) of the U.S. National Research Council (NRC). It presents a comprehensive and multidisciplinary examination of privacy in the information age, not directly focusing ubicomp applications, but with a wide enough view that looks towards current and future technological developments. The full text is available for free from http://books.nap.edu/catalog.php?record_id=11896#toc
- David Wright, Serge Gutwirth, Michael Friedewald, Elena Vildjiounaite, Yves Punie (Eds.), *Safeguards in a World of Ambient Intelligence*, Springer, Dordrecht, 2008
  Based on the results of an EU research project with the same title (SWAMI for short) the book illustrates the threats and vulnerabilities of ubicomp applications by means of four "dark scenarios." The authors analyze and identify safeguards to counter the foreseen threats and vulnerabilities, and make recommendations to policy-makers and other stakeholders on how to protect privacy in future ubicomp scenarios.

**Acknowledgements**

This chapter benefited from the thorough reading and the many comments on earlier drafts of the text by Aaron Quigley and John Krumm. Thanks! I am also indebted to the authors and copyright holders of the various figures and images I was kindly allowed to reprint.

## BIBLIOGRAPHY

1. Weiser, Mark: The Computer for the 21st Century. Scientific American 265(3), 94-104 (September 1991)

2. Warren, Samuel, Brandeis, Louis: The right to privacy. Harvard Law Review 4(5), 193-220 (December 1890)

3. Westin, Alan: Privacy and Freedom. Atheneum, New York, USA (1967)

4. Marx, Gary: Murky Conceptual Waters: The Public and the Private. Ethics and Information Technology 3(3), 157-169 (2001)

5. Solove, Daniel: A Taxonomy of Privacy. University of Pennsylvania Law Review 154(3), 477-560 (January 2006)

6. Solove, Daniel: Understanding Privacy. Harvard University Press, Cambridge, USA (2008)

7. Rössler, Beate: Der Wert des Privaten. Suhrkamp, Frankfurt/M, Germany (2001)

8. Kumaraguru, Ponnurangam, Cranor, Lorrie: Privacy Indexes: A Survey of Westin's Studies. Technical Report CMU-ISRI-05-138, Institute for Software Research International, Pittsburgh, PA, USA (2005) Available online at

http://reports-archive.adm.cs.cmu.edu/anon/isri2005/abstracts/05-138.html.

9. Cranor, Lorrie, Reagle, Joseph, Ackerman, Mark: Beyond Concern: Understanding Net Users' Attitudes About Online Privacy. Technical Report (1999)

10. TNS Emnid: Kundenkarten etablieren sich als feste Einkaufsbegleiter. Umfrage, Bielefeld, Germany (2006) http://www.loyaltypartner.com/fileadmin/upload/Presse/Studien/Kundenkarten_Deutschland_Emnid_LP.pdf.

11. Danezis, George, Lewis, Stephen, Anderson, Ross: How Much is Location Privacy Worth. Fourth Workshop on the Economics of Information Security, Harvard University (2005)

12. Solove, Daniel: The Future of Reputation: Gossip, Rumor, and Privacy on the Internet. Yale University Press, New Haven, USA (2007)

13. Cohen, Julie: Examined Lives: Informational Privacy and the Subject as Object. Stanford Law Review 52, 1373-1438 (2000)

14. Solove, Daniel, Schwartz, Paul: Information Privacy Law 3rd edn. Aspen, New York, USA (2009)

15. Altman, Irwin: The Environment and Social Behavior: Privacy, Personal Space, Territory and Crowding. Brooks/Cole, Monterey, USA (1975)

16. Palen, Leysia, Dourish, Paul: Unpacking "privacy" for a networked world. In : Proceedings of the 2003 Conference on Human Factors in Computing Systems, CHI 2003, Ft. Lauderdale, Florida, USA, April 5-10, 2003, New York, p.129.136 (2003)

17. Dourish, Paul, Anderson, Ken: Collective Information Practice: Exploring Privacy and Security as Social and Cultural Phenomena. Human-Computer Interaction 21(3), 319-342 (2006)

18. Grudin, Jonathan: Desituating Action: Digital Representation of Context. Human-Computer Interaction 16(2), 269-286 (December 2001)

19. Chaum, David: Untraceable electronic mail, return addresses, and digital pseudonyms. Communications of the ACM 24(2), 84 - 90 (February 1981)

20. Chaum, David, Fiat, Amos, Naor, Moni: Untraceable Electronic Cash. In : Proceedings of the 8th Annual International Cryptology Conference on Advances in Cryptology (CRYPTO 1988), New York, vol. LNCS 403, pp.319-327 (1988)

21. Sweeney, Latanya: k-Anonymity: a model for protecting privacy. International Journal on Uncertainty, Fuzziness and Knowledge-based Systems 10(5), 557-570 (2002)

22. Casassa Mont, Marco, Pearson, Siani, Bramhall, Pete: Towards Accountable Management of Identity and Privacy: Sticky Policies and Enforceable Tracing Services. In : Proceedings of the 14th International Workshop on Database and Expert Systems Applications (DEXA'03), September 1-5, 2003, Prague, Czech Republic, pp.377-382 (2003)

23. Agrawal, Rakesh, Kiernan, Jerry, Srikant, Ramakrishnan, Xu, Yirong: Hippocratic Databases. In : Proc. of the 28th Int'l Conf. on Very Large Databases (VLDB 2002), Hong Kong, China, pp.143-154 (2002)

24. Cranor, Lorrie: Web Privacy with P3P. O'Reilly, Sebastopol, USA (2002)

25. Hong, Jason, Landay, James: An Architecture for Privacy-Sensitive Ubiquitous Computing. In : Proceedings of the Second International Conference on Mobile Systems, Applications, and Services (MobiSys 2004), June 6-9, 2004, Hyatt Harborside, Boston, Massachusetts, USA., New York, pp.177-189 (2004)

26. Langheinrich, Marc: A Privacy Awareness System for Ubiquitous Computing Environments. In Borriello, Gaetano, Holmquist, Lars-Erik, eds. : Ubiquitous Computing, 4th International Conference (UBICOMP 2002), Göteborg, Sweden, September 29 - October 1, 2002, Proceedings., New York, vol. LNCS 2498, pp.237-245 (2002)

27. Bellotti, Victoria, Sellen, Abigail: Design for Privacy in Ubiquitous Computing Environments. In : Third European Conference on Computer Supported Cooperative Work, ECSCW'93, Milano, 13-17 September 1993, Proceedings, pp.75-90 (1993)

28. Ohkubo, Miyako, Suzuki, Koutarou, Kinoshita, Shingo: Cryptographic approach to "privacy-friendly" tags. In : RFID: Applications, Security, and Privacy. Addison-Wesley (2005)

29. Molnar, David, Soppera, Andrea, Wagner, David: A scalable, delegatable pseudonym protocol enabling ownership transfer of RFID tags. In : Selected Areas in Cryptography - SAC 2005, Kingston, Canada, August 2005., Heidelberg, Berlin, New York, vol. LNCS 2897, pp.276-290 (2005)

30. Rieback, Melanie, Gaydadjiev, Georgi, Crispo, Bruno, Hofman, Rutger, Tanenbaum, Andrew: A Platform for RFID Security and Privacy Administration. In : Proc. USENIX/SAGE Large Installation System Administration Conference, New York, pp.89--102 (2006)

31. Fishkin, Kenneth, Roy, Sumit, Jiang, Bing: Some Methods for Privacy in RFID Communication. In Castelluccia, Claude, Hartenstein, Hannes, Paar, Christof, Westhoff, Dirk, eds. : European Workshop on Security in Ad-hoc

and Sensor Networks (ESAS 2004), Berlin, Heidelberg, New York, vol. LNCS 3313, pp.42--53 (2005)

32. Langheinrich, Marc, Marti, Remo: Practical Minimalist Cryptography for RFID Privacy. IEEE Systems Journal, Special Issue on RFID Technology 1(2), 115--128 (December 2007)

33. Spreitzer, Mike, Theimer, Marvin: Providing Location Information in a Ubiquitous Computing Environment. In : Proceedings of the Fourteenth ACM Symposium on Operating Systems Principles (SOSP '93), New York, NY, pp.270--283 (1993)

34. Beresford, Alastair: Location Privacy in Ubiquitous Computing. Technical Report UCAM-CL-TR-612, Cambridge, UK (2005)

35. Beresford, Alastair, Stajano, Frank: Location Privacy in Pervasive Computing. IEEE Pervasive Computing 2(1), 46-55 (2003)

36. Beresford, Alastair, Stajano, Frank: Mix Zones: User privacy in location-aware services. In : Proceedings of the First IEEE International Workshop on Pervasive Computing and Communication Security (PerSec) 2004, Piscataway (2003)

37. Priyantha, Nissanka, Chakraborty, Anit, Balakrishnan, Hari: The Cricket Location-Support system. In : Proceedings of the 6th annual international conference on Mobile computing and networking (Mobicom 2004), New York, NY, pp.32-43 (2000)

38. Quigley, Aaron, Ward, Belinda, Ottrey, Chris, Cutting, Dan, Kummerfeld, Robert: BlueStar, a privacy centric location aware system. In : Position Location and Navigation Symposium (PLANS 2004), pp.684-689 (2004)

39. Krumm, John: A Survey of Computational Location Privacy. Personal and Ubiquitous Computing 13(Online First) (2008)

40. Kriplean, Travis, Welbourne, Evan, Khoussainova, Nodira, Rastogi, Vibhor, Balazinska, Magdalena, Borriello, Gaetano, Kohno, Tadayoshi, Suciu, Dan: Physical Access Control for Captured RFID Data. IEEE Pervasive Computing 6(4), 48-55 (2007)

41. Reimann, Robert: So You Want To Be An Interaction Designer. In: Cooper Interaction Design. (Accessed June 1, 2001) Available at: http://www.cooper.com/journal/2001/06/so_you_want_to_be_an_interacti.html

42. Lederer, Scott, Hong, Jason, Dey, Anind, Landay, James: Personal privacy through understanding and action: five pitfalls for designers. Personal and Ubiquitous Computing 8(9), 440-454 (November 2004)

43. Smith, Ian, Consolvo, Sunny, LaMarca, Anthony, Hightower, Jeffrey, Scott, James, Sohn, Timothy, Hughes, Jeff, Iachello, Giovanni, Abowd, Gregory: Social Disclosure of Place: From Location Technology to Communication Practices. In Gellersen, Hans, Want, Roy, Schmidt, Albrecht, eds. : Pervasive Computing. Third International Conference, PERVASIVE 2005, Munich, Germany, May 8-13, 2005. Proceedings, Berlin, vol. LNCS 3468, pp.134-151 (2005)

44. Friedman, Battya, Kahn, Peter, Borning, Alan: Value Sensitive Design and Information Systems. In : Human-Computer Interaction in Management Information Systems. M.E. Sharpe, Inc., New York (2006) 348-372

45. Freier, Nathan, Consolvo, Sunny, Kahn, Peter, Smith, Ian, Friedman, Batya: A Value Sensitive Design Investigation of Privacy for Location-Enhanced Computing. In: Workshop on Quality Value(s) and Choice: Exploring Wider Implications of HCI Practice (QVC) at CHI 2005, Portland, OR, USA. (Accessed April 3, 2005) Available at: http://www.eng.cam.ac.uk/~pw308/workshops/QVC/papers.html

46. Camp, L., Connelly, Kay: Beyond Consent: Privacy in Ubiquitous Computing. In Acquisti, Alessandro, De Capitani di Vimercati, Sabrina, Gritzalis, Stefanos, Lambrinoudakis, Costas, eds. : Digital Privacy: Theory, Technologies and Practices. Auerbach Publications, Boca Raton, USA (2007) 327-346

47. Felten, Ed: DRM, and the First Rule of Security Analysis. In: Freedom to Tinker. (Accessed March 19, 2003) Available at: http://www.freedom-to-tinker.com/blog/felten/drm-and-first-rule-security-analysis

48. Hong, Jason: Privacy risk models for designing privacy-sensitive ubiquitous computing systems. In : Proceedings of the 5th Conference on Designing interactive Systems: Processes, Practices, Methods, and Techniques (DIS '04). Cambridge, MA, USA, August 01 - 04, 2004., New York, pp.91-100 (2004)

49. Clarke, Roger: Privacy Impact Assessment in Australian Contexts. eLaw Journal 15(1), 72-93 (July 2008) https://elaw.murdoch.edu.au/.

50. Perrig, Adrian, Stankovic, John, Wagner, David: Security in wireless sensor networks. Communications of the ACM 47(6), 53-57 (June 2004)

51. Juels, Ari: RFID Privacy: A Technical Primer for the Non-Technical Reader. In Raicu, Daniela, Strandburg, Katherine, eds. : Privacy and Technologies of Identity: A Cross-Disciplinary Conversation. Springer, Berlin (2005) 57-74

52. Cranor, Lorrie: Security and Usability. O'Reilly, Sebastopol, USA (2005)

53. Roussopoulos, Mema, Beslay, Laurent, Bowden, Caspar, Finocchiaro, Giusella, Hansen, Marit, Langheinrich, Marc, Le Grand, Gwendal, Tsakona, Katerina: Technology-induced challenges in Privacy and Data Protection in Europe. A report by the ENISA Ad Hoc Working Group on Privacy and Technology, European Network and

Information Security Agency, Heraklion, Crete, Greece (2008)

54. Eagle, Nathan, Pentland, Alex: Reality Mining: Sensing Complex Social Systems. Personal and Ubiquitous Computing 10(4), 255-268 (2006)

55. Solove, Daniel, Rotenberg, Marc, Schwartz, Paul: Privacy, Information, and Technology. Aspen, New York, USA (2006)

56. Litman, Jessica: Information Privacy/Information Property. Stanford Law Review 52(5), 1283-1313 (2000)

57. Partridge, Kurt, Golle, Philippe: On Using Existing Time-Use Study Data for Ubiquitous Computing Applications. In McCarthy, Joe, Scott, James, Woo, Woontack, eds. : UbiComp'08, September 21-24, 2008, Seoul, Korea, New York, vol. 344, pp.144 -153 (2008)

58. McGraw, Gary: Software Security: Building Security In. Addison-Wesley Longman, Amsterdam, The Netherlands (2006)

59. Finkenzeller, Klaus: RFID Handbook 2nd edn. Wiley & Sons, Chichester, UK (2003)