Assessing Privacy by Design in Smart Cities: An analysis of the Stratumseind 2.0 project

Taissa de Lima Conde

ID: 2111551

Thesis Supervisor: Dr. Vlad Niculescu-Dincă

Second Reader: Dr. Els de Busser

Thesis submitted to the Faculty of Governance and Global Affairs of Leiden University in partial fulfilment of the requirements for the MSc Crisis and Security Management, supported by Deloitte Nederland.

03 March 2019

Abstract

The introduction of Information and Communication Technology (ICT) in the field of urban development propelled the application of technologies as 'smart' solutions for cities. Likewise, the concept of smart city transformed the notion of physical cities to a network of flows-systems that entangle the digital and physical world. Accordingly, the growth of smart city projects introduced a new dilemma for privacy in public spaces, and the increasing use of big data analytics denounced the potential risks to data privacy. Consequently, these (privacy) concerns addressed the unquestionable need for investigating whether there are sufficient guarantees for citizens' privacy in the context of smart cities. Furthermore, the focus on safeguarding citizen's privacy impelled the development of a new guideline on privacy by design (PbD) to support the employment of these projects. This thesis aims to assess the application of PbD by smart cities (projects) in the safeguard of data protection, encompassing both organizational and technical components of the architecture. Hence, this thesis validates that PbD is not fully incorporated in smart city projects and demonstrates the challenges with regard to multiple stakeholders ensuring privacy and security measures throughout the smart city architecture. Finally, it indicates further research on some aspects of the new guideline, such as the incorporation of legacy systems and a checklist evaluation, while suggesting more legal and architectural recommendations applicable to the demands of smart cities.

Keywords: Smart City, Privacy by Design (PbD), privacy, data protection, Big Data, public spaces.

TABLE OF CONTENTS

1 INTRODUCTION

The 21st century is marked by significant urban developments around the globe (Eremia, Toma, & Sanduleac, 2017). The increased number of people living in the cities propelled the fields involved in urban planning towards finding solutions to the challenges that emerged (e.g. "energy supply, waste management, transportations, environmental issues and security to mention a few" (Ståhlbröst, Padyab, Sällström, & Hollosi, n.d., p. 1)) through different means, particularly the Internet of Things (IoT). This phenomenon encouraged an Information Technology (IT) development focused on urban solutions defined as 'smart city' (Ståhlbröst et al., n.d.), mainly planned to Information and Communication Technology (ICT) infrastructures.

According to several authors (Anthopoulos, 2015; Batty, 2013; Eremia et al., 2017; Ståhlbröst et al., n.d.), the concept of smart cities is widely used to explain the integration of smart technologies to strengthen governance and enhance urban planning, economic growth and sustainability, whilst ensuring quality of life. One of the many revolutionary features of smart cities was the use of big data applied in urban systems, which entails huge amounts of data that can be collected, stored and processed in a short amount of time or even in real-time. For instance, one of Deloitte's reports on Smart Cities (2017) introduces the 'smart traffic control', which is characterized by a traffic control system with real-time information capable of optimizing and adjusting traffic flows. Similarly, these developments were responsible for stimulating a rapid adjustment within different public and private organizations.

As much as these innovative 'smart' technologies can solve problems, they also introduce new paradigms (Waelbers, 2011). Privacy issues in big data are a great source of concern with the technological development introduced with smart cities, particularly when it comes to data protection. Since most smart city projects rely on aggregation and real-time

analysis of data to be able to perform different activities, personal data became an asset to these processes, still, most projects lack data subjects' awareness of data collection. However, these technologies are designed as efforts to push societies towards a 'sustainable development through participatory governance' (Deloitte, 2015; Lacinák & Ristvej, 2017). The political and social relevance of the implementation of these technologies is then undisputable: it introduces a new form of integration within cities by shortening the distances between citizens and governments while offering more efficient solutions to support urban development (Swinhoe, 2018). Moreover, with the increased implementation of ICT's integrated systems with IoT and the massive volume of data that is needed to provide efficient and automated services (Swinhoe, 2018), several privacy dilemmas emerged concerning 'personal data'.[1]

Additionally, it is keen to understand the differences between privacy and data protection. Among several typologies and considerations, privacy can be described as: the right to autonomy, to a private life, to be let alone and to be in control of information about oneself; however, more than an individual (fundamental) right, privacy is also a social value (Diffie & Landau, 1998, p. 98; Smith, 2016). Data protection concerns the protection of "any information relating to an identified or identifiable natural (living) person, including names, dates of birth, photographs, video footage, email addresses and telephone numbers" (Smith, 2016). It aims to ensure that such personal data is processed – which entails the processes of collection, use and storage – fairly by both public and private sectors (Article 8, "Charter of Fundamental Rights of the European Union," 2012; Smith, 2016). Both concepts overlap since the notion of "data protection originates from the right to privacy and both are instrumental in preserving and promoting fundamental values and rights" (Smith, 2016); yet, only privacy is recognized as a universal human right.

Arguably, due to the constant technological development, particularly with information systems, and the challenges to data privacy, new developments on data protection regulations

became urgent (Ryz & Grest, 2016, p. 1). Hence, on the 25 May 2018 the EU enforced the General Data Protection Regulation (GDPR), which focus on data controllers that process big data and personal data by addressing their responsibility and accountability (Ryz & Grest, 2016). The legal framework established with the GDPR encourages "the adoption of the principles of 'privacy by default' and 'privacy by design'" (Ryz & Grest, 2016, p. 1) in order to ensure the rights to privacy and data protection while also informing individuals about how their data is processed. However, debates about the exponential increase of smart cities and the potential risks to data subjects still need close attention.

As these risks to data privacy emerged, Ann Cavoukian (2011) developed the pioneer concept of Privacy by Design (PbD). The framework is based on seven foundational principles and aims to cover most of the elements throughout the composition of a system/process in order to ensure privacy. PbD entails embedding security features into software or data management to safeguard personal privacy. It reflects an effort to integrate legal and technological approaches to mitigate the risks posed especially by big data and assure compliance with current regulations covering data protection (Wiese Schartum, 2016). However, the seven foundational principles were frequently described as vague and unrealistic as a practical guideline (Domingo-Ferrer et al., 2014; Gurses, Troncoso, & Diaz, 2011; Kroener & Wright, 2014; Perera, McCormick, Bandara, Price, & Nuseibeh, 2016; Spiekermann, 2012; Wiese Schartum, 2016).

In order to ensure security and privacy in every step of implementing smart city technologies, it is essential to investigate the arrangement of the multiple stakeholders involved in the projects. Therefore, this thesis aims to understand how PbD is being implemented in a smart city project in The Netherlands. A new guideline of PbD will be developed and used as a model to assess the data processing and the responsibilities of the stakeholders involved in the project: it is essential to understand the dynamics of the different stakeholders engaged in

smart cities' projects. Once these arrangements are explored, it will be possible to stress how PbD is applied and how these stakeholders comply with it. Likewise, it will be possible to determine to what extent PbD suits this set of cooperation and whether it is correctly applied in the technologies integrated in smart cities' projects. Therefore, this thesis aims to answer the following research question:

To what extent are smart cities applying Privacy by Design in the safeguard of data protection?

It is evident that the GDPR enforces compliance with a set of rules concerning data processing; however, how do the actors engaged in a smart city project ensure PbD? Considering these projects usually integrate multiple stakeholders, how are they ensuring data protection? The risks of controversial consequences caused by a system or technology should be considered; therefore, are these stakeholders aware of their responsibility throughout the process?

These questions aim to address concerns regarding different (emerging) smart cities technologies, despite the growing awareness and focus on the framework of PbD, partially stimulated by the recent implementation of the GDPR. To answer them, a case study design will be conducted on the 'Stratumseind 2.0 project' in Eindhoven, The Netherlands. The study will contrast the different roles of stakeholders in managing data sharing and addressing the potential impacts on data subjects under a new framework of PbD that will be developed. This study provides a more operationalized and comprehensive approach to projects embedding PbD and ensures that different stakeholders might be held accountable for the safeguarding of data protection.

The selection of the case is based on the necessity of an in-depth analysis on how PbD is concretely applied in smart cities. The project aims to reduce criminal behavior and support economic growth in a critical and important area of the city, combining resources and

knowledge from different stakeholders. In that sense, this study will potentially increase people`s trust on the capacity of smart cities to support economic development and transform urban environment. Besides, it could also stimulate more specific data protection regulations capable of keeping pace with the most recent technological developments applied in the public sphere (and support upcoming innovative technologies). Thus, the case is representative of a growing tendency of smart cities' technologies: developing smart technologies using big data as smart solutions for an efficient growth of the city.

Additionally, by incorporating several PbD strategies to the theoretical framework, it will be possible to address the critics on the subject and create a more assertive guideline for future projects. Finally, by contributing to the discussion regarding data protection in smart cities, the thesis will add a nuanced perspective to the literature, based on stakeholders` responsibilities to safeguard personal privacy.

The present thesis is structured as follows: in the second chapter I will provide the conceptualization of smart cities, big data and the difference between privacy and data protection. Moreover, I will present a discussion on the framework of PbD, contrasting the different perspectives and providing a guideline to assess the management and protection of data in different smart city projects. The third chapter introduces the methodology and the case study specifications. The forth chapter introduces the analysis and the discussion regarding the findings of the case. Finally, the last chapter presents the conclusions about the research question and engage in further discussion concerning the findings of this study.

2 THEORETICAL FRAMEWORK

This chapter introduces the relevance of the concepts of smart city and big data to this study. Moreover, it delineates the difference between privacy and data protection while addressing the debates on privacy concerns brought by smart cities. Moreover, it reflects on

the challenges for privacy in public spaces brought by smart cities. Finally, the discussion will be narrowed down to PbD, contrasting the numerous perspectives concerning the PbD framework and providing a more appropriate guideline to assess the data management in different smart city projects regarding data protection.

2.1 SMART CITIES

This section introduces the concept of smart city and its pioneer features brought by the incorporation of ICT technologies. The concept represents a growing tendency of infinite possibilities to urban development worldwide and exposes several challenges to privacy in public spaces.

The term Smart City is widely used to describe different perspectives and strategies to the planning of urban spaces. It describes a 'smart' urban development driven by the availability and quality of ICT, economic development and the importance of human capital, education and sustainability (Caragliu, del Bo, & Nijkamp, 2011). In other words, it defines a networked infrastructure based on ICT attributes and solutions in the urban space (Anthopoulos, 2015); also, it offers a constant monitoring of any aspect of urban life . There are different domains applicable in the integrative framework for smart cities (Anthopoulos, 2015). For instance, according to Neirotti et al. (as cited in Anthopoulos, 2015), the analysis on smart cities encompasses two domains: first, the soft (domain), which is about economy, government, etc.; and the second, the hard (domain) is about energy, transportation and others. The existence of various interpretations regarding the domains of analysis for smart cities addresses the multidisciplinary aspect of the subject and its focus on different perspectives (Anthopoulos, 2015).

This study explored the emerging role of ICT, particularly with regard to the challenging adoption of big data analytics. The 'smart' aspect in urban spaces describes the

development of digital technologies applied in the city dimensions and their interrelationship. Furthermore, a smart city is described as a 'system of systems', which evokes a "cross-domain sharing of information" (Osman, 2019, p. 620) and the large volume of data created by the interaction between humans and machines with the advance of digital technologies. Likewise, the so called 'network society' also alters the delimitation of public spaces: from "static places to a space of flows" (Timan, Newell, & Koops, 2017, p. 7). However, it is keen to understand that one of the challenges provoked by these technologies is the (mis)use of mechanisms to achieve questionable results, particularly when considering the use of integrated databases. For instance, Batty (2013) addresses the concern with the adoption of sensors streaming real-time data using precise geo-positioning and the integration of these databases to ensure an output with the expected value.

The concept of smart city entails "an intersection of city administration, citizen value creation, local business, ICT development and application, urban big data, economics, and sociology, among other" (Lim, Kim, & Maglio, 2018, p. 88). The enriching experience gained with the emerging market of ICT, particularly with big data and networks, surely offers more tools and opportunities to allow for better interaction and responses in cities, both in social and operational decision-making processes (Batty, 2013). However, this development also encompasses (known and unknown) risks to privacy and security, particularly issues regarding the processing of personal data and anonymization (Batty, 2013).

2.1.1 Big data

After exploring the several innovations introduced by smart cities with the incorporation of ICT technologies, this section dive into the inherent introduction of big data to the smart city scenario. Likewise, it will briefly elucidate the dilemmas concerning data protection, especially with the increasing use of big data analytics.

As technologies evolved, digital technologies invaded people's daily lives and the interaction between humans and machines produced "large and fast-growing volumes datasets which go beyond the abilities of commonly known data management systems to accommodate" (Osman, 2019, p. 260) data. Big data is defined as large or complex datasets developed to overcome the limits of "traditional data processing applications" (Mehta & Rao, 2016, p. 120) and process large volumes of "digital traces of human activity" (Lim et al., 2018, p. 86) – usually in the form of raw data. When applied to smart cities the data collected varies from urban assets to the different stakeholders, and the main value of big data comes with the capacity to extract valuable information, produce knowledge and positively affect both the city and stakeholders (Lim et al., 2018; Osman, 2019).

Big data analytics concerns the "process of probing big data set to reveal hidden patterns, unknown correlations and other important information that can be used to make ['successful'] decisions" (Kumar & Prakash, 2013, p. 14). Big data analytics or big chain value chain, "which is considered as one of the key enabling technologies of smart cities" (Osman, 2019, p. 260), introduced groundbreaking methods of data extraction and challenged researchers to develop sophisticated methods, techniques and platforms specifically designed to deal with big data – and, consequently, with smart cities' projects.

2.2 PRIVACY CONCERNS

After exploring the context of smart cities, this section reflects on the impact of privacy posed by different smart city technologies in public spaces. Firstly, the definitions of privacy and data protection are delineated and, secondly, this section elucidates the discussion concerning the safeguard of data protection in smart city projects by presenting the framework of PbD.

According to Moore (2008), several authors have described privacy in different terms along the years. The definitions vary from the 'right to be let alone' (Warren & Brandeis, 1890) and "the state of possessing control over a realm of intimate decision, which include decision about intimate access, intimate information, and intimate action" (Julia Inness, as cited in Moore, 2008, p. 412), to the development of typologies for privacy that can encompass "information control" (Alan Westin, as cited in Moore, 2008, p. 412), 'behavioral privacy', 'bodily privacy' and others (Koops, Newell, Timan, Chokrevski, & Gali, 2017). Still, the scope of privacy overlap with the scope of data protection, "but [there are] also some areas where their personal and substantive scope diverge" (Kokott & Sobotta, 2013, p. 228). In that sense, the Court of Justice of the European Union (CJEU) ensures this broader scope of privacy by adopting the definition of personal data by the EU Data Protection Directive, which was later replaced by the GDPR, and by the Data Protection Convention of the Council of Europe. However, the two rights are distinguished in the EU Treaties and in the Charter of Fundamental Rights of the EU, since Article 7 concerns the "Respect for private and family life", and Article 8 encompasses the "Protection of personal data" ("Charter of Fundamental Rights of the European Union," 2012, p. 326; Kokott & Sobotta, 2013; Smith, 2016).

The several definitions and different terms used to describe privacy reflect its inherent adjustability. According to Klitou (2012), an impossibility of a consensus about the concept of privacy is also a result of changes in values over time, the context in which privacy is considered and the different opinions among scholars. For instance, the Charter of Fundamental Rights of the EU does not define what privacy is; however, it focuses on describing the terms in which privacy can be applied by describing the scope of 'private life' (Psychogiopoulou, 2017). Likewise, the considerations about values play an important role in the discussion of privacy since it entails the need to analyze the societal context: it stresses the importance to embrace the pivot role of technology, social acceptance, social norms and political momentum

(Klitou, 2012). Therefore, in the context of 'surveillance society', 'risk society' and the urge for smart solutions using predictive systems, defining privacy appears even more challenging (Klitou, 2012); nevertheless, the social relevance of these debates is boosted by the need to define the purpose and limits of data in the era of big data.

Although the scope of privacy revolves around 'private life' and the aforementioned definitions, it does not specifically include "all information on identified or identifiable persons" (Kokott & Sobotta, 2013, p. 225) – which is considered the scope of data protection. In that sense, the formulation of the concept of data protection was a pioneer move to ensure the protection of personal data. Hence, data protection, as a fundamental right, entails "the protection of natural persons in relation to the processing of personal data" (Recital 1, "GDPR General Data Protection Regulation," 2016, p. 1). The right to the protection of personal data – provided by the GDPR, the Charter of Fundamental Rights of the EU and the Treaty on the Functioning of the European Union (TFEU) – reflects the importance of limitation of data and purpose associated to the processing of personal data, which includes specific risks that regulations should consider (Kokott & Sobotta, 2013). Therefore, the right specifies the principles and ensures appropriate purpose and limitation in which personal data can be processed and collected. Accordingly, the positive outcome of such process is that regulations mitigate most of the (known) risks caused by data-driven innovations whilst regulators can focus their efforts in protecting the society from other potential (unknown) risks (von Grafenstein, 2018).

As technologies evolved, privacy shifted from a personal good toward a societal value (Cavoukian, A., & Jonas, J., 2012). In that sense, data management became the center of new data protection regulations, such as with the creation of Fair Information Practices (FIPs) (Cavoukian, A., & Jonas, J., 2012, p. 7). The focus of the new regulations was: first, on the "purpose specification and use limitation for disclosure of personally identifiable information"

(Cavoukian, A., & Jonas, J., 2012, p. 7), second, on "user participation and transparency" (Cavoukian, A., & Jonas, J., 2012, p. 7) giving users an active "participatory role concerning the lifecycle and disclosure of their personal data" (Cavoukian, A., & Jonas, J., 2012, p. 8) and; finally, on "the need for strong security to safeguard the confidentiality, integrity and data availability as appropriate to the sensitivity of the information" (Cavoukian, A., & Jonas, J., 2012, p. 8).

The theoretical discussion about privacy proposed by Friedman (2000) stressed how different technologies allow for individuals' control of their own information. These technologies, within the category of Information Processing, were developed to "substantially affect the cost of obtaining information about other people, concealing information about oneself, and transacting in information" (Friedman, 2000, p. 204). With the technological development and the exponential engagement of individuals with the virtual world, informational privacy became the center of privacy discussions. Once the development of information processing technologies enabled an inexpensively collection of huge amounts of information, the increase availability of information online facilitated the growth of organizations collecting and processing them.

Particularly in the smart city context, privacy encompasses: the physical aspect, data subjects' control over their own data collected and shared to third parties, and awareness about the risks when personal data is wrongfully available to unauthorized actors (Cilliers & Flowerday, 2015). Moreover, the constant development of privacy rights anchored on the notion brought by Warren and Brandeis (1890), the 'right to be let alone', stimulated several new considerations regarding privacy, such as the definition of information privacy: "information privacy relates to the person's right to determine when, how and to what extent information about him or her is communicated to others" (Westin, 1967, p. 1). In that sense,

personal information is the product in discussion and its use and availability is the current challenge for privacy rights.

As stated by Friedman (2000), one implication of information processing`s development entails the capability of organizations to collect, store and use these information available worldwide. Another one entails the increasing interest in collecting dispersed information to be used in the future by private companies – with the purpose to facilitate voluntary transactions and "produce net benefits" (Friedman, 2000, p. 204) – and the difficulty it poses to define a limitation in order to avoid the use of such information for other purposes (Friedman, 2000).

However, the discussion on privacy introduces other issues when it enters the public sphere. Several authors (Koops et al., 2017; Nissenbaum, 1998; Timan et al., 2017) address the gap regarding privacy in public spaces by denouncing the false dichotomy between the notions of private and public; the dichotomy pushed the right to privacy towards the realms of the private space and excluded the public sphere from the scope of privacy (Nissenbaum, 1998). Notably, with the expansion of ICTs the collection of private information, which used to be inaccessible and constrained to the private sphere, intensified; for instance, "by connecting to wired or wireless networks (3G/4G, WiFi, etc.), which our devices often do automatically, even without our knowledge, we automatically bring all sorts of things into (virtual) public space" (Timan et al., 2017, p. 2). Accordingly, in the context of smart cities, the scrutiny of privacy concerns in public spaces becomes indispensable, since it explores the two facets of governments: the responsibility i. to guarantee individual's privacy, and ii. to ensure safety, which threatens privacy in both private and public spaces (Timan et al., 2017).

On the one hand, as innovations in ICT and the attractiveness of smart solutions to urban issues exponentially grows, the urge for understanding privacy issues emerges from big data analytics and the implementation of different technologies (van Zoonen, 2016). Arguably,

discussions regarding the "social effects of data collection and analysis" (van Zoonen, 2016, p. 472) provide different insights on how these new types of governance based on the 'politics of city data' threaten people's right to privacy when considering the gathering of sensitive personal information being collected. For instance, a framework of privacy concerns in smart cities could be analyzed in two dimensions: sensitiveness of data (personal or impersonal) and data collection purpose (services or surveillance), which should also include the processing of the data collected (van Zoonen, 2016).

On the other hand, these technologies substantially prevent individuals from exerting control over their personal information (Cavoukian, A., & Jonas, J., 2012). Although this process was minimized with the creation of a new regulation that aims to mitigate data breaches and restore individual's control over the lifecycle of their personal information, the GDPR (Ryz & Grest, 2016), big data still "generate enormous values to society" (Cavoukian, A., & Jonas, J., 2012, p. 13). Friedman (2000) suggests that at the same time information technologies are capable of processing and collecting information about many people, it "also makes it inexpensive to keep track of the conditions under which various pieces of information can be disclosed" (Friedman, 2000, p. 205). Moreover, the author developed his argument by engaging in anonymized transactions, which ensures that no party involved in the transaction get access to relevant information about the individuals. This approach entails a "combination of technologies for information processing and encryption" (Friedman, 2000, p. 205) to safeguard information and protect privacy from any type of interception. Therefore, designers and engineers should be aware that a more responsible innovation should keep pace with ethical considerations and embedded privacy settings (Cavoukian, A., & Jonas, J., 2012).

2.3 PRIVACY BY DESIGN

The efforts on inserting privacy into the technological development depicted the urge for developers to take into consideration their socio-ethical responsibility in the process. For instance, it addressed the importance of considering what impacts for privacy a new system would have. Thus, this section will elaborate on the emergence of PbD as part of the solution to the problem.

During the 90's, the category of Privacy-Enhancing Technologies (PETs) was formalized by Ann Cavoukian, the Ontario's Privacy Commissioner, the Dutch Data Protection Authority and the Netherlands Organization for Applied Scientific Research. The increase in interconnected information technologies and in the volume of personal data collected pushed the mindset on privacy protection from legal compliance to the incorporation of technologies that could enhance privacy as well (Cavoukian, 2010). This shift was introduced by the concept of PETs stressing how data protection regulations' practices "could be reflected in information and communication technologies to achieve strong privacy protection" (Cavoukian, 2010, p. 247) and proposing "technologies that minimize the processing of personal data" (Domingo-Ferrer et al., 2014, p. 1) by maintaining their trustworthiness as well. Although it took some time until the concept finally reached global reconnaissance, eventually it was incorporated by both privacy and information technology fields. As technology and privacy evolved and constantly (re)shaped and challenged societies, the stakes for data subjects became higher.

Finally, in 2009, during the conference *Privacy by Design: The Definitive Workshop* in Madrid, the seven principles of privacy by design were presented. PbD was proposed as one of the most relevant guidelines for data privacy at the time. The concept was an extended version of PETs and was formulated to address "the ever-growing and systemic effects of Information and Communication Technologies, and of large-scale networked data systems" (Cavoukian, 2011, p. 1). Similarly, to the PETs' objectives, PbD encompasses the need for preventative

measures for privacy by changing organization's mentality regarding privacy and incorporating these strategies as their default mode of operation, instead of mainly relying on remedy solutions provided by legislations. PbD was intended to fully safeguard personal data, particularly sensitive data, with the objective of ensuring privacy and individuals' control over their own personal data, while offering a "sustainable competitive advantage" (Cavoukian, 2011, p. 1) to organizations.

The importance of the PbD strategy is to ensure responsible information management with regard to strengthening personal data and maintenance of relationships in the context of technological development. It is extremely necessary for business to incorporate a privacy approach in order to ensure customer's trust and to keep generating business (Cavoukian, 2010). Indeed, to keep pace with the technological and business success, one must be able to demonstrate a privacy focus strategy and compliance with trusted privacy practices and technological procedures according to one's necessity. Thus, internal and external growth are mainly related to these requirements and can be translated into competitive advantages in the market – what Cavoukian called "Privacy Payoff" (2010, p. 249).

Additionally, the concept of PbD addressed the importance of moving beyond compliance and adopting preventative measures to enhance organizational accountability (Cavoukian, Taylor, & Abrams, 2010). The debate on the necessity of an "accountability-based regulatory structure" (Cavoukian et al., 2010, p. 407) was proposed in cases where societal objectives towards the protection of individuals from any harm need to be embedded in organizations. The authors state that PbD, as a conceptual model, can enable this level of privacy-protective control for management of information into every layer of a business process. Therefore, accountability becomes the governance model for organizations and stresses the importance of overcoming privacy and security risks through PbD (Cavoukian et al., 2010).

In that sense, PbD was intended to fully cover what Cavoukian (2010) called the "'trilogy' of encompassing applications" (2010, p. 249): first, IT systems, second, accountable business practices and, third, physical design and networked infrastructure. The differential perspective of PbD is the constant need to ensure personal privacy and to provide a solid foundation of trust. Incorporating the strategy of PbD relies on fully complying with the seven foundational principles developed by Cavoukian. Thus, the essential seven foundational principles proposed by the PbD approach are:

1. ***Proactive*** *not Reactive;* ***Preventative*** *not Remedial*

   The PbD approach is expressed in this principle by the change from reactive to preventative measures. It proposes an attitude that focus on anticipating privacy invasive events and preventing them from happening. Thus, PbD aims to minimize risks by focusing on prevention and avoiding remedy actions for privacy events; it "comes before-the-act, not after" (Cavoukian, 2011, p. 2). One of the examples of how to achieve this would be by adopting mechanisms to resolve privacy issues before they could evolve into a problem (Cavoukian et al., 2010).

2. *Privacy as* ***Default Setting***

   The default settings in this principle relates to ensuring any type of IT system or business practice have the maximum degree of privacy for personal data by default. The idea is that personal data must be automatically protected in any system; thus, individuals are not required to take action to protect their own information (Cavoukian, 2011). One example is to have a secure environment where the collection and processing of data would happen, ensuring consumers/customers' trust on the entire process (Cavoukian et al., 2010).

3. *Privacy **Embedded** into Design*

   This principle addresses the importance of embedding privacy "into the design and architecture of IT systems and business practices" (Cavoukian, 2010, p. 250). In that sense, it means that privacy should be considered by engineers since the first step of the designing phase. Likewise, preventative privacy-measures should be introduced into the core functionality as one crucial component to be delivered. Furthermore, since organizations' accountability can be assured when privacy is embedded into the design of different business processes (Cavoukian et al., 2010), delivering privacy as an integral part of the system assures both privacy and functionality.

4. *Full Functionality – **Positive-Sum**, not Zero-Sum*

   This principle evokes one of the concepts from game theory, namely the Positive-Sum. To begin with, Cavoukian cites two of the concepts within game theory: i. positive-sum, the sum of the outcomes of a situation must be greater than zero, and ii. zero-sum, the sum of the outcomes must be equal zero. In this model the variables in question are *functionality* and *privacy*, thus, the sum of the equation relates to the sum of changes in both variables. PbD aims to achieve a double 'win-win' model, where both functionality and privacy can be reinforced, avoiding the 'trade-offs' forced by the (false) dichotomy of privacy vs. security within the traditional model of zero-sum (Cavoukian, 2011). In other words, with PbD privacy and functionality can be equally enhanced and achieved without any trade-offs. Hence, organizations with structured mechanisms and rules to ensure individual privacy are aware of the risks involved in this process and are capable of generating economic value more appropriately (Cavoukian et al., 2010).

5. *End-to-End Security – **Full Lifecycle Protection***

   This principle addresses the importance of ensuring a proper lifecycle management of information. PbD seeks to embed privacy measures "throughout the entire lifecycle of

the data involved" (Cavoukian, 2010, p. 250), for instance, securing all the data will be destroyed within the expected time at the end. In order to be accountable, organizations must have privacy-controls in their business processes to ensure adequate assessment concerning the lifecycle of the data (Cavoukian et al., 2010).

6. ***Visibility*** *and* ***Transparency*** *– Keep it* ***Open***

   This principle is aimed at ensuring that every stakeholder within the trilogy must operate according to the "stated promises and objectives, subject to independent verification" (Cavoukian, 2010, p. 250). Independently of the technology or business practice, operations and components must be open (visible and transparent) to verification by both users and providers. Thus, to be accountable means to be answerable for its business processes and practices, but it also means to have the responsibility towards individuals to provide all the necessary information about each process (Cavoukian et al., 2010).

7. ***Respect*** *for User Privacy – Keep it* ***User-Centric***

   Lastly, the seventh principle of PbD addresses the importance of primarily keeping the interests of individuals by ensuring strong measures are in place, such as by privacy by default, meanwhile it assures "appropriate notice, and empowering user-friendly options" (Cavoukian, 2010, p. 250). Therefore, the principle seeks to maintain the interest of individuals as the main focus of the model, which is called user-centric. Likewise, the management of information must respect individual's privacy (Cavoukian et al., 2010).

Privacy by Design as a concept, along with the seven principles developed by Cavoukian, set the pace for innovative adaptable standards in privacy measure to several areas, such as health care and smart grids. Due to its adaptability, a wide range of disciplines and areas of service embraced it as part of their discussions; the principles introduced new

interpretations such as the use of PbD as an approach to tackle "operational and management issues" (Cavoukian, 2010, p. 250). Since the conceptualization of PETs and the development of PbD at least two main improvements were introduced: i. the focus on the positive-sum paradigm and ii. the possibility "to consider technology, business processes, management functions and other organizational issues in a comprehensive manner" (Cavoukian, 2010, p. 251), whilst embedding privacy in all the layers.

More than extending the trilogy applications and ensuring preventative measures during the entire process-model, PbD seeks to bring awareness to the need for a culture of privacy (Cavoukian, 2010). Ideally, it should be incorporated as an organizational approach to create positive privacy controls and business opportunities (Cavoukian, 2010). Thus, PbD seeks to integrate different parts of privacy protections such as legislations, privacy instruments, consumer/customer awareness, accountability, etc (Cavoukian, 2010).

2.3.1 Criticism of the Privacy by Design approach

Before diving into the critical debate and analyzing the PbD principles, it is keen to reflect on the reasoning behind the term 'Privacy by Design' itself. Since the principles of PbD are mostly focused at safeguarding data protection, why is the concept not named as 'data protection by design'? To answer this question, it is relevant to analyze Article 25 and the Recital 78 of the GDPR. They state that the implementation of appropriate technical and organizational measures by companies/organizations with regard to the processing of personal data "at the earliest stages of the design of the processing operations, in such a way that safeguards [both] privacy and data protection principles right from the start ('data protection by design')" ("GDPR General Data Protection Regulation," 2016; "What does data protection 'by design' and 'by default' mean?," n.d.). In that sense, the concept of PbD merges the urge to incorporate both privacy and data protection features to implement a desired functionality, which means achieving "successful privacy-friendly design of systems and services"

(Domingo-Ferrer et al., 2014, p. 4) in a broader spectrum of the field and promoting an extensive range of action to safeguard data subjects' rights.

The principles of PbD proposed a pragmatic approach to the development of privacy and security in technology by integrating several perspectives and methodologies involved particularly in information systems. Indeed, the importance of PbD and the awareness regarding the risks at stake were reflected on the GDPR and on other privacy practices. Still, the strategy is considered too vague to be applied as one practical guideline (Domingo-Ferrer et al., 2014; Gurses et al., 2011; Kroener & Wright, 2014; Perera et al., 2016; Spiekermann, 2012; Wiese Schartum, 2016).

The first principle entails using privacy standards through a comprehensive and proactive approach. PbD aims to cover most of the elements in the composition of a system/process in order to ensure data privacy and accountability. However, a gap was created between policy makers and engineers when interpreting PbD; due to their lack of knowledge in privacy combined with their of experience in engineering systems with privacy in mind and the vagueness regarding recommendations about how to achieve data protection (Gurses et al., 2011). Additionally, not only there are several definitions of privacy, but also legislations and regulatory frameworks may vary according to the region and context where they are implemented. Thus, adopting privacy regulations when developing a system also depends on the desirable and mandatory requirements of each context (Wiese Schartum, 2016).

Another issue concerning the gap between policy makers and engineers is the problem of communication and lack of clarity, which can be reflected on consumers privacy. Since PbD is a socio-technical solution, it is keen that both technical and non-technical strategies have explicit privacy requirements to be translated into systems (Gurses et al., 2011). Likewise, when delineating an effective and comprehensive guideline for smart cities projects it is crucial to consider that such context demands attention to its architectural aspect, usually in the form

of IoT structures. Since smart cities use IoT applications, they encompass both software and hardware components engaging in "multiple heterogenous nodes with different capabilities under different conditions" (Perera et al., 2016, p. 2). Thus, it is essential to provide a PbD framework as a systematic guidance to software engineers to assess such complex environments, since it can lead to more consistent results (Perera et al., 2016).

The second principle states privacy requirements must be into the default settings of systems in order to ensure the protection of personal data and, consequently, social trust. Nonetheless, the lack of guidance on how to achieve these privacy-friendly settings according to the several legal requirements makes this principle unclear. Hence, software engineers and those involved in the designing process ought to have in mind the importance of understanding which applicable privacy regulations they must follow in order to assure products' or systems' compliance since the beginning. In that sense, one of the most suitable solutions is the integration between areas and experts during the designing step, incorporating the expertise of different disciplines to achieve the desirable result in terms of privacy, security and functionality (Gurses et al., 2011).

The third principle concerns embedding privacy requirements into the designing phase of a system. The challenges of this principle entails organizations' commitment to privacy strategies – since personal data is one of their core assets – as well as the integration between stakeholders in the architectural decisions to assess privacy risks during systems development (Spiekermann, 2012). Particularly, software engineers, designers and other actors involved in the development or maintenance of information systems often lack the knowledge to understand how the principles ought to be applied during the development of a system (Gurses et al., 2011). The process of engineering systems should integrate several basic requirements – security, privacy and functional related, among others – and it should involve risk and threat analysis as well.

The fourth principle encompasses achieving privacy requirements without trading it for any security features, or vice-versa, when considering the functionality of the system. This notion addresses concerns regarding its feasibility and put into question its desirability for existing systems/processes. Likewise, by moving away from the dichotomy between security and privacy, it dives into the need for 'pragmatic trade-offs' that encompass the desirability of the outcome between security by design and privacy by design, namely functionality (Bier, Birnstill, Krempel, Vagts, & Beyerer, 2012). Besides, another negative consequence caused by the vagueness in definitions also applies to the mechanisms used to ensure data protection, namely data minimization. For instance, the Data Protection Directive (FIP) and the Article 6 of the General Directive of the EU restricts data collection and processing differently (Gurses et al., 2011; Kroener & Wright, 2014). This fuzziness concerning data minimization opened the debate about the consequences of its implementation. Thus, it entails the debate on privacy and functionality on the operational level, while discussing what should be the minimum amount necessary for the processing of personal data on the implementational level. (Gurses et al., 2011).

The fifth principle of PbD says it is essential to thoroughly analyze the data life cycle in order to ensure an adequate data management protection. The importance of understanding the IoT architecture present in smart cities is to consider how the data flows according to the type of IoT application, but also to reflect on how multiple stakeholders, when applicable, ensure data privacy requirements in this scenario. The data flow on the IoT application relies on the type of architecture used, which can be either centralized or decentralized. Irrespective of the type of architecture, each layer is composed by specific elements as exemplified in Figure 1 (Perera et al., 2016). Hence, the technical challenge for PbD is to ensure privacy protection capabilities in each layer of the net while reaching the results expected with several devices in place. Moreover, the non-technical challenge relies on assuring organizational accountability

within multiple stakeholders involved in this net, defining specifications according to responsibilities and roles.

The discussion regarding the data flow of the IoT architecture is relevant to explain how data moves along the cycle (life cycle phases). According to Perera et al. (2016), "data moves through five data life cycle phases" (2016, p. 3) and by presenting the cycles and their respective layers it is possible to define the actions that should be taken in each part. Hence, in the smart city scenario each phase demonstrated on Figure 1 represents one step to be assessed. Thus, this is how the PbD guidance can be provided to IT developers during the entire development process of a concept (Hoepman, 2012).

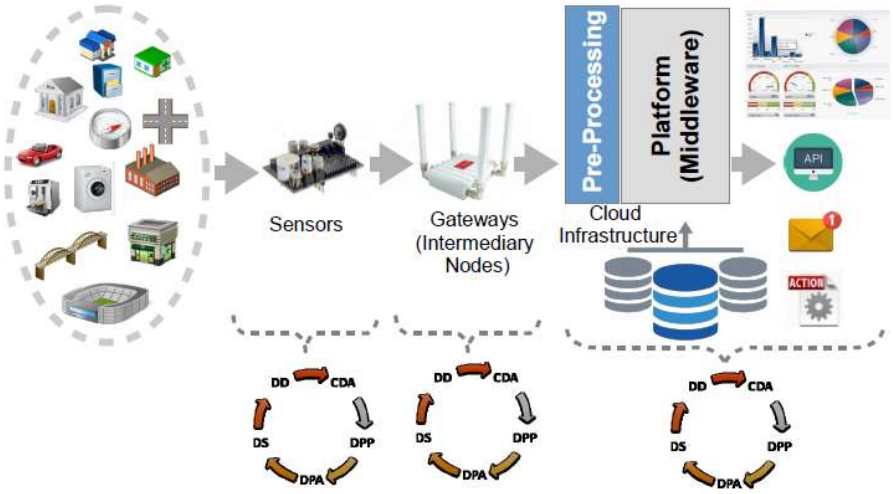

Figure 1. Typical Data Flow in IoT Applications. Reprinted from Privacy-by-Design Framework for Assessing Internt of Things Applications and Platforms in Germany, by Perera et al., November 2016, retrieved from http://dl.acm.org/citation.cfm?doid=2991561.2991566 Copyright 2016 of ACM Press.

The sixth principle addresses the need for transparency by keeping every process as visible and open as possible. Although there are several legal requirements regarding transparency, there is no clear guidance on how it should be done (Wiese Schartum, 2016). For instance, as Recital 39 of the GDPR states "any information and communication relating to the processing of those personal data be easily accessible and easy to understand" ("GDPR General Data Protection Regulation," 2016, p. 7), but it leaves room to discussion on how to do it.

Likewise, it is also crucial to give the same importance to accountability and transparency in the guideline; this argument is based on the fact that organizations must be certain of their responsibility to be answerable for their actions.

Lastly, the seventh principle concerns the data management centered in user's privacy and the empowerment of users. However, the principle lacks a realistic approach to privacy-design systems and its user-centric notions restricts users mainly to data subjects, excluding other potential ones, such as in the case of controller's personnel or an "automated system operated by a large number of data subjects" (Wiese Schartum, 2016, p. 70). Besides, Article 23 of the GDPR should be expanded in order to the encompass more than only controllers' systems when it comes to enforcing privacy by design to ensure privacy protections (Wiese Schartum, 2016)

Therefore, these critiques denounce the urge for updating the PbD principles developed by Cavoukian, since it is not clear what are the crucial requirements to the development of IT systems when considering privacy requirements and focusing on data protection legislations. Hoepman (2012) develops one of the most prominent strategies of privacy by design. His concept on PbD focus on a guidance for IT developers to be used "throughout the full software development life cycle" (Hoepman, 2012, p. 2). In that sense, the guidance works by providing a translation of the most important strategies: "Minimize, Hide, Separate, Aggregate, Inform, Control, Enforce and Demonstrate" (Hoepman, 2012, p. 2). Accordingly, Perera et al. (2016), developed a new framework of PbD that focused on the first strategy developed by Hoepman – Minimize -, validating their perspective on the fact that technologies, IoT applications in this case, should "achieve their goals with the minimum amount of data" (2016, p. 3). Hence, Perera et al. (2016) put their focus on minimization techniques, which leads to a new set of guidelines that ensures compliance through a rigid assessment of IoT applications in relation to privacy requirements.

Arguably, after reflecting on the seven foundational principles of PbD and its most prominent critiques, it is conceivable the proposition of a different guideline to be used as a practical model for PbD. This new model shall encompass a practical strategy to address the previous gaps in the field: what privacy by design is and how it should be applied in practice (Domingo-Ferrer et al., 2014; Perera et al., 2016; Wiese Schartum, 2016).

2.3.2 New Guideline on Privacy by Design

This new PbD guideline will develop a more appropriate and practical translation of privacy requirements through explicit and well-defined measures. It addresses the "social, legal and ethical concerns into systems requirements" (Gurses et al., 2011, p. 1), and, more importantly, the concerns on data minimization principles. Moreover, this guideline combines several approaches proposed about PbD while taking into account the most relevant criticism on the field. Thereby, the new PbD guideline combines parts of existing suggestions on PbD, consolidating legal and technical elements through a case-by-case approach.

1. ***Identification***

   This principle concentrates on proactively identifying the context of information systems or design processes, by proposing a series of evaluations as early as possible, originated from the first principle of Cavoukian. This principle emphasizes the importance of incorporating privacy considerations into every step of IT systems' and organizations' structures. Ideally, experts should be working together to provide support and guidance to ensure a constant privacy-oriented mindset since the idealization phase. This involves incorporating several narratives to discuss the potential ethical and social implications of information systems to society, which could be embodied in the form of internal privacy policy as well as training programs. In that sense, the ideal scenario is a proactive behavior that expects 'over-fulfilling' the systems and processes with privacy requirements (Wiese Schartum, 2016).

This principle aims to cover different settings and structures, both organizational and technological, by taking into account the systems', actors' and regulations' requirements of each case. Since there is no 'one solution fits all', the goal is to provide tailored analyzes for technology's and regulation's contexts to ensure that the most appropriate measures can be applied.

2. ***Architectural Privacy - Embedding Security***

This principle addresses the importance of assessing the scope of the architectural platform, taking into account both new and existing systems, in order to assist and enforce information security and privacy techniques: it combines the second, third and fifth principles of Cavoukian and defines specific rules on how data management and the privacy requirements must be implemented. To incorporate a higher number of structures, particularly when considering smart cities platforms, the reference architecture used is inspired on the model presented in Figure 2. Therefore, the architectural layers encompass: a) Abstract; b) Service, c) Concrete, and d) Platform.

This principle focus on the incorporation of security and privacy techniques in every layer of the architecture, pursuing "confidentiality, accessibility and integrity" (Wiese Schartum, 2016, p. 156). Thus, it ensures technical and organizational mechanisms for privacy protection of information systems without loss of functionality; moving beyond the limited notion of privacy-friendly by default – when data subjects input information that can impact their own privacy (Wiese Schartum, 2016).

Incorporating mechanisms to comply with this principle should encompass information security by design and privacy by design as well, for instance, by applying transparency-enhancing techniques and intervenability-enhancing techniques (Wiese Schartum, 2016). To achieve this requirement, some of the mechanisms are: authentication, secure communication, identity management, chain aggregation,

knowledge discovery based aggregation, geography based aggregation, chain aggregation, time-period based aggregation, category based aggregation, data anonymization, encrypted data processing, encrypted data storage, and reduce data granularity (Perera et al., 2016; Wiese Schartum, 2016).

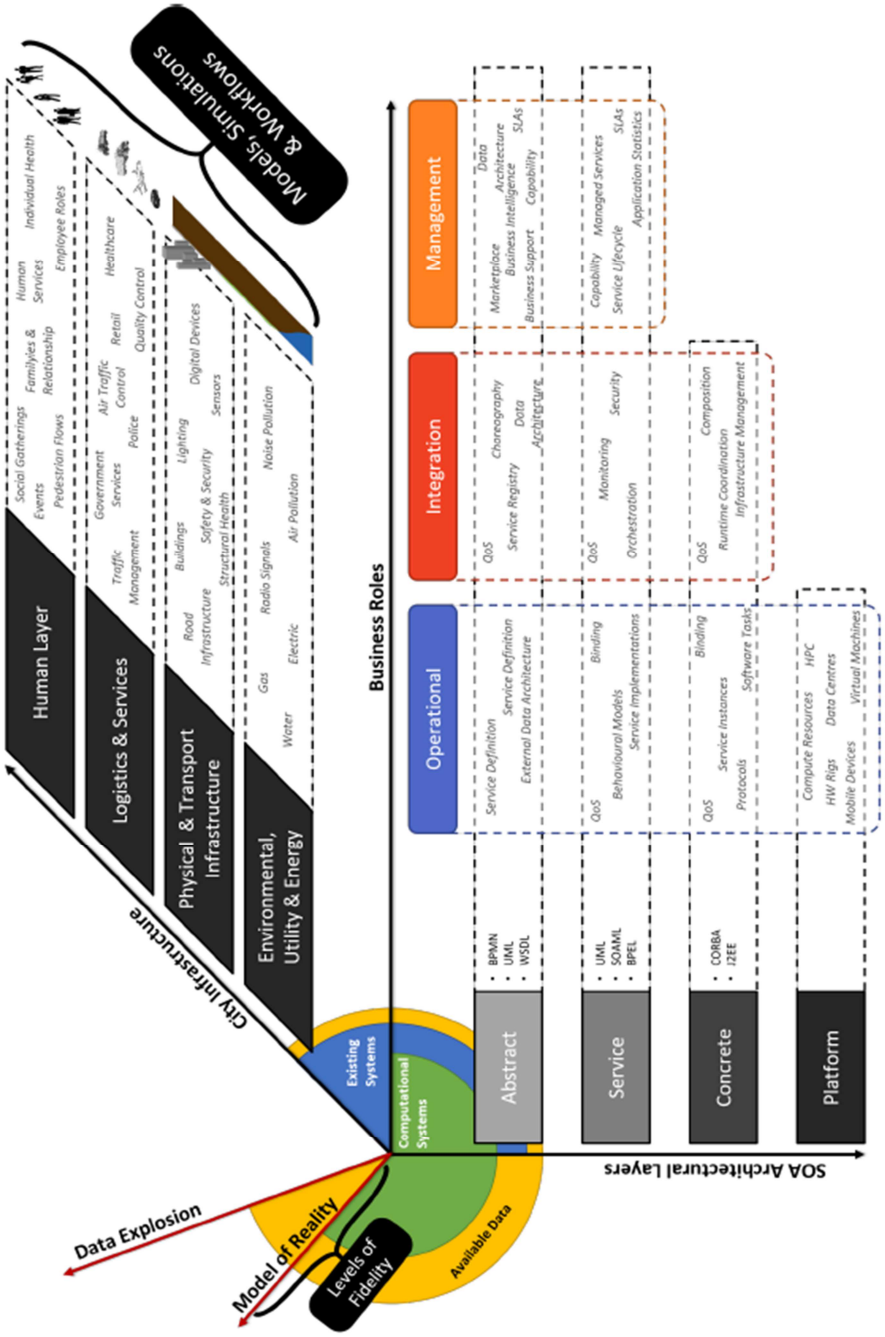

Figure 2. Service-Oriented Reference Architecture for Smart Cities. Reprinted from Service-Oriented Reference Architecture for Smart Cities in Leeds, UK, 2017, retrieved from http://eprints.whiterose.ac.uk/113342/ Copyright 2016 IEEE.

3. ***Minimum Data, Maximum Privacy***

This principle reflects on the relation between data, functionality and privacy. Avoiding the positive-sum proposition of Cavoukian's fourth principle, this principle also calls for an integration of security by design and privacy by design. This principle evokes the need for a clear guidance on how to combine privacy and security strategies when thinking about functionality, which is the output of the two strategies. Moreover, the focus of such combination should be on data minimization, simultaneously ensuring data utility and data management control based on privacy requirements. Thus, it allows for existing IT structures to incorporate PbD principles into any phase of the system in order to prevent business activities' or business systems' disruption; however, it is important to note that in the case of legacy systems this solution may not work properly. For instance, some of the techniques that can be used to ensure data minimization are: minimize data acquisition, minimize the number of data sources, minimize raw data intake, minimize knowledge discovery, minimize data storage and minimize data retention period (Perera et al., 2016).

4. ***Transparency and Accountability***

This principle entails the responsibility to inform data subjects, demonstrate and be accountable for every system and activity involving the processing of personal data, based on Cavoukian's sixth principle. This principle follows the considerations of Article 12 of the GDPR: "transparent information, communication and modalities for the exercise of the rights of the data subject" ("GDPR General Data Protection Regulation," 2016, p. 39), encompassing any measure to ensure the exercise of data subjects rights. Since data subjects may only exert agency when they are aware and, informed of and, if required, have consented to the processing of their personal data, it's crucial the employment of categorical information contents (Wiese Schartum,

2016). Additionally, as mentioned during the previous sub-section, accountability must be taken into account in combination with transparency in order to establish unambiguous procedures and actions; consequently, keeping in mind that visibility and answerability are interdependent and beneficial to organizations, besides increasing social trust (De Fine Licht, Naurin, Esaiasson, & Gilljam, 2014).

Some of the examples to achieve this principle are: compliance – policies, laws, regulations, privacy framework, guideline, etc –, demonstration of data flow diagrams, data management, certification and legitimacy of systems and processes, standardization of security and privacy practices (ISO), logging and auditing (Hoepman, 2012; Perera et al., 2016).

5. ***Data Subject Control***

This principle argues for a shift from the concept of data subject rights to the notion of data subject control, contrasting with the seventh principle proposed by Cavoukian. As such, combining some elements of informational privacy, it suggests "privacy interest exists in restricting access or controlling the use of information about that aspect of human life" (Koops et al., 2017, p. 569). Hence, it explains that since privacy legislations support data subject rights by imposing several requirements to ensure individuals' agency over their own personal data, organizations merely comply with them out of obligation. However, organizations should also comply with legislations to ensure data subjects control, and not only to avoid backlashes for non-compliance.

Additionally, it is essential to data subjects to have a central role in being informed about organizations' processes through transparency and accountability. The positive consequences of such change in mindset would encompass an increase in social trust and, particularly, the strengthen of organizations' legitimacy, which entails the

> "acceptance of decision-making procedures" (De Fine Licht et al., 2014, p. 113) due to the general belief that an organization is 'appropriate and just' to perform them.
>
> In order to safeguard data subjects' rights, organizations must ensure individuals' agency over their own personal data, as well as the power to fully exercise their rights. Consequently, this principle aims to move beyond a mindset, avoiding the idea of ensuring data subject rights to what it's defined here as promoting 'data subject control', by focusing on data subjects' agency over their own information. A precise and well-known example would be through Information disclosure proposed by Perera at al. (2016), which suggests an active information about acquisition, processing and dissemination of personal data at 'any stage of the data life cycle' (2016, p. 5).

This guideline should be incorporated within (any of) the phases of IT systems' development: idealization, elaboration and implementation. It is crucial to consider that depending on each case's context the mechanisms to be applied might change according to what is required. Besides, the mechanisms proposed are not related to one specific principle nor to one specific phase, since each case has its own particularities and needs. Hence, a tailored analysis incorporates the different specifications of the case and provides a clear guidance for developers on how, where and when to apply privacy-preventive mechanisms.

2.4 THE STRATUMSEIND 2.0 PROJECT

The choice for the Stratumseind 2.0 project – a living lab – in Eindhoven is based on the innovative use of information systems combined with other technologies to propose smart solutions to the city. As a unique case, it proposes different discussions concerning privacy by design and the development of different systems applied in the urban space, although it is not fully consolidated yet. The initiative started as a proposition by the city of Eindhoven to reduce criminal behavior at the longest pub street in the Netherlands, with the involvement of multiple

stakeholders. Then, the city of Eindhoven decided to introduce technology as part of the solution for the economic and security agenda.

As a combination of different projects organized by the DITSS (Dutch Institute for Technology, Safety & Security), the smart city project combines several other actors providing their own expertise in each area. The project is partially a living lab – organized by the DITSS – aimed to provide smart solutions and improve the quality of life in the city through the incorporation of smart sensors. Moreover, they implemented smart ICT-based lightening systems initially designed to deal with people's unwillingness to leave their homes to go to the city's famous pub street, Stratumseind, due to high levels of aggressiveness. The objective of the project was to reduce criminal behavior and support economic growth in a critical and (still) important area of the city, combining resources from different stakeholders to implement the achieve the results. For instance, Philips in responsible for implementing the lightening systems that can be controlled in the Stratumseind 2.0 base; Sorama support with the sound intelligence analysis – visualizing and localizing –; and Axis provides the CCTV cameras combined with ViNotion crowd management software capable of interpreting the images.

The project, as a living lab, monitors several high-tech materials implemented to collect data from the Stratumseind. There are "five telephoto cameras and several sound meters and 22 LED lamp posts in the street" (Merlijn van Dijk, 2018) , and these devices can influence the mood of people by using light. The preciseness of these devices rely on an effective data collection; for instance, the sound cameras can differentiate the sound of a "gunshot, fireworks and breaking glass" (Merlijn van Dijk, 2018). Besides the living lab, the company Atos works in a joint cooperation (CityPulse pilot) by offering the real-time analysis – control center – of the street. In a different setting, Atos provide the real-time intelligence to analyze the information provided by the several sensors on the street. Likewise, they are responsible for

promptly sending the information to de Dutch Police in Eindhoven whenever an event is triggered.

Based on the level of complexity, an analysis of this project through a practical PbD guideline may provide an impact assessment and compliance overview of the several systems. Indeed, the uniqueness of smart cities projects instigated a new guideline that could address the specifications of the case while offering general strategies to other cases. The new PbD guideline proposed in this study addresses the importance of incorporating specifications to each assessment within the framework to mitigate gaps and vulnerabilities caused by different factors, such as differences in data protection regulations.

Accordingly, the new guideline will be used to assess the Stratumseind 2.0 project and to test the following hypothesis:

Hypothesis: The Stratumseind 2.0 project does not fully comply with the PbD guideline in the safeguard of data protection.

By exposing potential risks to data protection and analyzing how stakeholders handle PbD when cooperating in the project, this analysis will recommend best practices on data protection mechanisms for similar smart cities` projects.

3 RESEARCH DESIGN AND METHODOLOGY

This thesis examined how smart cities' technologies comply with the PbD model previously presented. In that sense, a single case study was conducted due to its usefulness in allowing an in-depth evaluation from a singular case and for offering generalized understandings for similar cases (Simons, 2015). A single-case study offers a "richly described and evidence-based [research], in the form of observations and perspectives of stakeholders and participants, significant incidents, narratives and critical analysis of any relevant documents" (Simons, 2015, p. 176). Since smart cities are a growing tendency that offers

several solutions specially to urban development – economic, social and political – it is keen to understand the importance of such context and investigate the different risks for data privacy within a smart city environment. Therefore, this study aimed to address the gaps on data privacy present in data protection frameworks, primarily with PbD, when considering the implementation of new technologies that process or collect personal data.

While there are several examples that can be assessed through the lenses of PbD, a smart city project as a case study addressed interesting layers for the analysis, such as stakeholders' responsibilities, the risks to data protection in public spaces and how to ensure privacy and security controls in such a complex collaboration. Therefore, to understand this phenomenon, it was essential to observe the use of different IT systems in IoT applications and the challenges they pose to ensuring individual's privacy.

3.1 JUSTIFICATION OF RESEARCH DESIGN

Despite the potential positive impacts of the project in terms of safety, livability and attractivity of smart cities, a new (use of) technology poses challenges to data privacy, both in terms of risks and adequacy to the technology (Siggelkow, 2007). Therefore, in order to understand this process, a case study is appropriate to demonstrate the importance of a relatively recent phenomenon with smart cities because it contributes as a reference to future applications of similar settings. In that sense, "cases allow the evaluator to learn intricate details of how a treatment is working, rather than averaging the effect across a number of cases" (Kennedy, 1979, p. 663). Moreover, it is worth to note that when focusing on a contemporary event, adopting as strategy an explanatory case study can be more relevant in obtaining insightful information concerning the topic (Yin, 1994).

Moreover, understanding the challenges and effects for individual's privacy in public spaces helped evaluate how the topic evolved in the legal framework, particularly with the

application of a practical guideline for the use of PbD. In that sense, it is essential to understand the consequences of the implementation of information systems to society, as well as how these systems might impact data privacy in future projects.

3.1.1 Logic of Case Selection

This thesis aimed to pursue a theory testing method of the PbD guideline to assess the Stratumseind 2.0 project. The living lab is based on the innovative use of information systems combined with other technologies to propose smart solutions to the city. As a unique case, it proposed new discussions about privacy by design and the development of different systems applied in the urban space, although the project was not fully consolidated yet. The initiative started as a proposition by the city of Eindhoven to reduce criminal behavior at the longest pub street in the Netherlands, with the involvement of multiple stakeholders.

This study addressed the challenges of privacy by design within and across organizations while developing such technologies, which entails organizations' and systems' accountability. The common factor that links this case to similar projects is the application of information technologies in public spaces, which presents new implications for data protection and raises unparalleled concerns. For instance, how technological features and governance models born from private and public collaborations ensure data privacy through anonymization. The logic of this case selection reflected its potential as a new type of collaboration in public spaces involving information technologies and the interesting dimensions developed by the project regarding PbD. Thus, with these considerations in mind, the case allowed the use of the practical guideline of PbD to understand how the project incorporated privacy-preventive measures since it started.

3.2 OPERATIONALIZATION

This study is operationalized through a content analysis founded on the new guideline of PbD developed in this research. As a smart city project the Stratumseind 2.0 project can pose unprecedent risks to data privacy. Hence, the new guideline on PbD will be contrasted with the data privacy risks posed by the smart city project in Eindhoven that were extensively discussed in the previous chapter. The new PbD guideline outlines the 5 principles used to define the categories and the most suitable indicators for the content analysis, as shown in Table 1. Therefore, the content analysis will ascertain whether the Stratumseind project fully comply with the new PbD guideline to safeguard data protection. Furthermore, it facilitates the identification of vulnerabilities that can be further assessed by the stakeholders involved, if that is the case, offering suggestions on best practices. This analysis covers paragraphs in the multiple documents about the case and interviews as well. The choice for paragraphs was based on the characteristics of the analysis, by focusing on specific terms and contexts. Accordingly, this analysis will provide results categorized into 'covered' and 'not covered'.

Table 1

Codebook used in this analysis to categorize the five principles of the new PbD guideline

Code	***Category***	***Definition***	***Examples***	***Indicators***
1	**Identification**	This principle concentrates on proactively identifying the context of information systems, by proposing a series of evaluations as early as possible and emphasizes the importance of incorporating privacy	"This type of data collected for this project is…" "These technologies are used for…" "The systems in place perform…"	Paragraphs related to applying identification processes concerning the context of each case, such as the type of data collected, the type of technologies, the systems in place, the

		considerations into every step of IT systems' and organizations' structures.		applicable regulations etc.
2	**Architectural Privacy - Embedding Security**	This principle addresses the importance of assessing the scope of the architectural platform – new and existing structures – to assist and enforce information security and privacy techniques.	"The type of architecture is…" "The reference-based architecture relies on the use of …" "These systems are embedding techniques for…" "The platform entails a data flow structured as…"	Paragraphs about how techniques are applied to each layer of the structure.
3	**Minimum Data, Maximum Privacy**	This principle reflects on the relation between data, functionality and privacy while also calling for an integration of security by design and privacy by design.	"The techniques for data minimization used are…" "The functionality of the systems supports maximum privacy in…"	Paragraphs that encompass the interrelation between embedded privacy requirements and systems' functionality.
4	**Transparency and Accountability**	This principle entails the responsibility to inform data subjects, demonstrate and be accountable for every system and activity involving the processing of personal data.	"The organizations inform data subjects about…" "Reports have been published…" "Data subjects are aware and consented the…"	Paragraphs that describe procedures and practices related to information and communication of the processing of personal data.

			"The privacy standards were applied…"	
5	**Data Subject Control**	This principle argues for a shift from the concept of data subject rights to the notion of data subject control.	"Individuals are aware of their rights to…" "Individuals have the agency to request…"	Paragraphs that entail the responsibility to support data subjects' rights by promoting control.

3.3 METHODS OF DATA COLLECTION

This project relied on the collection of data from primary and secondary sources, since to conduct an integrative case study it is crucial to obtain enough information. In that sense, the different actors involved in the case were contacted to provide official reports and documents about the practices regarding the data processing in place and the application of privacy by design. Additionally, this data was supported by semi-structured interviews of two experts involved in the project. Finally, the triangulation of data was used as a tool to strengthen the analysis. The data collection of this study was conducted between October 22, 2018 and December 30, 2018.

3.3.1 Official Documents and Reports

Documents and publications developed by the Stratumseind 2.0 project, along with secondary actors involved in the collaboration, were used to investigate the undertaking processing of data and the several technologies implemented. The focus on this particular case provides an opportunity to analyze reports regarding the (ongoing) results of the technologies' application and the management of information in place. Thus, the analysis relied on: two reports published by the municipality of Eindhoven (Gerwen, 2013; Merlijn van Dijk, 2018), two documents retrieved from DITSS (Kanters, 2013, 2017; Kanters & Gemeente Eindhoven, 2015), one report from a third party (Mol, Khan, Aalders, & Schouten, 2015) and, lastly, a few

articles and reports about the case study published by Atos (Atos Group, 2015; "Atos uses Big Data analytics for safer streets," 2015; "Intelligent City Management," 2015).

3.3.2 Semi-Structured Interviews

The aim of conducting semi-structured interviews with open-ended questions was to thoroughly investigate the role of professionals involved in the smart city project in Eindhoven, allowing the interviewees to elaborate on their own experiences. The semi-structured interview strategy entailed initiating the interviews with a brief introduction about the topic being discussed. Moreover, it focused on obtaining information about the responsibilities and roles of each actor involved, thereby supporting an in-depth analysis of the entire project. Therefore, interviews provided clarification on stakeholders' participation in the Stratumseind 2.0 project, regarding the mechanisms in practice that ensure data protection within the framework of privacy by design. After the interviews were conducted, the transcriptions were made to facilitate interpreting the analysis.

The first interview conducted took place in Eindhoven at the Living Lab office: the interviewee, Tinus Kanters, is the project leader of the Living Lab Stratumseind. The interview elucidated the application of the several systems and sensors, while addressing the level of cooperation in the project, and the implementation of privacy requirements into the project.
The second interview was conducted through Skype: the interviewee, Albert Seubers, is the director of Global Strategy IT in cities at Atos, which is the company responsible for developing the CityPulse Pilot. The interview explored the real-time analysis propelled by the CityPulse software and the challenges for data protection. Furthermore, the interviews were used to support the content analysis of the other documents (Yin, 1994).

3.3.2.1 Ethical Considerations.

An 'informed consent' (Allmark et al., 2009) was given to the participants during the interviews, and the statements provided the following remarks:

i. A brief description of the content of the thesis;
ii. Brief personal introduction as researcher and explanation about the participants' rights as subjects of the study;
iii. The use of process model of consent: confirmation of voluntary participation and that the interview may be terminated at any time (Allmark et al., 2009);
iv. A summary of the findings after the interview and potential termination of the participation (if requested);
v. An acknowledgement regarding potential implications with the thesis' results;
vi. Anonymity (if requested);
vii. An acknowledgement of the benefits of the study and their contribution (if requested);

Moreover, the files obtained with the interviews, including recordings and transcripts, were securely stored and anonymized, to ensure a safe disclosure of these data and minimize reliability issues.

3.4 VALIDITY ISSUES

To avoid construct validity issues, multiple sources were used to support the findings by comparing interviews with the different documents and articles particularly related to the implications to data privacy. When considering the issues of reliability, protocols for interviews were provided to the interviewers and a description of the procedures for the analysis of the other documents was specified as well (Yin, 1994). The research was restricted by the stakeholders' willingness to engage: although the main data source was concentrated on the documents retrieved from them, there was limited access to such information online. Some of

the experts provided relevant information regarding the use of IT systems in smart cities; however, the collection of data about different technologies' applications can still be considered limited. Besides, the triangulation of the data from multiple sources and methods was applied to minimize the internal validity issues – in terms of credibility – and the external validity issues – in terms of qualitative boundaries for generalizations (Løkke & Sørensen, 2014).

In addition, applying a case study research design might have presented constraints in terms of replicability. The features and variables in the case selected can be replicable in other scenarios, but the focus of this thesis – privacy implications – might largely differ depending on the legal and political context. Nonetheless, when considering the generalization concerns due to the case's specifications in terms of population and unit of analysis, there was sufficient theoretical background to support the analysis, providing adequate considerations applicable to similar structures (Yin, 1994). In other words, the theory of PbD addressed the concerns regarding "context and particularity" (Simons, 2015, p. 173) by incorporating these factors into the framework. The PbD guideline stated that in order to be effective, context and particularity were the main factors to be analyzed during the initial phase of any systems' designing. Therefore, by integrating these two essential variables into both the analysis and the methodology itself, the framework of PbD ensured replicability on the basis of its case-by-case and holistic approach (Simons, 2015).

4 ANALYSIS

The present chapter provides an assessment of the PbD principles in the Stratumseind 2.0 project. Moreover, the first part of the analysis outlines the stakeholders' cooperation in the project while evaluating the most relevant aspects regarding data protection, which means

considering the roles and responsibilities in the Stratumseind 2.0 project. Next, the project will be contrasted with PbD principles.

- Stakeholders

As previously mentioned in section 2.3, the Stratumseind 2.0 project – and the joint collaboration with the CityPulse pilot – consists of a cooperation network that incorporates the quadruple helix to improve safety and the quality of life in Eindhoven. The quadruple helix encompasses governments, businesses, knowledge institutes and citizens, and is the structure in which the project is being developed in Eindhoven, supporting an integrated ecosystem to exchange knowledge with the several on-going smart city projects around the country (T. Kanters, personal communication, October 22, 2018). It is important to notice that, although there are multiple actors involved in the project, only the most prominent ones are mentioned in this analysis, as shown in Table 2.

The expansion of the project can be perceived as part of the strategy adopted to engage several actors into the different field labs developed along the years. The cooperation developed in the regions of Eindhoven, Amsterdam and Rotterdam, which is called Brainport, focus on “health, mobility, energy, food and safety” (Kanters, 2017, p. 2) with the goal to use technology to become an example of smart society.

Ever since the project started, the discussion on data ownership appeared as one of the biggest challenges, proposing questions such as “who owns the data coming from sensors on the street?”. The project encompasses multiple stakeholders, therefore, the legal boundaries for privacy and for contractual agreements evolved as a complex challenge to the cooperation for those involved. Particularly regarding the cooperation with the police, the limitations for the living lab’s range of action became more evident and demanded well-defined strategies to ensure compliance with privacy requirements.

Table 2

Roles and Responsibilities of stakeholders in the Stratumseind 2.0 project and CityPulse Pilot

Stakeholders		Roles	Target
Public Sector	Municipality of Eindhoven	Street Manager and "Stratumseind 2.0" Economic Affairs and Safety Coordinator of projects Intermediary between public and private partners Area Development	Administrate the Living Lab Awareness campaigns Policies for Economy and Safety in particular Management, maintainance and cleaning of the public space Stimulate and facilitate promising initiatives Enforcement and supervision of rules Client assignment and direction Structural cooperation with partners Regular administrative coordination on progress and development direction
	Police of Eindhoven	Network Inspector	Public order and safety Supervision and enforcement in the public area Improve collaboration with catering and porters Response for safety Assistance
Companies	Atos DITSS Light House ViNotion Axis Communication Sorama Munisence Icen Others	Consortium leader Project Leader Suppliers	Real-time analysis of the data collected Use the data collected for research Test facility for sensors, dataharvesting, privacy and system-architecture Technological development Big data analytics Research
	Philips	Supplier	Lightining sensors Research
Educational Institutions	Tilburg University	Guidance	Ensuring adequate data collection and compliance to privacy regulations
	Eindhoven University of Technology	Research	De-escalate project: influence behavior
	Amsterdam University Fontys University Intelligent Lighting Institute Others	Project management	Research Smart Lighting
Citizens	Residents, Customers/Visitors	Participants	Participate in awareness campaigns Informing and signaling with regard to the quality of life and safety of the environment Reaction to different experiments

For Tinus Kanters, the project leader of Stratumseind 2.0, and for the municipality of Eindhoven, one of the main propositions of the project is to develop an open data library to divulge results obtained with all the projects, ensuring (governmental) transparency (T. Kanters, personal communication, October 22, 2018). However, when dealing with such a complex infrastructure considering all technological components (sensors) from different suppliers, an (on-going) discussion on what was missing in the privacy domain emerged with this new scenario. Promptly, it escalated to the national level, incorporating the Ministries of Justice and Internal Affairs and the National Privacy Authorities, stablishing round-table discussions concerning the data retrieved in public spaces in the country (T. Kanters, personal communication, October 22, 2018). Thus, the efforts were concentrated on clarifying the boundaries of each sector – public and private – with regard to data management in public spaces, particularly in safeguarding citizens' privacy in public spaces.

Accordingly, citizens were inserted as a center piece in these debates, as well as in technological developments of the project, meaning that individuals were (and should be) informed, when necessary, of every step of this process (T. Kanters, personal communication, October 22, 2018). Likewise, the living lab addressed the importance of integrating citizens to the experiments in order to sustain social trust. For instance, informing residents and visitors at the Stratumseind ensured social acceptance, creating a hospitable scenario for the lighting experiment, and the sound and even cameras application.

Thereafter, certain technological developments reached the public sphere, shifting the social mindset from acceptance to the invasion of privacy. This critical reaction emerged when the living lab decided to initiate the experiment with fragrances to influence people's behavior on the streets, similarly to the lighting poles in place. The new step into adopting sensors that would spread the fragrance (of orange) were deemed too intrusive and people started questioning such level of interference; however, the technique has been (mostly) used in

private spaces for quite some time, such as in shopping malls and stores (T. Kanters, personal communication, October 22, 2018). The interesting point explored in this critical reaction is people's 'unwillingness to be observed or interfered in public spaces', as described in the definition of 'behavioral privacy' proposed by Koops et al. (2017). Furthermore, the repercussion instigated a new debate regarding the ethical dimension behind the application of information technologies, stressing the importance of transparency and openness to citizens.

Moreover, the importance of understanding the interaction of stakeholders in the smart city project relies on the applications of privacy requirements, which means addressing the complexity of incorporating privacy into every i. layer of the systems, sensors, activities and any other component (technical components), but also ii. strictly defining which actors must comply with privacy requirements (organizational components). Since there is no guideline on developing smart city projects, most of the current projects are pioneers in the field.

Overall, the smart city project in Eindhoven was developed with the focus on safeguarding citizens' privacy, implementing several strategies related to PbD. Thus, the following part presents an assessment on the PbD principles contrasted in the Stratumseind 2.0 project:

1. Identification

The application of the first principle concerns the identification of the context in which the project took place and the technological instruments implemented to achieve the projects' objectives. Likewise, it means defining what will be applied, in what ways, and, especially in the smart city context, by whom. Accordingly, the first information provided by DITSS (Kanters, 2017) regarding the application of this principle entails defining the mission of the project as a test facility, and an enabler for other field labs, for new technological solutions and the several technical components required for such implementation, such as smart sensors, smart data and others. Additionally, they have expanded the project using a "tailor-made

context brokers/event processors" (Kanters, 2017) strategy, promoting a tailored application of technologies according to its context, which means the range of action and limitations of the stakeholders; for instance, in the case of the living lab, the responsibilities rely on the aspect of a research lab that collects, gather and analyze the data obtained from the sensors. Likewise, the project focus on the concept of quadruple helix, which indicates the involvement of companies, educational institutes, governments and citizens in the projects, and as well as the Brainport region, meaning the regions of Amsterdam, Eindhoven and Rotterdam working together in an ecosystem supporting an exchange of knowledge and experience of smart city projects (T. Kanters, personal communication, October 22, 2018).

Also, the project provides clarification on the functionalities of the several implemented technologies – lighting, video analysis, sound detection and others – to achieve the 'integrated Smart Crowd Management', safety and security (Mol et al., 2015, p. 10). Moreover, it depicts the incorporation of privacy strategies, including PbD, from the beginning, specifically with a privacy policy for Eindhoven, combining the efforts of several actors to ensure an adequate incorporation of privacy requirements. For instance, during the interview, Kanters (personal communication, October 22, 2018) addressed the importance of incorporating privacy by design since the beginning of the project in both organizational and technological elements to avoid privacy breaches. Additionally, he explained the importance of defining the adequate and appropriate mechanisms in order to demand compliance from companies in the project that may not necessarily be concerned with ensuring individuals' privacy.

Nonetheless, the analysis of the documents from the municipality of Eindhoven only provided superficial information regarding the technical components used in the project. The one technical description offered by the municipality entails the mission of the project in the field of crowd management, justifying the combination of data from multiple sources (Merlijn van Dijk, 2018). Still, the objectives are specified and the documents provide a detailed

description of data (sources), such as noise levels, temperature, visitor numbers, etc (Gerwen, 2013, p. 7). Furthermore, the report about the plan of approach for the project provides a detailed description of the actors' responsibilities and goals (also mentioned in Table 2) (Gerwen, 2013). Likewise, a description of how the cooperation was stablished is also present: the report explains the requirements that companies had to comply with in order to participate in the project, such as the type of technology to be used, how it would be explored and the specific functionalities as public utility (T. Kanters, personal communication, October 22, 2018).

Finally, in the case of the CityPulse pilot, the application of this principle concentrates on specifying the sources to be used in the big data analytics, in order to support the police with informed decisions about events that may happen at Stratumseind. This project is coordinated by Atos, which combines "both big data analytics and real-time analytics with privacy guarantees" ("Atos uses Big Data analytics for safer streets," 2015, p. 1), analyzing data from sound sensors, security cameras, microphones and social media, to create intelligent patterns that recognize triggers in specific situations and alert the police control room ("Intelligent City Management," 2015). Thus, the main goal of the project was to support authorities with a predictive system and avoid the escalation of certain events, using "'on the ground' information with data gathered online to create a powerful picture of the street" ("Atos uses Big Data analytics for safer streets," 2015, p. 2), providing a well-structures outlining of the first principle.

Two critical aspects related to the application of this principle are the challenge of data ownership, which escalated to the national level with round-table discussions, and the outlining of a system architecture, which incorporates elements such as technical user experience, open data and software to the discussion (T. Kanters, personal communication, October 22, 2018).

2. Architectural Privacy – Embedding Security

The second principle encompasses the application of techniques for security and privacy in each layer of the architectural platform. Firstly, it is important to identify the scope of the architecture used in the Stratumseind 2.0 project, which is inspired by the FIWARE (Open Source Platform for our Smart Digital Future) architecture, supporting a smart city platform "as a data/knowledge hub and non-intrusive, open to third parties" (Kanters, 2017, p. 31). The structural services performed by the living lab are "sensors, infrastructure, social network, actors, monitoring, big data storage, 3D models, organization and people, etc" (Kanters, 2017, p. 14). The data is gathered from a range of existing sources and the sensors are 'info aligned', enabling a cross-reference analysis of the data, when required, as well as the collaboration between partners when conducting the analysis. The sound analysis, provided by the Sorama sound imaging system, depicts sound level and sound spectrum, detecting the mood in the street, levels of stress and, finally, an aggression event; when an aggression is detected, an alert is sent to the authorities (T. Kanters, personal communication, October 22, 2018).

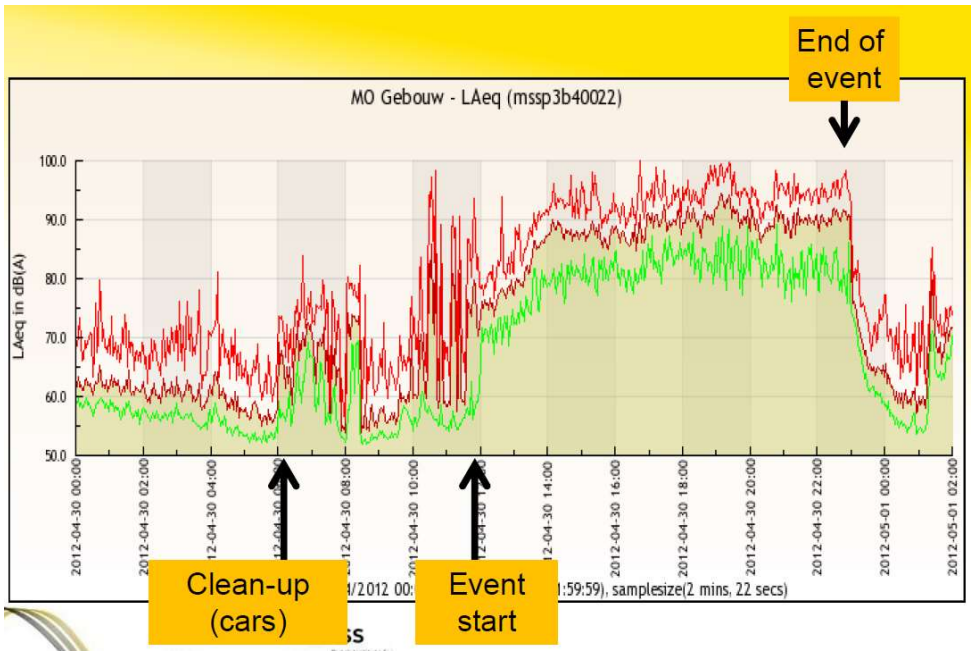

Figure 3. Sound analysis, detection of anomalies. Reprinted from Living Lab Stratumseind, DITSS, 2017, retrieved from https://www.bnsp.nl/wp-content/uploads/2017/10/Presentation-LL-sept..pdf of © Tinus Kanters.

The camera network, supported by ViNotion, provides an anonymized system that outputs the counting of visitors in the street and performs movement detection, revealing anomalies such as people running or moving in groups (Kanters & Gemeente Eindhoven, 2015).

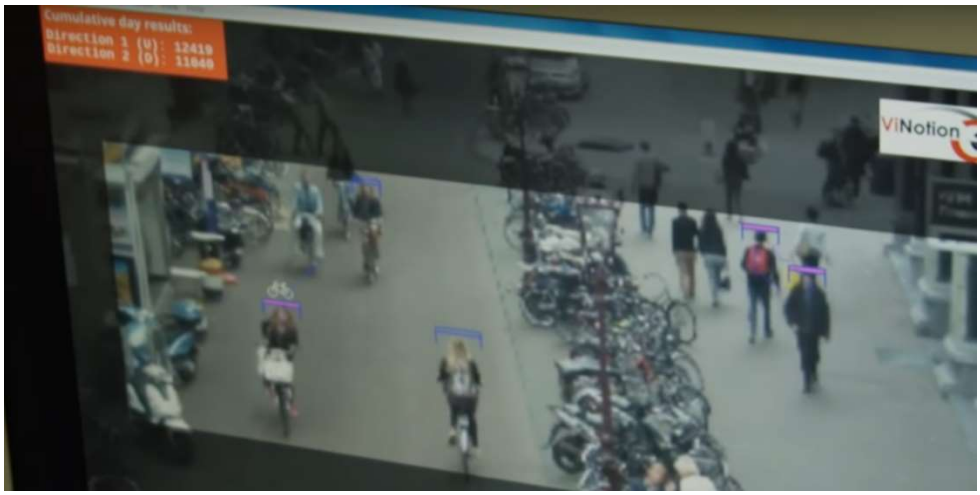

Figure 4. Video Camera (anonymized) system. Reprinted from Stratumseind 2.0: Big Brother's surveillance and interference in people's lives, YouTube, 2018, retrieved from https://www.youtube.com/watch?v=IiVEs6GIJT8.

For this principle, despite the differences in activities by the living lab (research lab) and the CityPulse pilot (real-time analysis), the description is centered in the structure of the platform that supports all projects. In that sense, all the sensors and data sources for the video, five audio sensors and social media analysis were connected wirelessly in conformity with open standards, besides the continuous effort put into developing a standardized API for this project based on the FIWARE architecture (Mol et al., 2015). A netcentric information gathering is used, with a dashboard presenting information from all sensors, including sensors that are not used anymore due to potential privacy breaches, such as MacAddress readers. Moreover, the challenge of system architecture encompasses the lack of regulations concerning the designing of digital cities.

Additionally, some of the systems incorporated by the project were also developed by citizens, such as the network of air quality measurements in Eindhoven (T. Kanters, personal

communication, October 22, 2018). In that sense, the involvement of different actors proposes a well-defined organizational arrangement based on the "different levels of access to personal information" (T. Kanters, personal communication, October 22, 2018) and frequent update requests enforcing security measures; notably, the participation of the national police represents a challenge regarding data management. Hence, they introduced the concept of context broker, which is incorporated in the systems architecture and works by collecting a pre-defined dataset and presenting the requested information on the interface (T. Kanters, personal communication, October 22, 2018). Thereafter, by defining the context and providing the specification into the context broker, it is possible to avoid privacy, security and safety risks while complying with current applicable regulations.

3. Minimum Data, Maximum Privacy

This principle argues for a maximum compliance with privacy by using the minimum amount of data and focuses on assuring privacy requirements into systems' functionalities. In this case, one of the first concepts exploited in the Stratumseind 2.0 is the modular data harvesting, enabling the "(re)use of data for multipurpose", with anonymized datasets. One example of this principle application is the use of cellphone data from Vodafone, which passes through a strict process that exclude any type of identifiable data (also the origin is not indicated when fewer than fifteen visitors come from the same region) before sending it to the living lab, supporting an analysis of the origin of visitors at Stratumseind (Merlijn van Dijk, 2018, p. 50). Moreover, every footage is anonymized, not stored or accessible to third parties, presented in the form of data statistics regarding the people counting (Merlijn van Dijk, 2018). This is also applicable for Bluetooth and MacAddress readers that are no longer used due to the lack of techniques capable of masking the identifiable information retrieved from them.

Primarily, the sensors for camera and sound used personal data in the beginning of the project: at first, a system blurred individuals' faces retrieved in the cameras' image and

presented in the dashboard, until the living lab realized it was possible to revert the action. Similarly, it was possible to determine where sound decibels came from and inform about both noise levels and origin locations from the sound sensors, which was also interpreted as part of an identification of individuals (Merlijn van Dijk, 2018). The sound systems now define normal patterns and the range of the signal, differentiating firecrackers from glass breaking and gun shots, while also identifying stress by analyzing characteristics of peoples' voices.

Thus, after some years, the project was able to prevent systems from collecting or processing personal data. Likewise, social media analysis was incorporated, using public posts that could provide a picture of the mood of the environment, with a system capable of listening to wordings ("Atos uses Big Data analytics for safer streets," 2015). Overall, the maximum data retrieved is the strictly necessary for operating a service, such as "to get out the signal to a service that something is happening over there, could be air pollution, or traffic control or anything" (A. Seubers, personal communication, November 30, 2018).

Regarding the organizational dimension when considering this principle, every stakeholder was responsible for embedding PbD into their systems, which facilitated the strategy of avoiding the storing of personal data (A. Seubers, personal communication, November 30, 2018). Another challenge is ensuring that the outcomes of techniques applied for each sensor goes to their right place, for instance, "police information only goes to the police" (T. Kanters, personal communication, October 22, 2018). Besides defining these limitations, the techniques applied in the data gathered facilitated business intelligence on crowd management, which was also relevant for predictions incorporated into the police systems. Hence, the goal is enabling systems to interact with each other and perform the desired services with the minimum data.

4. Transparency and Accountability

This principle states that the rights to transparency and accountability must be better addressed, seeking to approach actors' responsibility to be transparent and to demonstrate accountability for its systems' activities in the processing of personal data. In this case, the strategy is centered on privacy by design and focuses on open standards (Kanters, 2017). Indeed, the smart city platform used as reference model is described as "non-intrusive and open to third parties" (Kanters, 2017), allowing visibility of processes and openness to the assessment by third parties. Additionally, the smart city platform based on the FIWARE architecture proposed by the EU is also part of the project strategy to ensure accountability, since every layer is visible allowing an assessment of every aspect, when adopting a reference model in such complex environment of (potential) sensitive data sources – cities and citizens (T. Kanters, personal communication, October 22, 2018). Moreover, certifications and ISO norms were introduced as part of the municipality directive to any technology implemented in the city enforcing information security and adequate processes in place.

Besides the architecture, introducing "API agreement, open data principles and IoT principles" also promotes compliance to a set of high standards for business to participate in smart city projects in the Netherlands (T. Kanters, personal communication, October 22, 2018). However, part of the challenge is the political issue regarding pushing companies to accept several contractual terms regarding data ownership, for instance, that are not part of regulations yet (T. Kanters, personal communication, October 22, 2018). In that sense, while private companies insist to avoid the responsibility of informing people of certain actions, such as influencing behavior by smell, the living lab initiated the experiment with fragrances and immediately started to be questioned about their legitimacy to influence people's behavior with such approach (T. Kanters, personal communication, October 22, 2018).

The development of the project also focus on the replicability and expansion of solutions tailored to sectors and countries, "personalized anywhere, anyhow, while maintaining the strictest security and privacy levels" ("Atos uses Big Data analytics for safer streets," 2015). Before expanding the solution to other situations, the project is supported by parties like the University of Tilburg, which ensures the designing of systems and analyze the privacy regulations applicable to the project (A. Seubers, personal communication, November 30, 2018).

Particularly with the living lab, citizens' awareness is an essential element to strengthen the project's continuity. They concentrate on disseminating the project in multiple communication channels, offering signings and informing citizens about the range of the project around the country, aiming to increase acceptance and trust. For smart city projects, citizens' reaction and acceptance becomes an important factor to be taken into account. Hence, it is keen to mitigate risks exposed in the smart city context in order to increase citizens' support by being "technologically compliant with standards, market-supported and backed up by a flourishing community" (Mol et al., 2015, p. 10). As a smart city project, it is essential to constantly inform people on the streets about the analyzes and results, addressing the need to comply with the principles of transparency and accountability.

Moreover, one of the points elaborated is the process of decision-making with automated systems. The limits between human and automated decision-maker are explored particularly when it comes to addressing the risks of accountability, responsibility and interpretation; thereby, this point also addresses the importance of the GDPR's provision on "Automated individual decision-making, including profiling" (Article 22, "GDPR General Data Protection Regulation," 2016) in the safeguard of data subjects' rights when it comes to automated decision making. However, in order to mitigate the aforementioned risks, the decision still relies on providing information to humans to support their decision-making

process. Thus, to fully encompass accountability, it is extremely necessary to assess the process that entails the data translated into information.

5. **Data Subject Control**

This principle concerns the responsibility of actors to proactively inform and notify data subjects when processing personal data to conform with applicable legal requirements. In this case, since no more personal data is being collected, there are no requirements to comply with. However, an example of a proactive incorporation of data subjects' rights – data subject control – is the importance of informing people through communication and signs that data processing is taking place.

4.1 DISCUSSION

After the assessment of the PbD principles in the project, this section will present an overview clarification of the conclusions drawn during the analysis, determining whether the principles are fully implemented.

The first principle addresses the importance of being able to understand the context in which privacy requirements will be applied, which also entails defining organizational and technological coordination. In other words, it means that each organization and each type of technology demands specific measures, requirements in specific steps. In that sense, the challenge for smart cities starts with ensuring a network that tackles both dimensions and their components. Moreover, despite the increasing and extensive effort of the project to incorporate a holistic and tailored approach to the development of the project, the lack of legislations and clarifications on how to coordinate the cooperation hampers the application of the first principle.

The analysis elucidated a substantial attitude towards privacy, notably with PbD incorporated into the technical components, since the beginning of the project. On the other

hand, when analyzing the documents from the municipality of Eindhoven (Gerwen, 2013; Merlijn van Dijk, 2018), it seemed to lack technical specifications regarding the project, providing superficial information on how the objectives would be accomplished. However, the lack of specific technical information is compensated by an extensive explanation regarding the roles of the multiple stakeholders involved in the project (Table 2). This suggests a new kind of division between the main actors– organizational expertise and technical expertise – promoting a new approach to the integration of expertise and coordination in the smart city structure. Hence, this ecosystem is setting the pace for future socio-technical solutions in the context of smart societies.

Additionally, the discussion on data ownership introduced a new mindset regarding data management and data processing. It proposes reliability for future initiatives by enforcing privacy by design when developing such projects, addressing the responsibility towards data privacy and citizens integrity. In a certain level, it also explores the ethical dimension which suggests that there are (social) limits to technological interference, also explored with the definition of behavioral privacy. For instance, this limitation is explored in the cases of data analysis and the translations into decisions, whether it was an automated or human made decision, since it incorporates the potential risks and ethical consequences caused by an error – ethical and accountable dimensions of the process. Hence, outlining a framework for designing digital cities incorporating privacy must address all the dimensions and components that encompass the smart city project, ensuring not only (data) privacy but also safety to say the least; the implementation of the first principle can be defined as fully implemented.

The second principle addresses the architectural layers and decisions with regards to sensors/systems. Arguably, with the implementation of privacy by design strategies, the project relies on achieving their results without risking individual's data privacy. For instance, the video analysis by ViNotion can "even indicate people running, or faces all turning in the same

direction" (Mol et al., 2015, p. 52) , but the data "will be blurred at the source which excludes personal or facial recognition" ("Atos uses Big Data analytics for safer streets," 2015, p. 2) due to the anonymization technique embedded into the CPU's systems directly. Moreover, the sounds sensors and sound imaging analysis incorporating machine learning to rigorously detect and recognize specific sound levels and anomalies to prevent events from happening. The Stratumseind 2.0 project validates the opportunity to an efficient cooperation between different parties to increasing safety and livability, including between the police and companies, while safeguarding citizens' privacy in the process. Moreover, the project is considered an innovative initiative in the country, impelling the continuous effort on (re)defining and (re)discovering boundaries that are not yet delineated, propelled by the ever-growing challenges concerning systems architecture and the lack of legislations for designing smart cities – the digital layer in the public domain (Timan et al., 2017). More specifically, this discussion entails defining the boundaries on data collection and the use of sensors from any party on public spaces. Thus, most of the systems were designed or enhanced to sustain security and privacy, including anonymization techniques to avoid the identification of individuals, this principle was adequately implemented.

The third principle argues for the minimum use of data and maximum privacy sustaining a system's functionality. In this case, it addresses the constant effort in avoiding personal information to perform the services, even when it could indicate further developments or detailed results. Despite the use of personal information in the initial stage of the project, the preciseness of services and systems, according to each system and project, demonstrate a clear application of privacy features and a high level of maturity in terms of compliance as well. However, the lack of specification on certain datasets, data sources and data (re)use leaves room to uncertainty. Moreover, it specifically addresses the risk posed by the integration of several databases lacking appropriate security and privacy measures in place. Thus, the efforts

to avoid the collection and processing of personal data are visible, but the lack of clarity with regards to the smart city project's purpose leaves a gap: the multipurpose of data may broaden the opportunities for data analysis, however, it does not categorically specify the type of data (re)used.

The application of the forth principle in this project offers the opportunity to understand the use of data collected from sensors that averts personal data, such as the sounds sensors and the cameras network – a new perspective for a new context. The project demonstrates a set of strategies centered in PbD and translated into the 'smart city infrastructural information', the building block – architectural reference – used for future projects, and it ensures a high level of accountability by following open standards and supporting adequate levels of privacy and security measures. Accordingly, the inclusiveness of citizens in the development of systems for the city, or in the debates concerning the implementation of the project, addresses the importance of the organizational elements in the legitimacy of the project. For instance, the debate regarding the fragrance experiment can describe the contrast between making people aware of a certain situation and introducing the notion of invasion of privacy in public spaces, while also elaborating on the mindset behind the reaction to fragrances in each scenario: does the type of influence in behavior (reduce aggressive behavior vs increase consumption) expected with the fragrances relevant to peoples' perception of the situation or is the lack of awareness the reason behind the critical reaction to one situation but not the other?

Additionally, a deficiency in providing the disclosure of the results obtained in the project and the burden to retrieve online reports or documents about the project may represent an obstacle to ensuring transparency in general. The issue regarding disseminating both 'data'[2] and 'information'[3] relies on the risk brought by the combination of different open datasets, which enables the retrieving of potential personal information. Still, the biggest challenge for the city of Eindhoven is to define how to disclosure these datasets. Therefore, this principle is

not fully implemented, notably due to the lack of guidance and regulations that sustain the development of smart city projects, but the project brought this discussion to a national level and is setting the pace for future projects with regards to privacy, security, safety, etc.

The fifth principle begins by discussing the importance of guaranteeing data subjects' rights and ensuring social trust, which would strengthen data subjects' control – over personal information, when applicable –; also, the impact caused by openness and visibility increase citizens' acceptance to the project's legitimacy. This principle refers to the processing of personal information; therefore, it does not apply while the project ensures the exclusion of personal data and the potential identification of individuals. Still, it is crucial to understand that data subjects, especially for smart cities, have the control over their information according to the rights exposed in regulations, and stakeholders' responsibility is to ensure citizens' privacy. More importantly, this case outlined that both informational privacy and behavioral privacy encompass citizens' control in the public space, covering privacy beyond the dichotomy of private and public spheres.

Overall, the project demonstrates an ever-growing effort to embed privacy requirements and other measures into the project. It was proven a good example for the application of the PbD guideline and support the standardization of a single framework for smart cities indeed. Likewise, it addressed the importance of an appropriate integration between both organizational and technological components for privacy when delineating a smart city project. The privacy mindset and the guidance for organizational or technical expertise substantially impelled the project to the national level, facilitating future initiatives in the country while setting high standard to smart city projects. Moreover, it exposes the efficacy of incorporating a holistic and tailored approach when applying PbD: the cooperation between educational institutes and companies proved extremely reliable and effective.

5 CONCLUSION

This thesis analyzed to what extent PbD suits this set of cooperation and whether it is correctly applied in the technologies integrated in smart cities' projects. The analysis confirmed the hypothesis that the Stratumseind 2.0 project does not fully comply with the PbD guideline in the safeguarding of data protection, by providing a full description of what was applied and what missed. Thus, in response to the research question presented in this study: *To what extent are smart cities applying Privacy by Design in the safeguard of data protection?,* it was determined that the PbD principles were applied in most of the steps, but were not fully implemented. In this regard, according to the findings, the Stratumseind 2.0 project ensures the application of the principles of identification and architecture privacy – embedding security. The principles of minimum data, maximum privacy and transparency and accountability were determined as partially applied: they were substantially implemented in the project, still, the analysis exposed some missing aspects related to these principles. Lastly, the fifth principle of the new PbD guideline for data subjects' control was not applicable to the project since no personal data was being used; however, it suggested the inclusion of aspects that incorporate individuals' privacy in public spaces and, consequently, provided a new perspective concerning individual's control. Furthermore, the project requires some development on PbD strategies, but it can be considered a great example of adequate practices and substantial incorporation of privacy measures in smart city projects.

In this thesis, a new guideline on PbD was developed, which incorporates practical elements and provides clear guidance for the designing of technologies. The previous frameworks, inspired in the seven foundational principles of Cavoukian, often lacked technical guidance and practical solutions, leaving a gap in the translation on how engineers and other participants should apply the framework. Thus, the new model on PbD has proved to be

consistent to its objective of providing a 'holistic and tailored approach' as part of a clear and practical guideline that encompasses different contexts and applications worldwide.

Accordingly, the PbD principles' inherent characteristics of a comprehensive approach address external validity while ensuring generalization for different smart city projects. However, as a single case study, the analysis may have developed a limited testing of the framework; thus, further validation by future researches should grant the confirmation of its replicability in other projects. Still, a challenge for the new model in PbD entails assuring the designed and incorporated privacy measures will effectively persist over time. It means ensuring that privacy measures must be structured adequately and, consequently, be simply incorporated in the PbD steps. Likewise, to ensure the appropriate application of PbD, a checklist evaluation should be incorporated during the processes. Therefore, this checklist should validate whether all privacy measures are correctly performed and implemented.

Another challenge for the new PbD guideline is applying privacy measures into legacy systems, which means uncertain and (potentially) costly-ineffective results. PbD is aimed to the designing of new systems and offers several constraints when it comes to adapting it into legacy systems, particularly when considering the risks to systems' functionality. Nonetheless, further research could refine the framework developed in this study by investigating how the aforementioned checklist could be incorporated in the PbD strategy and how legacy systems could be better incorporated into the discussion. Accordingly, the findings propose numerous aspects for further research; for instance, they suggest an examination of the consequences introduced by new regulations concerning data ownership and the impacts of a unique model for a smart city platform for organizational and technical coordination in bigger arrangements.

Overall, this thesis confirmed the importance of a practical guideline on PbD to promote a stronger privacy mindset and application of specific privacy and security requirements, particularly in the context of smart cities, where protecting citizens' rights is indispensable.

The growing tendency of smart cities requires the development of higher standards to ensure privacy, security and data protection, considering the risks at stake for individuals. The potential greatness of smart city projects for urban development – social, political, economic, technological – is undisputable; thus, it is essential to set the pace for future projects in this (ongoing) revolution and define high levels of privacy and security requirements to safeguard individuals' privacy.

REFERENCES

Allmark, P., Boote, J., Chambers, E., Clarke, A., McDonnell, A., Thompson, A., & Tod, A. M. (2009). Ethical Issues in the Use of In-Depth Interviews: Literature Review and Discussion. *Research Ethics*, *5*(2), 48–54. https://doi.org/10.1177/174701610900500203

Anthopoulos, L. G. (2015). Understanding the smart city domain: A literature review. *Transforming City Governments for Successful Smart Cities*, 9–21. https://doi.org/10.1007/978-3-319-03167-5_2

Atos Group. (2015). *Big data analytics for safer streets*. Retrieved from https://www.youtube.com/watch?v=kkXxCkQMkDo&feature=youtu.be

Atos uses Big Data analytics for safer streets. (2015, July 21). Retrieved January 10, 2019, from https://atos.net/en/2015/press-release/general-press-releases_2015_07_21/pr-2015_07_21_01

Batty, M. (2013). Big data, smart cities and city planning. *Dialogues in Human Geography*, *3*(3), 274–279. https://doi.org/10.1177/2043820613513390

Bier, C., Birnstill, P., Krempel, E., Vagts, H., & Beyerer, J. (2012). How Is Positive-Sum Privacy Feasible? In N. Aschenbruck, P. Martini, M. Meier, & J. Tölle (Eds.), *Future Security* (Vol. 318, pp. 265–268). Berlin, Heidelberg: Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-33161-9_39

Caragliu, A., del Bo, C., & Nijkamp, P. (2011). Smart cities in Europe. *Journal of Urban Technology*, *18*(2), 65–82. https://doi.org/10.1080/10630732.2011.601117

Cavoukian, A. (2010). Privacy by design: the definitive workshop. A foreword by Ann Cavoukian, Ph.D. *Identity in the Information Society*, *3*(2), 247–251. https://doi.org/10.1007/s12394-010-0062-y

Cavoukian, A. (2011). Privacy by Design: The 7 Foundational Principles. *Information and Privacy Commissioner of Ontario*, 2.

Cavoukian, A., & Jonas, J. (2012). *Privacy by design in the age of big data*. Information and Privacy Comissioner of Ontario, Canada. Retrieved from www.privacybydesign.ca

Cavoukian, A., Taylor, S., & Abrams, M. E. (2010). Privacy by Design: essential for organizational accountability and strong business practices. *Identity in the Information Society*, *3*(2), 405–413. https://doi.org/10.1007/s12394-010-0053-z

Charter of Fundamental Rights of the European Union. (2012, October 26). European Union. Retrieved from http://www.europarl.europa.eu/charter/pdf/text_en.pdf

Cilliers, L., & Flowerday, S. (2015). The Relationship Between Privacy, Information Security and the Trustworthiness of a Crowdsourcing System in a Smart City. In *Proceedings of the Ninth International Symposium on Human Aspects of Infromation Security & Assurance* (p. 13). University of Fort Hare, South Africa.

Council of Europe. (1981). *Convention for the Protection of Individuals with regard to Automatic Processing of Personal Data* (European Treaty Series No. ETS 108) (p. 9). Strasbourg. Retrieved from https://rm.coe.int/1680078b37

De Fine Licht, J., Naurin, D., Esaiasson, P., & Gilljam, M. (2014). When Does Transparency Generate Legitimacy? Experimenting on a Context-Bound Relationship. *Governance*, *27(1)*, 111–134.

Deloitte. (2015). *Smart Cities – A Deloitte Point of View, Version 1.0* (p. 86). Retrieved from www.deloitte.nl/govlab

Deloitte. (2017). *Smart City | Smart Nation* (pp. 1–44). Retrieved from smartcity.deloitte.com

Domingo-Ferrer, J., Hansen, M., Hoepman, J.-H., Le Métayer, D., Tirtea, R., Schiffner, S., … European Network and Information Security Agency. (2014). *Privacy and data protection by design - from policy to engineering.* Heraklion: ENISA. Retrieved from http://bookshop.europa.eu/uri?target=EUB:NOTICE:TP0514111:EN:HTML

Eremia, M., Toma, L., & Sanduleac, M. (2017). The Smart City Concept in the 21st Century. *Procedia Engineering*, *181*, 12–19. https://doi.org/10.1016/j.proeng.2017.02.357

Friedman, D. (2000). Privacy and Technology. *Social Philosophy and Policy*, *17*(02), 186. https://doi.org/10.1017/S0265052500002168

Gerwen, E. van. (2013, September). "Stratumseind 2.0" Plan van Aanpak 2013-2017. Retrieved from https://eindhoven.raadsinformatie.nl/document/1047462/1/document

Gurses, S., Troncoso, C., & Diaz, C. (2011). Engineering Privacy by Design. *Computers, Privacy & Data Protection*, 25.

Hoepman, J.-H. (2012). Privacy Design Strategies. *ArXiv:1210.6621 [Cs]*. Retrieved from http://arxiv.org/abs/1210.6621

Intelligent City Management for Eindhoven CityPulse – Case Study. (2015). Atos. Retrieved from https://atos.net/wp-content/uploads/2016/06/atos-ph-eindhoven-city-pulse-case-study.pdf

Kanters, T. (2013, August). *Living Lab, onderdeel van Stratumseind 2.0*. Eindhoven. Retrieved from https://www.google.com/url?sa=t&rct=j&q=&esrc=s&source=web&cd=6&ved=2ahUKEwiIhLr-luvfAhWFbFAKHW0kAWUQFjAFegQIBhAC&url=https%3A%2F%2Fhetccv.nl%2Ffileadmin%2FBestanden%2FOnderwerpen%2FGrote_steden%2Fnetwerkdagen_eindhoven_living_lab.pdf&usg=AOvVaw38Wzz_lA-5lISYSP0bhv8r

Kanters, T. (2017, August). *LivingLab Stratumseind*. Eindhoven. Retrieved from https://www.bnsp.nl/wp-content/uploads/2017/10/Presentation-LL-sept..pdf

Kanters, T., & Gemeente Eindhoven. (2015). *Living Lab Stratumseind*. Eindhoven. Retrieved from https://www.google.com/url?sa=t&rct=j&q=&esrc=s&source=web&cd=2&cad=rja&uact=8&ved=2ahUKEwiIhLr-luvfAhWFbFAKHW0kAWUQFjABegQICBAC&url=https%3A%2F%2Fveiligheidsalliantie.nl%2Faction%2F%3Faction%3Ddownload%26id%3D301&usg=AOvVaw1mXGvoLQVvrx04GvEFpsDl

Kennedy, M. M. (1979). Generalizing From Single Case Studies. *Evaluation Quarterly*, *3*(4), 661–678. https://doi.org/10.1177/0193841X7900300409

Klitou, D. (2012). *Privacy-Invading Technologies: Safeguarding Privacy, Liberty & Security in the 21st Century* (Doctoral thesis). Faculteit der Rechtsgeleerdheid, Leiden University. Retrieved from http://hdl.handle.net/1887/20288

Kokott, J., & Sobotta, C. (2013). The distinction between privacy and data protection in the jurisprudence of the CJEU and the ECtHR. *International Data Privacy Law*, *3*(4), 222–228. https://doi.org/10.1093/idpl/ipt017

Koops, B.-J., Newell, B. C., Timan, T., Chokrevski, T., & Gali, M. (2017). A Typology of Privacy, *38*(2), 93.

Kroener, I., & Wright, D. (2014). A Strategy for Operationalizing Privacy by Design. *The Information Society*, *30*(5), 355–365. https://doi.org/10.1080/01972243.2014.944730

Kumar, S., & Prakash, A. (2013). Role of Big Data and Analytics in Smart Cities. *International Journal of Science and Research (IJSR)*, *5*(2), 12.

Lacinák, M., & Ristvej, J. (2017). Smart City, Safety and Security. *Procedia Engineering*, *192*, 522–527. https://doi.org/10.1016/j.proeng.2017.06.090

Lim, C., Kim, K.-J., & Maglio, P. P. (2018). Smart cities with big data: Reference models, challenges, and considerations. *Cities*, *82*, 86–99. https://doi.org/10.1016/j.cities.2018.04.011

Løkke, A.-K., & Sørensen, P. D. (2014). Theory Testing Using Case Studies, *12*(1), 9.

Mehta, B. B., & Rao, U. P. (2016). Privacy Preserving Unstructured Big Data Analytics: Issues and Challenges. *Procedia Computer Science*, *78*, 120–124. https://doi.org/10.1016/j.procs.2016.02.020

Merlijn van Dijk. (2018). *Stratumseind: Eindhoven's data street - Innovation Origins*. Retrieved from https://innovationorigins.com/stratumseind-eindhovens-data-street/

Mol, C., Khan, O., Aalders, R., & Schouten, N. (2015). A Spotlight on Smart City Eindhoven - How can Eindhoven become a Smart City faster? Venturespring B.V.

Moore, A. (2008). Defining Privacy. *Journal of Social Philosophy*, *39*(3), 411–428. https://doi.org/10.1111/j.1467-9833.2008.00433.x

Nissenbaum, H. (1998). Protecting Privacy in an Information Age: The Problem of Privacy in Public. *Law and Philosophy*, *17*(5/6), 559. https://doi.org/10.2307/3505189

Osman, A. M. S. (2019). A novel big data analytics framework for smart cities. *Future Generation Computer Systems*, *91*, 620–633. https://doi.org/10.1016/j.future.2018.06.046

Perera, C., McCormick, C., Bandara, A. K., Price, B. A., & Nuseibeh, B. (2016). Privacy-by-Design Framework for Assessing Internet of Things Applications and Platforms. In *Proceedings of the 6th International Conference on the Internet of Things - IoT'16* (pp. 83–92). Stuttgart, Germany: ACM Press. https://doi.org/10.1145/2991561.2991566

Psychogiopoulou, E. (2017). The European Court of Human Rights, privacy and data protection in the digital era. In M. Brkan & E. Psychogiopoulou, *Courts, Privacy and Data Protection in the Digital Environment* (pp. 32–62). Edward Elgar Publishing. https://doi.org/10.4337/9781784718718.00010

Regulation (EU) 2016/679 of the European Parliament and of the Council 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 the Directive 95/46/EC (General Data Protection Regulation). (2016, April 5). Official Journal L119. Retrieved from https://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX:02016R0679-20160504&from=EN

Ryz, L., & Grest, L. (2016). A new era in data protection. *Computer Fraud & Security*, *2016*(3), 18–20. https://doi.org/10.1016/S1361-3723(16)30028-8

Siggelkow, N. (2007). PERSUASION WITH CASE STUDIES. *Academy of Management Journal*, 5.

Simons, H. (2015). Interpret in context: Generalizing from the single case in evaluation. *Evaluation*, *21*(2), 173–188. https://doi.org/10.1177/1356389015577512

Smith, J. (2016, November 11). Data Protection [Text]. Retrieved February 28, 2019, from https://edps.europa.eu/data-protection/data-protection_en

Spiekermann, S. (2012). The challenges of privacy by design. *Communications of the ACM*, *55*(7), 38. https://doi.org/10.1145/2209249.2209263

Ståhlbröst, A., Padyab, A., Sällström, A., & Hollosi, D. (n.d.). DESIGN OF SMART CITY SYSTEMS FROM A PRIVACY PERSPECTIVE, 16.

Stratumseind 2.0: Big Brother kijkt niet alleen mee maar beïnvloedt je ook - YouTube. (2018, April 8). Retrieved January 13, 2019, from https://www.youtube.com/watch?v=IiVEs6GIJT8

Swinhoe, D. (2018). Why GDPR means Smart Cities need to move on from an 'Open Data' approach | IDG Connect. Retrieved October 21, 2018, from https://www.idgconnect.com/abstract/29380/why-gdpr-means-smart-cities-open-data-approach

Timan, T., Newell, B., & Koops, B.-J. (2017). *Privacy in Public Space*. Edward Elgar Publishing. https://doi.org/10.4337/9781786435408

van Zoonen, L. (2016). Privacy concerns in smart cities. *Government Information Quarterly*, *33*(3), 472–480. https://doi.org/10.1016/j.giq.2016.06.004

von Grafenstein, M. (2018). *The Principle of Purpose Limitation in Data Protection Laws: The Risk-based Approach, Principles, and Private Standards as Elements for Regulating Innovation*. Nomos Verlagsgesellschaft mbH & Co. KG. https://doi.org/10.5771/9783845290843

Waelbers, K. (2011). *Doing good with technologies: taking responsibility for the social role of emerging technologies*. Dordrecht ; New York: Springer.

Warren, S. D., & Brandeis, L. D. (1890). The Right to Privacy. *Harvard Law Review*, *4*(5), 193–220.

Westin, A. F. (1967). *Privacy and Freedom.* United States: New York: Atheneum Press.

What does data protection 'by design' and 'by default' mean? (n.d.). [Text]. Retrieved March 2, 2019, from https://ec.europa.eu/info/law/law-topic/data-protection/reform/rules-business-and-organisations/obligations/what-does-data-protection-design-and-default-mean_en

Wiese Schartum, D. (2016). Making privacy by design operative. *International Journal of Law and Information Technology*, *24*(2), 151–175. https://doi.org/10.1093/ijlit/eaw002

Yin, R. K. (1994). *Case Study Research - Design and Methods* (2nd ed., Vol. 5). SAGE Publications.

Zins, C. (2007). Conceptual approaches for defining data, information, and knowledge. *Journal of the American Society for Information Science and Technology*, *58*(4), 479–493. https://doi.org/10.1002/asi.20508

Footnotes

[1] "'Personal data' means any information relating to an identified or identifiable natural person ('data subject')" ("GDPR General Data Protection Regulation," 2016, p. 33) and corresponds with the concept provided by the Council of Europe (1981) for the protection of individuals with regard to automatic processing of personal data.

[2] "Data are facts that are the result of observation or measurement." (Landry et al., as cited in Zins, 2007, p. 486)

[3] "Information is meaningful data or data arranged or interpreted in a way to provide meaning." (Zins, 2007, p. 486)

Appendix

Appendix I: Interview Details

ORGANIZATION - PROJECT	TITLE	DATE OF INTERVIEW
DITSS – Stratumseind 2.0 (Living Lab)	Project Leader	October 22, 2018
Atos – CityPulse Pilot	Director of Global Strategy IT in cities	November 30, 2018

Appendix II: Operationalization

Table 1

Codebook used in this analysis to categorize the five principles of the new PbD guideline (for interview and content)

Code	***Category***	***Definition***	***Examples***	***Indicators***
1	**Identification**	This principle concentrates on proactively identifying the context of information systems, by proposing a series of evaluations as early as possible and emphasizes the importance of incorporating privacy considerations into every step of IT systems' and organizations' structures.	"This type of data collected for this project is…" "These technologies are used for…" "The systems in place perform…"	Paragraphs related to applying identification processes concerning the context of each case, such as the type of data collected, the type of technologies, the systems in place, the applicable regulations etc.
2	**Architectural Privacy - Embedding Security**	This principle addresses the importance of assessing the scope of the architectural platform – new and existing structures – to assist and enforce information security and privacy techniques.	"The type of architecture is…" "The reference-based architecture relies on the use of …" "These systems are embedding techniques for…" "The platform entails a data flow structured as…"	Paragraphs about how techniques are applied to each layer of the structure.

3	**Minimum Data, Maximum Privacy**	This principle reflects on the relation between data, functionality and privacy while also calling for an integration of security by design and privacy by design.	"The techniques for data minimization used are…" "The functionality of the systems supports maximum privacy in…"	Paragraphs that encompass the interrelation between embedded privacy requirements and systems' functionality.
4	**Transparency and Accountability**	This principle entails the responsibility to inform data subjects, demonstrate and be accountable for every system and activity involving the processing of personal data.	"The organizations inform data subjects about…" "Reports have been published…" "Data subjects are aware and consented the…" "The privacy standards were applied…"	Paragraphs that describe procedures and practices related to information and communication of the processing of personal data.
5	**Data Subject Control**	This principle argues for a shift from the concept of data subject rights to the notion of data subject control.	"Individuals are aware of their rights to…" "Individuals have the agency to request…"	Paragraphs that entail the responsibility to support data subjects' rights by promoting control.

Appendix III: Structuring the Data

Labeling used for interviews and content

Labels	Coloring
Identification	These are marked yellow
Architectural privacy - embedding security	These are marked green
Minimum data, maximum privacy	These are marked red
Transparency and accountability	These are marked blue
Data subject control	These are marked grey
Stakeholders' roles and responsibilities	~~These are marked crossed~~

Appendix IV

Table 2

Roles and Responsibilities of stakeholders in the Stratumseind 2.0 project and CityPulse Pilot

Stakeholders		Roles	Target
Public Sector	Municipality of Eindhoven	Street Manager and "Stratumseind 2.0" Economic Affairs and Safety Coordinator of projects Intermediary between public and private partners Area Development	Administrate the Living Lab Awareness campaigns Policies for Economy and Safety in particular Management, maintainance and cleaning of the public space Stimulate and facilitate promising initiatives Enforcement and supervision of rules Client assignment and direction Structural cooperation with partners Regular administrative coordination on progress and development direction
	Police of Eindhoven	Network Inspector	Public order and safety Supervision and enforcement in the public area Improve collaboration with catering and porters Response for safety Assistance
Companies	Atos DITSS Light House ViNotion Axis Communication Sorama Munisence Icen Others	Consortium leader Project Leader Suppliers	Real-time analysis of the data collected Use the data collected for research Test facility for sensors, dataharvesting, privacy and system-architecture Technological development Big data analytics Research
	Philips	Supplier	Lightining sensors Research
Educational Institutions	Tilburg University	Guidance	Ensuring adequate data collection and compliance to privacy regulations
	Eindhoven University of Technology	Research	De-escalate project: influence behavior
	Amsterdam University Fontys University Intelligent Lighting Institute Others	Project management	Research Smart Lighting
Citizens	Residents, Customers/Visitors	Participants	Participate in awareness campaigns Informing and signaling with regard to the quality of life and safety of the environment Reaction to different experiments

Appendix V

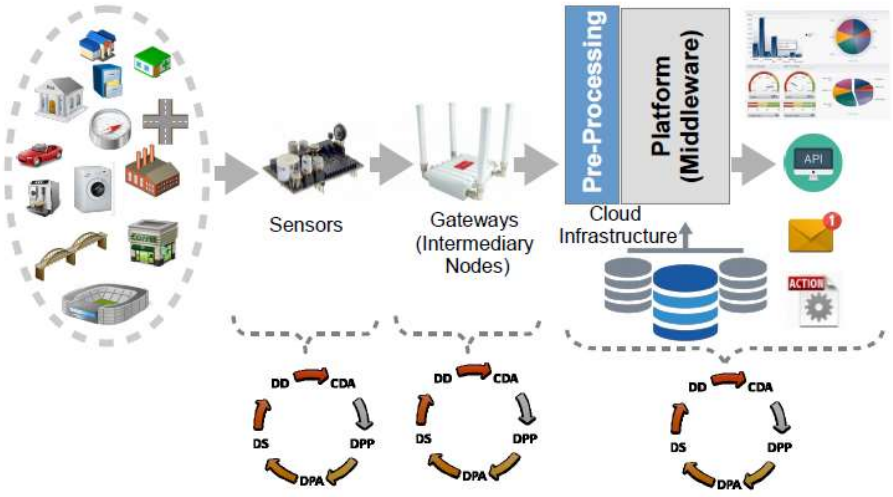

Figure 1. Typical Data Flow in IoT Applications. Reprinted from Privacy-by-Design Framework for Assessing Internt of Things Applications and Platforms in Germany, by Perera et al., November 2016, retrieved from http://dl.acm.org/citation.cfm?doid=2991561.2991566 Copyright 2016 of ACM Press.

Appendix VI

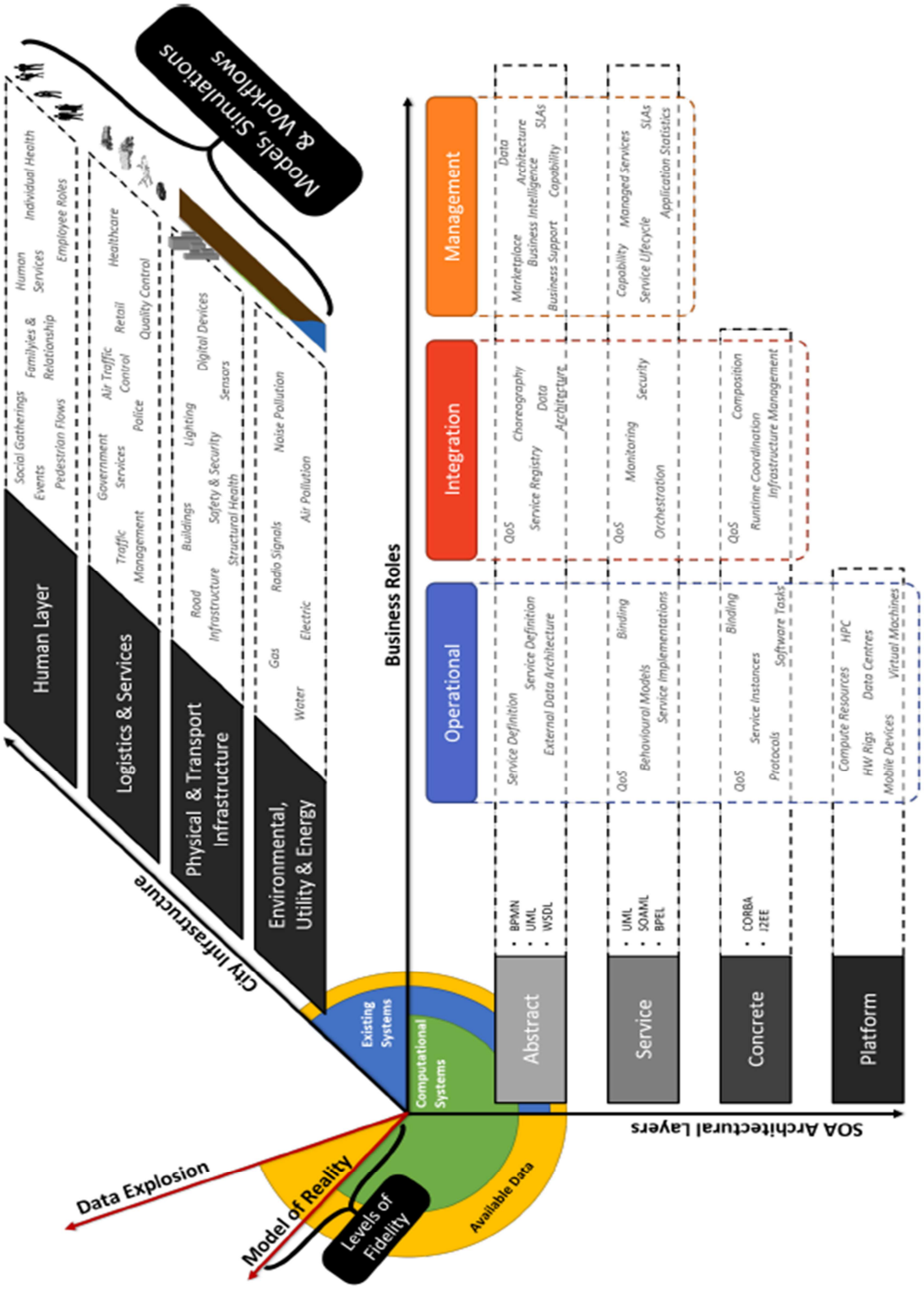

Figure 2. Service-Oriented Reference Architecture for Smart Cities. Reprinted from Service-Oriented Reference Architecture for Smart Cities in Leeds, UK, 2017, retrieved from http://eprints.whiterose.ac.uk/113342/ Copyright 2016 IEEE.

Appendix VII

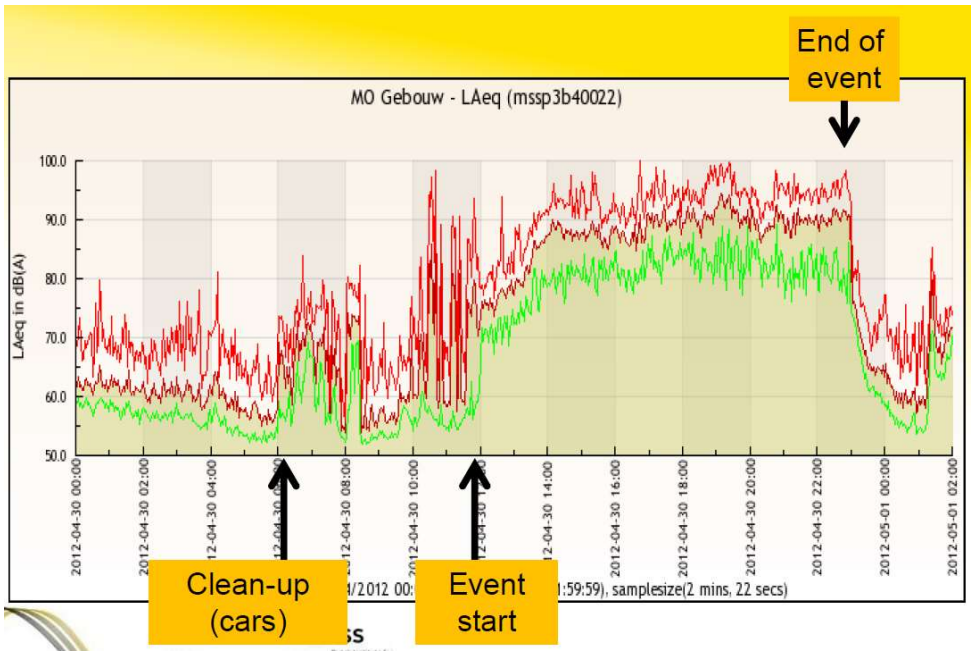

Figure 3. Sound analysis, detection of anomalies. Reprinted from Living Lab Stratumseind, DITSS, 2017, retrieved from https://www.bnsp.nl/wp-content/uploads/2017/10/Presentation-LL-sept..pdf of © Tinus Kanters.

Appendix VIII

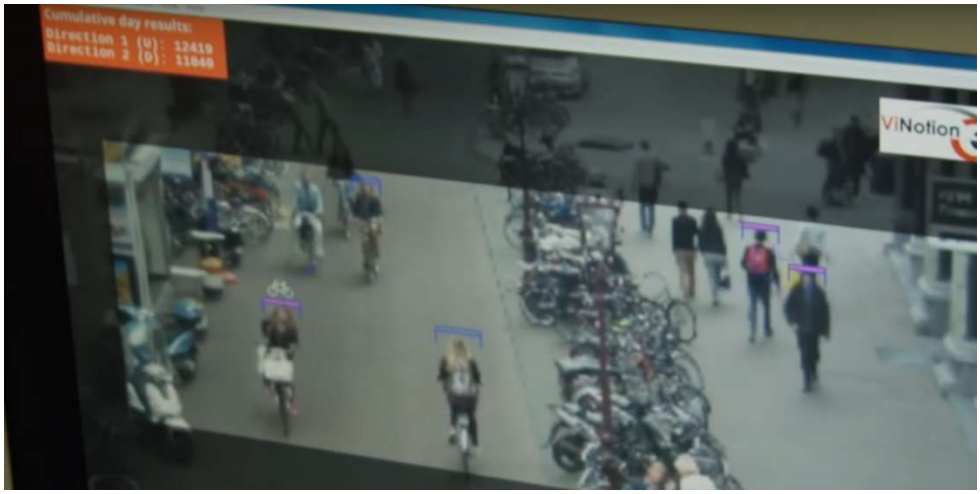

Figure 4. Video Camera (anonymized) system. Reprinted from Stratumseind 2.0: Big Brother's surveillance and interference in people's lives, YouTube, 2018, retrieved from https://www.youtube.com/watch?v=IiVEs6GIJT8.