15

Information Technology, Privacy, and the Protection of Personal Data

Jeroen van den Hoven

Information technology allows us to generate, store, and process huge quantities of data. Search engines, satellites, sensor networks, scientists, security agencies, marketers, and database managers are processing terabytes of data per day. A good part of these data are about persons – about their characteristics, their thoughts, their movements, behaviour, communications, and preferences – or they can be used to produce such data[1]. All countries and cultures in the present and past have constrained access to certain types of personal data in some way or the other (see Moore 1984). There are etiquettes, customs, artefacts, technologies, or laws, and combinations thereof, which prevent or proscribe against the use or dissemination of personal information. Walls, curtains, doors, veils, sealed envelopes, sunglasses, clothes, locked cabinets, privacy laws, secure databases, cryptographic keys, and passwords all serve the purpose of preventing individual persons to acquire and use information about other persons. The issues are often discussed in the context of a specific social sector or professional domain, such as health care, social or homeland security, search engines, marketing, or policing. More specifically the issues can be concerned with camera surveillance, the monitoring of Internet communications, the retention of Internet traffic data, the disclosure of passenger lists to security agencies, the availability of individual medical information in the public health system, the linking and matching of databases in social security to detect

[1] There is a widely accepted convention to distinguish between data (raw data), information (meaningful data), and knowledge. My main concern here is with data or the raw material that can be used and interpreted by a variety of methods and tools to serve many different purposes. This paper explores the moral foundations of data protection. In many cases, not much depends on whether we use 'data' instead of 'information'. It is important to realize though that, even when no meaning can be assigned to data (because there is too much data, and it is too difficult to interpret them), it does make sense to think about protecting them because they may start to make sense when new tools and techniques are applied to them, or when they are combined with other data.

fraud, and sifting and trawling through financial databases in order to find suspect transactions. In the near future, they may be about who has access to the scans and digital images of our brains or about who can track and trace tagged personal belongings, everyday objects, and consumer products.

Ethical issues concerning information about persons are typically cast in terms of *privacy.* Privacy is construed as a need, a right, a condition, or an aspect of human dignity. Sometimes it is construed as intrinsically valuable; sometimes it is construed as deriving its value from other sources, for example, from the fact that it is conducive to autonomy or freedom (see Shoeman 1984, Wagner DeCew 1997, Roessler 2005, Nissenbaum 2004, Solove 2002, Velleman 1996, and Nagel 1998). The largest part of privacy research is concerned with the moral justification of a right to privacy. There is little agreement about the most adequate moral justification[2], but there is consensus among privacy scholars about the fact that privacy is important and that privacy is vague, fuzzy, and hard to explicate or pin down (Wagner DeCew 1997).

In public debates about privacy at the beginning of the twenty-first century, there are roughly three positions. First, there is the view that we should stop worrying about privacy, because there is so much personal information available and everyone can know almost everything about everyone, if one would bother to make the effort.

Every credit card payment, Internet search, mobile telephone call, and every movement of a tagged object spawns data about its use and user (van den Hoven 2006). Our life worlds have turned into ambient intelligent environments (Aarts and Encarnacao 2006) which soak up, process, and disseminate personal data. There is so much information that the idea of constraining or controlling the flow in conformity with moral considerations, laws, and regulations is absurd (Safire 2005; Spinello 1997; Quittner 1997).

Second, there is the view that Western democracies cannot afford the high levels of individual privacy that they now attempt to provide. Even if it were technically feasible, high levels of privacy are undesirable[3]. Proponents of this view often also argue along utilitarian lines that modern societies involve large numbers of free moving individuals, exhibit high degrees of mobility, social and institutional complexity, and anonymity, which facilitate free-riding in the form of criminal behaviour, fraud, and tax evasion. In order to mitigate the adverse effects of anonymity and mobility, information about individuals should be made available to governments. It is assumed that groups and communities benefit significantly from knowledge about

[2] For a recent and comprehensive discussion of philosophical accounts, see Roessler (2005). Already, in 1986, Thomas Perry indicated that privacy was a battleground for different positions in applied ethics (Perry 1986).

[3] The most influential proponent of this idea seems to be communitarian thinker Amitai Etzioni (2005). For a short position paper, see Etzioni (1996): '. . . giving up some measure of privacy is exactly what the common good requires'.

their members. Third, another position is to argue that there are good moral reasons to protect individuals from Big Brother, data-greedy companies, and snooping fellow citizens. There are good moral reasons to justify a potent regime of individual rights which constrains access to information about individuals.

In discussions about privacy, those who represent the second view are mostly communitarians. Communitarians emphasize with respect to privacy that it offers a degree of anonymity that facilitates antisocial behaviour, whereas liberalists emphasize the importance of the individual's right to privacy. Michael Walzer has observed in this context that liberalism is, therefore, plagued by free-rider problems, 'by people who continue to enjoy the benefits of membership and identity although no longer participating in the activities that produce these benefits, whereas Communitarianism is the dream of a perfect free-riderlessness' (Walzer 1995). Or as Thomas Nagel has put it: 'Those of us who are not political communitarians want to leave each other some space . . . ' (Nagel 1998, p. 20).

Ever since the debate about privacy was triggered by an article of Warren and Brandeis titled 'The Right to Privacy' at the end of the nineteenth century, the confusion and conceptual controversy concerning privacy has grown and one account of privacy after the other has been suggested (Warren and Brandeis 1890). I will not deal with the possible answers to the question as to what the best conceptual analysis of the term 'privacy' is because we can do without such an analysis and still articulate what bothers us about others having access to information about us that we did not volunteer. The analogy with contemporary epistemology may serve as a cautionary tale. The quest for the best analysis of knowledge has ended in a wild goose chase for necessary and sufficient conditions for the truth of 'John knows that p'. Several decades of philosophical research in epistemology have failed to yield a comprehensive and noncontroversial conclusion. In the case of knowledge, it is not problematic that all resources are used in theoretical epistemology, because there are not many hotly debated practical moral issues concerning knowledge.

Privacy research is different, however, in the sense that we are confronted by urgent and hard privacy problems in practice, and we need to justify and account for legal, practical, technical, and political decisions every day. Privacy issues are high on the political agenda and are hotly debated in software engineering and systems development, health care, e-government, criminal justice, law enforcement, marketing, and e-commerce. In the case of privacy, we need to repair our boat at sea and we cannot take it out of the stormy waters to study it for an indefinite period of time, with uncertain or highly controversial outcomes.

The central role given to the concept of *privacy* in our thinking about the moral issues concerning the protection of personal data obfuscates practical solutions to the everyday problems we encounter in law, public policy, and software engineering concerning them. It lands us in the middle of the controversy between liberal and communitarian political philosophies and the associated conceptions of the Self. Because this controversy cannot be easily decided in the favour of either point of view, I propose to address the central question of the moral problem underlying privacy issues head on: why should we protect personal data; what moral reasons do we have to protect personal data? I would like to construe this question on a par with questions, such as 'why should we protect nuclear reactors, medieval manuscripts, babies, and bird sanctuaries?' In each of these cases, we have good reasons to constrain access, think about visiting hours, stipulate how different persons or groups ought to behave in the vicinity of these entities, and how they may interact with them. In each of these examples, protection takes on a different form and has a different rationale. What would count as a good moral reason to protect personal data and what type of reasons would justify putting limits to the freedom of others to get access to them? This I will discuss in Section 2. First I will discuss our long-lasting interest in personal data.

1. WHY PERSONAL DATA WILL ALWAYS BE IN DEMAND

Personal data will always be in demand. We will continue to amass personal data in the future and questions concerning their protection are therefore unlikely to subside. First, I distinguish between reasons that governments and nongovernmental parties may have to gather information about individuals. Second, I distinguish between reasons for acquiring information about a person that are primarily directed at the interest of the data subject and reasons primarily concerned with the interests of others than the data subject. This gives us four types of reasons for data collection, which help us to understand the logic which drives the accumulation of personal data, now and in the future.

First, government agencies may want to have access to data about citizens to serve them better. In order to provide better services, they will have to know a good deal about them. Government agencies could alert individual citizens to the fact that they are entitled to benefits they did not know about. This type of proactive service delivery to citizens has become more common in recent years. In the Scandinavian countries, citizens seem to be at ease with the idea that the government thinks on their behalf about their welfare and citizens seem to be comfortable with fewer impediments for government to find out details about their individual lives.

Second, the same logic applies to commercial parties and companies. Commercial parties want to be able to serve their customers or clients better. The more they know about them, the better they can fine-tune their propositions to their preferences and needs. Attention of consumers is scarce and commercial proposals therefore need to raise the immediate interest of potential customers. Many customers have no problems with alerts by businesses which draw their attention to bargains in the areas they are interested in and many seem willing to volunteer personal information in exchange for bargains.

Third, companies or commercial parties also have strictly prudential reasons to collect or accumulate personal data, both about their customers, their partners in transactions and their employees. These reasons are not at all concerned with the interests of the data subjects. Transactions between private parties always present chances for exploitation. Other parties can break their promises, break the contract, or buy without paying. In these cases, adequate information about the partners one is dealing with, for example, information about credit risks or commercial past performance are thought to be extremely helpful in gauging the risks associated with the transaction. Perhaps even more fundamental; in order to be able to trust parties, re-identification of individual partners is a necessary condition for building up a reliable picture of someone's track record in interactions. In game theory and the study of iterated prisoners' dilemmas, the (re)identification of players in consecutive rounds is simply assumed. In the real world, however, (re)identification is often a practical problem. Computer applications concerning the identity and relevant properties of individuals are widely used to counter the problem of reliable identification and authentication of persons in private interactions. Information technology holds up the promise that it can deal with the knowledge deficit we are often confronted with in our dealings with strangers and people we know very little about.

The deployment of information technology in the relation between employer and employee may be accounted for in terms of the Principal–Agent theory, according to which the Principal, for example, an employer, always has the problem of making sure that the employee (the agent) is doing what he or she ought to do, when he or she is out of sight. The Principal, therefore, has to make so-called agent cost and has to monitor the agent and check what he or she is doing. This accounts for the incredible explosion of workplace monitoring and surveillance by means of logging, CCTV cameras, smart badges, and black boxes in cars.

Finally, government also has reasons to try and get information on citizens, which are not primarily and directly concerned with the individual interests of individual citizens about whom information is collected, but are primarily concerned with the public good. One of the central tasks of government is the production and maintenance of public goods. One of the central problems of the management of public goods is managing the access to public goods and more specifically the problem of excluding those who are benefiting from the public good without contributing to the maintenance and reproduction of the public good. This category of individuals is referred to as 'free riders'. The containment of free riders is a central task for government. Free riders can thrive and exist only if they are anonymous.[4] If they can be identified government can affect their pay-off matrix and their self-interested calculation. Identifying information is thus very helpful to governments as managers of public goods.

These four types of 'logic of the situation' explain why a range of actors in the government and the market sector, will engage in massive computer-supported data collection, and will continue to do so for reasons both concerning the good of the data subject or the good of others than the data subject. They will always welcome and use new developments in information technology that may support their attempts to reach these goals.

What do people object to when they object to gathering personal data in these and other cases? We need to distinguish between different objects of concerns. When a man who enters an almost empty restaurant picks the table right next to me, there are several things I may object to. First, I may not at all be concerned with *my* personal data, but rather with *his* personal data. I may in other words not be concerned with what this person is learning about *me*, but rather with what I am learning about *him*. I just don't want to know the things that I am about to find out about him. A further concern may be the fact that my choice not to learn anything about anybody at that moment is preempted by his decision to sit next to me. I will hear what he orders, smell his aftershave, hear him turn the pages of his newspaper and hear his mobile phone conversation. A perceptual relation is imposed upon me, because he chose to move into my perceptual field without my consent. In that sense, the setting is turned into a source of personal data about the intruder. Data are stored in my brain, and I may forget about them immediately or may remember them later.

A second possible object of my discontentment in the restaurant may be that this person has manoeuvred himself into a position where he is now able to acquire data about *me*, which can be passed on – about what I was having for dinner that evening, what I was wearing, and so forth. He may decide to tell others, or, in secret, make video footage of me munching my garlic bread. Even if I would know that the merely onlooking person would not be recording or storing information in an external information carrier, or would not be able – as a result of a rare brain disease – to retain the data acquired beyond the specious present, I could still feel uncomfortable and awkward because the imposed perceptual relationship heightens my awareness of myself and forces an external – and not freely chosen – perspective upon me.

The ethics of data protection is first and foremost about the second and third type of grievance of the lonely diner. These grievances come under the heading of informational privacy or tort privacy and need to be distinguished

[4] De Jasay (1989) observes: '... it is not non-exclusion that makes retaliation impossible ..., but anonymity of the free-rider'.

from the first problem sketched above (the right 'to be left alone') and also from what has been termed 'decisional or constitutional privacy', that is, the right to decide without government interference – for example, the right to decide in which kinds of sexual behaviour to engage between consenting adults in the privacy of one's bedroom, or to decide to have an abortion or to use contraceptives. Sandel refers to the latter as the *new privacy* and to the former as the *old privacy* (Sandel 1989). In Sandel's classification, data protection is about the old privacy.

2. PERSONAL DATA

Personal data are and will remain a valuable asset, but what counts as personal data? If one wants to protect X, one needs to know what X is.

Before we start answering this question a couple of things need to be observed about personal data. First of all, personal data are multiple-realizable, that is, they may be stored in different places and in different media. They may be generated and acquired by different types of information processors, whether human or artificial and silicon-based, or a combination thereof. Second, data may be generated by means of a variety of methods, techniques, and artefacts. Individuals may be monitored by cameras, by persons using binoculars, by scanners which track RFID tagged items they carry around, a discussion may be overheard, or agencies may trawl and sift through databases. And, finally and importantly, data do not have a meaning separate from the context in which they are used.

So when is someone else acquiring and storing information *about* me? When is someone processing my personal data? Let's consider the following two claims C1: '*X* is in restaurant *A* at time *t1*' and C2: '*Y* is in Restaurant *A* at time *t1*'. Is C2 *about X*? C2 presents itself obviously as information about *Y*. When looked at in isolation '*Y* is in the Restaurant at time *t1*' does not tell you anything about *X*, but when combined with C1, it does provide information about *X* which was not contained in C1. Good detective stories often present information to the reader which is seemingly irrelevant to the crime or to the biography of the protagonist, but later turns out to be, in an unexpected sense, *about* the murderer or his victim. As the story unfolds and the plot unravels, the insignificant piece of information is situated in a context where it suddenly picks out an individual. We suddenly see how the insignificant and seemingly irrelevant piece of information suddenly applies to the protagonist.

We are introduced to people and get to know people under certain descriptions and modes of presentation or 'guises' as H.-N. Castaneda has called them (Castaneda 1989). Some individuals present themselves in misleading ways or 'dis-guise' themselves. Personal data believed to be applicable to a person are often stored in different mental files because there were different representations or self-presentations of the person (let us call them m1 and m2), while there is no belief to the effect that m1 = m2. New beliefs about the person under mode of presentation m1 can thus not be added to the beliefs filed under m2[5].

The practical importance of identification may be illustrated by a discussion of attempts to uncover tax evaders[6]. The government registers all bank accounts in the owner's true name. Two names for the same account holder would ideally be codesignative. Tax fraud, however, is about having accounts under different names so that the noninterchangeability of the names blocks the tax office's access to the person holding accounts under different names.

> 1. The tax authorities know that taxpayer x has more than $1.000.000 in the bank
> 2. Taxpayer x also has an additional $500.000 in an account under the false name 'y'
> 3. Since x=y, the tax authorities know that y has more than $1.000.000 in the bank
>
> By mere manipulation of symbols we have arrived at statement (3), which is clearly false.... Although 'x' and 'y' are both names for the same tax evader, they are not interchangeable because each name is associated with a different sense, with a different way in which the reference is given. (Ortiz Hill 1997, p. 117)

Being able to recognize and (re)identify people unambiguously and in a coordinated way is a very important feature of human life. Without the ability to know 'who is who', and without the ability to tell people apart and physically locate them and thereby to 'arrest' them, modern nation states cannot get off the ground and could not be sustained (see Caplan and Torpey 2001). The state's monopoly of violence and the exercise of legitimate power over individual citizens could not be effective if the state were unable to identify individuals in a straightforward referential sense, that is, in a way that allows for physical encounters, such as arrests and imprisonment.

'The man at the table next to me', 'a former colleague', 'the guy with the awful aftershave', 'John', 'John Smith', 'the guy who always takes the 9.45 train and gets off at central station', 'the owner of a blue Ford', 'the person who collected 200 Euro at the teller machine in the Centre of Amsterdam at 14:21:11 at August 1 2005', 'the person on the CCTV tape who put two orange boxes in the trunk of a blue Ford', 'the owner of a bank account 1234567', 'the person on Flight Q1 from Sydney to London on 2 October 2005 in seat 55c', 'the idiot with the baseball cap', 'the guy who dumped Alice' these descriptions could all be about different persons, but they could

[5] John Perry (2002) observes that 'What unifies files, and makes two beliefs about the same person relevant to each other, is not that they are about the same person, but that they contain the same notion or linked notion.'

[6] I have taken this example from Claire Ortiz Hill (1997, pp. 116–117).

also be about the same person (in which case the descriptions provide a lot of information about John Smith).

I could rapidly expand my knowledge of the persons who chose to sit next to me in the restaurant by making one or two identifications, for example, 'the person next to me is identical with the guy who dumped Alice' and 'the guy who dumped Alice owns a blue Ford'. Ruth Gavison's anecdote is instructive in this context:

> Consider the famous anecdote about the priest who was asked, at a party, whether he had heard any exceptional stories during confessionals. 'In fact', the priest replied, 'my first confessor is a good example, since he confessed to murder'. A few minutes later, an elegant man joined the group, saw the priest, and greeted him warmly. When he asked how he knew the priest, the man replied: 'Why, I had the honour of being his first confessor'. (Gavison 1980)

Gavison presents this anecdote in the context of a discussion of the various problems associated with the clarification of the notion of privacy. It all starts with the requirement 'that for a loss of privacy to occur, the information must be 'about' the individual' (Gavison 1980). It is essential to the ethics, law, and technology of data protection to be able to articulate what counts as worthy of protection and what does not, in other words to be able to articulate and define what counts as personal data, and what does not. The legal definition is not of much help here. According to the very influential EU data- protection laws, personal data are characterized as follows: '"personal data" shall mean any information relating to an identified or identifiable natural person ("data subject"); 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' (European Parliament 1995). The priest initially did not provide personal data according to the standard legal definition of personal data.

There is a basic ambiguity here, however, that needs to be brought out because it is relevant to ethical discussions of the protection of personal data. Keith Donellan distinguished between *referential* use of descriptions and *attributive* use of descriptions[7]. 'The owner of a blue Ford living in postal code area 2345' could have more than one individual satisfying the description, and the user of these descriptions may not have a particular individual in mind; he just thinks about the owner of a blue Ford 'whoever he is'. 'The owner of a blue Ford', however, could also be used referentially, when we have a particular person in mind or in attendance. 'The man sipping his whisky' (pointing to the person at a party) is used referentially, and is *about* the person the speaker mistakenly thought was drinking whisky,

[7] First suggested in my 'Information Technology and Moral Philosophy' (PhD thesis, Erasmus University Rotterdam, 1995; see also Donellan 1966).

even when it turns out he is having apple juice instead of whisky, and there is, strictly speaking, no one over there sipping his whisky.

The important thing to note is that both attributively used and referentially used descriptions figure in epistemic and doxastic strategies to collect information on people and directly or indirectly help to expand our knowledge about them. Both represent *identity-relevant information.* One may open a mental or another type of file on a person under the label 'the murderer of Kennedy', in the same way crime investigators do, in the hope to find out more information about this person who ever he is, or turns out to be. These initially nondescript identifications may eventually lead to a physical encounter (i.e., arrest or interrogation) later. The history of a particular criminal investigation is at the same time the history of filling the file[8] with identity-relevant information.

The referential reading of 'personal data', 'identity' and 'identifiability' of the EU data-protection laws leads to an unduly narrow construal of moral constraints on the use of personal data. As a result, attributively used descriptions could go unprotected. This seems a major weakness of data-protection regimes, because we know that large amounts of data are used attributively in marketing and homeland security investigations, for example, and are the stepping stones to find out about people. One could have a file on an owner of a blue Ford, and add a long list of descriptions, all used attributively, but one piece of information added to the rich and anonymous file could suddenly make the data set uniquely referring.

It may well be the case that given the prominence and importance of identity management technology, RFID technology, profiling and data mining, and genetic profiling, we need to have a new look at the dominant referential interpretation of personal data. Instead of defining the object of protection in terms of referentially used descriptions, we need to define the object of protection in terms of the broader notion of 'identity relevant information' (van den Hoven and Manders, 2006).

3. MORAL REASONS FOR PROTECTING PERSONAL DATA

Having qualified the object of protection and, thereby, extended the scope of data protection, from referentially used personal data to both referentially and attributively used identity-relevant data, I will now discuss four types of moral reasons for engaging in data protection, that is, moral justifications

[8] The world of personal data contains objective representations, singular reference to persons by means of descriptions and proper names, passwords, personal identification numbers, and user names. We use descriptions when we are not presented or acquainted with persons, when there is no *de re* or perceptual thought. According to John Perry (2002) thinking of a person by name or description can be construed in terms of 'calling up a file on that individual'.

for constraining actions[9] regarding identity-relevant information. These moral reasons provide the grounds to have principles like those of the 1995 EU data-protection act and the Organisation for Economic Cooperation and Development (OECD) principles in place. These legal regimes give to individuals autonomy and the right to control their personal data.

3.1. Information-Based Harm

The first type of moral reason for thinking about constraining the flow of identity-relevant data is concerned with the prevention of harm. Information is a very useful thing for criminals and crooks to have. A random attack in the street on an anonymous individual does not require information, but a bank robbery requires a good deal of intelligence and information, planning, and foresight. Some harms could not have been inflicted (or at least not as easy) if certain information would not have been available. Let's refer to this type of harm as 'information-based harm'. Cybercriminals and malevolent hackers use databases and the Internet to get information on their victims in order to prepare and stage their crimes. One of the most pressing problems is 'identity theft' and identity fraud, which brings high risk of financial damages and emotional distress. One's bank account may get plundered and one's credit reports may be irreversibly tainted so as to exclude one from future financial benefits and services. Stalkers and rapists have used the Internet and online databases to track down their victims. They could not have done what they did without access to electronic resources and without accessing some of the details of their victim's lives.

In an information society, there is a new vulnerability to harm done on the basis of personal data – theft, identity fraud, or straightforward harm on the basis of identifying information. Constraining the freedom to access information of persons who could cause, threaten to cause, or are likely to cause information-based harm can be justified on the basis of Mill's Harm Principle. Protecting identifying information, instead of leaving it in the open, diminishes epistemic freedom of all to know, but also diminishes the likelihood that some will come to harm, analogous to the way in which restricting access to firearms diminishes both freedom and the likelihood that people will get shot in the street. In information societies, identity-relevant information resembles guns and ammunition. Preventing information-based harm clearly provides us with a strong moral reason to limit the access to personal data. Arguments against central databases with personal data of individual citizens in The Netherlands often makes reference to World War II when the Nazis occupied The Netherlands and found a well-organized population registration very conducive to their targeting and deportation of the Jews in Holland. It would be strange, however, to claim that the Nazis violated the

[9] These actions include: generation, acquisition, processing, and dissemination.

privacy of the Jews. A better description seems to say that they used insufficiently protected personal data to take people out of their houses and send them to the concentration camps. Access to personal information made possible the most horrible of all harms. This is the first thing we want to prevent, and we do it by effectively protecting identity-relevant information.

There is, of course, a broad range of harms to individuals that can be inflicted on the basis of personal information. Someone's career may be systematically corroded by the piecemeal release of selected information. This may start to add up in the eyes of others, and lead to serious reputational harm.

Another type of harm could be the harm that lies in classifying people in such a way that their chances of getting some good are diminished. Being classified as Muslim in many Western countries implies a reduced chance of getting a job. Accumulative information-based harm would refer to the releasing snippets of identity-relevant information at different occasions on the basis of which others may eventually form a rich and comprehensive picture of a person and inflict harm on him or her.

3.2. Informational Inequality

The second type of moral reason to justify constraints on our actions with identity-relevant information is concerned with equality and fairness. More and more people are keenly aware of the benefits the market for identity information can provide them with. If a consumer buys coffee at the modern shopping mall, information about that transaction is generated and added to his file or profile. Many consumers now begin to realize that every time they come to the counter to buy something, they can also *sell* something, namely, the information about their purchase or transaction, the so-called transactional data. Likewise, sharing information about ourselves on the Web with Web sites, browsers, and autonomous agents may pay off in terms of more and more adequate information (or discounts and convenience) later. Many privacy concerns have therefore been and will continue to be resolved in *quid pro quo* practices and private contracts about the use and secondary use of personal data. But, although a market mechanism for trading personal data seems to be kicking in on a global scale, not all individual consumers are aware of their economic opportunities, and if they are, they are not always in a position to trade their data or pursue their interests in a transparent and fair market environment so as to get a fair price for them. The use of RFID chips in consumer products in shops, the use of extensive cross-domain consumer profiling combined with dynamic pricing may facilitate price discrimination.

Consumers do not always know what the implications are of what they are consenting to when they sign a contract for the use of identity-relevant information. We simply cannot assume that the conditions of the developing market for identity-relevant information guarantees fair transactions by independent standards. Constraints on the flow of personal data need to be put in place in order to guarantee equality of arms, transparency, and a fair market for identity-relevant information as a new commodity.

3.3. Informational Injustice

A third and very important moral reason to justify constraints on processing of identity-relevant information is concerned with justice in a sense which is associated with the work of Michael Walzer. Michael Walzer has objected to the simplicity of Rawls's conception of primary goods and universal rules of distributive justice by pointing out that 'there is no set of basic goods across all moral and material worlds, or they would have to be so abstract that they would be of little use in thinking about particular distributions' (Walzer 1983, p. 8). Goods have no natural meaning; their meaning is the result of sociocultural construction and interpretation. In order to determine what is a just distribution of the good, we have to determine what it means to those for whom it is a good. In the medical, the political, and the commercial sphere there are different goods (medical treatment, political office, money) which are allocated by means of different allocation criteria or distributive practices: medical treatment is allocated on the basis of need, political office on the basis democratic election, and money on the basis of free exchange. What ought to be prevented, and often is prevented as a matter of fact, is dominance of particular goods. Walzer calls a good *dominant* if the individuals that have it, because they have it, can command a wide range of other goods. A monopoly is a way of controlling certain social goods in order to exploit their dominance. In that case, advantages in one sphere can be converted as a matter of course into advantages in other spheres. This happens when money (commercial sphere) could buy you a vote (political sphere) and would give you preferential treatment in health care (medical), would get you a university degree (educational), and so forth. We resist the dominance of money – and other social goods for that matter (land, physical strength) – and we think that political arrangements allowing for it are unjust. No social good *x* should be distributed to men and women who possess some other good *y* merely because they possess *y* and without regard to the meaning of *x*.

What is especially offensive to our sense of justice is, first, the allocation of goods internal to sphere *A* on the basis of the distributive logic associated with sphere *B*, second, the transfer of goods across the boundaries of separate spheres and third, the dominance and tyranny of some goods over others. In order to prevent this, the 'art of separation' of spheres has to be practised and 'blocked exchanges' between them have to be put in place. If the art of separation is practised effectively and the autonomy of the spheres of justice is guaranteed then 'complex equality' is established. One's status in terms of the holdings and properties in one sphere is irrelevant – ceteris paribus – to the distribution of the goods internal to another sphere.

Walzer's analysis also applies to information. The meaning and value of information is local and allocation schemes and local practices that distribute access to information should accommodate local *meanings* and should, therefore, be associated with specific spheres.

Many people do not object to the use of their personal medical data for *medical* purposes, confined to the medical sphere, whether these are directly related to their own personal health affairs, to those of their family, perhaps even to their community or the world population at large, as long as they can be absolutely certain that the only use that is made of it is medical, that is, to cure people from diseases. They do object, however, to their medical data being used to classify them or disadvantage them socioeconomically, to discriminate against them in the workplace, refuse them commercial services, deny them social benefits, or turn them down for mortgages or political office.

They do not mind if their library search data are used to provide them with better *library* services, but they do mind if these data are used to criticize their tastes and character. They would also object to these informational cross-contaminations when they would benefit from them, as when the librarian would advise them a book on low-fat meals on the basis of knowledge of their medical record and cholesterol values, or when a doctor asks questions on the basis of the information that one has borrowed a book from the public library about AIDS.

We may thus distinguish a third moral reason to constrain actions regarding identity-relevant information: prevention of 'informational injustice', that is, disrespect for the boundaries of what we may refer to, following Michael Walzer, as 'spheres of justice' or 'spheres of access'. What is often seen as a violation of privacy is often more adequately construed as the morally inappropriate transfer of personal data across the boundaries of what we intuitively think of as separate 'spheres of justice' or 'spheres of access' (van den Hoven 1999; van den Hoven and Cushman 1996).

A couple of illustrations are in order. When government agencies, such as social security agencies, outsource part of their operations to commercial banks, the part of the bank that will take care of the public tasks needs to be separated from the commercial branches. Software protections are put in place, referred to as 'Chinese Walls', which separate the commercial from the public social security sphere. In this way, a Walzerian *blocked exchange* for personal data is implemented, and the *art of information sphere separation* is put into practice.

We have seen a similar normative logic of spheres being operative in constraining cookies to retrieve information across the boundaries of top level domains. We do not mind if the .com site we visit collects information about our search profile on that particular site. We may not even mind if .com sites exchange information. We probably would mind if .com sites used information from .org sites or .gov sites, or vice versa. The lessons learned from the so-called DoubleClick case[10], where clickstream data were collected by cookies working across sites in different top-level domains seem to confirm these Walzerian intuitions about blocked exchanges between spheres.

A Walzerian account along these lines also accommodates the idea, incorporated in many legal data-protection regimes in the world, of 'purpose specification and use limitation' which ensures that information is not used outside the area for which informed consent was given by the data subject. Helen Nissenbaum has introduced the term 'contextual integrity' to refer to these Walzerian type constraints (Nissenbaum 2004). According to Nissenbaum, the benchmark of privacy is contextual integrity. She distinguishes between norms of appropriateness and norms of flow or distribution. Contextual integrity is maintained when both types of norms are upheld, and it is violated when either of the norms is violated.

Nissenbaum does not provide an account of the nature of the context boundaries. A Theory of Sphere or Context Boundaries is crucial because boundaries are disputed, fuzzy, in flux and deemed important. Without such an account, the idea of separate spheres or contexts is practically empty. Wiegel, Lokhorst, and van den Hoven provide a reconstruction of the idea of a boundary between two spheres in terms of a list of deontic statements about which actions with data are (not) permitted or (not) obligatory (Wiegel, van den Hoven, and Lokhorst 2005; van den Hoven and Lokhorst 2002). Per case, or type of case, we need to draw the boundaries and argue for the deontic constraints that we want to impose. In the context of software engineering, this comes down to a specification of a fine-grained authorization matrix and role-based access management scheme. In the design of a hospital information system for example, difficult privacy issues may be resolved by deciding in which situation which professionals can do what to which types of information. Can the janitor print electronic patient records? No. Can the nurse change lab tests? No. Information maps are thus drawn up and 'privacy issues' are addressed in detail. Moral arguments about privacy are given a distributed treatment and, instead of discussing 'The Privacy Issue in Health Care' in abstracto, we address more tractable and more precise questions[11].

3.4. Moral Autonomy and Moral Identification

A fourth type of moral reason for constraining the flow of identity-relevant information could be referred to as *moral autonomy*, that is, the capacity to

[10] For case descriptions, see: http://www.epic.org/privacy/doubletrouble.

[11] Wiegel, van den Hoven, and Lokhorst introduce the idea of 'deontic isographs' – boundary lines, composed of lists of implementable deontic statements which apply equally and, therefore, connect positions in the information landscape to positions with equal deontic status, between which data may flow. See also Moor (2006).

shape our own moral biographies, to present ourselves as we think fit and appropriate, to reflect on our moral careers, and to evaluate and identify with our moral choices, without the critical gaze and interference of others and without a pressure to conform to the 'normal' or socially desired identities. We want to be able to present ourselves and be identified as the ones we identify with.

David Velleman, in his analysis of shame and privacy, draws attention to self-presentation as a constitutive feature of moral persons, namely their capacity and need for self-presentation. What it means to be a person is, according to Velleman, to be engaged in self-presentation. Persons 'have a fundamental interest in being recognized as a self-presenting creature' (Velleman 2001). Failures of privacy and the accompanying emotion of shame are not so much about disapprobation concerning what is revealed when others get access to information we did not volunteer, but are about disqualification of the person who failed to prevent the revelation. Teenagers are very open in their interactions and communications on the Web 2.0. Nudity and explicit material may sometimes leak out of their circle of chat friends. The content, it seems, is not what embarrasses them, but the fact that they failed to manage their public face, and that, as a result, their carefully cultivated identity was spoiled. Therefore, the realm of privacy, according to Velleman, is the central arena for threats to one's standing as a social agent.

A *moral* person is thus characteristically engaged in self-presentation, but, at the same time, she experiences the normative pressures which public opinion and moral judgements of others exert. When information about Bill becomes available, it facilitates the formation of beliefs and judgements about Bill. Beliefs and judgements about Bill, when he learns about them, when he suspects that they are made, or fears that they are made, may bring about a change in his view of himself. They may induce him to behave and feel differently than he would have done without them. This is what Berlin calls 'the most heteronomous condition imaginable'[12]. What others know about you can radically affect your view of yourself, although seeing yourself as others see you does not necessarily make your view of yourself more true or more adequate (Benn, 1988).

Stereotyping is an extreme case of casting people and preempting their choice to present themselves. Modern individuals who have cast aside the ideas of historical necessity, and who live in a highly volatile socioeconomic environment, confront a great diversity of audiences in different roles and settings, the rigging of one's moral identity by means of public opinion, beliefs, and judgements of others is felt as an obstacle to 'experiments in living', as Mill called them. The modern individual wants to be able to

[12] 'I cannot ignore the attitude of others with Byronic disdain, . . . for I am in my own eyes as others see me, I identify myself with the point of view of my milieu' (Berlin 1969).

determine himself morally or to undo his previous determinations on the basis of more profuse experiences in life or on the basis of additional factual information. Some have argued that privacy creates a time out from social morality, in order to engage in ever new experiments in living. Privacy covers purely self-regarding acts and, therefore, implies a right to nonjustification (Monteiro 2004). As Newton Garver aptly put it:

> Contemporary freedom and choice go farther than Mill suspected – we all choose our identities, and make that choice from among a heterogeneous set of data, . . . we rarely choose our nationality, sex or religion, but we do choose to make these data part of our identity. (Garver 1990)

The conception of the person as being morally autonomous, as being the author and experimentator of his or her own moral career, provides a justification for constraining others in their attempts to engineer and directly or indirectly shape the subject's identity, either by stereotyping, or by the application of identity-management tools and techniques. Data-protection laws thus justifiably provide protection against the fixation of one's moral identity by others. They do so by requiring informed consent for the processing of identity-relevant information. If there are domains where for obvious reasons individuals in well-ordered societies cannot be allowed to write their own biographies from cover to cover, they at least should be allowed to write those parts that are amenable to it and individuals should be given an opportunity to authorize the parts that were, or had to be, written by others.

A further explanation for the importance of respect for moral autonomy may be provided along the following lines. Factual knowledge of another person is always *knowledge by description.* The person himself, however, does not only know the facts of his biography, but he is the only person who is *acquainted* with the associated thoughts, desires, emotions, and aspirations. However detailed and elaborate our files and profiles on a particular individual may be, we are never able to refer to the data subject as he himself is able to do. We may only approximate his knowledge and self-understanding[13].

Bernard Williams has pointed out that respecting a person involves 'identification' in a very special sense, which I refer to as 'moral identification'.

> . . . in professional relations and the world of work, a man operates, and his activities come up for criticism, under a variety of professional or technical titles, such as 'miner or 'agricultural labourer' or 'junior executive'. The technical or professional attitude is that which regards the man solely under that title, the human approach that which regards him as a man who has that title (among others), willingly, unwillingly, through lack of alternatives, with pride, etc. . . . each man is owed an effort at

[13] Russell says, 'When we say anything about Bismarck, we should like, if we could, to make the judgement which Bismarck alone can make . . . In this we are necessarily defeated' (Russell 1978, p. 31).

> identification: that he should not be regarded as the surface to which a certain label can be applied, but one should try to see the world (including the label) from his point of view. (Williams 1973)

Moral identification thus presupposes knowledge of the point of view of the data subject and a concern with what it is for a person to live that life. Persons have aspirations, higher-order evaluations, and attitudes, and they see the things they do in a certain light. Representation of this aspect of persons seems exactly what is missing when identity-relevant data are piled up in our databases and persons are represented in administrative procedures, are profiled or construed in statistical terms. Thomas Nagel observes that 'To really accept people as they are requires an understanding that there is much more to them than could possibly be integrated into a common social space' (Nagel 1998, p. 16).

The simple identifications made on the basis of our data fall short of accepting and respecting the individual person, because they will never match the identity as it is experienced by the data subject. It fails because it does not conceive of the other on his or her, own terms.

Because we feel we have inaccessible qualitative aspects of our own private mental states – that is, that we have hopes and purposes and there is something that it is like to have them which cannot be known from the outside – we insist on epistemic modesty on the part of others in their claims to know who we are and in their involvement in the management of our identities. Moreover, we see ourselves as our own moral projects, subject to moral development and capable of moral improvement, so the result of the management of our identities seems a premature fixation of what is an essentially dynamic project.

An outsider's understanding of a person needs to include, ideally, not only the objective representations, but also what he wants or hopes to be, his gratitude or pride or shame or remorse, and how the person interprets them. These conditions are conditions of the whole person. The very object of the outsider's interpretation ought to aim at representing and understanding the person's second-order as well as first-order attitudes, which is not only difficult, but impossible in principle.

> ... the apprehension of the mind of another person may thus only count as knowledge to the extent that it can approximate to this kind of awareness... such an approximation can never be more than a very distant one. (Moran 2001, p. 154)

When a person is considered as 'one person among others', his attitude and self-directed reactive attitudes (his shame or shamelessness) expresses the kind of person he is. It is the sort of thing we take into account in determining how we feel about him. Moran argues that not doing so would be wrong because it would be failing to respect the 'total evidence' of the case. For responding to what he did with shame, pride, or gratitude constitutes a new fact about him, which is morally salient and provides part of the total evidence of who he is. Anything less would not only be wrong, but also epistemically irresponsible (Moran 2001, pp. 182–183).

Respect for privacy of persons can thus be seen to have a distinctly epistemic dimension. It expresses an acknowledgement that it is impossible to know other persons as they know and experience themselves. Even if we could get it right about persons at any given point in time, by exhibit of extraordinary empathy and attention, then it is highly questionable whether the data subject's experience of himself, as far as the *dynamics* of the moral person is concerned, can be captured and adequately represented. The person conceives of himself as dynamic and as trying to improve himself morally. The person cannot be identified, not even in the sense articulated by Bernard Williams, with something limited, definite, and unchanging. The person always sees itself as becoming, as something that perhaps even has to be *overcome,* not as a fixed reality, but as something in the making, something that has to be improved upon. As Gabriel Marcel puts it, the individual's motto is not *sum* (I am) but *sursum* (higher)[14].

CONCLUSION

I have argued that data about persons are very important and will remain important and much sought after in the future. I have argued that personal data need to be construed in a broad sense to include attributively used descriptions. I have provided four moral reasons for the protection of personal data. The first three reasons (concerning avoiding harm, preventing exploitation in markets for personal data, and preventing inequality and discrimination) can be shared by both liberals and communitarians; they both oppose inflicting harm, exploitation, and discrimination. The fourth reason, however, invokes the essentially contested liberal self. It is the liberal self that wants to decide what to think of itself and what to make of him or herself. And how others should identify him or her, preferably identify with what he himself identifies with. There is probably always over-determination of these moral reasons at stake, so that these reasons can be invoked simultaneously in the moral discussion about data protection.

This analysis opens up a space of potential agreement between parties that are usually deeply divided concerning 'privacy issues'. The analysis provides three central and weighty reasons to engage in the protection of personal information in the light of new technologies, which they can share. So, instead of arguing over necessary and sufficient conditions of 'privacy', we can actually think about designing smart schemes of justified and implementable deontic constraints on flows of personal data.

[14] This is also suggested by Isaiah Berlin: '. . . what I may seek to avoid [is to be] insufficiently recognized, . . . a statistical unit without identifiable . . . purposes of my own.' (Berlin 1969).

References

Aarts, E., and Encarnacao, J. L. (Eds.). 2006. *True Visions. The Emergence of Ambient Intelligence.* Berlin/Heidelberg: Springer.

Bach, K. 1987. *Thought and Reference.* Oxford: Oxford University Press.

Benn, S. 1988. *A Theory of Freedom.* Cambridge, UK: Cambridge University Press, pp. 277, 288 (resp).

Berlin, I. 1969. *Four Essays on Liberty.* Oxford: Oxford University Press, p. 156, n. 1.

Caplan, J., and Torpey, J. C. (Eds.). 2001. *Documenting Individual Identity: The Development of State Practices in the Modern World.* Princeton: Princeton University Press.

Castaneda, H.-N. 1989. *Thinking, Language and Experience.* Minneapolis: University of Minnesota Press.

De Jasay, A. 1989. *Social Contract, Free Ride: A Study of the Public Goods.* Oxford: Clarendon Press.

Donellan, K. 1966. Reference and definite descriptions. *Philosophical Review, 75*, 281–304

Etzioni, A. 1996. Less privacy is good for us. *The Responsive Community, Summer*, 11–13. Available at www.gwu.edu/~ccps/etzioni/M28.pdf.

Etzioni, A. 2005. Limits of privacy, in A. I. Cohen and C. H. Wellman (Eds.), *Contemporary Debates in Applied Ethics.* Oxford: Blackwell, pp. 253–262.

European Parliament. 1995. *Directive 95/46/EC* on The protection of individuals with regard to the processing of personal data and the free movement of such data, adopted October 24 1195, *Official Journal, 281*, pp. 31–50, Brussels.

Garver, N. 1990. Why pluralism now? *Monist,* 7, 388–410.

Gavison, R. 1980. Privacy and the limits of law. *Yale Law Journal, 89*, 421–471. Reprinted in Schoeman, F. (Ed.). 1984. *Philosophical Dimensions of Privacy.* Cambridge, UK: Cambridge University Press, pp. 346–402.

Monteiro, N. P. 2004. No privacy, no poetry, no progress. Department of Political Science, University of Chicago, Political Theory Workshop, 18 October 2004. Available at www.pwt.uchicago.edu/monteiro002.pdf.

Moor, J. H. 2006. The nature, importance, and difficulty of machine ethics. *Machine Ethics, 21*, 4, 18–21.

Moore, B. 1984. *Privacy: Studies in Social and Cultural History.* New York: M. E. Sharpe.

Moran, R. 2001. *Authority and Estrangement. An Essay on Self-Knowledge.* Princeton: Princeton University Press.

Nagel, T. 1998. Concealment and exposure. *Philosophy and Public Affairs, 27*, 1, 3–30.

Nissenbaum, H. 2004. Privacy as contextual integrity. *Washington Law Review, 79*, 1, 119–158.

Ortiz Hill, C. 1997. *Rethinking Identity and Metaphysics, On the Foundations of Analytic Philosophy.* New Haven: Yale University Press.

Perry, J. 2002. *Identity, Personal Identity and the Self.* Indianapolis: Hackett Publishing Company, p. 195.

Perry, T. D. 1986. *Professional Philosophy. What it is and Why it Matters.* Dordrecht: D. Reidel Publishing Company.

Quittner, J. 1997. The death of privacy. *Time,* 25 August, p. 18.

Roessler, B. 2005. *The Value of Privacy.* Cambridge, UK: Polity Press.

Russell, B. 1978. *The Problems of Philosophy.* Oxford: Oxford University Press.

Safire, W. 2005. Goodbye to Privacy. *New York Times,* April 10.

Sandel, M. J. 1989. Moral argument and liberal toleration: Abortion and homosexuality. *California Law Review, 77*, 521–538.

Schoeman, F. (Ed.). 1984. *Philosophical Dimensions of Privacy.* Cambridge, UK: Cambridge University Press, pp. 346–402.

Solove, D. J. 2002. Conceptualizing privacy. *California Law Review, 90*, 4, 1087–1155.

Spinello, R. 1997. End of privacy. *America, 176*, January 4–11, 9–13.

van den Hoven, M. J. 1999. Privacy and the varieties of informational wrongdoing. *Australian Journal of Professional and Applied Ethics, 1*, 1, 30–44.

van den Hoven, M. J. 2006. Nanotechnology and privacy: the instructive case of RFID. *International Journal of Applied Philosophy, 20*, 2.

van den Hoven, J., and Cushman, R. 1996. Privacy, Health Care Data and Information Technology, Conference Report. *Journal of Information, Law, and Technology (JILT), 3.* Available at http://www2.warwick.ac.uk/fac/soc/law/elj/jilt/1996_3/hoven/

van den Hoven, M. J., and Manders, M. 2006. Identiteitsmanagement en morele identificatie' (Identity management and moral identification). *ANTW, 98*, 2, 111–128.

Velleman, J. D. 1996. Self to self. *Philosophical Review, 105*, 1, 39–76.

Velleman, J. D. 2001. The genesis of shame. *Philosophy and Public Affairs, 30*, 27–52.

Wagner DeCew, J. 1997. *In Pursuit of Privacy: Law, Ethics, and the Limits of Technology.* Ithaca and London: Cornell Press.

Walzer, M. 1983. *Spheres of Justice.* Oxford: Blackwell.

Walzer, M. 1995. Critique of liberalism, in Amitai Etzioni (Ed.), *New Communitarian Thinking.* Richmond: University Press of Virginia, p. 63.

Warren, S., and Brandeis L. D. 1890. The right to privacy. *Harvard Law Review, 4*, 193–220.

Wiegel, V., van den Hoven, M. J., and Lokhorst, G. J. C. 2005. Privacy, deontic epistemic action logic and software agents: An executable approach to modeling moral constraints in complex informational relationships. *Ethics and Information Technology, 7*, 4, 251–264.

Williams, B. 1973. *Problems of the Self.* Cambridge, UK: Cambridge University Press.