Translating privacy into digital designs: technical strategies to counter everyday surveillance

Seda Gürses[1] and Jason Pridmore[2]

[1] ESAT/COSIC and IBBT
K. U. Leuven
Belgium
sguerses@esat.kuleuven.be
[2] The DigIDeas Project
Infonomics and New Media
Hogeschool Zuyd
The Netherlands
J.H.Pridmore@hszuyd.nl

Abstract. Privacy policies, data protection frameworks, and privacy (enhancing) technologies exist as a conglomeration of strategies that seek to counter issues arising from the proliferation of digitally mediated surveillance in everyday life. Systems designers and policy makers have proposed the introduction of privacy related measures into Information Technology (IT) systems at their inception and during their development in recent years as an effort to limit potential harms created by the intensification of digitally mediated surveillance. These attempts to design in privacy from the outset are interpreted and implemented in different ways, often in the form of lists of 'design principles' and desired system properties. In this paper, we suggest that the process of designing systems that maintain privacy are reliant upon particular translations of privacy; translations that address privacy and subsequently surveillance and related issues, in different ways. That is, they re-interpret, re-represent and re-appropriate the notion of privacy in ways that fit particular objectives for differing stakeholders in their use of new technologies. This paper provides a taxonomy of three significant proposals to maintain privacy in systems design and discusses their usefulness and drawbacks in different contexts in relation to digitally mediated surveillance. First, the paper looks at the privacy by design through pre-emption. This largely security engineering perspective relies heavily on data minimization, cryptography, and distributed architectures to reduce and limit potential surveillance use. Second, the paper examines proposals for privacy by design that focuses on data protection compliance. These proposals are largely derived from legal and policy oriented organizations that draw on existing laws and legal rulings to ensure that technical infrastructures fit with legal frameworks. Last, a third group of proposals focuses on user-centricity. This perspective attempts to make privacy control more visible and accountable by creating (the perception of) individualized means for data management controlled by users/customers/citizens. All of these proposals to design into the system privacy features, techniques and practices attempt to counter increasingly intrusive and ubiquitous

potentials for surveillance in different contexts. Each begins to capture some of the implications of contemporary systems of surveillance, and yet fail at others. In some cases, the proposals turn the concept of privacy upside-down, proposing strong surveillance as a precondition to accountability and oversight, two important principles of data protection. Others propose transparency mechanisms, yet this transparency only applies to the users and not to the data collecting organizations and their processing practices. We argue that a dynamic and nuanced approach toward privacy as a mechanism to minimize and reduce surveillance is necessary to ensure that the desire to design in privacy legitimately counters the intrusive potentials of new digitally mediated surveillance. We see the resulting taxonomy as helpful for developing systems of surveillance that are less intrusive and more transparent, while also enabling critical engagement with these practices.

1 Introduction

The development of nearly ubiquitous forms of information technologies (IT), almost universally available and accessible, has produced the emergence of and potential for more and more pervasive forms of digitally mediated surveillance. In this context, privacy policies, data protection frameworks, and privacy (enhancing) technologies exist as a conglomeration of strategies that seek to counter issues arising from the proliferation of surveillance in everyday life. In order to resolve the potential harms and concerns that can arise from surveillance, these tools are employed not as an antidote to surveillance [30], but rather as a means for limitation. Unfortunately, all too often privacy concerns in system design are an afterthought, and approaches to ensuring privacy in many circumstances occur on an ad hoc or post-hoc basis. Though this has not universally been the case, growing concerns about digitally mediated surveillance has led to the development of several proposals and sets of practices that seek to design into systems specific features, factors, and functions that at the outset serve to ensure privacy. These approaches fit within the rubric of value sensitive design, in which a given value, such as the value of (and right to) privacy, is accounted for in a principled and comprehensive manner throughout the design process [13]. With the rise of digitally mediated surveillance and its implications becoming an increasingly central concern in social discourse, at least as is presented in the media, discussions about how to incorporate privacy into system design has become an ever more pressing issue. The proposals and practices for incorporating privacy into design rely on certain design principles and system properties that integrate intentional choices to preserve or enhance privacy directly into the system architecture. These provide strategies to reduce the potential for abuse or unwanted surveillance that may occur through the everyday use of (usually) large-scale information architectures. However, the translation of privacy into design principles and system properties is not straightforward. The concept of privacy has to be interpreted and implemented into systems design by engineers and researchers that rely upon less than clear policy directives and their own

perceptions and understandings of privacy. The implementation of privacy into systems design is a difficult task given that there is no clear and precise conception of the problem of privacy no pregiven essence [25] - that is universally accepted.

What occurs in the process is similar to that which has occurred with approaches to privacy: particular configurations of the privacy problem draw boundaries around what should be done and how [25]. That is, specific perspectives and readings of privacy are systematically integrated into system design both explicitly and implicitly based on an idea of privacy that is held in common by certain groups of designers and policy makers. The conception of privacy and the problem it serves to solve is transferred into the development and design of these systems (both in principles and practice). In what follows, we seek to articulate the ways in which the problem of privacy has been integrated into different design principles and system properties, acknowledging that these are broad strokes that seek to be representative of the emphasis in approach. As initiatives to incorporate values including privacy into design are diverse and multifaceted, we acknowledge that this taxonomy is limited. However, the description of these approaches are essential for understanding the potential and current limitations of what we see as a well intentioned approach to integrating value sensitivity into system design in order to reduce the harmful effects of digitally mediated surveillance.

2 Taxonomy

We suggest that approaches to integrating privacy into system design maintain three fundamental emphases in practice, that of pre-emption, compliance, and user-centricity. It is important to note that these are simply emphases in approach, and not mutually exclusive practices. Rather these emphases serve to focus the development and implementation of privacy into design principles and system properties in specific ways. Pre-emption, compliance and user-centricity are seen as solutions to the problem of privacy, forming the basis through which privacy sensitive and aware technological designs emerge.

However, the following sections trace how an orientation to privacy, something already contested and negotiated, pre-structures the subsequent development of systems that seek to limit privacy invasive potentials. This orients practices towards ensuring data confidentiality in some cases, adhering to legal requirements in others, and shifting the locus of data control in yet others. While these may be seen as design principles, intended ways to develop the system, they are achieved through desired system properties, techniques and tools such as anonymous communication or access control. The taxonomy described below is intended to help conceptualize the differences in approach in order to more systematically conceptualize how privacy can act to counter the harms of current and future forms of digitally mediated surveillance.

2.1 Pre-emption

Amidst a myriad of definitions, privacy has been defined as "the right to be let alone" [32]. This definition is connected to an individualistic liberal tradition in which an assumed a priori self is granted a sphere of autonomy free from both governmental and societal intrusions [24]. It is a definition that has also been popularly used by computer security researchers in computer science whom have translated this concept into the need for an autonomous (digital) sphere in which data about persons is protected so that it cannot be accessed by unauthorized others. This interpretation of privacy is inextricably linked with data confidentiality and in this context privacy is seen as ensuring that personal information is not accessible to a greater public. Privacy is lost at the point in which personal data becomes public.

Given a conception of the privacy problem that seeks to ensure that personal data does not become public in an identifiable manner, these researchers have focused their efforts on particular design strategies that reduce these risks. This has led to the development of privacy technologies concerned with data confidentiality and minimization, including Privacy Enhancing Technologies (PETs). The main objective of such technologies is to enable the use of information based services while either limiting the collected information (data minimization) or concealing the origin (and destination) of transmitted information (anonymous communications).

The approach toward the privacy problem in this case is hinged on the idea that once information about a person exists in a digital form (data), it is difficult to provide technical guarantees that the distribution of this data can be controlled. Maintaining privacy by constraining access to data is a highly complicated matter: inherent in digital systems is the fact that any data can be replicated and widely distributed, and the copy is not distinguishable from the original. These two features of digital systems make it impossible to identify the origin of a copy that is distributed without authorization. This is especially the case in our networked world where "it [is] virtually impossible to tell where one computer stops, and another begins" [25]. Therefore the choices available to keep data private are either to significantly reduce data capture or to apply various cryptographic building blocks and other security mechanisms to achieve system properties like anonymity, unlinkability, unobservability of traffic data and communications content confidentiality.

One prominent privacy solution addressing the confidentiality of the conditions of a communication, i.e., the origin and destination or identifying correspondents, is called an anonymizer. Anonymity is achieved when a person is not identifiable within a limited set of users, called the anonymity set [23]. Such a system delinks the identity of a person from the traces of her activities in information systems. The desired side effect is that no observing party can link the content of the communication to the communicating parties (unless this is volitional), and no other parties can observe who is communicating with whom. The current state-of-the-art system, Tor, counts thousands of volunteer relays

and hundreds of thousands of users, and focuses research effort on how to scale the system beyond this.

Other confidentiality approaches depend upon the underlying architecture of the system to be optimized to achieve data minimization using cryptographic protocols and distributed architectures [2]. Within such a system users may be identified, but the objective is to enable the desired functionality without collecting more data than is deemed necessary about the persons in centralized databases.

While computer security engineers are doing significant work to pre-empt this version of the privacy problem, data mining and knowledge discovery researchers have proposed a parallel privacy approach: Privacy Preserving Data Publishing (PPDP). While this approach does not pre-empt centralized databases as in the above cases, it does apply the concept of protecting privacy by guaranteeing the indistinguishability of people after data collection. While the objective for anonymous communications is to create an architecture in which it is impossible to link people with their communications, the objective for PPDP and other database anonymization practices rely on an anonymizing party not to reveal that link to the world.

Database anonymization seeks to provide analysts of personal information databases (microdata) with methods to analyze and infer certain information from the database, but reduces or eliminates the potential to infer certain other information (information that could lead to privacy breaches through the identification of individuals, e.g.,[31, 17, 19, 27]. By applying these database anonymization methods, the database owners may maintain the economic utility of the database while avoiding subsequent harms to individuals represented in the database.

A further alternative to database anonymization and PPDP is possible in interactive systems with a secure query interface, known as differential privacy. Differential privacy limits those that want to benefit from the utility of the database so that they may query it in a privacy preserving manner, with the database remaining under the control of the database owner. Differential privacy serves to define the problem of privacy on the basis of statistics that show how much perturbation is necessary to prevent additional inferences about specific individuals despite numerous queries to the database, regardless of their participation in the database. Hence, surveillance remains in tact: the results of statistical analysis are available, however the risk to an individual incurred by joining (or leaving) the database is minimized [11].

Anonymizers, database anonymization, differential privacy and other forms of privacy enhancing technologies are an attempt to pre-empt the potential for privacy breaches and invasions. At their basis they recognize the ease at which digital information can be replicated and distributed as indistinguishable from original content and see the involuntary publication of this data as the fundamental privacy problem. This perspective on privacy is therefore translated into techniques and tools that can be used to redress these problems, however the

attempts at pre-emption are never seen as complete or definitive solutions to the overall privacy problem.

2.2 Limitations of Pre-emption

The intention for anonymous communications is to keep the identity of the persons in information systems confidential, unless this is volitionally disclosed to the communication party. This translation of the privacy problem, however, does not take into consideration that even with the use of anonymous communication, unlinked traces of that communication can and do subsequently become public. These solutions assume that unlinking the (unidentifiable) traces a person produces in daily interactions with digital technologies is a desirable and sufficient protection of the communications of that individual.

This perspective, at first sight, is in line with data protection legislation that presupposes that anonymity is sufficient to protect individual privacy. However, by definition data protection does not protect anonymous data it only applies to personal data, data that can be (likely or reasonably) identifiable by being linked back to a person. Legally, data protection does not and technically cannot apply to digital traces produced through anonymous communications. Applying data protection to anonymous communications would paradoxically require that an identifier is left behind. This has become a fundamental tension between the requirements of Data Protection regimes, the objectives of those engaged in pre-emptive strategies, and the perceived desires of users. In order to implement an effective and comprehensive data protection regime, the implementation of an extensive surveillance and tracking infrastructure is necessary.

In this process, anonymous communications uphold the economic logic of data protection by enabling free data flows. Anonymous communications make it difficult to link traces back to the specific people, but it does not impede upon collecting data traces for surveillance purposes. For most forms of digitally mediated surveillance, data may not necessarily be individually identifiable as persons in the population are substitutable, but this data can still be processed to infer categories that are useful in influencing, controlling, manipulating, entitling and protecting that population [20]. By removing personally identifiable information or dismantling the link between disclosures and the individuals, hence inevitably stepping outside of data protection, anonymous communications in turn produce conditions that are necessary for free data flows, unlimited processing, and hence, population surveillance.

A comparable tension exists also between database anonymization methods and data protection principles. Recent years we have seen a number of de-anonymization attacks demonstrated on database anonymization techniques. These attacks cast serious doubt on whether data-sets containing relational data between users (for instance, ratings of movies, product reviews, etc.) are ever completely safe to release, despite the application of elaborate anonymization techniques. It hence becomes unclear which legal category anonymized data-sets fit and whether data protection must apply to all data-sets that may likely

reasonably be linked to individuals [12]. This likewise makes the use of PPDP methods for anonymization of databases increasingly questionable.

Similarly, the protection offered by anonymous communication systems are weakened when used persistently, as they eventually can disclose long term communication partners. Anonymous communications should instead be seen as a tactical protection that can be used in the same manner only for a limited time period. Overall, the increasing application of data mining and the ability to link information from multiple sources raises the likelihood of re-identification of users or linking of traces left behind using anonymous communications back to the individual senders.

These limitations points out that the deployment of these technologies to achieve anonymous communications or anonymize databases should be practiced with great caution. However, one could move a step further and suggest that regardless of their mathematical sophistication, success and failures, the use of anonymized data sets and attempts to ensure the unlinkability of communication data would only create a false sense of autonomy. In these cases, there is little to distinguish a person ontologically from his representation in databases. In the case of consumers, for instance, they are constituted by language and the language governing this space is constituted by (marketing) databases [33]. Anonymizers or database anonymization cannot offer an intervention into that language and its governance. In this context, pre-emptive solutions are limited in their effects: while ensuring to differing degrees personal privacy, the potential for this conception of the privacy problem to limit the potential harms and risks of digitally mediated surveillance is insufficient.

However, dispensing of anonymizers and anonymization techniques because of their shortcomings is clearly not in the best interest for reducing these harms and risks. Instead, the increase in the collection of massive amounts of personal data, the need to participate in networks, together with the increasing popularity of anonymizers underline the challenges to and the relevance of developing pre-emptive privacy solutions in a digitally mediated surveillance world. These challenges demand at the very least, that researchers advance new solutions that are robust enough against the vulnerabilities that occur as a result of advances in data mining. As Phillips argues, the importance of anonymity for the negotiation of the public and private boundaries for social issues is undeniable [24]. However, as he also suggests, we should not confuse their importance with the urge to define privacy absolutes that reinforce normative boundaries.

2.3 Compliance

The goal to create and maintain a privacy by design approach in the development of new systems is increasingly voiced by policy makers, especially privacy commissioners. It is less frequently discussed by system designers as the privacy by design approach is concerned with building systems that are primarily compliant with various forms of privacy legislation, most importantly data protection legislations. However, this emphasis on privacy by design does have an active following in the engineering community, with computer scientists recognizing the

need for privacy and security solutions to be developed in compliance with privacy and data protection regulations from the start of the systems engineering process.

One of the first and most prominent advocates of the term privacy by design is Ann Cavoukian, the Information and Privacy Commissioner of Ontario. In her articulation of how privacy by design can be accomplished she names seven guiding principles [7]. These principles were later widely adopted as a resolution by other prominent policy makers at the 32nd Annual International Conference of Data Protection and Privacy Commissioners meeting in Israel. These principles are:

- Proactive not reactive, Preventative, not Remedial
- Privacy as the default
- Privacy Embedded into Design
- Full functionality - Positive Sum not Zero Sum
- End-to-end security - Lifecyle Protection
- Visibility and Transparency
- Respect for User Privacy

Cavoukians [7] concept of privacy by design extends to the trilogy of 1) IT Systems, 2) accountable business practices, and 3) physical design and networked infrastructure. The principles listed in the document apply to this trilogy, and hence demand a holistic approach to privacy.

These principles can be seen to extend the fundamentals of data protection that form the basis of most data protection acts. Much like these principles, the approach is one that attempts to remain technology (and organization) neutral, by making few or no statements about relevant engineering practices or the application of privacy by design solutions. Despite its comprehensiveness, it is not clear from these documents what privacy by design actually is and how it should be translated into the engineering practice. Many of the principles are tautological; they include the term privacy by design in the explanation of the principle itself. The main point is that compliance with data protection and the adoption of privacy solutions should not be an afterthought, but carried out from the beginning of systems development and organizational management. What this means is left up to the interpretation of the reader.

From our vantage, this approach is indicative of a conception of the problem of privacy as needing to ensure the free ow of personal information through various technologies and organizational structures while constraining its collection and processing via regulatory procedures and instruments introduced in the system design. This becomes clear in proposals by several policy makers that seek to assist organizations in perceiving potential privacy risks, mitigating these through technology while also achieving compliance. For example, the Privacy Impact Assessments (PIA) [16] list a number of activities that are expected to guide organizations in the following tasks: developing organizational processes to capture privacy risks; using these privacy risks to drive the design of organizational processes as well as the underlying information systems; allowing

organizations to justify and document the privacy risks expected from the data collection that they will not mitigate; and processes for ensuring compliance with relevant local and international legislation.

It is anticipated that by applying the privacy impact assessment organizations will be able to identify privacy risks to individuals as well as related liabilities, avoid damages to the reputation of the organization, costs resulting from bolt-on solutions, and instill public trust and confidence in their projects and products. Of course it is not simply PIAs that are part of privacy by design, but the intentions for privacy by design is to maintain this key point: privacy by design through compliance is incentivized. It is a means to reduce the legal, financial and reputation risks that would result from potential data breaches and breaches to the privacy of the data subjects.

Relevant technologies and practices for privacy by design include those that are seen as means for pre-emption described above: compliance includes PETs, anonymous communication, anonymous credentials, and database anonymization. However, in this case, the emphasis within the privacy by design approach is on technologies that digitally mimic a bureaucratic intake and data management procedure in information systems. User-friendly policy languages, sticky policies coupled with elaborate access control models, as well as the user centric models described in the next section, are part and parcel in these recommendations. Within technology research, the corresponding focus on privacy by design includes ways to interpret data protection and other privacy legislation, e.g., HIPPAA, as a set of rules that can be used to automate compliance with legislation within the system [21]. Some researchers seek to generate privacy requirements from privacy policies [4], and from data protection legislation [14], while others seek to integrate compliance into systems engineering methods [10].

As suggested, there is much overlap in the privacy by design approach with the pre-emption approach mentioned above and the user-centric approach described below. The distinction that we are making here is on the impetus for and the focus of designing in privacy and in how the problem of privacy is understood. The compliance approach sees the issues of privacy as regulatory ones, governing the development of systems so that these maintain the rights and opportunities of citizens, consumers, travellers, and so on. However, it is recognized that regulation is limited: it puts a significant burden on those bodies that must evaluate such practices. Instead, the compliance approach uses the concept of privacy by design to translate certain principles and practices into the conception and development of digitally mediated surveillance systems at the outset. This may use the best of pre-emptive practices and user centric practices all with the intention of complying with data protection principles and privacy legislation. This limits the need for later interventions that can be seen as constraining opportunities for data flow and use. Thus a compliance approach seeks to balance both the innovative and efficient use of data within informational capitalism [6] and the protection of personal information. Once again, the focus in this approach is largely individualistic, but is significantly empowered to alter corporate practices. Attempts to incorporate privacy into design from

the outset should be lauded, however more efforts need to be made to ensure that privacy by design takes into account the difficulties of controlling digital information given the ease of replication and distribution, the risks of creating centralized databases given these properties, and the complications arising from surveillance practices that are not transparent.

2.4 Limitations of Compliance

There are some significant limitations to compliance oriented approaches and privacy by design. While some legal scholars may infer data minimization and confidentiality from existing legislation, e.g., see Kuner [18], and from the Data Protection Directive [12], the focus in these documents is on informing people about the collection and processing of their data. They provide some justification and options with respect to which data will be collected, and they articulate the legal obligations of data collection and processing organizations. In a technical sense, privacy by design is mainly about integrating and adding mechanisms to the system to assist organizations in remaining compliant. For example, sticky policies from this perspective can be seen as a set of instructions that indicate how to process data in order to remain complaint.

However, if the conditions for data access are met, there are no technical means to guarantee that the obligations for privacy are fulfilled after data access. Policy obligations may restrict the use of the data for a specific purpose only and may permit access for the given purpose and the data collector or processor may handle the data according to these policies. Yet these mechanisms do not mitigate the risks of abuse or potential leaks of data resulting from vulnerabilities in the technical system, whether this occurs in the process of communicating with the data collector, or from through security weaknesses in databases held by the data collectors.

Part of the problem is that attempting to ensure compliance through technical mechanisms, such as requirements for consent, opt-in and opt-out choices and elaborate access control mechanisms, disregards the basic properties of digital information mentioned above they are easy to replicate, distribute and manipulate, and as such are difficult to safeguard in the way commonly imagined in the different data protection legislation. Compliance all too often limited to a layer of bureaucratic mechanisms for accountability transposed onto digital systems. Even when this is integrated into the system at the outset, as is the intention with privacy by design, the limitations of this accountability are determined by the vulnerabilities of the systems at hand. It is hence no surprise that organizations state that any loss of data due to security leaks from vulnerabilities beyond those security measures required to be compliant are not their responsibility [15]. In this way, privacy as compliance can be seen as a series of symbolic activities in information systems to assure consumers confidence as well as the free flow of information in the marketplace as suggested above [9, 8, 16]. In the worst case, compliance becomes a way for organizations that practice a trust us, we do no evil philosophy that pays lip service to privacy and data protection legislation.

Moreover, given the intensity of the personal information economy and the absence of an explicit and precise data minimization principle, the compliance approach is susceptible to a reversal of its intended effect. Data collectors can articulate the purpose specification to include any data of their liking [15], eliminating the need to consider data minimization. Even further, the interpretation of the Data Protection Directive (or Fair Information Practices) to collect all data of interest is acquiesced as long as individuals are provided with control over the collected information, i.e., informed consent and subject access rights (this is discussed below in the section on user centricity). In the personal information economy in which data has become an all important resource [22], there is little desire on the part of corporations to reduce the collection of information that they see as aiding them in producing appropriate services and opportunities to consumers [26].

The fact that the existing principles of data protection legislation are not effective in communicating data minimization as an important principle has been confirmed in the findings from privacy roundtables held by the FTC [9]. These findings state that i) there is increasing collection and use of consumer data, ii) that the companies that collect data [...] share the data with multiple entities, and iii) that this growth in data collection and use [is] occurring without adequate concern for consumer privacy.

In order to mitigate these problems, the FTC recommends incorporating privacy into design and compliance with the Fair Information Practices, specifically simplified choice and transparency. By implementing these recommendations, the FTC hopes to increase trust in the data collectors, and create confidence in the market. Though there may be significant differences in the case of the FTC in the United States from EU and other privacy enforcing bodies, this perspective is mirrored in other discussions such as with privacy impact assessments in which individuals are expected to be more likely to contribute their data to trusted organizations. However, given the symbolic nature of privacy solutions for making information systems compliant and the fact that data parsimony is not a primary objective, the compliance approach to privacy risks being associated with similar commercial attempts of the past that target increasing trust, while not mitigating the risks associated with large centralized databases, as is the case with the Truste seal [28]. The limitations of the compliance approach are especially problematic in the context of large-scale systems that have become mandatory, such as road tolling systems and smart energy systems, or de facto mandatory systems like telecommunications (e.g., mobile phones). Each of these systems require large databases that contain highly sensitive information. The dimensions of these databases, as well as their resulting attractiveness to companies, governments, and other organizations, make the risks associated with them very high. Compliance with data protection may help to detect data security breaches, but it is limited in preventing the risks associated with this information collection.

The compliance approach seems to further disregard existing computational capabilities. Though new technological means can lead to surges in the collec-

tion and processing of data, little recognition is given to the fact that technical mimicry of daily activities can be enabled in unintuitive ways and in some cases with much less data than is necessary in the analogue world. Interactions in online systems may require less data then their analogue counterparts for a number of reasons. First, organizations may find that when they transpose some of their workflows to the digital realm, certain information is not needed. For example, Schaar [29] points out that during the development of the ELENA project, not all data fields used on the conventional form were necessary in the digital workflow. He warns against "the tendency to reproduce increasingly complicated bureaucratic systems exactly in information technology [...]". This is almost paradoxical, since data protection in the compliance approach is about creating complex bureaucratic systems expected to enhance data protection and hence privacy.

Second, the information that is necessary to enable the digital equivalents of workflows may be simplified and completed with much less data using the mathematical and computational capabilities of technologies that often transcend the boundaries of our intuition. For example, lossy data compression algorithms that underlie file types like mp3, allowing the downsizing of image and sound files. These algorithms have produced almost indispensable functionality on the internet like voice-over-IP, online broadcasting, etc. While it is not intuitive how a sound file that is substantially reduced in size sounds indistinguishably similar to the original, we have come to accept mp3s as a desirable computational given.

When applied during data collection and processing, mechanisms using unintuitive concepts can allow the data that would usually be found to be adequate, relevant and not excessive in relation to the purpose to be further minimized before collection in centralized databases. For example, while a specific credential may encode the date of birth of a subject, a zero-knowledge protocol is able to prove that the subject is over the age of 18 without revealing the actual date of birth, or any other information [1]. Likewise, with existing research results, systems can be developed where individuals are identified, but it is not possible to observe their activities. For example, Private Information Retrieval (PIR) mechanisms allow authorized users to query a database, but keep the database owner from seeing which database items have been queried. The current compliance documents neither require nor promote the use of such computational capabilities in building compliant systems.

Clearly we are not suggesting that compliance is unimportant. Rather that there is often a disconnect between the intentions of a compliance approach and what happens in practice, and see that all too often the competing tension between enabling the free flow of information and the need to ensure a privacy that mitigates the harms and risks of digitally mediated surveillance often is resolved in favor of the former rather than the latter. The lack of a data minimization principle and an articulation of specific tools that can be integrated into privacy by design practices are examples of this. In order to achieve this, privacy by design again must be about more than individual privacy and the control of digital information (this is articulated further in Section 2.6). It should also be

used as a tool that serves to question the development of more intrusive forms of surveillance that are in fact compliant with legislation and data protection principles.

2.5 User Centricity

Privacy can be defined not only as a matter of the concealment of personal information, but also as the ability to control what happens with ones information. This is in part because the revelation of data is necessary and beneficial under many circumstances, and that control over that data may help to prevent abuses in its collection and processing. This idea is expressed [..] in the term informational self-determination by the German Bundesverfassungsgericht [5] and informational self-determination is also expressed in numerous international guidelines for data protection, e.g., EU Data Protection Directive.

In this context, privacy ensures personal control over the collection, processing, distribution and deletion of ones personal data. Current identity management (IDM) designs seek to accomplish this by putting the user in control. IDM systems allow users to establish and secure identities, describe those identities using attributes, follow the activities of their identities, and delete identities. Using these functions, these systems are expected to address identity theft and provide users with oversight over the collection and processing of their personal data. Research within this approach to privacy seeks to develop cryptographic schemes and protocols to enable strong authentication with anonymous or pseudonymous credentials coupled with, ideally, the anonymous communication systems discussed in Section 2.1. Further, IDM systems depend on policies that are used to negotiate, define and enforce access control rules with respect to revealed information. These core technical aspects are currently being included in commercial projects.

The first widely deployed identity management system was Microsoft Passport, a service which allowed users to sign-up and log in to an account once and then use this account with other cooperating services on-line. The passport system was also the earliest failure, because the system was shunned by third-parties as it locked them in a relationship with a single identity provider that of Microsoft. Interestingly, the ability of the identity provider to observe all browsing habits and interactions of users, and the resulting privacy and surveillance concerns were seen only as a marginal issue. An industry response was to launch Liberty Alliance, now the Kantara Initiative, a consortium that promoted open standards for federated identity management. The emphasis was placed on the fact that different identity providers can inter-operate and make claims about a user (data subject). These claims are relayed by the data subject to relying parties that accept those claims and grant the subject some privileges. This form of third party architecture, a ménage à trois, has since become the dominant pattern for building user-centric IDM systems.

However, the privacy and surveillance concerns raised by MS Passport are not necessarily alleviated with the federated identity management model. While this model introduces multiple identity providers and relies on the user to mediate

communications between them, a cooperating identity provider and relying party can still collude and infer all of a users activities. Thus, the federated model relies on these two parties not cooperating to violate privacy, a trust based system that is, from a technical perspective, a rather weak form of privacy. In addition, given that short-term economic interests so far have overridden privacy and compliance concerns, the sharing of datasets between data collectors and third parties is likely to be a desirable business practice also among identity providers. Under these circumstances, collusion and linking of user activities across the federated identity management landscape remains a substantial risk.

There are a couple of solutions to the privacy risks inherent to the federated identity management architecture. First, users may cache the tokens representing the claims and hide the number of times they use a service, although they will not be able to hide the fact that they use the service. Further, they may carefully tailor identities at multiple identity providers so that even if these collide, they will not be able to link these different identities. However, these solutions once again burden the user with making up for the shortcomings of the system with respect to the promise of control over separation of identities and audiences.

A solution that further mitigates the existing privacy risks depends on the use of what is called anonymous credentials and zero knowledge proofs in the context of IDMs. Based on an elaborate public key infrastructure, when a user registers with an identity provider, she receives a credential from the service provider containing the necessary claims (in this case called attributes). Each time the user wants to interact with a service provider, she will use the credential to prove a logical statement about the attributes in it. She will not hand over the credential, but only prove the absolutely necessary information to the service provider and nothing else, e.g., she will be able to prove that she is over 18 and has residency in a given city. Assuming that anonymity is also guaranteed at the network layer, the service provider learns nothing other than the fulfillment of an attribute by the user.

In this architecture, the identity providers only know about the credentials they provide to the user, while the service providers only learn the minimally necessary information, addressing each of the privacy risks raised with the federated and monolithic identity provider models. As a result, the identity provision and identity use are unlinkable. These schemes can prevent others from breaking privacy regardless of computational abilities. However, this scheme, it is argued, can easily fall prey to fraud due to the anonymity of the users. Hence, particularly within the European research agenda, it has become customary to introduce a functionality to allow a trusted third party to de-anonymise transactions if there is ever such a need.

These user-centric models, which have evolved and developed from lessons learned from the Microsoft passport case and others, have translated the privacy problem into one that places the responsibility for ones data on the individual consumer or user. The objective is to provide simplified solutions that empower and enable consumers to reveal and conceal the information they want to in the context of their choice. However, this user focus is limited in addressing some

of the harms and risks created by digitally mediated surveillance particularly as this user control and empowerment is more illusive than concrete.

2.6 Limitations of User Centricity

The dominant IDM architectures, including their privacy-enhanced variants, place the user in the middle of all communications, mediating the flows of certified attributes from the identity provider to the relying party. Arguably, this position is the ultimate privacy safeguard, as the user is fully in control of information flows. However, in many cases, instead of being a privacy safeguard, this position of the user merely gives an illusion of control.

This is perhaps best illustrated by the analogue to the first widely deployed IDM system, that of Passport. Citizens are issued passports certifying their citizen status of the issuing country, serves to certify some of their attributes, and binds the carrier to a biometric identifier. Passports are an identity card presented at various checkpoints, yet none of these presentations are particularly empowering. Instead, one may argue, there is a double inconvenience of not only being subject to registration and controls, but also to presenting documentation that is part of the identity machinery. While current implementations of IDM systems are less coercive, this only serves to give users a greater illusion of control. Special interfaces allow users to select what they will reveal and to halt transactions if they feel that too much information is required. However, choice at this level, as well as the option to abstain from using a service or online space, does not represent true control. In this case, much like national identification schemes, this IDM only allows a choice between showing a passport or equivalent ID, or abstaining from travelling. In practice the balance of power between the organization requiring identification/authentication and the data subject dictates how much information about the one or the other is exchanged. This is hardly ever the subject of individual negotiation. It again interprets privacy through the lens of the familiar bias, the individualistic approach to privacy with an illusion of personal choice.

Despite inconveniences, IDM architectures are favoured within e-government circles. This may be related to the power of these architectures in which tokens provided by the identity provider mediate all transactions between subject and service providers, even when the user is in the middle of communications. This mediation, through the tokens provided by the identity provider, offers great potential for surveillance. Even in the case of privacy enhanced proposals for IDMs using anonymous credentials, the proposals for trusted third parties that can reveal identities in case of anomalies raises concerns that these only provide a false sense of control over anonymous revelations, which can be revoked at will. However, its proponents often veil this weakening of privacy properties with terms such as accountability and fraud detection, while others denounce it as escrow or simply surveillance by design.

This escrow or surveillance by design critique is based on the reliance on third parties to initiate tracing and identity revocation. The assumption is that third parties are judicial authorities or law-enforcement bound by the rule of law

and privacy legislation. In practice though, there is no tradition of the judiciary holding and managing cryptographic keys. Rather, these third parties are likely to be national security agencies that are commonly entrusted by government with holding and managing such keys. Escrow mechanisms are orthogonal to selective disclosure credentials, and one could see an IDM infrastructure free from surveillance by design. However, this would require moving away from some of the open-ended scenarios for IDM that makes it impossible to predict the effectiveness of any custom-built abuse prevention measures short of full identity escrow. Given a specific security goal, proportionate and appropriate abuse prevention mechanisms that might preserve privacy can be devised, e.g., like double-spending prevention for e-cash, blacklisting of users without revealing their identities, and reputation systems to prevent spam. Yet, it is impossible to build those for the open ended IDM systems: an off switch for privacy becomes the only acceptable mechanism to ensure that any abuse could be traced.

The vision of general purpose privacy-preserving IDM is one of user choice and control, their deployment could have perverse and opposite effects. An alternative to the reversal of a privacy infrastructure into a powerful and self-sustaining surveillance system may lie in reducing the relationship between the three parties back to two. This can be done by allowing the users to continue self-certifying their identity and attributes, as they currently do using the Internet. For some important aspects of identity this is fully sufficient: the OpenID mechanism, as well as self-issued credentials, can ensure for example that two otherwise anonymous transactions are performed by the same person. Research also shows that there is much potential in crowd sourcing for authentication, e.g., a closed network vouches for the identity claims of an individual.

If giving users control is the objective, it is probably best served, not by allowing multiple providers, but by allowing user data to be fully under the control of the user. This allows easy and cheap migration between services, as well as local processing. Data protection regimes, as well as many privacy solutions that aim to support them, are poor substitutes for this simple architecture in which users and the software under their control should be able at all times to fully access, process, copy, and delete information that is held on their behalf to provide a service. Moreover, the technical limitations of IDM systems have to be recognized, specifically that IDM Systems cannot offer any protection to the data once it is in the possession of third party service providers. Control in this case is reduced to users being able to choose which entity can violate their privacy though malice or carelessness.

The proliferation of service providers, and the ephemeral nature of the interactions with many of them, makes it a management nightmare for users to keep track of where their information is being processed. IDM systems offer no solution to this. Further, the inherent difficulties in anonymizing complex interactions makes information held in pseudonymous profiles sensitive. While IDM systems might help keep track of authentication relationships, actual user data stored in services is not mediated through the system, and this requires an audit relationship with all the relevant service providers. The proposals for

functionality to enable subject access rights opens the possibility for users to participate in an army of system auditors, a responsibility smoothly shifted from governments to crowds. Nevertheless, the service providers can forego this functionality by limiting the scope of "personal data" and applying what they claim to be anonymization, which usually does not go beyond simple de-identification. Thus, even when it comes to enforcing strict data protection requirements, IDM systems are only part of the solution. Finally, IDM systems do not recognize how people actually construct and play with their fluid identities when interacting with each other, as well as with private or government services. The current, monolithic, approach to identity management presents inherent limitations also in that respect.

Finally, in order for IDMs to take off, they first need to be "bootstrapped": there needs to be an intake process for which transactions in the off-line world need to occur in order to collect "authentic" data about the users and to bind these users to their data. There is recognition among IDM circles that the best resource for authentic data are government agencies and public services. Governments hold an authority in generating authentic data, coupled with law enforcement to detect and punish those who deceive or refuse to participate in the data collection systems, e.g., identity cards, population registers and tax declarations. Where pressures in financing public services exist, government agencies are expected to participate in the data economy. It is at this crossing that the interest of companies behind IDMs to turn government services into profit, the interest in governments to outsource their services, their parallel interest in population surveillance and, surprisingly, open government data initiatives find themselves in harmony.

User-centric approaches to designing privacy into system development such as is the case with IDM systems have significant advantages over the pre-emptive and compliance approaches when it comes to mitigating the harms and risks of digitally mediated surveillance. Yet these proposals also falter in both similar and distinctive ways from the others. The individualistic perspective does little to alleviate the very social harms present in surveillance practices, and the control that it seeks to achieve is more elusive than these systems seem to suggest. From our vantage, true user control would look different than the third party design inherent in most IDM systems. Further, the role of government agencies is a problematic one, as an entity that is intended to protect individual privacy, but, due to escrow, may maintain agencies that are also enable to override this with limited oversight.

3 Users, Design, Privacy and Digitally Mediated Surveillance

Our development of the above taxonomy takes for granted that inherent in the increase in digitally mediated surveillance are certain harms and risks to citizens, consumers, patients, users, employees, and so on. Although these harms and risks may vary dependent upon context, if we wish to reduce these, we need to

think more clearly about the means and mechanisms for this. While it is not the antidote for surveillance, the notable solution of an increasingly complex and multifaceted understanding of privacy is crucial [3]. However, we must be clear that we are using the appropriate (privacy) tools for the job, and what is made clear in the above taxonomy is that these vary based on how the problem (of privacy) has been defined and what these approaches seek to address. Our intention here is that a discussion about how privacy is integrated into design and a taxonomy of how this is interpreted/translated into the development of new systems is necessary for us to recognize more appropriate solutions and how these solutions are always predicated on how we define the problem.

Each of the technological approaches has different advantages and disadvantages, and each modifies and shifts the power relations inherent in digitally mediated surveillance. What is important is that these choices are understood and made explicit as we move forward in research, policy and deployment. Perhaps the best way in which we can illustrate this is to conceptualize what happens to the individual user based on the approach to the privacy problem defined above.

In the case of pre-emption, the user is assumed to be active; a person that comprehends the technical sophistication and has an awareness of the risks associated with every disclosure in relation to existing databases and surveillance practices. This person can afford to some degree to unlink in a networked world, where the ability to prove connectedness is more important than personal attributes. She controls her transactions and disclosures through technical mechanisms that she has employed in order to avoid the potential of control asserted by surveillance systems that she participates in, a somewhat paradoxical but sometimes necessary relationship to control. This sophisticated approach to privacy protect her individually from explicit infringements against her personal data, however, shows the limited effectiveness of anonymous communications against the social sorting inherent in surveillance remains [33]. Though this does not undermine the need for anonymity and the necessity of such technologies within the information landscape, it shows how pre-emptive tools are useful in certain contexts (anonymous communications) and not others (the social sorting of surveillance).

In the compliance approach, the user is less of a focus. Instead, regulatory bodies focus on organizations that are expected to introduce what we see as a layer of bureaucratic mimicry within information systems. This serves to provide users a degree of trust regarding their use of their data and alleviates the burden of the user having to make risk analyses and disclosure decisions for transactions with each and every organization. However, the trust inherent in privacy by design is useful in some contexts, but does not mitigate the possible harms that may arise from the flexibilities of digital information systems. The transparency and accountability sought in privacy by design may stem from corporate desires to comply with regulation, but in practice, these approaches may be a way of increasing the users' trust in the market with limited guarantees against ongoing and recurrent problems of leaks and abuses of large databases. Further, this compliance is largely focused on revealing where the data is collected and where

it resides, and as long as legislation is complied with, the approach leaves the problems associated with surveillance and profiling practices untouched. Hence, these measures allow the users to take action only post-hoc, as in the case of abuse. While the punishments of data abuse may be severe, it is limited to the mishandling of individual personal data. Protections afforded to the user are only based on what is agreed between policy makers, regulators, governments and corporations. Given the focus on the free flow of information that is necessary for the market to function, this approach leaves the individual users dependent on the negotiation of these different powers and their consensus on how much power should be afforded to individual users.

Finally, in the user centric approach, the user is located in the centre of a ménage à trois relationship and is provided with "control". This control is again limited to transparency over the users own data, and does not directly address the social sorting and other issues inherent in surveillance. The approach is also the most explicit in treating personal data as the property of the user: personal data is constructed as something that the user has and can give away (and take back) in return for services. However, the projected market for identity providers depends on the emergence of a market for quality profiling and data mining on strongly authenticated data and this desire for authentication is part of the problem. In order to address identity theft problems, user-centric approaches force users into a situation where a transaction needs to be bound to authenticated attributes, which makes them liable for the given activity that they partake in. Putting together profiling and strong-authentication across services may lead to an increase in surveillance systems that most privacy technologies can be seen to provide protection against. Identity management affords strong authentication but ends up indicating that every user transaction needs to be authenticated. This solution once again burdens the user with providing (binding) attributes and partaking in a large surveillance system that serves to secure corporate and governmental financial systems. If single sign-on solutions are established as popular applications on the market, the federated and the privacy enhanced versions of identity management will prove themselves more reasonable alternatives. However, the choice to move in this direction demonstrates a process under which new surveillance is created in the name of privacy.

By focusing on the user, we have attempted to illustrate a bit more clearly the implications of how these approaches to differently defined privacy problems can be understood. We have shown that approaches to design in privacy are dynamic and nuanced, and that in different ways these each reduce the harms and risks associated with digitally mediated surveillance while at the same time either not addressing other surveillance issues or creating new ones. The intention of our taxonomy is not to suggest that privacy in design is pointless, rather to enable critical engagement with these practices in ways that allow us to more systematically conceptualize how privacy can counter some of the harms and risks inherent in current and future forms of digitally mediated surveillance.

Acknowledgements This work was supported in part by the European Community's Seventh Framework Programme (FP7/2007-2013) under grant agree-

ment n. 201853 (The DigIDeas Project) and n. 216287 (TAS3 -Trusted Architecture for Securely Shared Services), 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 IWT SBO Project on Security and Privacy for Online Social Networks (SPION).

References

1. Claudio A. Ardagna, Jan Camenisch, Markulf Kohlweiss, Ronald Leenes, Gregory Neven, Bart Priem, Pierangela Samarati, Dieter Sommer, and Mario Verdicchi. Exploiting cryptography for privacy-enhanced access control. *Journal of Computer Security*, 18(1), 2009.
2. Josep Balasch, Alfredo Rial, Carmela Troncoso, Christophe Geuens, Bart Preneel, and Ingrid Verbauwhede. PrETP: Privacy-preserving electronic toll pricing. In *19th USENIX Security Symposium*, pages 63–78. USENIX Association, 2010.
3. Colin J. Bennett. In defence of privacy: The concept and the regime. *Surveillance and Society*, 8(4), 2011.
4. Travis D. Breaux and Annie Antòn. Deriving semantic models from privacy policies. In *6th IEEE International Workshop on Policies for Distributed Systems and Networks*, pages 67–76, 2005.
5. 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.
6. Manuell Castells. *The Rise of the Network Society: Economy, Society and Culture*. Oxford University Press, 2000.
7. Ann Cavoukian. Privacy by design: The 7 foundational principles. Information and Privacy Commissioner of Ontario, Canada, 2009.
8. European Commission. Communication from the commission to the european parliament, the council, the economic and social committee and the committee of the regions: A comprehensive strategy on data protection in the european union. Technical report, European Commission, October 2010.
9. U.S. Federal Trade Commission. Protecting consumer privacy in an era of rapid change: A proposed framework for businesses and policymakers. Technical report, Federal Trade Commission, December 2010.
10. Luca Compagna, Paul El Khoury, Alzbeta Krausova, Fabio Massacci, and Nicola Zannone. How to integrate legal requirements into a requirements engineering methodology for the development of security and privacy patterns. *Artificial Intelligence and Law*, 17(1):1 – 30, 2009.
11. Cynthia Dwork and Adam Smith. Differential privacy for statistics: What we know and what we want to learn. *Journal of Privacy and Confidentiality*, 1(2), 2009.
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. *Official Journal of the European Communities*, (L. 281), November 1995.
13. Batya Friedman, Peter H. Kahn Jr., and Alan Borning. Value sensitive design and information systems. In P. Zhang and D. Galletta, editors, *Human-computer interaction in management information systems: Foundations*, pages 348 – 372. Armonk, New York; London, England: M.E. Sharpe, 2006.

14. Paolo Guarda and Nicola Zannone. Towards the development of privacy-aware systems. *Information and Software Technology*, 51(2):337 – 350, 2009.
15. Seda Gürses, Ramzi Rizk, and Oliver Günther. Sns and 3rd party application privacy policies and their contruction of privacy concerns. In *ECIS 2010*, 2010.
16. U.K. Information Commissioner. Pia handbook, 2009.
17. Daniel Kifer and Johannes Gehrke. l-diversity: Privacy beyond k-anonymity. In *IEEE 22nd International Conference on Data Engineering (ICDE'07*, 2006.
18. Christopher Kuner. *European Data Protection Law: Corporate Compliance and Regulation, Second Edition*. Oxford University Press, 2007.
19. Ninghui Li and Tiancheng Li. t-closeness: Privacy beyond k-anonymity and -diversity. In *IEEE 23rd International Conference on Data Engineering (ICDE'07*, 2007.
20. David Lyon. *Surveillance Society: Monitoring Everyday Life*. Buckingham, UK: Philadelphia, Pa: Open University, 2001.
21. Aaron Massey, Paul Otto, Lauren Hayward, and Annie Antón. Evaluating existing security and privacy requirements for legal compliance. *Requirements Engineering*, 15:119–137, 2010. 10.1007/s00766-009-0089-5.
22. 6 Perri. The personal information economy: Trends and prospects for consumers. In Susanne Lace, editor, *The Glass Consumer: Life in a Surveillance Society*. Bristol: The Policy Press, 2005.
23. Andreas Pfitzmann and Marit Hansen. Anonymity, unobservability, and pseudonymity: A consolidated proposal for terminology. Technical report, Technical University, Dresden, 2008.
24. David J. Phillips. Privacy policy and PETs. *New Media and Society*, 6(6):691–706, 2004.
25. Irma Van Der Ploeg. *The Machine-Readable Body*. Maastricht: Shaker, 2005.
26. Jason H. Pridmore. Consumer surveillance: Customers, choices and cumulative disadvantage. In Lyon David, Kevin Haggerty, and Kirstie Ball, editors, *The International Handbook of Surveillance Studies*. Taylor and Francis Books, (forthcoming).
27. David Rebollo-Monedero, Jordi Forné, and Josep Domingo-Ferrer. 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.
28. N.J. Rifon, R. LaRose, and S.M. Choi. Your privacy is sealed: Effects of web privacy seals on trust and personal disclosures. *Journal of Consumer Affairs*, 39(2):339 – 362, 2005.
29. Peter Schaar. Privacy by design. *Identity in the Information Society*, 3:267–274, 2010.
30. Felix Stalder. Privacy is not the antidote to surveillance. *Surveillance and Society*, 1(1), 2002.
31. Latanya Sweeney. k-anonymity: a model for protecting privacy. *International Journal on Uncertainty, Fuzziness and Knowledge-based Systems*, 10(5):557–570, 2002.
32. S. Warren and L. Brandeis. The right to privacy. *Harward Law Review*, 4:193–220, 1890.
33. Detlev Zwick and Nikhilesh Dholakia. Whose identity is it anyway? comsumer representation in the age of database marketing. *Journal of MacroMarketing*, 2003.