IDIS (2010) 3:539–563
DOI 10.1007/s12394-010-0073-8

PETs and their users: a critical review of the potentials and limitations of the privacy as confidentiality paradigm

Seda Gürses

Received: 13 July 2009 / Accepted: 21 August 2010 / Published online: 13 November 2010

Abstract “Privacy as confidentiality” has been the dominant paradigm in computer science privacy research. Privacy Enhancing Technologies (PETs) that guarantee confidentiality of personal data or anonymous communication have resulted from such research. The objective of this paper is to show that such PETs are indispensable but are short of being the privacy solutions they sometimes claim to be given current day circumstances. Using perspectives from surveillance studies we will argue that the computer scientists’ conception of privacy through data or communication confidentiality is techno-centric and displaces end-user perspectives and needs in surveillance societies. We will further show that the perspectives from surveillance studies also demand a critical review for their human-centric conception of information systems. Last, we rethink the position of PETs in a surveillance society and argue for the necessity of multiple paradigms for addressing privacy concerns in information systems design.

Keywords Privacy · Confidentiality · PETs · Surveillance studies

Introduction

> My concern has been and will continue to be directed solely at striking a balance between democracy and technology, between the advantages of computerization and the potential safeguards that are inherent in it and the right of every citizen to the protection of his right of privacy (Gallagher 1967).

S. Gürses (✉)
ESAT/COSIC/HMDB and IBBT, K.U. Leuven, Heverlee, Belgium
e-mail: seda@esat.kuleuven.be

These are the words of Representative Cornelius Gallagher who through a series of hearings in 1966 brought the problem of computerization and its threats to privacy to public consciousness. Interestingly, today many of the arguments that Gallagher gave in his hearings still prevail. At the time, Gallagher was responding to the plans of the U.S. government to introduce a National Data Center that would collect information about every U.S. citizen. During those hearings Gallagher articulated important challenges to the then nascent scientific discipline of computer science. Shortly after these hearings Gallagher was invited as a keynote speaker at the American Federation of Information Processing Societies' 1967 Spring Joint Computer Conference (JCC) in New Jersey, U.S.A.[1] Subsequently, his challenges were echoed in the Communications of the ACM (Titus 1967), historically one of the most prominent journals within the computer science community.

Given that special historical constellation, we can state that 1967 was the year in which privacy as a research topic was introduced to the field of computer science.[2] A session on privacy problems was introduced at the conference and a total of 5 papers were presented on the topic of privacy.[3] Privacy was never explicitly defined in these papers but assumed to be confidentiality of data, the breach of privacy then meaning the leakage of data to unauthorized principals in military and non-military systems. Although Representative Gallagher defined privacy as the right to freedom of action, speech, belief or conscience and listed potential risks of information collection to these (Gallagher 1967), the transposition of the concept to research in computer science was driven by existing mechanisms. Privacy was hence limited to data confidentiality and secrecy of communications.

Research on privacy enhancing technologies (PETs) that guarantee some type of data confidentiality and hence user privacy have been an important research topic ever since.[4] The data that is kept confidential using PETs may be stored data, communicated data, or the conditions of a given communication. The last one guarantees the anonymity of the sender and/or receiver of the communication and in some cases only the confidentiality of the relationship between the sender and the receiver.[5] The data is kept confidential from an

[1] The series of conferences commenced in 1961 and were dissolved in 1990 (Room 2007).

[2] Although in the papers presented at the Spring JCC the suggested privacy and security solutions have parallels to the much longer standing tradition of research on cryptography and communications security, we in this paper start our account of privacy research in computer science with the explicit introduction of the term "privacy" at the Spring JCC Conference.

[3] Three of the authors were from the RAND Corporation (Ware 1967a; Petersen and Turn 1967), one from M.I.T (Glaser 1967) and one from a company named Allen-Babcock Computing (Babcock 1967).

[4] Later, other sub-fields in computer science have proposed other types of PETs that often rely on the contractual negotiation of personal data revelation. These are sometimes called "soft privacy" tools and a review of such PETs can be found in (Wang and Kobsa 2006). Such tools are not the focus of this paper since they are not based on the same assumptions that PETs for preserving data confidentiality, also known as "hard privacy", build upon.

[5] In relationship anonymity, the sender and receiver of messages may be known, but not the relationship between any pair of senders and receivers.

adversarial which may be the communication partner, e.g., service provider, data holder or an attacker in the environment. Such PETs are designed to enable anonymous speech, hide communication content and guarantee the unlinkability of electronic activities in a networked world.

In the last ten years, various authors in the emerging field of surveillance studies have raised critiques with respect to PETs and what can be called the privacy as confidentiality paradigm. Surveillance studies is a cross-disciplinary initiative to understand the rapidly increasing ways in which personal details are collected, stored, transmitted, checked, and used as a means of influencing and managing people and populations (Lyon 2002). Surveillance is seen as one of the defining features that define and constitute modernity. In that sense, surveillance is seen as an ambiguous tool that may be feared for its power but also for its potential to protect and enhance life chances. Departing from paranoid perspectives on surveillance, the objective of these studies is to critically understand the implications of current day surveillance on power relations, security and social justice.

Some of the surveillance studies scholars show that, given modern day conditions, computer scientists' conception of privacy through data or communication confidentiality has been mostly techno-centric and displaces end-user perspectives. Orlikowski (2007) defines techno-centricity as a perspective which is mainly interested in understanding how technology leverages human action, taking a largely functional or instrumental approach. Within such approaches, engineers and others tend to assume unproblematically that "technology is largely exogenous, homogenous, predictable, and stable, performing as intended and designed across time and place". This perspective tends to put technology in the center, blend out cultural and historical influences, and produces technologically deterministic claims.

The dichotomous opposite of techno-centricity is human-centricity, a perspective which focuses on how humans make sense of and interact with technology in various circumstances. In human-centric accounts, technology is understood to be different depending on meanings humans attribute to it and through the different ways people interact with it. This approach takes "cultural and historical contexts into consideration, but has a tendency to minimize the role of technology itself" (Orlikowski 2007).

Our objective in this paper is to study how materiality—in our case surveillance and privacy technology—and the social are constitutively entangled. Meaning that PETs and privacy related social practices in spaces of ubiquitous surveillance cannot be seen as ontologically distinct matters but constitute each other: social practices in spaces subject to ubiquitous surveillance are constituted by existing surveillance practices, technologies and by PETs, whereas PETs are the product of humans, their own social practices and conceptions of how surveillance is made effective and can be countered. Through a deeper understanding of this constitutive entanglement we can be more precise and constructive with our critique of PETs. We can also make explicit, evaluate and reconfigure the assumptions that designers of PETs rely upon. Our long term objective is use the analysis of these constitutive entanglements as a point of departure to explore critically informed approaches to privacy research within computer science.

In order to accomplish our objective, in Section "Privacy as data confidentiality and anonymity" we give an overview of computer science research on privacy as data confidentiality and anonymity. Here we also list some of the assumptions common to this type of research. Next, in Section "Surveillance society and PETs" we survey perspectives from surveillance studies for analyses of current day conditions and how these escape techno-centric narrations of PETs. Then, in Section "Revisiting PETs" we follow up the critiques of PETs with an evaluation of both, the consequences of these critiques for computer scientists and designers of systems, as well as the human-centricity in the different surveillance perspectives. Last, we propose the application of multiple privacy research paradigms that could be utilized to improve information systems design and that make apparent the constitutive entanglement of technologies and social practices in the context of privacy.

Hence, our main contribution with this paper is a critical interdisciplinary review of PETs that explores their potentials and limitations in a world of rapidly developing technology, ubiquitous surveillance, as well as changing perceptions, legislation and practices of privacy. Similar critique of PETs have been written in the past (Phillips 2004; Stalder 2002; Tavani and Moor 2001). We have included the first two of these perspectives in the later sections. In addition to reviewing existing critiques, the contributions of this paper are valuable for the following: First, confidentiality as *the* way to preserve privacy is a recurring theme in privacy debates generally and in computer science specifically. Therefore, re-examining whether this holds given changing social and technical conditions is important. Second, we have some additional arguments in the section on "the information perspective" which provides interesting challenges to privacy research in computer science. Further and most important of all, our argument is not that anonymity and confidentiality tools should be done away with. On the contrary, we believe that they need to be re-positioned with respect to the effects of ubiquitous surveillance in order to adapt the assumptions and claims of PETs to changing user requirements in their socio-technical contexts.

Privacy as data confidentiality and anonymity

Personal data as the focus of PETs

In those initial papers on privacy presented at the 1967 Spring Joint Computer Conference (Glaser 1967; Babcock 1967; Petersen and Turn 1967; Ware 1967a, b) the researchers had noticed the importance of sensitive data. Yet, defining what actually counted as "private" sensitive data was difficult to define. The authors distinguished between military and non-military systems, with the latter meaning industrial or non-military governmental agencies. Their objective was to devise systems that would avoid intentional or accidental disclosure of confidential sensitive data belonging to other users (Ware 1967a). In defining how to classify private information, a pragmatic approach was to ask how the military classification of document confidentiality i.e. confidential, highly-confidential, top secret, could somehow be mapped onto sensitive personal data. The authors also discussed, if such a mapping was unrealistic when a central authority and the necessary discipline to enforce such schemes were lacking in non-governmental/private computer networks (Ware 1967b).

In all the papers, confidentiality, a concept that played an important role in military think, was chosen as the main paradigm for privacy research. The authors were aware that military concepts could not be applied to society at large, hence all the papers attempted to distinguish security (military) from privacy (non-military). Yet, they had few other frameworks than military to take as a reference.

Ever since, there has been much progress in "non-military" contexts defining what should count as *informational privacy* (Gutwirth 2002; Solove 2006; Nissenbaum 2004) and which data are personal (EU Directive 95/46/EC (EU 1995), Art. 2 (a)) and sensitive (EU Directive 95/46/EC (EU 1995), Art. 8). Various data protection legislations have defined the category "personal data" and "sensitive data" as subject to privacy protection. An example is the EU Directive that states: *Personal data* are "any information relating to an identified or identifiable natural person [...]; an identifiable person is one who can be identified, directly or indirectly, in particular by reference to an identification number or to one or more factors specific to his physical, physiological, mental, economic, cultural or social identity" (EU Directive 95/46/EC (EU 1995), Art. 2 (a)); *sensitive data* include data "revealing racial or ethnic origin, political opinions, religious or philosophical beliefs, trade-union membership, and the processing of data concerning health or sex-life" (EU Directive 95/46/EC (EU 1995) Art. 8).

Important for our purposes is the emphasis these definitions put on identity. Identity is assumed to be unique for each natural person and data traces can be linked to these uniquely identifiable persons. This emphasis on identity is coupled with a relative under-specification of content that counts as personal or sensitive data. The U.S. terminology requires the protection of *personally identifiable* data. In contrast, personally identifiable data refers to a data set that is well defined. The standard types of personally identifiable data are profile data describing individuals, including name, address, and health status.

It is probably in line with this focus on identity and identifiability that, in comparison to the contributions at the 1967 Spring Joint Computer Conference, from the 80s onwards PETs focusing on the unlinkability of data[6] gained

[6] Unlinkability may refer to the unlinkability of the identity of the person from the traces she leaves behind, or the unlinkability of multiple traces, or simply of two (digital) items of interest such that these items of interest are no more and no less related after the observation of an (attacker) than they are related concerning (the attacker's) a-priori knowledge (Pfitzmann and Hansen 2008).

in popularity. Hence, in computer science identifiability includes inferring the link between an individual and some data probabilistically.

The construct of probabilistic identification would require the expansion of which data would count as personal or sensitive data in data protection legislation.[7] We discuss the different definitions of probabilistic identification and its implications in more detail in the next section.

Anonymity as a privacy enhancing mechanism

Once data about a person exists in digital form on a networked system it is very difficult to provide that person with any guarantees on the control of that data. This lack of control has contributed to personal data leakages and undesirable secondary use of personal data on the Internet,[8] often leading to privacy breaches. Much of the personal data collected using current day technologies represent activities of individuals assumed to be private or shared by a few prior to their digitization. Not digitizing these activities and avoiding the exchange of such digitized data would avoid further parties acquiring knowledge of these activities. But, this would substantially limit many of the technologies we use daily and is often undesirable for reasons we will account for in later sections.

A weaker form of preserving privacy is then to keep personal data confidential from a greater public. This would require the use of cryptography and secure systems: the prior has so far proven to be inaccessible for most Internet users, while the latter is extremely difficult to accomplish in networked systems.

A yet another form of privacy can be achieved through what is called anonymity. Anonymity is achieved by unlinking the identity of the person from the traces of his/her activities in information systems. Anonymity keeps the identity of the persons in information systems confidential but is not primarily concerned with how public the traces consequently become. This is also paralleled in data protection legislation which by definition cannot and does not protect anonymous data (Guarda and Zannone 2009).

In technical terms, anonymity can be based on different models. In communications, anonymity is achieved when an individual is not identifiable within a limited set of users, called the anonymity set (Pfitzmann and Hansen 2008). An individual carries out a transaction anonymously, if an observer cannot

[7] The consequences of probabilistic identification as legal evidence has been picked up by Braman (2006) but is beyond the scope of this paper. Similarly, in Wills and Reeves (2009) the authors problematize some of the heuristics used to probabilistically infer information about individuals based on the characteristics shared by their network vicinity. The authors argue that the heuristics are comparable to notions like "birds of a feather flock together", "judge a man by the company he keeps", or "guilty by association".

[8] The focus of this paper is on the Internet based technologies and privacy concerns. We are aware that breaches also occur with devices that are off-line and on networks other than the Internet. Further, mobile technology has opened up a whole new set of questions about the feasibility of keeping location data confidential, or what location privacy can mean. Whether our analyses in this paper also hold for these problems is beyond the scope of this paper.

distinguish her from others. The observer, often also called the adversary, may obtain some additional information (Diaz 2005). This means that the observer captures probabilistic information about the likelihood of different subjects having carried out a given transaction. The observing party may be the communication partner or some other party with observation capabilities or with the ability to actively manipulate messages. Depending on the observer's capabilities different models can be constructed with varying degrees of anonymity for the given anonymity set. Exactly what degree of anonymity is sufficient in a given context is dependent on legal and social consequences of a data breach and is an open question (Diaz 2005). TOR[9] and JAP[10] are some popular and established examples of applications that enable anonymous communication on the Internet. From here on, we refer to such anonymity systems as anonymizers.

Technically the capabilities of anonymizers are limited by the size of the anonymity set, the powers of the observer and by the content of the communication. If the communication content includes personally identifiable information, then the communication partners can re-identify the source person. Worse, if the communication content is unencrypted then intermediary observers may also re-identify persons (or institutions) as was shown by Dan Egerstad.[11] Another technical vulnerability can stem from the persistence of communication, meaning after observing multiple rounds of communication, the degree of anonymity may decrease and the probability of re-identification increases.

In databases that store personal data, the objectives and conditions for establishing anonymity sets are somewhat different. The end of the 90s saw the development of a new research field called privacy preserving data mining (PPDM) or publishing (PPDP). The objective of PPDM is to find methods to publish databases holding records of personal data such that certain information can be inferred from the database, while forestalling the inference of certain other information (information that could lead to privacy breaches). Ideally, the so published database cannot be used to identify individuals uniquely or in a small set of individuals.

Early results showed that simple de-identification is not enough for database anonymization. In her seminal work Sweeney (Sweeney 2002) showed that she could re-identify 85% of the persons in anonymized hospital reports using the publicly available non-personal attributes i.e. gender, age and zip code. Sweeney proposed methods to achieve k-anonymity in databases: A database listing information about individuals offer k-anonymity protection, if the information contained for each person cannot be distinguished from $k - 1$ individuals whose information also appears in the database. Later work in the field proved that k-anonymity alone, or any improvements of the same idea as

[9] http://www.torproject.org/

[10] http://anon.inf.tu-dresden.de/index_en.html

[11] In 2007 Dan Egerstad set up a number of TOR exit nodes (a popular anonymizer http://www.torproject.org) and sniffed over 100 passwords from traffic flowing through his nodes. The list included embassies and government institutions (Paul 2007).

provided in (Kifer and Gehrke 2006; Li and Li 2007; Rebollo-Monedero et al. 2008; Domingo-Ferrer and Torra 2008), cannot provide absolute guarantees against re-identification (Dwork 2006).

The results of the two branches of research on anonymous communication and database anonymization point to research challenges and limitations. PPDM research results show that the scope of personal data cannot be restricted to a limited set of data e.g., name, birthday, address and that the scope of what counts as personal data may be difficult to determine. The results from traffic analysis and anonymous communications show that the information that can be inferred by analyzing the conditions of communication, even if the content of the communication is hidden, may lead to undesirable revelation of personal data. If anonymizers are used to hide who is communicating with whom, however, without encryption at some point in the communication path, an analysis of the unencrypted content may lead to the identification of the anonymous communication partners. Anonymous and encrypted communication addresses all these problems, although it may also be vulnerable to re-identification if the communication is persistent.

These are known technical limitations which may be addressed through the development of improved methods that better mitigate the vulnerabilities. However, these results also demand that the scope of personal data be expanded and the privacy as confidentiality paradigm be re-evaluated. They invoke additional research questions about the usability and practicability of PETs and PPDM methods in a world of multiple data sources, unlimited aggregation and possible inferences. We will come back to some of these issues later. For now, let us look at some of the assumptions that underlie PETs.

Anonymity and confidentiality in the internet: assumptions of PETs

There are some basic assumptions of the privacy as confidentiality paradigm that inform the logic of PETs. These assumptions are often not made explicit in research papers on privacy technologies but are articulated between the lines. The assumptions vary between technical and socio-technical ones and can be listed as follows:

1. *There is no trust on the Internet* Here, the absence of trust refers to a number of things. First, given its current architecture, the users of the Internet can never be sure if a communication has been established with the correct party, i.e., as the popular analogy states, whether the partner they are communicating with is a dog, and if this dog/non-dog is receiving the sent messages; and if the message is received. Further, the users cannot determine whether those messages remain unchanged. Next, given an unprotected communication channel and unencrypted messages, there are no guarantees against third parties eavesdropping on that communication. Last, once the communication has been established and message or data transfer has taken place, there are no guarantees that the transmitted data will not be distributed to further parties.

2. *Users are individually responsible for minimizing the collection and dissemination of their personal data* By suggesting that through the use of PETs privacy can be enhanced or protected, an implicit claim is that personal data on the Internet originates from the users themselves. If the users want to protect their privacy then they should protect their data individually. This is especially the case because, see assumption (1) there is no trust on the Internet.
3. *If they know your data, then they know you* Although this is not necessarily limited to computer science papers, when discussing privacy and the effects of data collection knowing data about individuals gets mixed up with knowing those individuals. Knowing a user or individual here refers to knowing some or all of the following: intentions, causalities, and reached objectives.[12]
4. *Collection and processing of personal data, if used against individuals, will have a chilling effect* The personally identifiable data distributed on the Internet may be used against individuals in unwanted and unclaimed ways. Especially if we make the assumption (3) if they know your data, they know you, then it follows that massive collection of data may have severe consequences. Misuse of personal data in repressive, discriminating, or simply undesirable ways may have a chilling effect. Hence keeping personal data confidential is a secure way to avoid such chilling effects.
5. *Technical solutions should be used to protect privacy instead of relying solely on legal measures* Data may leak legally or illegally since it is difficult to control data and (1) there is no trust on the Internet. In order to avoid such leakage of data, technical solutions like anonymity and confidentiality should be preferred over legal protection.

These assumptions, although coherent and strong in their arguments, can be criticized for being techno-centric. In order to show why, we will first look at some of the findings of surveillance studies with respect to life in a surveillance society. When appropriate, we will list critiques of PETs based on the privacy as confidentiality paradigm, as they are articulated in the different surveillance studies perspectives. In Section "Revisiting PETs" we will return to these assumptions.

Surveillance society and PETs

In the following, we give a short overview of some of the voices that we identify as surveillance studies perspectives that position themselves with respect to the

[12] There are numerous court cases in which digital data are used as evidence in ways which claim much more than the data seems at face value to represent. For example, a court in the U.S. accepted pictures from a social network site of a young woman enjoying a party a number of weeks after a car accident with casualties. The picture was used as proof that she lacked remorse (Wagstaff 2007).

privacy as confidentiality paradigm or PETs. Each of the perspectives investigates surveillance spaces and their affects in our lives and on our understanding of privacy. We sum up the different perspectives in the daily, marketing, political, performative and information perspectives. We will shortly account for each of these perspectives and their critique of PETs.

The daily perspective on surveillance

A typical critique of Internet users is that they do not care for their privacy. Even though different types of PETs are available to them, and in different studies users express their concerns for the privacy of their data, when they do make use of systems, these concerns evaporate and PETs are rarely utilized (Berendt et al. 2005). This low adoption of PETs has been contributed to the usability problems that are associated with PETs. Studies like "Why Johnny can't encrypt?" have argued exactly for this point (Whitten and Tygar 1999). In other cases, the users have been accused of being insensitive to the abuse of their data, naive, or simply ignorant of the risks that they are taking (Gross and Acquisti 2005) allowing the surveillance of every aspect of their daily lives.

In comparison, surveillance studies conceptualize individuals as being overburdened with the duty of protecting their individual privacy against powerful (governmental) institutions. In describing why individual protection of privacy is an unlikely solution to the surveillance problem, Felix Stalder underlines the consequences of the fact that our societies are increasingly organized as networks underpinned by digital information and communication technologies (Stalder 2002). He argues:

> In a network, however, the characteristics of each node are determined primarily by its connections, rather than its intrinsic properties, Hence isolation is an undesirable option.

As an example of the networked society, Stalder talks about how when going to the social services, we have to show information about housing and work, while at work we have to give bank information, which provides us with a credit card, which is a precondition to renting a car, etc. Therefore, it is the connectedness that provides individuals with access to various systems rather than their relationship one-by-one with these institutions.

Therewith, Stalder problematizes PETs that make use of anonymity or even unlinkable pseudonyms[13] for daily encounters. Wide spread use of anonymity and unlinkable pseudonymity tools burdens the individuals, more so then offering them tools of protection. Instead, Stalder argues that the burden

[13] Unlinkable pseudonyms refer to systems in which users identify themselves with different pseudonyms for different sets of transactions. For an observer it should not be possible to identify if two pseudonyms are coming from the same user. If implemented in an infrastructure with a trusted third party distributing the pseudonyms, then these can be revoked, ideally only under certain (legally defined) conditions.

to protect their privacy should be taken off the individual's shoulders and accountability should be asked of database holders.

The marketing perspective on surveillance

One of the main concerns with the collection of personal data is that of categorization and consequent social sorting practices. A typical example of discriminatory categorization is the use of geodemographic systems, which have been critiqued in previous studies (Curry and Phillips 2003; Graham 2005). Therewith, marketers and companies can decide on desirable and non-desirable customers, excluding parts of the population. The latter are no exceptions to existing economic models and are called *dead weight loss* (Bauer et al. 2006).

Anonymity offers little protection against these systems. The classification of individuals with respect to marketing categories (or for that matter governmental surveillance categories for crime suspects) is not necessarily based on the unique identity of persons but rather on attributes that they carry. In that sense, these systems continue to work even if the individuals using the systems are anonymized. What is important for these systems are behavioral data (patterns of user activities over time) and user attributes. A newcomer to the system does not have to be identified uniquely, it is enough if their behavior can be matched, given a set of categories.

Exactly this critique is picked up and taken a step further by Zwick and Dholakia, marketing specialists critical of existing geodemographic systems and consumer databases (Zwick and Dholakia 2003). Zwick states that the control of the consumer to determine his digital identity is minimal. Because, Zwick claims:

> Implicit in the conceptualization of all of these tactics is the assumption that the consumer self is ontologically distinct from its representation in the electronic market-space. Yet, from a poststructuralist perspective, the subject cannot be conceived in this way. Because the consumer is constituted by language and the language governing the electronic market space is constituted by databases. The consumer (as a meaningful cultural representation, not as a body) does not exist outside this constitutive field of discursive power. Hence, the consumers digital identity is his or her real identity because marketing is targeted toward the consumer profile rather than the real person.

The authors argue that knowledge is a function of linguistic power and linguistic power in the mode of information resides with database technologies. Hence, very much like Stalder the authors propose that a struggle for consumer identity needs to be fought at the level of the database.

The authors' critique of PETs is severe. They claim that PETs offer "customers" only a false perception of autonomy. The consumer categories cannot be manipulated. Indeed, as an alternative the authors argue that consumers must be given direct access to customer databases in order to ensure that he or she regains a viable voice in the act of his or her constitution as a customer. The customer has to be enabled in (co)-authoring their own identity.

The political perspective on surveillance

Within the privacy as confidentiality approach, there is an assumption that the protection of those issues, activities, opinions deemed to be private and the private sphere is ultimately a good thing. Phillips (2004) produces a critique of this normative approach based on feminist and queer theory. He says that:

> Some feminist scholars have argued that it is this creation of a private, domestic sphere apart from the public realm, that is the privacy problem. Certain populations and issues are relegated to this private sphere, particularly women and issues of sexuality. The public/private distinction then serves as a tool of social silencing and repression. From this perspective, the most important privacy issues are not those of freedom from intrusion into the domestic realm, but instead of the social construction of the public/private divide itself.

If we take a look at newspaper headlines in mainstream media on possible privacy breaches in the context of new technologies then Phillips' pointer becomes even more interesting. More often than not, these articles are about: drinking and drug habits or sexual preferences being visible to future employers(social networks), bodies becoming visible publicly (the airport scanners), severe illnesses or dissident opinions becoming available to public, etc. In all these horror stories of lives destroyed through the revelation of personal information, it is also possible to find traces of backlashes against political struggles of the last decades, e.g., politics of sexuality, labor rights, anti-discrimination struggles. We are not arguing that all these matters should be public and considering them private is always a matter of repression, but rather that the debates on privacy and the development of technologies that collect and process personal data are not devoid of the political and societal interests of dominant groups to emphasize certain concerns while silencing others. Further, privacy is not uniformly available and uniformly valued. The need for privacy can change depending on the context, as in the case of abuse or violence in the private sphere. For those who have little public power, the apparent invasion of privacy can sometimes seem welcome (McGrath 2004).[14]

[14] There are cases, where this is exactly turned around as in the case of Federal Record Keeping and Labeling Requirements which require secondary producers to be responsible for the record keeping procedures primary producers gather when they produce sexually explicit material (DoU 2008). The lack of public power of those who produce sexually explicit images leads to both an intrusion of the effected person's privacy and to silencing of those who want to publish explicit material. So, the argument is not that the private is always repressive and the public intervention is always a positive one, but that both the private and public needs to be negotiable and questionable, especially when the lives or livelihoods of those with meager public power are disputed.

Phillips' critique shows that through new technologies and their ability to make things visible and invisible we are forced to re-consider what should remain public and private. In expectation of such reconsiderations, in the EU, the notion and protection of privacy has since long included a social dimension, i.e., a freedom from undue interference in the development of one's social identity. But it still remains open, who decides when and how the boundaries are settled between the public and private when building information systems.

Ideally, such privacy related design decisions need to be taken through a social and democratic process and not determined solely by (technical, business and legal) experts. Even then, the power relations between the different participating and non-participating stakeholders have to be considered.

Last, Phillips suggests that instead of inscribing into systems absolute values of what should remain private or public, systems that allow users to negotiate when and if they want to keep their information private and public should be considered. From the political perspective, it is therefore advisable to build technologies that enable individuals or communities to safely and strategically push the boundaries between the public and the private.

The performative perspective on surveillance

The importance of performativity in surveillance space becomes easily evident in the interventions of many artists whose works have often inspired surveillance studies authors. Stephen Mann, in his series of performances "My Manager" wears a visible camera in stores and restaurants where customers are not permitted to take photographs, yet where CCTV surveillance is practiced. When approached by security staff, Mann tells them that "my manager" insists he wears the camera to ensure that he is not wasting his time during his errands. This is not photography, he explains, since the signals are being beamed offsite where they will be turned into images (McGrath 2004). Mann blames everything on "My Manager" and watches to see how the power relationships shift, how not taking responsibility for surveillance becomes ridiculous, and managers do all of a sudden appear to question his intervention.

McGrath in his book "Loving Big Brother" very much appreciates the wit and passion of such daily performative interventions. His analysis of surveillance space avoids a value judgment about the morality or desirability of surveillance technology per se. As a strategy, McGrath demands a shift form anti-surveillance to counter-surveillance. He points to the impossibility of controlling surveillance space itself i.e. once surveillance space exists, it can be used by unexpected parties in unexpected ways. In that, he accepts that there can be no trust on the Internet but derives different conclusions. Although performances like Mann's, or the use of cameras at demonstrations against police brutality (Lewis 2009), or even the coincidental recording of the police brutality practiced against Rodney King in the U.S.A. (BBC 2002) are valuable examples of reversing the gaze, McGrath focuses on counter-surveillance that goes beyond the reversal strategy. Instead, he proposes counter-surveillance that opens a space for all sorts of reversals in relation to how the gaze and its imagery may be experienced.

McGrath argues that proliferation of surveillance will produce discontinuities in experience of surveillance and produce excess. The more the surveillance proliferates, and the more surveillances start competing, McGrath argues the more we will see a battle over meaning. If we accept that surveillance space is in suspense, that the way we will take up surveillance, the way we will be affected, and the way we will respond are in suspense, then radical possibilities for counter-surveillance pop-up. And, he argues that the focus of these counter-surveillance strategies should be on deconstructing and subverting the tyranny of meaning given to surveillance material.

But, how? Mcgrath picks up on the theories of the poststructuralist Mark Poster, who states that: "we are already surrounded by our data bodies in surveillance space." He gives examples of artists' works and everyday surveillance uses that highlight the discontinuities of the surveillance narrative. It follows from the examples that these data bodies are neither simple representations of ourselves, nor straight falsifications, but hybrid versions of ourselves susceptible to our interventions. McGrath is aware that this multiplicity of selves will be distorted and exploited by the consumer-corporate system. But, he concludes, that the real danger lies in disengaging with the surveillance space.

> [T]he emergence of surveillance culture is nothing less than a challenge to our consciousness. [...] we ignore the circulating, multiple, hybrid versions of ourselves at our peril. If we deny their relation to us in an attempt to maintain the integrity of a unified self—rooted in rights of privacy—we risk surrendering any control, any agency, in relation to our lives and society.

According to these arguments, recent forms of practiced surveillance like reality shows e.g., Big Brother, and even most articulations of web based social networks can be seen as a realization of the fact that we live in a surveillance saturated society. By participating in these programs and environments we are discovering and exploring what to do about it. In that sense, McGrath questions any assumptions on surveillance data being representative for what people are, feel, intend, or achieve, etc. Instead, he encourages members of our societies to enter surveillance space, to experience its affects and to challenge any narratives that are limited to those of control and authority, or that try to monopolize what data means and how it can be used.

The information perspective on surveillance

In most surveillance systems data receive meaning because of their relationality. A single piece of data about a single person says very little unless there is a set of data to compare it to, a framework to give it some meaning. In the age of statistical systems, the collection of data sets is what makes it possible to make inferences on populations, to evolve categorizations of these populations and to practice social sorting. This is also what Phillips describes as surveillance:

> Surveillance is the creation and managing of social knowledge about population groups. This kind of privacy can easily be violated if individual observations are collated and used for statistical classification, which applied to individuals makes statements about their (non)-compliance with norms, their belonging to groups with given properties and valuations, etc.

In that sense any data, by its potential to be aggregated, has many data subjects, or better said always carries the potential of pointing to a data population. Individual decision making on personal data always effects all correlated subjects (Rouvroy 2009). Individualized concealment of data or unlinking of identities from traces provides little or no protection against breaches based on statistical inferences or discriminatory categorization for all correlated data subjects. Any individual can be categorized, as long as enough information has been revealed by others and aggregated in databases. The other way around, it is impossible to guarantee semantic security in statistical databases. Meaning it cannot be guaranteed that access to a statistical database would not enable one to learn anything about an individual that could not be learned without access (Dwork 2006). Hence, not only is the categorization of individuals based on attributes a problem, but also the analysis of statistical databases may reveal additional information about individuals. Similarly, recent studies have shown that attributes revealed by a user's friends or group affiliations may be used to infer the user's hidden attributes (Zheleva and Getoor 2009).[15]

Further, much information is revealed about individuals by virtue of their associations with others. By now, we all carry digital devices that accumulate data about our environments, as well as providing information about us to those environments. We disseminate this information on the social web or simply among our ecology of devices. We talk loudly on our cell phones about ourselves and others, reveal pictures of family, friends or visited locations, archive years worth of emails on multiple backup devices from hundreds of persons, map out our social networks which reveals information about all those in the network. In each of these cases it is evident that it is not only individuals that reveal information about themselves, but we all participate in multiple kinds of horizontal and vertical information broadcasts and surveillances. We collectively produce data: we produce collaborative documents on wikis, take part in online discussions and mailing lists, comment each others pictures etc. All of this data is relational and makes sense in its collectivity. It also works the other way, in the networked world that Stalder describes, much information is collected about us and is linked in order to give us access to systems.

[15] As we mentioned earlier, such inferences are problematic because a) they are likely to gain in popularity without a deeper understanding of what such mathematical inferences mean b) they are based on normative assumptions with respect to "like people network with each other" which can be oppressive or overly-judgmental as discussed in Wills and Reeves (2009).

Hence, looking at data as snippets of individual contributions to digital systems actually misses the actual value of that data in its collectivity and relationality. It does not recognize the publics we create and share, one of the greatest promises of the Internet. Individualizing participation in a surveillance society makes it difficult to develop collective counter-surveillance strategies, and limits our engagements with surveillance systems to individual protections of our actions.

The depiction of information privacy and data protection analogous to individual property rights actually exacerbates the problem of a contested public sphere. Although the rise of the social web can be celebrated as a long expected gift or cooking pot economy, not only its privatization through large companies, but also the privacy debates contribute to articulations against seeing information on the web as a public good.[16] The logic of privacy and private ownership has created the false perception that data in its singularity is of outmost value and is controllable. Privacy through confidentiality and anonymity can follow this logic and can be detrimental to collective critical engagement in surveillance systems.

Revisiting PETs

Personal data is an undefined category that is widening ever since PPDM has introduced the concept of quasi identifiers. This means that practically all data are always potentially linkable to an individual and hence are personal data. If we apply that directly to the logic of privacy as confidentiality, any data may reveal something about the individual, and hence needs to be kept confidential.[17] In the different surveillance perspectives, the authors discuss how the burden of keeping their personal data is inconvenient, undesirable, and often impossible.

We first return to the assumptions of PETs using the critiques in the different surveillance studies perspectives and provide an analysis for when these assumptions tend to be techno-centric. Then, we consider the human-centric assumptions underlying the surveillance studies perspectives, and, if relevant, argue when PETs based on the privacy as confidentiality paradigm may nevertheless be desirable, and even indispensable.

[16] It is therefore no surprise that in the latest uproar against the new Terms of Use of Facebook users have argued for a radical deletion of their profile to include the deletion of all their contributions to other profiles and all their correspondences to other people. The protection of individual privacy in such instances is valued over the integrity of the discussions forums, mailboxes of friends, posted photographs etc.

[17] Legally, this would mean that all data that can be related to persons will fall under the data protection regime. The usefulness of the category "personal data" when its scope is so widened is a worthy topic of investigation both legally, as discussed in Ohm (2009), and technically, as discussed in Shmatikov and Narayanan (2010).

Returning to the assumptions of PETs

1. *There is no trust on the Internet* This is a technical fact. Especially the underlying design of the Internet makes it very difficult to give guarantees on the security of communication. However, the absence of technical measures to make certain guarantees gets conflated with social definitions of trust. Exactly what social definitions of trust may be and how these categories get conflated is beyond the scope of this paper. However, this assumption should not stand alone as the only articulation of what trust is on the Internet.
2. *Users are individually responsible for minimizing the collection and dissemination of their personal data* As we argued in the information perspective to surveillance, the protection offered through the confidentiality of personal data, even if personal data were a controllable category of information, is limited. Anonymously collected data does not protect against surveillance systems and the reflexes of their controllers to manage and sort populations. Categories created through the databases using de-identified data can easily be used to classify individuals by virtue of mapping some of their attributes and their behavioral data to categorical descriptions. The affectivity of these categories depends on the power of those holding this data and their success in making constitutive claims. Therefore, assuming that keeping personally identifiable data confidential or unlinking individuals identities from their data traces may protect individuals from social sorting and marketing systems does not hold. This will remain the case, as long as the power of those data controllers are not questioned. Moreover, individuals are often not the only source of data and may not be able to control the revelation of data about them by others.
3. *If they know your data, then they know you* Surveillance data is often a place holder. It points to something that has happened or that has been, it often looses a sense of sender and receiver. Data looses the intentions behind its creation and starts to float in digital systems as data bodies. Stating that knowing data is equivalent to "knowing a person" reinforces the power of such data to stand for some "reality" or "truth". Rather, it is necessary to scrutinize the uses of data and deconstruct attempts to monopolize its meaning. When accountability is of concern, e.g., medical data, financial data, other technical mechanisms should be put into place that guarantee the necessary proofs that a certain data validly stands for something. In any case, to claim truth to surveillance data in order to argue for data confidentiality as privacy protection should be practiced sparingly.
4. *Collection and processing of personal data, if used against individuals, will have a chilling effect* It is true that collection and processing of data, if used against individuals can have a chilling effect. McGrath shows in most of his examples that there could also be other effects. By now, there are multiple occurrences in which people have publicized data in order to

protect themselves against the breach of their privacy.[18] In ubiquitously surveilled spaces, confidentiality and anonymity can actually make somebody suspect or devoid of protection.[19] These examples point out that in a surveillance society information may have both a chilling effect as well as an empowering effect.

5. *Technical solutions should be used to protect privacy instead of relying solely on legal measures* Given the amount of surveillance measures that have been installed and the amount of information collected about each person by their friends and organizations they are affiliated with, it is unrealistic to expect that technical measures can be applied realistically to actually keep many of our daily interactions confidential. Hence, we must search for a combination of all three, technological solutions, legal protections and social practices to respect those activities that we would like protected. A solely techno-centric approach is unrealistic, is bound to overwhelm any individual, and often too quick to dismiss many social contracts that we enjoy in everyday life.

Re-positioning the use of PETs:

Although we have shown that many of the assumptions underlying PETs are problematic, we still see an important value in the opportunities they have at offer. Especially given their vulnerable position as a result of politics of hyper-security—based on another assumption, that only criminals and people who have something to hide use PETs—it is necessary to be precise with our critique.

The human-centricity in the surveillance perspectives sketched above lies in the fact that they tend to dismiss the ability of PETs to enable any alternative digital behavior and to produce multiple (unexpected) affects. The solutions suggested by the different perspectives also often delegate the problems to the social: in the form of accountability, performativity and liability. We would like to explore the potential constitutive role that PETs play or can play in a surveillance society. Hence, we revisit PETs from each of the surveillance perspectives that we have summarized above in order to explore their limitations as well as their potentials.

We agree that putting the responsibility of re-establishing privacy should not solely lie on individuals. Further, it is paradoxical to suggest that the solution to

[18] Anne Roth started a blog (http://annalist.noblogs.org) in which she documented the everyday activities of her family after her partner, Andrej Holm, was arrested in Germany in 2007 under the accusation that he had engaged in a terrorist association. After the arrest, the family found out that their family had been subject to police surveillance for over a year. Similarly, New Jersey artist Hasan Elahi started documenting every minute of his life on the Internet after the FBI mistakenly detained him at an airport (http://trackingtransience.net/). Both of these persons made the assumption that keeping their lives public and visible protects their freedoms when government authorities choose to threaten these freedoms.

[19] If using PETs makes someone suspect, then it may be necessary to think of PETs of higher orders i.e., PETs that are used to keep the users of PETs anonymous.

potential undesirable instances of control should lie in users having to control their actions and data all the time. Therefore, visions of large scale anonymous or pseudonymous systems where users constantly hide their attributes and connections (Chaum 1985) and have to re-establish other forms of relatedness through digital reputation (Dingledine et al. 2002) are inconvenient and undesirable in a networked world.

Nevertheless, disabling the users' options to exercise some control over revealing their personal data if they want to is just as undesirable. We would hence argue that the accountability that Stalder demands of data processing systems should include making it mandatory for service providers to grant anonymous and/or pseudonymous access to their services and to practice data minimality for basic services. The networked society should continue to offer access to those who do not want to be so engrained in the existing networks i.e., not being very well networked should not lead to individuals being excluded from services. Therefore, PETs that enhance privacy through anonymity and confidentiality—or where better suitable unlinkable pseudonyms—should be integrated where possible. Nevertheless, their mere existence should not be enough to relieve data controllers and processors of their responsibilities and accountabilities. This should also have a legal effect: we should consider, if and how anonymized data can also be legally protected without requiring identifiability as a condition for protection.

We very much agree with the critique of Zwick and Dholakia with respect to the constitutive force of these marketing databases. PETs, and more specifically anonymizers, as long as they do not hide users' behavioral data, do not protect against the categorization discriminations based on marketing databases. Further, if we accept that categories of desirable and undesirable customers are constituted by the owners of those databases, then even hiding behavioral data may not protect against such constitutive forces. Hence, suggesting PETs can protect the privacy of individuals against aggressive marketing databases goes beyond burdening individuals, but actually produces a false perception of autonomy, and maybe even an illusion of control: if my data is important to me, then I can protect it, if I want to give away my data for a utility, I can do so. Such utility arguments are not viable in existing databases, given the categorization powers exercised by their owners.

Further, Zwicks and Dholakia's proposal for customer agency is limited to the engagement of the "customer" with the database. Their suggestions are to allow individuals to correct their profiles, which may go as far as deleting oneself from the marketing databases. But in a networked world, as Stalder argues, deletion will often not be an option. This is coupled with the technical unfeasibility of guaranteeing that any traces of data are deleted. In that sense, although their critique of PETs is substantial, by resurrecting a notion of "autonomy" based on access and control of individual database entries, the authors re-install a false sense of autonomy.

Instead, agency with respect to databases may be found in making visible the established relationality of data. Phillips states that the only thing that remains private in current surveillance systems are the methods through which these discriminatory classifications are created and used. By that he refers to the databases and algorithms used for data mining in those databases. It is through engagement with surveillance methods that it is possible to actually determine possible unwanted discriminations as well as desirable effects. It is also then possible to understand what becomes known to a specific public e.g. government, marketers, through the aggregation of individual revelations of information. Such transparency practices could also demystify the effects of data collection and processing. Ideally, we can then actually collectively as societies or communities discuss if and how desirable the newly created private or public spaces are.

Significant is also the role PETs can play in the negotiation of the public and private. Tools such as anonymizers and anonymized data publishing methods may allow users to challenge boundaries between the private and public. Anonymous speech has always been an important tool in democracies and an extremely helpful tool against repressive regimes and legislation. Nevertheless, the ultimate goal is not to limit the articulation of such opinions to anonymized spaces, but to make it possible to state such opinions in public, and to safely hold public discourse on issues that are repressed. In that sense, PETs are not normative tools for how we should ultimately communicate if we want to have privacy, but can play an indispensable role in negotiating the public and private divide.

For McGrath rather than re-establishing privacy and public assembly absolutes within a relativistic culture, engagement with surveillance is key. Instead, he suggests that: ownership of imagery and data selves; freedom of image and data circulation; the multiplicity and discontinuities of data experience; and, the emotional instability of security systems should be the center of our focus. Prima facie, it is possible to conclude that PETs solely enable the development of anti-surveillance strategies, and hence offer limited points of intervention in a surveillance society. But, given their recent popularity to circumvent repressive governments i.e. as a way to reach Internet sites that have been blocked by governments, there may be more value in those tools than visible at first sight. Even the mere existence of PETs may actually force governments, law enforcement and marketers to spell out their discomfort with anonymized data. It may force those parties to be transparent about their desire for total surveillance.

Discussion

We have argued in the last section that despite the problems with the assumptions underlying PETs, and despite the legitimate critiques articulated in the surveillance perspectives, PETs can and should be part of privacy research. The more we engage in surveillance space, the more we will find diverse uses for PETs. But, it was also one of our objectives in writing this paper to give legitimacy to multiple approaches to “privacy design” in computer science. We are especially interested in investigating approaches that do not work solely with the privacy as confidentiality paradigm and are able to integrate techno-centric and human-centric perspectives as constitutive others. We believe that a broader vision of privacy and a deeper understanding of surveillance could help both users and computer scientists to develop systems that support multiple kinds of privacies and data protection practices. For all of this, it is clear that an interdisciplinary approach is imminent.

In the past years, other approaches to privacy in computer science have started flourishing. Two of these can be shortly listed as follows:

- *privacy as control* A wider notion of privacy, appearing in many legal codifications, defines the term not only as a matter of concealment of personal information, but also as the ability to control what happens with it. This idea is expressed in Westin's (1970) definition of (data) privacy: the right of the individual to decide what information about himself should be communicated to others and under what circumstances and in the term "informational self-determination" first used in a German constitutional ruling relating to personal information collected during the 1983 census (Bundesverfassungsgericht 1983). Some examples of research in this area are accounted for under the title of identity management systems and trust based systems with (sticky) privacy policies.
  The technical mechanisms developed for the privacy as control paradigm rely on and make use of data protection legislation. This approach is hence valuable for improving the accountability of organizations collecting and processing surveillance data, a point emphasized in the different surveillance studies perspectives. Specifically, accountability of data controllers and processors are increased through individual oversight of information practices, as well as through the oversight capabilities it provides to regulatory and/or independent bodies like data protection authorities. Technologies and mechanisms under the privacy as control paradigm are based on the building stones offered by the techniques developed in the privacy as confidentiality paradigm, e.g., anonymity and unlinkability of different identities, but are not limited to it.
- *privacy as practice* Most existing privacy mechanisms are preemptive. Few mechanisms provide users with the ability to engage with data that is revealed or is no longer under the control of that individual user. It can help users and user populations immensely to know what data exists about them, to understand how it travels and how it is used, and to comprehend ways of improving privacy practices in the future. The users can then either make requests for amendments to existing data about themselves, which is the idea of informational self-determination, or reconfigure their settings and change their interactions to strategically reveal or conceal data in the future. There is little but valuable research done on systems that support users strategic revelation and concealment based on what is already known. Palen and Dourish argue that "privacy management in everyday life involves combinations of social and technical arrangements that reflect, reproduce and engender social expectations,

> guide the interpretability of action, and evolve as both technologies and social practices change" (Palen and Dourish 2003). Following the same lines of thought, (Lederer et al. 2004) suggest improving privacy sensitivity in systems through feedback that improves users' understanding of the privacy implications of their system use. They add that this can then be coupled with control mechanisms that allow users to conduct socially meaningful actions through them. On a similar line of thought, Liu et al. (2006) suggest the use of mechanisms like the identity Mirror and Nguyen (2002) the use of privacy mirrors. Hansen (2008) suggests combining such features with identity management mechanisms that implement privacy as control.

We see the way forward in combining all three paradigms: privacy as confidentiality, control and practice. The development and deployment of these systems should ideally be accompanied with studies investigating their social reception, use, and even rejection. We believe that a broader vision of privacy, a deeper understanding of surveillance, a sharp understanding of systems security, and user involvement in defining system priorities, can help both users and computer scientists to develop systems that support multiple kinds of privacies and data protection practices.

Last but not least, there is a need to consider other categories of data then just personal data. We have explained in Section "The information perspective on surveillance" the importance of the relationality of information. If we accept that data are relational, then we have to reconsider what it means to think within the framework of data protection that only protects "personal" information. For legal frameworks this can be the challenge of dealing with data sets that are co-created by many, that have multiple data subjects, or that are controlled by many. In domains where profiling and surveillance are common practice, addressing relational information includes the evaluation of the appropriateness of applying Article 15 of the EC Data Protection legislation which grants "individuals the right not to be subject to a decision ... which is based solely on automated processing". For computer scientists, it requires thinking of collaborative tools, anywhere from new forms of access control to methods for negotiating the visibility, availability and integrity of data owned and shared by many.

Conclusion

In this paper, we have shown some of the problems that arise with the privacy as confidentiality paradigm in a surveillance society. We have accomplished this by first studying how PETs based on data confidentiality and anonymity function and making explicit the techno-centric assumptions they rely on. We then used surveillance studies perspectives to show how these assumptions can be shaken by our current day conditions. We last returned to the surveillance studies perspectives to step out of some of their human-centric assumptions. In doing that we explored the potentials and limitations of PETs in a surveillance society.

Finally, we sketched two other paradigms that can be used to address informational privacy concerns within computer science. We also pointed out the importance of understanding the relationality of data in statistical systems and our networked world. In doing so, we argued that conceptions of surveillance and privacy technologies are constitutively entangled and that these can be integrated in developing privacy designs. A third element that is constitutively entangled with both technology and surveillance but that was not the focus of this paper are legal codings and theories. We touched on such legislation, theories and visions when they were relevant to our discussion. We believe it is through the combination of all three paradigms: privacy as confidentiality, privacy as control and privacy as practice, that we can both: recognize users abilities, wit and frustrations in navigating in surveillance space; and, develop critical and creative designs with respect to privacy.

Acknowledgements This work was supported in part by the Concerted Research Action (GOA) Ambiorics 2005/11 of the Flemish Government, by the IAP Programme P6/26 BCRYPT of the Belgian State (Belgian Science Policy), and by the European Community's Seventh Framework Programme (FP7/2007-2013) under grant agreement ną 216287 (TAS^3—Trusted Architecture for Securely Shared Services).[20] The authors also wish to thank Carmela Troncoso, Claudia Diaz, Nathalie Trussart, Katrin Borcea-Pfitzmann, Thibaut D'alton, Serge Gutwirth, Bettina Berendt, Andreas Pfitzmann, Brendan van Alsenoy, Sarah Bracke, Manu Luksch, Aaron K. Martin and the anonymous reviewers for their valuable and critical comments.

References

Babcock JD. A brief description of privacy measures in the rush time-sharing system. In: AFIPS '67 (spring): Proceedings of the Spring Joint Computer Conference. 1967. p. 301–2.

Bauer M, Fabian B, Fischmann M, Gürses S. Emerging markets for RFID traces. 2006. http://arxiv.org/abs/cs.CY/0606018. Accessed 23 Sept 2010.

BBC. Flashback: Rodney king and the LA riots. 2002. http://news.bbc.co.uk/1/hi/world/americas/2119943.stm. Accessed 23 Sept 2010.

Berendt B, Günther O, Spiekermann S. Privacy in e-commerce: stated preferences vs. actual behavior. Commun ACM. 2005;48(4):101–6.

Braman S. Tactical memory: the politics of openness in the construction of memory. First Monday. 2006;11(7). http://firstmonday.org/htbin/cgiwrap/bin/ojs/index.php/fm/article/viewArticle/1363. Accessed 23 Sept 2010.

[20] The information in this document is provided "as is", and no guarantee or warranty is given that the information is fit for any particular purpose. The above referenced consortium members shall have no liability for damages of any kind including without limitation direct, special, indirect, or consequential damages that may result from the use of these materials subject to any liability which is mandatory due to applicable law.

Bundesverfassungsgericht. BVerfGE 65, 1 – Volkszählung. Urteil des Ersten Senats vom 15. Dezember 1983 auf die mündliche Verhandlung vom 18. und 19. Oktober 1983 – 1 BvR 209, 269, 362, 420, 440, 484/83 in den Verfahren über die Verfassungsbeschwerden. 1983. http://www.servat.unibe.ch/dfr/bv065001.html. Accessed 23 Sept 2010.

Chaum D. Security without identification: transaction systems to make big brother obsolete. Commun ACM. 1985;28(10):1030–44.

Curry MR, Phillips D. Privacy and the phenetic urge: geodemographics and the changing spatiality of local practice. In: Lyon D, editor. Surveillance as social sorting: privacy, risk, and automated discrimination. London: Routledge; 2003.

Diaz C. Anonymity and privacy in electronic services (thesis). K.U. Leuven; 2005.

Dingledine R, Mathewson N, Syverson P. Reputation in privacy enhancing technologies. In: CFP '02: Proceedings of the 12th Annual Conference on Computers, Freedom and Privacy. 2002.

Domingo-Ferrer J, Torra V. A critique of k-anonymity and some of its enhancements. In: Third International Conference on Availability, Reliability and Security. ARES 08; 2008.

DoU. Department of Justice, 28 CFR part 75 revised regulations for records relating to visual depictions of sexually explicit conduct; inspection of records relating to depiction of simulated sexually explicit performance; final rule. Fed Regist. 2008;73(244).

Dwork C. Differential privacy. In: ICALP (2). 2006. p. 1–12.

EU. 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. Off J Eur Communities. 1995;31:31–9. http://europa.eu.int/eur-lex/en/consleg/main/1995/en_1995L0046_index.html.

Gallagher CE. The computer and the invasion of privacy. In: Proceedings of the 5. SIGCPR Conference on Computer Personnel Research; 1967. p. 108–14.

Glaser EL. A brief description of privacy measures in the multics operating system. In: AFIPS '67 (spring): Proceedings of the Spring Joint Computer Conference; 1967. p. 303–4.

Graham S. Software-sorted geographies. Prog Hum Geogr. 2005;29(5):562–80.

Gross R, Acquisti A. Information revelation and privacy in online social networks. In: WPES '05: proceedings of the 2005 ACM Workshop on Privacy in the Electronic Society; 2005. p. 71–80.

Guarda P, Zannone N. Towards the development of privacy-aware systems. Inf Softw Technol. 2009;51(2):337–50.

Gutwirth S. Privacy and the information age. Rowman and Littlefield Publishers. 2002.

Hansen M. Linkage control—integrating the essence of privacy protection into identity management. In: eChallenges; 2008. p. 1585–92.

Kifer D, Gehrke J. l-diversity: privacy beyond k-anonymity. In: IEEE 22nd international conference on data engineering; 2006.

Lederer S, Hong JI, Dey AK, Landay JA. Personal privacy through understanding and action: five pitfalls for designers. Personal Ubiquitous Computing. 2004;8(6):440–54.

Lewis P. Video reveals G20 police assault on man who died. Guard. 2009. http://www.guardian.co.uk/uk/2009/apr/07/video-g20-police-assault. Accessed 23 Sept 2010.

Li N, Li T. t-closeness: privacy beyond k-anonymity and -diversity. In: IEEE 23rd international conference on data engineering; 2007.

Liu H, Maes P, Davenport G. Unraveling the taste fabric of social networks. IJSWIS. 2006;2(1): 42–71.

Lyon D. Editorial. Surveillance studies: understanding visibility, mobility and the phenetic fix. Surveillance & Society. 2002;1(1):1–7.

McGrath J. Loving big brother: performance, privacy and surveillance space. London: Routledge; 2004.

Nguyen DH. Privacy mirrors: understanding and shaping socio-technical ubiquitous computing. Technical Report. 2002. http://smartech.gatech.edu/handle/1853/3268. Accessed 23 Sept 2010.

Nissenbaum H. Privacy as contextual integrity. Wash Law Rev. 2004;79(1):119–58.

Ohm P. Broken promises of privacy: responding to the surprising failure of anonymization. Tech. rep., University of Colorado Law School. 2009.

Orlikowski WJ. Sociomaterial practices: exploring technology at work. Organ Stud. 2007;28(9): 1435–48.

Palen L, Dourish P. Unpacking "privacy" for a networked world. In: CHI '03: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. 2003. p. 129–36.

Paul R. Security expert used TOR to collect government e-mail passwords. Ars Technica. 2007. http://arstechnica.com/security/news/2007/09/security-expert-used-tor-to-collect-government-e-mail-passwords.ars. Accessed 23 Sept 2010.

Petersen HE, Turn R. System implications of information privacy. In: AFIPS '67 (spring): Proceedings of the Spring Joint Computer Conference; 1967. p. 291–300.

Pfitzmann A, Hansen M. Anonymity, unobservability, and pseudonymity: a consolidated proposal for terminology. Tech. rep., Technical University, Dresden. 2008. http://dud.inf.tu-dresden.de/Anon_Terminology.shtml. Accessed 23 Sept 2010.

Phillips DJ. Privacy policy and PETs. New Media Soc. 2004;6(6):691–706.

Rebollo-Monedero D, Forné J, Domingo-Ferrer J. From t-closeness to pram and noise addition via information theory. In: PSD '08: Proceedings of the UNESCO Chair in Data Privacy International Conference on Privacy in Statistical Databases; 2008.

Room ICSP. Computer society history committee names top 60 events (1946–2006). 2007. IEEE website, http://www.computer.org/portal/cms_docs_annals/annals/images/top60.pdf. Accessed 1 May 2009.

Rouvroy A. Technology, virtuality and utopia. In: Reading panel on autonomic computing, human identity and legal subjectivity—legal philosophers meet philosophers of technology, CPDP 2009; 2009.

Shmatikov V, Narayanan A. Myths and fallacies of "personally identifiable information". Commun ACM. 2010;53(6):22–6.

Solove DJ. A taxonomy of privacy. Univ PA Law Rev. 2006;154(3):477.

Stalder F. The voiding of privacy. Sociol Res Online. 2002;7(2). http://www.socresonline.org.uk/.

Sweeney L. k-anonymity: a model for protecting privacy. Int J Uncertain Fuzziness Knowl-Based Syst. 2002;10(5):557–70.

Tavani HT, Moor JH. Privacy protection, control of information, and privacy-enhancing technologies. SIGCAS Comput Soc. 2001;31(1):6–11.

Titus JP. Security and privacy. Commun ACM. 1967;10(6):379–81.

Wagstaff E. Court case decision reveals dangers of networking sites. Daily Nexus News. 2007. http://www.dailynexus.com/article.php?a=13440.

Wang Y, Kobsa A. Privacy enhancing technologies. In: Gupta M, Sharman R, editors. Handbook of research on social and organizational liabilities in information security. Hershey, PA: IGI Global; 2006.

Ware WH. Security and privacy in computer systems. In: AFIPS '67 (spring): Proceedings of the Spring Joint Computer Conference; 1967a. p. 279–82.

Ware WH. Security and privacy: similarities and differences. In: AFIPS '67 (spring): Proceedings of the Spring Joint Computer Conference; 1967b. p. 287–90.

Westin AF. Privacy and freedom. Atheneum; 1970.

Whitten A, Tygar J. Why Johnny can't encrypt: a usability evaluation of PGP 5.0. In: SSYM'99: Proceedings of the 8th Conference on USENIX Security Symposium; 1999.

Wills D, Reeves S. Facebook as a political weapon: information in social networks. Br Polit. 2009;4(2):265–81.

Zheleva E, Getoor L. To join or not to join: the illusion of privacy in social networks with mixed public and private user profiless. In: WWW '09; 2009.

Zwick D, Dholakia N. Whose identity is it anyway? comsumer representation in the age of database marketing. J Macromark. 2003;24(1):31–43.