Chapter 6
Privacy and Social Values in Smart Cities

Leonardo A. Martucci, Simone Fischer-Hübner, Mark Hartswood and Marina Jirotka

6.1 Introduction

The appeal of smart cities is the exploitation of information technology to better manage and plan the utilization of resources of a urban area. The benefits of smart cities are obtained by collecting and processing data from the city public services and utility companies, such as traffic information and water consumption, and from the city dwellers and its visitors. Smart cities ideally involve real-time data collection, processing and intervention, allowing public services to adapt to new conditions and constraints as they appear, and also to be better planned. For instance, road speed limits can adapt to traffic conditions or air pollution, public transport can be better monitored, allocated and redistributed, and law enforcement officials relocated more efficiently. In this chapter, we look at the personal data collected and processed by collective adaptive systems (CAS) and the internet of things (IoT), which are key information sources for enabling smart cities. The role of the IoT is to collect data and act locally while the CAS aggregates and processes the data and allows for people and machines to complement each other and operate collectively to achieve their, possibly conflicting, goals. The goal of this chapter is to provide an overview about

L.A. Martucci (✉) · S. Fischer-Hübner
Karlstad University, Karlstad, Sweden
e-mail: Leonardo.Martucci@kau.se

S. Fischer-Hübner
e-mail: Simone.Fischer-Hubner@kau.se

M. Hartswood · M. Jirotka
University of Oxford, Oxford, UK
e-mail: Mark.Hartswood@cs.ox.ac.uk

M. Jirotka
e-mail: Marina.Jirotka@cs.ox.ac.uk

© Springer International Publishing Switzerland 2017

V. Angelakis et al. (eds.), *Designing, Developing, and Facilitating Smart Cities*,
DOI 10.1007/978-3-319-44924-1_6

personal data protection for smart cities by looking into the techno-legal requirements and challenges using a Privacy by Design (PbD) focused on data minimization approach.

The term "smart" in smart cities refer to the use of information technology, especially data gathering, communication and analysis, to help society by promoting efficiency in services and rational use of resources with the ultimate goal of enhancing quality of life. All the personal data processed in applications designed to support smart cities need to be handled according to (local) social and legal requirements. As computer systems, algorithms, data and devices become increasingly closely coupled to people, as individuals and collectives, significant privacy challenges arise.

In this chapter, we look at the privacy challenges in smart cities from a point of view of the data collection and processing and the CAS harmonization and coordination aspects. We list the social legal principles behind smart cities from an European-centric perspective, and list the involved challenges to privacy using a urban car pooling scenario as our case study. We illustrate the privacy challenges with a privacy impact assessment (PIA) of a car pool (ride share) application developed within the EU FP7 *SmartSociety* project (cf. [28] for general information on the project and [20] for previous *SmartSociety* work on privacy in CAS).[1]

The remainder of this chapter is organized as follows. Section 6.2 briefly introduces the background on privacy on smart cities. The application scenario on digital transport that we use throughout this chapter is introduced in Sect. 6.3. Section 6.4 outlines the legal requirements and the derived set of privacy-related technical requirements. The privacy-enhancing methods, procedures, and technologies for fulfilling the requirements are presented in Sect. 6.5. Section 6.8 discusses the limitations of the existing solutions and concludes the chapter.

6.2 Background: Privacy, Social Principles, and Smart Cities

The network-enabled sensors and actuators that constitute most of the IoT often have limited resources, such as processing power and memory. It is auto-sufficient for small-scale interventions on the local scope, such as for heating, ventilation, and air conditioning (HVAC) climate control. However, the main benefits of IoT are not on the local scope but on the global one. The IoT is a collective of interconnected data sensors an actuators which underpin larger and more ambitious projects and initiatives, including smart cities.

The increase in the number and type of devices connected to the Internet means that IoT has the potential to collect data in volumes that are many orders of magnitude greater than is possible today. This data will be increasingly intimate, as it will emerge from everyday uses of technologies leading to whole swathes of mundane

[1] http://smart-society-project.eu.

activity being newly interconnected with the digital realms. The technological foundations of IoT are blind to the nature of data that it collects and transfers, i.e., there is no distinction if the data that it handles is personal information or not. For example, if we consider two equal network-enabled sensor devices, e.g., two GPS beacons, one may process personal information (a person's location) while the other may not (a parcel's location). Hence, the context in which in the IoT devices are immersed and the purpose of the collected data are cornerstone to the question of privacy.

Threats to privacy can be very significant and aspects regarding where the data is captured, centralized, processed is of great importance for understanding the impact on privacy and the possible countermeasures, especially when the data is to be shared or forward to a government and corporate entities administering a smart city.

6.2.1 Social Principles and IoT

We already understand the potential for existing data collection and algorithmic profiling to regulate society [10]. Existing IoT applications already demonstrate how the capacity for regulation can be intensified via directly embedded norms and bridging directly between corporate interests and peoples' everyday activities. "Driving black box" technologies, like that shown in Fig. 6.1,[2] use sensors and algorithms to monitor and evaluate drivers in order to regulate individual driving patterns in exchange for preferential insurance rates. Similarly, digital medicines potentially connect peoples' use of medication directly to pharmaceutical interests [25], and smart meters connect peoples' use of domestic appliances to the interests of energy suppliers [4]. There are arguments made in each of these cases about societal benefits including safer roads, more effective medication regimes and more sustainable energy use. But by the same token there are also sinister overtones of control and important questions to answer about which norms and values are embedded in these systems, and who has a say in how these are selected.

Multiple interests may be served by IoT applications yet those who control the infrastructure and own the data have a significant advantage in embedding their interests above others. A satire and critique of the control potential of the IoT has been created by *Thing Tank*, a research project on IoT, in the form of a video of an elderly person living independently in an IoT world where his fork advises about what to eat, his bed dictates his waking and sleeping routine, and his walking stick regulates his daily exercise.[3] The video makes apparent the impersonal control non-present relatives since the IoT "assistive" technologies evidently stand proxy for the relatives' own anxieties, responsibilities and desire for control. It also shows too how these forms of non-consensual control provokes rebellion and resistance as the elderly person finds clever ways of circumventing each of these mechanisms turn. In a subtle

[2] https://www.ingenie.com/how-it-works.

[3] The "*Uninvited Guests*" is short film produced by *Superflux* and commissioned by *Thing Tank*. https://vimeo.com/128873380.

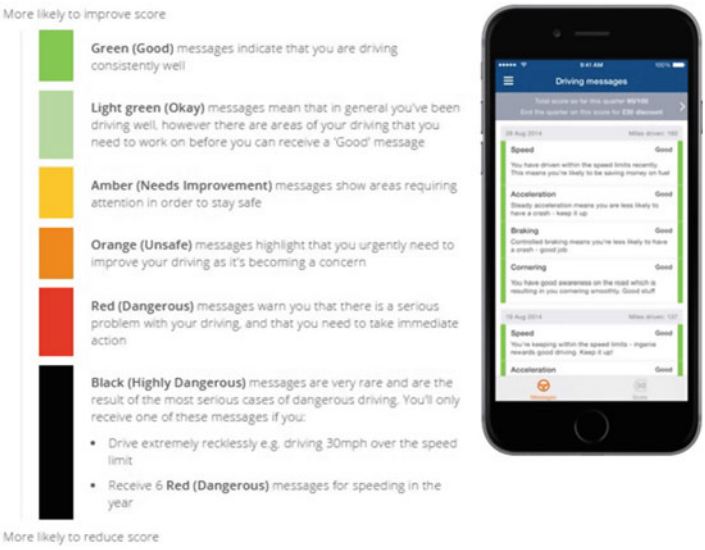

Fig. 6.1 A driving style tracker application

way, the video also expresses the value and pleasure obtained from certain freedoms that the overt regulation of the devices seems to deny. It encapsulates the conflicting values of the elderly person, his relatives and how these are entangled within wider cultural tropes about responsible lifestyle choices and personal freedoms. What is really absent in a centralized, technical and bureaucratic IoT future depicted by this video is any attention to creating space where these values can be negotiated.

6.3 Application Scenario: Urban Car Pooling

This chapter's use of digital transport as a focal example to explore privacy concerns is motivated by *SmartSociety*'s development of a car pool application, the Smart Share, to test and showcase how *SmartSociety* components can be used to build smart city applications. A privacy impact assessment (PIA) of the Smart Share application was conducted, and its architecture was conceived following a (PbD) approach focusing on data minimization. We aim, in this section, to draw lessons for Smart City IoT applications more generally. In this section, we more broadly introduce the concept of digital transport, show how real-world applications, such as Uber, can be problematic from a privacy perspective, before outlining the Smart Share, the *SmartSociety*'s own digital transport solution for smart cities.

6.3.1 Digital Transport and Privacy

Within IoT enabled smart city the vision for digitally augmented transport networks is to drive economic growth whilst mitigating problems of congestion and environmental sustainability [15]. The dynamics of real-world movements of people and vehicles are sensed, modeled and influenced via digital transport solutions comprised of sensors, algorithms and data connected by digital networks. Such solutions include, driverless cars [36], intelligent transport systems [13], personal transport advice [35], new business models, including collective utilization of spare capacity [34] and intelligent traffic management systems [19]. Alongside the undoubted social, environmental and personal benefits of digital transport there are risks too, and in this section we explore some of the risks to privacy posed by digital transport.

Travel is a vital social, economic and cultural activity so it is unsurprising that the journeys we choose or are obliged to make are also very revealing about ourselves. Our geographic location is a clue to what we are doing, and our pattern of journeys revealing of our activities and identities.

The risks of releasing our travel data is powerfully demonstrated series of media stories concerning the inappropriate use of information by the lift-sharing service Uber about journeys taken by passengers. An Uber senior executive (reportedly) threatened to reveal aspects of journalists' private lives deduced from Uber journey data as a punishment for their negative reporting of Uber. These threatened disclosures concerned journey signatures that may indicate an affair or a one-night stand. Other media stories report staff members casually accessing, circulating and commenting on journeys made by Uber passengers.[4] Subsequently Uber has stated that these practices and uses of its data are contrary to its privacy policies, and the Uber executive making the threatening remarks has apologized.[5] Although Uber could not persist with the impression of its being a playground for inquisitive employees or a vehicle for the senior executive to exact vengeance on critical journalists, these reports make clear the extent of the power attained over users from accumulating journey data. Uber's huge surveillance potential would be even further amplified should it achieve its goal of being a universally preferred option for any and every journey we might make. This spells out very clearly some of the risks to privacy of IoT in smart cities applications.

In addition to these direct implications for personal privacy, digital transport systems have a range of further risks that are linked to privacy concerns. These include the potential for new "digital divides" [37] where opting out of non-privacy-friendly systems may lead to inequality of opportunity. There are democratic risks, what were matters of public policy are transferred to private corporations. For example: supply, demand, price, quality and safety, of hire cars under Uber become (either directly

[4]Z. Tufekci and B. King. "*We Can't Trust Uber*". In: The Opinion Pages, NY Times, Dec. 7, 2014. www.nytimes.com/2014/12/08/opinion/we-cant-trust-uber.html.

[5]M. Isaac. "*Uber Executives Comments Leave Company Scrambling*". In: Bits, a NY Times Blog, Nov. 18, 2014. http://bits.blogs.nytimes.com/2014/11/18/emil-michael-of-uber-proposes-digging-into-journalists-private-lives/.

or diffusely) regulated within the Uber system, and no longer by local authorities. Often these types of regulation are driven by access to personal data and may work in ways to impinge on user autonomy, as with the example of "driver black boxes" given earlier. Issues around autonomy can be complicated. On the one hand, digital transport solutions promise user-centered services customized to personal need, but on the other hand they also afford delivery of finely tuned incentives to shape how transport options are chosen [12]. A simple reading is that systems with access to personal data have an equal potential to enhance or diminish user autonomy.

From a privacy perspective, exactly the same types of personal data needed for personalization may also be used to drive incentives. Thus, ommiting data to avoid incentives also restricts realizing benefit from personalization. Further types of privacy guarantee need to be built into the system, but these may lead to additional complications. For example, if individuals are (reliably) offered "opt-outs" from incentives then individual choice may conflict with the collective benefit of digital transport. Incentives need to be fairly implemented and carefully adjusted to encourage benign aims, such as carbon reduction, and avoid that those who opt out to be seen as selfish individuals undermining a common good. These types of consideration lead to important questions about those occasions where the value of privacy supports or undermines other social values expressed within the system. These are complicated questions about how privacy relates to democratic mechanism for setting system goals, how the interests of the system are made transparent and shown to be fair, and how we accept the balance between individual autonomy and collective good.

6.3.2 The Smart Share and Its Components

SmartSociety's Smart Share is a ride sharing, car pooling, application that supports drivers to fill spare capacity in their cars by enabling them to advertise the space to potential passengers. Passengers are able to search and signal their interest to participate in advertised trips. The Smart Share was designed to benefit from the following parts and components from *SmartSociety*:

- **Sensor fusion**. Use of IoT, such as sensors in mobile devices owned by the driver and passengers, to deduce information about the ride, including when it started, completed and who actually took part in it. This helps the system understand about the rides that were completed, which may feed back into reputation systems, incentives and algorithm optimization.
- **Peer profile**. It stores personal data about drivers and passengers. This includes identify information and preferences for taking rides. The release of personal data from the Peer Profile is controlled following a user defined privacy policy [22].
- **Social orchestration**. The matching algorithm that brings drivers and passengers together. "Ride Plans" are created for all the permutations of possible rides given driver and passenger constraints, with options for drivers and passengers to accept or decline rides.

- **Incentives**. They provide means of encouraging system uptake and meeting of specific goals, such as maximizing car occupancy, or use of less congested routes. It helps to meet global objectives, such as promoting sustainability, as well as improving how people meet their individual goals.
- **Reputation and provenance**. It encourages good behavior of Smart Share participants, such as timeliness and the quality of the ride. Driver and passenger reputation is visible when rides are negotiated. Provenance tracks the actions of any entity within the system, be it an algorithm or a person, to provide transparency and accountability of the actions of both algorithms and people.
- **Gamification elements**. It includes elements to make participation more enticing, such as achievement badges and the platform virtual currency.
- **Programming framework**. It provides a way of programmatically assembling the resources (including people) needed for some task to be accomplished within *SmartSociety*. In the case of Smart Share, the "task" is the ride that is collaboratively undertaken by drivers and passengers.
- ***SmartSociety* architecture**. It provides the coupling between all components in order to provide application programmers with the resources to extend the Smart Share application and other tools developed using the Smart Society platform.

6.4 Legal Aspects

In this section, we discuss the legal aspects and requirements pursuant the EU General Data Protection Regulation (GDPR) [17] concerning the processing of personal data in IoT and smart cities.[6] We first introduce the legal definition of personal data in Sect. 6.4.1, and emphasize the question around hardware identifiers, which are relevant to the discussion of personal data in IoT. Section 6.4.2 summarizes the general legal requirements, and the challenges to meetings such requirements are presented in Sect. 6.4.3.

6.4.1 Personal Data

Data Protection legislation only applies to data that classifies as personal data. The GDPR defines personal data as "any information relating to an identified or identifiable natural person (data subject)", who "can be identified directly or indirectly, in particular by reference to an identifier such as a name, an identification number,

[6]The EU GDPR was passed by the European Parliament in Dec. 2015, entered into force on 24 May 2016 and shall apply in all EU member sates from 25 May 2018. The GDPR was chosen as our reference for many reasons: (*a*) it applies to data controllers or processors located in the EU, and to any organization processing personal data of EU residents, (*b*) it reflects the basic privacy principles of the OECD privacy guidelines and (*c*) of the US Federal Trade Commission's (FTC) Fair Information Practice Principles (even going beyond them).

location data, online identifier or to one or more factors specific to the physical, physiological, genetic, mental, economic, cultural or social identity of that person."

The definition by the GDPR makes clear (as also stated previously by the Art. 29 Working Party [2]) that also unique device numbers, such as MAC addresses or RFID tag codes, can also be considered as personal data of users that can be associated with these devices (usually the device holders), with the consequence that these users could be uniquely (and secretly) profiled under these device identifiers by observers, even though the observed users may not be identifiable by name. For instance, an RFID tag in a watch that a person usually wears could be used within a supermarket to profile that user as a returning customer (cf. [2]).

The question whether MAC addresses or RFID codes constitute personal data or not can change over the lifetime of the respective devices or tags. Furthermore, in the context of IoT and smart sensing, as point out by the Art. 29 Working Party [1], individuals can often be identified with the help of data that originates from "things" and that may discern the life style of individuals and families, e.g., data generated by centralized control of lighting or heating in smart home applications.

6.4.2 Privacy Requirements

In this section, we present the legal requirements for a CAS computing platform that accommodate data protection and is designed upon a privacy preserving framework. Basic legal privacy principles, especially those by GDPR, are needed in order to identify the privacy threats as part of a Privacy Impact Assessment for a CAS computing platform and comprise the following ones listed in Table 6.1.

6.4.3 Challenges for Meeting the Legal Requirements

In the context of IoT and smart cities, fulfilling the legal requirements listed in Sect. 6.4.2 requires several challenges to be addressed, as noted by the Art. 29 Data Protection Working Party for IoT [1].

First, the current (in)security of IoT devices and platforms, which often have constrains concerning their battery and computational resources, is commonplace. IoT devices, from light bulbs and kettles to Barbie dolls, have flawed to none security mechanisms implemented.[7, 8, 9] Security is a fundamental legal requirement

[7] B. Ray. "Securing the Internet of Things–or how light bulbs can spy on you". Apr. 22, 2013. The Register. http://www.theregister.co.uk/2013/04/22/iot_security/.

[8] D. Pauli. "Connected kettles boil over, spill Wi-Fi passwords over London". Oct. 19, 2015. The Register. http://www.theregister.co.uk/2015/10/19/bods_brew_ikettle_20_hack_plot_vulnerable_london_pots/.

[9] I. Thomson. "Goodbye, Hello Barbie: Wireless toy dogged by POODLE SSL hole". Dec. 4, 2015. The Register. http://www.theregister.co.uk/2015/12/04/wireless_barbie_slipshod_security/.

Table 6.1 Privacy requirements derived from the EU GDPR

Privacy requirements	Description
Compliance with General Data Processing Principles, data protection by default (Art. 5, 25)	Key privacy principles are to be ensured, and must be enforced by the controller by appropriate technical and organisational measures by default, particularly: *Purpose specification & binding:* Personal data must be collected for specified and legitimate purposes and may later only be used for those purposes *Data minimization:* The amount of personal data and the extent to which they are collected and processed should be minimized, i.e., in particular if possible data should be anonymised or pseudonymised
Lawfulness of personal data processing (Art. 6, 7) & content	Lawfulness of processing to be ensured by an unambiguous informed consent, contract or legal obligation. The data subject shall have the right to withdraw his or her consent at any time
Lawfulness of processing special categories of data (Art. 9)	Lawfulness of the processing of "sensitive" personal data (such as data related to health, ethnicity, political opinions) must be ensured by explicit consent or special legal basis
Compliance with the right to be informed (Art. 14)	A data subject is to be provided with required privacy policy information including the identity of the data controller[a] and data processing purposes as well as the period for that the data will be stored at the time when the data is collected from the data subject
Compliance with transparency rights (Art. 15)	The data subject has the right to access their data (unless this adversely affects the privacy rights of others) and receive information about data processing purposes, data recipients or categories of recipients, the data retention period, the right to lodge a complaint with a supervisory authority as well as meaningful information about the logic involved on any automated processing including profiling, and the significance/envisaged consequences of such processing
Compliance with rights to rectification, erasure and restricting data processing (Art. 16, 17, 18)	The data subject can exercise the right to correct or delete their data, the right to restrict its processing, and the right to be forgotten in a timely manner
Compliance with the right to object (Art. 21, 22)	It must be ensured that the data subject has the right to object to the processing of their data, especially in the case of automated individual decision making, including profiling
Security of processing (Art. 25, 32)	It must be ensured that suitable security measures, including data minimization and pseudonymization, are implemented

[a]A data controller is a "person, public authority, agency or any other body which … determines the purposes and means of the processing of personal data". A data processor "processes personal data on behalf of the controller" [16].

for protecting, as presented in Sect. 6.4.2, and addressing the security problems IoT caused by flawed design is evidence to the lack of proper security testing.

Second, sensors are designed to collect data implicitly (without an explicit, case by case, consent) in an unobtrusive manner, which poses challenges to transparency. Moreover, fusion of sensor data allows further (sensitive) personal details may be derived, such as personal habits (as illustrated in the satire from *Thing Tank*, in Sect. 6.2.1), the driving style (as in the example shown in Fig. 6.1), and physical condition. The implicitly collected data and derived data can be quite diverse and with different purposes attached to each piece of information, which makes the process of obtaining informed consent difficult and cumbersome, and the end-user task of correctly setting their individual fine-grained privacy preferences for different types of data and purposes grueling.

6.5 Privacy by Design and Privacy Impact Assessment

Privacy by Design (PbD) is a framework for embedding privacy into the design and architecture of IT systems [6]. Its objective is for privacy to become an essential property of all of the components of an IT system. PbD claims a full life-cycle protection of personal information, as its guidelines advocate for all personal data to be securely collected, stored, used and destroyed. In theory, it is applicable even for evolving systems, as privacy properties are constantly analyzed and addressed following the evolutionary steps of the development of a the IT system.

PbD is based on a collection of loosely defined guiding principles, which include a proactive approach to privacy, and the promotion of user-centric systems, visibility and transparency. The absence of proper formalization allows for confusion and even intentional abuse [21]. A strategy to avoid such pitfalls is to link PbD to general privacy principles, such as data minimization.

Embedding privacy in the design of IT systems and applications requires a comprehensive evaluation of the collected personal data and its use. This evaluation is provided by the Privacy Impact Assessment (PIA), a systematic process for evaluating the effects of data processing on privacy [11]. It consists of multiple procedural and sequential steps that are related to: the characterization and use of information, retention of data, internal and external sharing and disclosure, notice, access, redress, correction, technical access to information, security aspects, and technologies involved. A PIA provides means of understanding privacy-related concerns regarding the adoption and deployment of new technologies and services, and also helps to mitigate risks to business [11].

A PIA can be summarized as five procedural step: (*a*) a check for the need of a PIA, (*b*) the identification of personal data in the application, (*c*) the identification of existing countermeasures, (*d*) the listing of the existing privacy threats, and (*e*) a recommendation for additional countermeasures.

A PIA can be tailored to specific technologies and applications. In the context of IoT, the PIA framework for RFID applications [18] is a relevant example. The framework specifies the need of the PIA, its scale, criteria and elements for assessment, including privacy goals derived from the EU Directive 95/46/EC [16] and a list of

privacy risks related to RFID, e.g., collection of personal data exceeds purpose and secret data collection by the RFID operator.

The scope of smart cities and IoT applications is much broader than the scope of RFID applications, which parts are also a subset of IoT applications. Nevertheless, the RFID PIA framework offers a set of processes and guidelines that could be adapted to IoT. An IoT PIA framework targeting the privacy requirements listed in Sect. 6.4.2 identifies threats to personal data and lists the appropriate controls and mitigation measures to avoid or minimize them, such as privacy-enhancing technologies (PETs). The drawback is such an IoT PIA framework would be too general, i.e., it would include a too large spectrum of threats and countermeasures, which would make it not useful in practice as PIAs are application specific.[10]

6.6 Countermeasures: Privacy-Enhancing Technologies

Legislation offers a list of legal definitions and privacy requirements (see Sect. 6.4.2), and the fines, reparations and penalties related to the legal infringements. Legislation, however, does not offers the technological means to enforce the data protection requirements. Hence, privacy in smart cities should not rely on legal measures only, but also with the support of computer and network security tools and mechanisms to enforce legal privacy principles. These privacy tools and mechanisms are generally referred to as Privacy-Enhancing Technologies (PETs). PETs can be divided into three categories, according to their specific goals.

The *first category* comprises PETs for enforcing the legal privacy principle of data minimization by minimizing or avoiding the collection and use of personal data of users or data subjects. PETs in this class provide a subset of the following:

- *Anonymity*, which is defined as an individual not being identifiable within a set of individuals, such as a collective.[11]
- *Unlinkability*, which means that two items of interest, such as individuals, objects and actions, cannot be sufficiently distinguished if they are related or not by a third party, e.g., an attacker.[11]
- *Pseudonymity*, which refers to the use of pseudonyms as identifiers. Pseudonyms are identifiers other than an individual's real names. Pseudonyms can be classified according to their degree of linkability to the individuals holding them, from a simple substitute to an individual's name, i.e., a nickname or a mobile phone number, to short-lived pseudonyms that are used for a single transaction or operation only.[11]

[10]The RFID capabilities and the scope of its applications are narrow enough to produce a PIA with a (non-exhaustive) set of 15 potential threats and five groups of countermeasures.

[11]The definition of the terms *anonymity*, *unlinkability*, *pseudonymity*, *unobservability* in this chapter follows the Pfitzmann and Hansen terminology [30].

- *Unobservability*, which means that an item of interest is undetectable, i.e., an attacker is not to sufficiently distinguish if an item of interest exist or not, and individuals involved in the item of interest are anonymous.[11]

A PbD targeting data minimization would plan and enforce this first category of PETs by default (as postulated by the GDPR, Art. 23). PETs in this category can be further classified depending whether data minimization is achieved on the network (data communication) level or the application level. Examples for PETs for achieving data minimization at the network level are anonymous communication protocols, which are based on either specialist nodes for forwarding network traffic, e.g., Mix Nets, DC Nets and Tor [7, 8, 14], or distributed solutions, where all devices in the network forward data on the behalf of others [26, 32]. On the application level, PETs include anonymous payment schemes [9], privacy-preserving digital identifiers [5, 27], oblivious data transfer (OT) [31], and obfuscation schemes [29].

Data minimization is the best strategy for protecting privacy because it decreases or avoids personal data from being processed. Nevertheless, there are many occasions in daily life when individuals have to, need to, or want to reveal personal data. For instance, when online shopping, an individual would reveal an address for billing and delivering goods, or when people willingly disseminate personal information because they wants to introduce themselves and interact with an online audience, such as on social network. In these cases, the privacy of the individuals concerned still needs to be protected by adhering to legal privacy requirements, which are covered by the second category of PETs.

The *second category* of PETs comprises technologies that enforce legal privacy requirements, such as informed consent, transparency, right to data subject access, purpose specification and purpose binding and security, in order to safeguard the lawful processing of personal data. Electronic privacy policies are PETs in this second category. They are statements that allow to describe how personal data is processed, by whom, for what purposes, and can be mathematically and logically formalized into machine readable, purpose-specific privacy policy languages. The PrimeLife Policy Language (PPL) [33] is a privacy policy language that falls into this second category of PETs. It can be used to enhance (ex-ante) transparency for users and to derive so-called sticky policies, which "stick" to the user's personal data to define allowed usage and obligations to be enforced a the service provider requesting the data and any third party to whom the data is forwarded.

The *third category* of PETs comprises technologies that combine PETs of the first and second categories, such as identity management systems.

6.7 A PbD Case Study: The Smart Share PIA

To illustrate how PbD can be included in the design of IoT applications for smart cities, we summarize the PIA for the *SmartSociety*'s car pool application, the Ride Share, introduced in Sect. 6.3.2. To conduct the Smart Share PIA we followed the gen-

eral framework for RFID applications [18] and the guidelines of the British Information Commissioner's Office (ICO) PIA code of practice [23]. In Sect. 6.5, we learned that the RFID PIA framework can be adapted to IoT and applications for smart cities. To broader the scope of the RFID PIA framework, we used elements from the general structure of the ICO PIA code of practice.

As described in Sect. 6.5, a PIA has five procedural steps. The first three PIA steps are presented in Sect. 6.7.1, the fourth step, concerning the encountered privacy threats, are summarized in Sect. 6.7.2, and last step, with the recommended additional countermeasures, is outlined in Sect. 6.7.3.

6.7.1 Personal Data and Existing Countermeasures in Ride Share

The first step in a PIA process is an initial assessment to identify the need of a PIA. For this evaluation, it is necessary to verify the objectives of the Ride Share application, and which of its component parts process personal data. The Ride Share is a car pooling application, and its explicit objective is to provide a ride sharing service. Other objectives that are less explicit are related to the use of the Ride Share as a testing platform for the *SmartSociety* components. The Ride Share processes personal data to test and evaluate its algorithms and protocols and keep records for data provenance and for its reputation system. Furthermore, Smart Share aims to release (anonymized) data sets to the general public. The collection and processing of personal data in Smart Share justify the need of a PIA.

The second step involves the identification of personal data in Ride Share, and personal data information flows in the application. This step requires a deep understanding of the system, its interfaces and its implementation details. In Smart Share, we first identified the personal data inputs, which happen either during registration phase or during operation phase. In the user registration phase, Ride Share has three mandatory fields (user name, email address, and phone number) and optional fields, such as a photo. Additional personal data collected / processed during operation phase include geographical location (departure and arrival addresses), date and time, smoking habits, tolerance for domestic animals, history of rides taken and shared, user feedback, and reputation.

We classified the personal data types collected according to their processing purpose and the source of the personal data. The personal data was organized according to the following data processing categories: (*a*) functional purpose, which means that the data is required to providing the explicit objective of the application, i.e., offer a platform for car pooling, (*b*) accountability, which includes provenance and reputation services, (*c*) statistical analysis, and (*d*)assessment and testing of the *SmartSociety* components. The possible sources of personal data were: user input, sensor data, or output from a third party application.

The third step identifies the countermeasures in place, such as PETs embedded in the *SmartSociety* components and tools and procedures that are independent of the platform, which are application specific, such as mechanisms and interfaces for deleting and exporting personal data, and for redressing incorrect or inaccurate data. This step also includes the list of the security mechanisms that guarantee that personal data is stored, processed and communicated securely. It also includes the procedures of obtaining user consent, the transparency mechanisms that allows users to check accuracy of their personal data, and contact information.

6.7.2 Privacy Threats

In this section, we present the fourth step of the PIA process, which lists the privacy threats in *SmartSociety*'s Smart Share application in Table 6.2. The privacy threats relate to the list of privacy requirements presented in Table 6.1.

6.7.3 Additional Countermeasures

When executing the PIA for Smart Share, we identified a series of common pitfalls that may occur when IoT-based applications are designed and implemented in a distributed and collaborative way.

The first three steps of the PIA (Sect. 6.7.1) demonstrated the importance of the privacy awareness concerning the definitions around personal information and the general need for application designers and software engineers to know the basics around data protection legislation, or to be supported by someone that is equipped with this knowledge. In the case of the developing team of Smart Share, it had, in general, no proper formal privacy-awareness, which led to delays in the implementation of the Smart Share, which is evident by the few identified privacy and security controls in place. In the fourth step of the PIA (Sect. 6.7.2), the list of privacy threats in relation to the privacy requirements identified in Table 6.1 were presented.

In the final step of the PIA, we present a recommendation for additional countermeasures. The non-exhaustive list of the suggested additional countermeasures included: (*a*) clear purposes for processing personal data, (*b*) well-defined consent forms, (*c*) means to withdraw consent, (*d*) a specified, limited duration for storing personal data, (*e*) data encryption (*f*) the use of pseudonyms for location, reputation, incentives and in the case of entangled data, which would allow individuals to access their personal data even if it is related to personal data of other individuals.

This list of recommendations resulted in a series of security and privacy enhancements in later versions of the Smart Share and of the *SmartSociety* components.

A variant of PPL, called A-PPL [3], was integrated to the Peer Profile component of *SmartSociety*. The Peer Profile allows users to define privacy policies to their personal data items. It also allows for semantic data obfuscation, i.e., personal data is

Table 6.2 Privacy threats in *SmartSociety*'s smart share

Privacy threats	Description
Imprecise terms in the informed consent and the right to withdraw (lawfulness of personal data processing threat)	The terms used in the Smart Share's consent form are not precisely outlined and what personal data is released to each of the different organizations involved in the Smart Share is not clearly defined. Abstract and imprecise purposes lead to the collection of much more data than strictly needed. For example, refining a purpose of "accountability" to "accountability of user profiles" would reduce the amount of personal data to be recorded There are no procedures to withdraw consent in Smart Share
Collection and storing of personal data beyond what is strictly needed (data minimization threat)	There is no limited duration defined for storing personal data collected by the Smart Share application
User profiling (data minimization threat)	Provenance makes it possible to link personal data to activities. These relationships can be used to create user profiles that include personal data beyond the original purpose/need of the application
Vague purposes and function creep (purpose binding threat)	The personal data processed is potentially excessive or irrelevant. It concerns especially data items processed for the purpose of "accountability" and "statistical analysis," which are broad and ill-defined concepts. Vaguely defined purposes allow personal data to be processed for purposes unintended at design time. For example, the purpose of "accountability" could be misinterpreted or extended to allow otherwise unauthorized parties to have access to the data
Unauthorized access to information (security of processing threat)	The *SmartSociety* platform has no access control mechanisms in place. None of the platform components encrypt data, which may allow for personal data to be read by unauthorized users
Processing of inferred personal data without consent (threat to processing of special categories of data and to lawful processing)	Information collected from sensors, including geographical location, may lead to conclusions regarding an individuals habits, life style and social connections. The personal data collected in Smart Share allows for revealing who is traveling with whom, their whereabouts, and other semantic information, such as visits to shrines, political demonstrations, etc
Obstacles to deletion of data (threat to rights to rectification, erasure and restricting data processing)	There are no automatic means to delete personal data entries. The deletion of personal data from the social orchestrator and peer manager can be performed manually upon request to the Smart Share data support team. There are no available means to delete data from the provenance and reputation servers

(continued)

Table 6.1 (continued)

Privacy threats	Description
Multiple data subjects: Obstacles to access and delete data (transparency rights threat and threat to the rights of rectification and to erasure)	Sharing rides are collective actions that involve two or more individuals. Hence, personal data from multiple individuals is entangled in Smart Share. Access requests for personal data might be denied because it may disclose data about other individuals
Limited transparency to data subjects (transparency rights threat)	There are no means to guarantee the right of data subjects to access and / or amend, to the full extent, all the personal data items that are processed by the Smart Share. The data subjects have access to personal data that is available on their user profile, and may amend their email address, but have no means to access or correct information related to past rides, or any provenance and reputation data. Furthermore, the consequences of processing personal data are often difficult to foresee. It is thus difficult for data controllers to inform individuals about all possible consequences of personal data disclosure
Obstacles to the implementation of technical measures, such as pseudonymization (security of processing threat)	The use of transaction (one time) pseudonyms hamper the *SmartSociety* incentives, provenance and reputation components, which rely on individuals' history of past interactions with the system and user profiles

semantically obfuscated after an ontology-based obfuscation mechanism [24]. Personal data in the peer manager is now stored encrypted, and the communication between *SmartSociety* components is secured using TLS 1.2. In Smart Share, it led to planning properly designed consent forms, and the inclusion of contact points and procedures for redressing incorrect data. Smart Share now provides a check-box that is mandatory for users to select to point out their informed consent. The registration page includes a link that points to the privacy policy.

6.8 Conclusions

In this chapter, we discussed the social impact and privacy aspects of smart cities that are based on IoT and CAS for collecting and processing data about individuals, taking the Smart Share application of the *SmartSociety* project as an example. A wide scope of privacy issues are raised, in particular by the nontransparent manner of data collection via IoT and challenges to secure IoT technology due to performance constraints, as well as challenges enforcing data minimization via pseudonymization due to the need to link data for the purpose of accountability, such as provenance data or reputation data. Also, in smart cities, data, such as data about reputation scores

of a rater and ratee, or data about shared car rides may refer to more than one data subject, who may have conflicting privacy preferences in regard to the handling of those data.

Furthermore, we outlined the process of a PIA for Smart Share that we conducted within the scope of the *SmartSociety* project and highlight the broad scope of threats that we determined, for we had to specify mitigation measures in a subsequent step. Challenges for conducting a PIA are not only posed by the complexity of smart city applications based on CAS, but also by their inherent dynamic structures: applications based on CAS can dynamically include new types of machines as peers, which may change the type of personal data collection or processing. In such situations, a new or revised PIA may be needed.

Acknowledgements This research was funded by SMARTSOCIETY, a research project of the Seventh Framework Programme for Research of the European Community under grant agreements no. 600854.

References

1. Article 29 Data Protection Working Party: Opinion 8/2014 on the on Recent Developments on the Internet of Things (2014). http://ec.europa.eu/justice/data-protection/article-29/documentation/opinion-recommendation/files/2014/wp223_en.pdf
2. Article 29 Data Protection Working Party: Working document on data protection issues related to RFID technology (2005). http://ec.europa.eu/justice/data-protection/article-29/documentation/opinion-recommendation/files/2005/wp105_en.pdf
3. Azraoui M, Elkhiyaoui K, Önen M, Bernsmed K, De Oliveira AS, Sendor J (2015) A-PPL: an accountability policy language. In: Data privacy management, autonomous spontaneous security, and security assurance, pp 319–326. Springer
4. Borges F, Martucci LA, (2014) iKUP keeps users' privacy in the smart grid. In: CNS, (2014) IEEE Computer Society. NY, USA, New York
5. Camenisch J, Lysyanskaya A, (2002) A signature scheme with efficient protocols. security in communication networks: third international conference (SCN, (2002) Lecture Notes in Computer Science, 2576 (2003)). Springer. Amalfi, Italy, pp 268–289
6. Cavoukian A (2009) Privacy by design. White paper, Information and Privacy Commissioner of Ontario
7. Chaum DL (1981) Untraceable electronic mail, return addresses and digital pseudonyms. Commun ACM 24(2):84–88
8. Chaum DL (1988) The dining cryptographers problem: unconditional sender and recipient untraceability. J Crypt 1(1):65–75
9. Chaum DL (1992) Achieving electronic privacy. Sci Am 267(2):96–101
10. Cheney-Lippold J (2011) A new algorithmic identity soft biopolitics and the modulation of control. Theory, Culture Soc 28(6):164–181
11. Clarke R (2009) Privacy impact assessment: its origins and development. Comput Law Secur Rev 25(2):123–135
12. Deloitte: Disruptive trends for smart mobility. http://www2.deloitte.com/uk/en/pages/business-and-professional-services/articles/transport-in-the-digital-age.html (2015)
13. Dimitrakopoulos G, Demestichas P (2010) Intelligent transportation systems. Vehicular Technology Magazine 5:77–84
14. Dingledine R, Mathewson N, Syverson P (2004) Tor: the second-generation onion router. USENIX-SS 2004. USENIX Association, Berkeley, CA, USA, pp 303–320

15. Earnst & Young: Routes to prosperity: How can smart transport infrastructure can help cities to thrive. http://www.ey.com/Publication/vwLUAssets/EY-routes-to-prosperity-via-smart-transport/$FILE/EY-routes-to-prosperity-via-smart-transport.pdf (2015)
16. European Commission: 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 L No.281 (1995)
17. European Commission: Regulation (EU) 2016/679 of the European Council and Parliament of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation), Official Journal of the European Union, L 119/1
18. European Union Norm: Privacy and data protection impact assessment framework for RFID applications, Appendix to the Opinion 9/2011 on the revised industry proposal for a privacy and data protection impact assessment framework for RFID applications (2011)
19. Figueiredo L, Jesus I, Machado J, Ferreira J, Carvalho J (2001) Towards the development of intelligent transportation systems. Intell Transp Syst 88:1206–1211
20. Fischer-Hübner S, Martucci LA (2014) Privacy in social collective intelligence systems. In: Social collective intelligence, pp 105–124. Springer
21. Gürses S, Troncoso C, Diaz C (2011) Engineering privacy by design. Computers, Privacy & Data Protection 14:
22. Hartswood M, Jirotka M, Chenu-Abente R, Hume A, Giunchiglia F, Martucci LA, Fischer-Hübner S (2014) Privacy for peer profiling in collective adaptive systems. In: Privacy and identity management for the future internet in the age of globalisation, pp 237–252. Springer
23. ICO UK: Conducting privacy impact assessments code of practice, v. 1.0. Technical report, Information Commissioner's Office (ICO), UK (2014)
24. Iwaya L, Giunchiglia F, Martucci LA, Hume A, Fischer-Hübner S, Chenu-Abente R (2015) Ontology-based obfuscation and anonymisation for privacy—a case study on healthcare. In: Proceedings of the 10th IFIP summer school on privacy and identity management. Springer
25. Jara A, Alcolea A, Zamora M, Skarmeta A, Alsaedy M (2010) Drugs interaction checker based on IoT. In: Internet of things (IOT), pp 1–8. IEEE
26. Martucci LA, Andersson C, Fischer-Hübner S (2006) Chameleon and the Identity-anonymity paradox: anonymity in mobile ad hoc networks. In: IWSEC 2006, pp. 123–134. IPSJ
27. Martucci LA, Kohlweiss M, Andersson C, Panchenko A (2008) Self-certified sybil-free pseudonyms. In: Proceedings of the 1st ACM conference on wireless network security (WiSec'08), pp. 154–159. ACM Press
28. Miorandi D, Maltese V, Rovatsos M, Nijholt A, Stewart J (2014) Social collective intelligence. Springer
29. Mowbray M, Pearson S (2009) A client-based privacy manager for cloud computing. In: ICST COMSWARE 2009, p 5. ACM
30. Pfitzmann A, Hansen M (2010) A terminology for talking about privacy by data minimization: Anonymity, Unlinkability, Undetectability, Unobservability, Pseudonymity, and Identity Management v.034. http://dud.inf.tu-dresden.de/literatur/
31. Rabin MO (2005) How to exchange secrets with oblivious transfer. IACR Cryptology ePrint Archive p 187
32. Reiter M, Rubin A (1997) Crowds: Anonymity for Web Transactions. In: DIMACS Technical report, pp 97–115
33. Trabelsi S, Neven G, Raggett D (eds) (2011) PrimeLife Public Deliverable D5.3.4 – Report on design and implementation
34. Trivett V, Staff S (2013) What the sharing economy means to the future of travel. Report, New York (Skift, p 7
35. Tumas G, Ricci F (2009) Personalized mobile city transport advisory system. Inform Commun Technol Tourism 2009:173–183
36. UK Department for Transport: The pathway to driverless cars. Summary Report and Action Plan (2015)

37. Velaga N, Beecroft M, Nelson J, Corsar D, Edwards P (2012) Transport poverty meets the digital divide: accessibility and connectivity in rural communities. J Transp Geogr 21:102–112