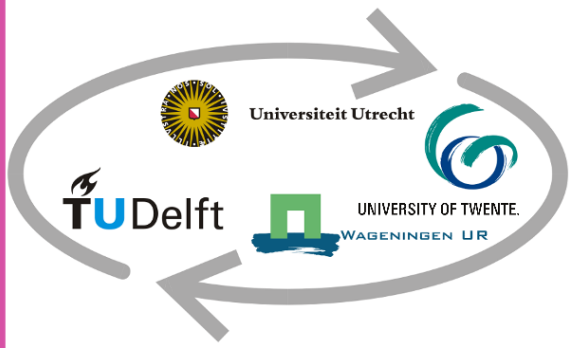
GIMA
Geographical Information Management and Applications

GIMA MSc thesis

Data use versus privacy protection in public safety in smart cities

The use of big, open and geographic data for achieving public safety objectives in the Dutch smart city context and the influence of European privacy protection here upon

Date: 2/27/2015
Student: Iris van de Kerk
Supervisors: dr. ir. Bastiaan van Loenen & prof. dr. Jaap Zevenbergen



February 2015
GIMA MSc Thesis

Contact information:
Iris van de Kerk
Iris.vandekerk@gmail.com

Summary

This research explores the influence of the European Data Protection Directive and ePrivacy Directive on the use of big, open and geo-data in smart city programs focussing on public safety in the Netherlands.

Many data can be considered personal data, due to the broad definition of personal data in the Data Protection Directive, recital 26 on de-anonymising data and technological innovations. Personal data is data that relates to an individual. Personal data can only be processed lawfully if the rules listed in article 6 of the Data Protection Directive are complied with. Among these rules is purpose limitation: the purpose for data processing needs to be specified, explicit and legitimate and data cannot be further processed for purposes that are considered incompatible.

However, since the smart city is a vague concept, difficulties are expected to arise in processing personal data according to the rules of the Data Protection Directive, most notably in purpose limitation. Two major theoretical problems are identified. First, the use of big data (by nature) and the use of open data (by definition) conflict with privacy protection regulation. In the field of purpose limitation, it is tempting to further process big data in ways that are incompatible, and open data is by definition not collected for specified, explicit and legitimate purposes. Open data, and to a lesser extent also big data, can thus not be personal data. Second, problems are expected to occur in personal data use for smart city purposes in general, since the smart city is a vague concept.

It is interesting to research to what extent these theoretical issues become evident in practice. Therefore, four smart city case studies were selected, with a focus on public safety: Eindhoven (Stratumseind 2.0), Almere (StraatKubus), The Hague and Zwolle. From theory, the relevance of big, open and geo-data for smart cities became evident and together with privacy protection, these are the main subjects addressed in the in-depth interviews.

The first theoretical problem could not be found in practice, since in the four case studies there was no personal big or open data being used. However, in all case studies personal data is being processed, which is either geo-data or 'other' data. The ePrivacy Directive lists rules for processing mobile phone location data, which is geo-data that is often used in smart city programs. The rules of article 6 including purpose limitation should thus be taken care of, but in only one case personal data is indeed being processed according to the rules: in Almere. The second theoretical problem did become evident in practice in the four case studies: the more concrete programs in Eindhoven and Almere are further progressed in taking care of privacy, while in the broader programs of The Hague and Zwolle it turned out to be difficult to define objectives to justify all personal data use and to do this well enough to enable lawfully processing personal data.

The solution Almere uses is the Gedragsrichtlijn StraatKubus, in which all demands from article 6 of the Data Protection Directive are satisfied. A formal solution is the only way to process personal data lawfully: documentation to justify personal data use, addressing purpose limitation and all demands of article 6 of the Data Protection Directive. However, the solution of Almere is not perfect, since there is limited space for flexibility and innovation, and the amount of people with access to the data is limited. Zwolle addresses these elements more flexible, but does not seem to satisfy all rules for processing personal data. A mix-up of these cases would be close to an ideal solution, but it needs to be researched if this is possible in practice.

There are different issues identified in taking care of privacy protection, but in all case studies a lack of clarity on, among others, what is and is not allowed and how processing personal data should be dealt with is an issue that became evident.

Two solutions were suggested as an example to overcome the identified problems. First, the process of exploring the possibilities for a more balanced solution by establishing a panel of stakeholders could be established. Second, the outcomes should be translated and communicated into clear and easily understandable documentation, in order to address the lack of clarity.

Acknowledgements

This research could not be established without the help of a number of people. It would not have been possible to research data use and privacy protection in smart city programs without the help of the respondents of the four case studies. I would like to thank all respondents for their time and for their important contribution to this research. In many cases, the interviews started with questions, turned into conversations, and ended in interesting discussions on privacy protection and innovations. For me, this highlights the importance and relevance of the issues addressed in this research. I hope that this thesis can be a useful contribution.

I would also like to thank my supervisor Jaap Zevenbergen for his feedback and for providing guidance in the chaos of all concepts included in this research. Furthermore, his last-minute feedback was very useful for finalizing my thesis.

Above all, I would like to thank my thesis supervisor Bastiaan van Loenen for his fast and clear feedback, for giving me new insights and for keeping me motivated and on track. Thank you for introducing me to a completely new field of research which has certainly gained my interest.

Table of Contents

1. Introduction

1.1 Introduction

In today's world, cities are facing many socioeconomic challenges. An important trend with large consequences is urbanisation (Bélissent et al., 2010 ; Bettencourt, 2013 ; Gil-Garcia et al., 2013 ; Naphade et al., 2011). Worldwide, the number of people will increase and it is expected that by 2050, about 70.0% of people will live in cities (Bélissent et al., 2010 ; Naphade et al., 2011). This puts lots of pressure on, among others, urban infrastructure, safety, healthcare, education and resources (Bélissent et al., 2010 ; Bettencourt, 2013 ; Gil-Garcia et al., 2013 ; Naphade et al., 2011). Furthermore, cities play a large role in achieving more environmental sustainability. Reducing greenhouse gas emissions, waste management and sustaining water, energy and food supply are important environmental challenges for cities (Naphade et al., 2011). Another important change in urban life is the rise of Information and Communication Technology (ICT). As Naphade et al. (2011, p.32) state: "*ICT advances have revolutionized all aspects of life*". ICT solutions contribute to a better understanding of the modern urban world and can be used to improve and attempt to resolve the many challenges cities are facing (Bettencourt, 2013). It helps to create more efficiency, reduce costs and increase the quality of urban life (Naphade et al., 2011). Innovations in planning, management and operations can be achieved by using ICT (Naphade et al., 2011).

ICT is thus a valuable tool to improve aspects of urban life. So-called smart city programs use ICT to make cities 'smarter'. The term smart city is a holistic concept and can include many components (see figure 1).

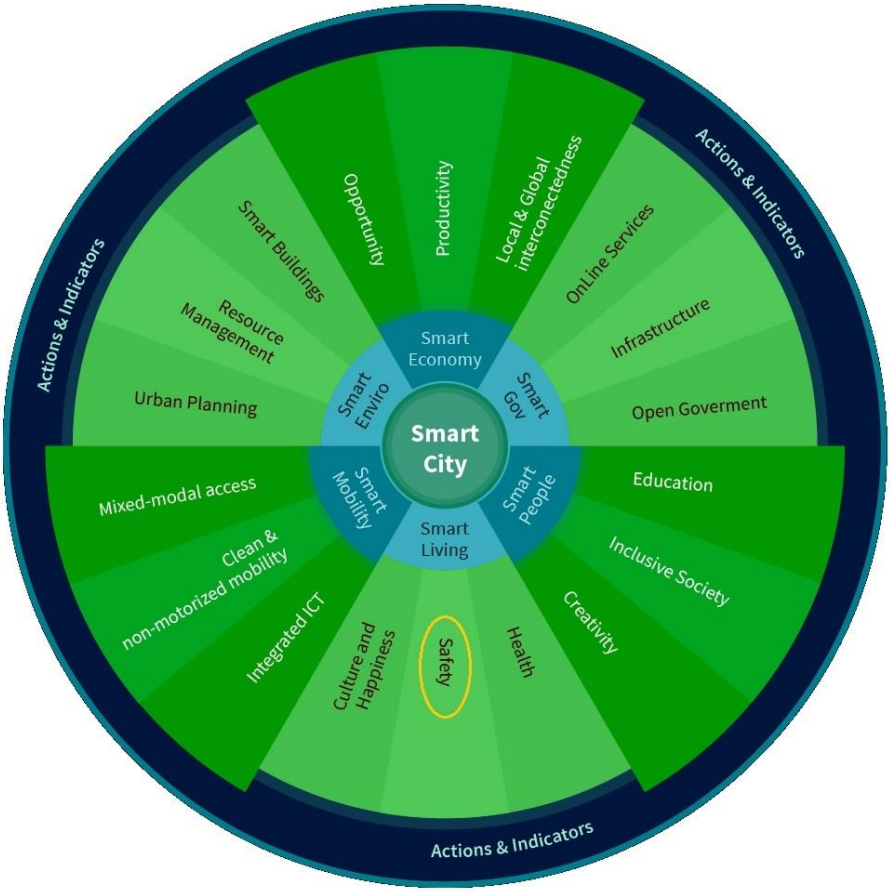


Figure 1: The smart city wheel by Boyd Cohen.
Source: Hoevenaars, 2013.

One of the components of the smart city is 'safety'. Public safety is the focus of this research, since, among other reasons, the smart city is a broad concept that needs to be further specified and safety is a subject that is important for every citizen (for more reasons, see paragraph 1.3.3, p. 10).

To ensure safety, smart city programs are established, in which ICT is a major component. Data is needed as input for ICT innovations (Gil-Garcia et al., 2013). Recently, in both data gathering and use innovations took place. Governments are working on making data freely available, so-called open data. In the Dutch public sector, the sector in which smart city programs are established, data is increasingly made publicly available and used. Big data is another data trend. It concerns enormous amounts of data, often collected by sensors (Analysys Mason Limited, 2010 ; Batty, 2013 ; Batty et al., 2012 ; Schaffers et al., 2011).

There is a downside to these developments. Data use for smart city purposes can conflict with privacy protection. Privacy protection is taken care of in directives (among others) in the European Union, which are implemented by the member states (Batty, 2013 ; European Commission, 2014c ; Gutwirth et al., 2013 ; Kulk & Van Loenen, 2012a ; Kulk & Van Loenen, 2012b). The most important privacy and data protection directives are Directive 95/46/EC on personal data (Data Protection Directive) and Directive 2002/58/EC and 2009/136/EC on data in the electronic communications sector (ePrivacy Directive) (EU Directive 95/46/EC, 1995 ; EU Directive 2002/58/EC, 2002 ; EU Directive 2009/136/EC, 2009). Data easily becomes personal data, due to the broad definition of personal data in these directives. Furthermore, a geographic component in data (geo-data) enables to link different datasets to a location, and thereby data leads easier to individuals (Van Oortmarssen & De Vries, 2014). Examples of personal geo-data are addresses, cadastral identification numbers and in some cases location coordinates. Personal data should be processed according to a number of rules (EU Directive 95/46/EC, 1995). Personal data, among others, needs to be collected for well-specified purposes and it needs to be kept for a limited amount of time. The definition of open data and the nature of big data conflict with these rules (Kulk & Van Loenen, 2012a ; Tene & Polonetsky, 2013). The smart city is a broad concept, and if its purposes are not further specified, personal data cannot be processed lawfully. It is thus expected that privacy issues occur in data use for smart city purposes:

There are two main privacy problems that can be expected in achieving smart city objectives. First, data can easily become personal data and in this case, open data (by definition) and big data (by nature) cannot be processed according to the rules listed in the Data Protection Directive. Second, the purposes for which personal data is being processed should be clearly defined and the smart city concept itself is too vague to enable processing personal data.

These developments lead to a quest for more clarity in the field of data use in smart cities in relation to privacy protection issues (see figure 2).

Figure 2: Visualisation of the components of this research.

1.2 Scientific and societal relevance

Smart city is a term that increasingly is being used, for example to market innovative products or for city marketing (Bélissent et al., 2010 ; Harrison & Donnelly, 2010 ; Hollands, 2008 ; Peek & Toxler, 2014 ; Philips, 2014 ; Siemens, 2014). Furthermore, the smart city concept can contribute to a better understanding of the modern urban world and can be used to improve and attempt to resolve the many challenges cities are facing (Bettencourt, 2013).

Big data and open data are also popular concepts in today's society (Analysys Mason Limited, 2010 ; Batty et al., 2012 ; Batty, 2013 ; Schaffers et al., 2011). So far, there is little literature that focuses on both data and smart cities, with the exception of big data in smart cities (Batty, 2013 ; Bettencourt, 2013). The importance of data is evident, but most articles on smart cities focus on the concept itself, on grading cities on their 'smartness' or on the role of ICT (for examples see Caragliu et al., 2011 ; Giffinger et al., 2007 ; Gil-Garcia et al., 2013).

This research does not only combine the concepts smart city and big, open and geo-data, but it also focuses on privacy protection. Privacy and data protection is a matter of concern for a lot of Europeans. Dutch people are in general very protective about their identity, but disclosing personal information is also being seen as part of modern life. Dutch people are, for example, not so concerned about gathering data in public spaces (77.0% is not concerned, compared to 62.0% of all Europeans) (Special Eurobarometer 359, 2011).

The three major subjects of smart city, data and privacy combine both public organisations and private companies as well as scientists, lawyers and citizens. The subject of this research thus provides an interesting topic for different groups. This research will contribute to knowledge on the role of privacy protection in data use for smart city purposes from a theoretical perspective, but also from a practical perspective. In society, this research will contribute to more knowledge on dealing with privacy issues in smart city programs, from both a theoretical as well as a practical perspective. In this way, this research makes it no longer necessary for people working on smart city programs to start from scratch on the issue of privacy. Especially for smart city programs starting in the near future, this research provides interesting information. Data use for smart city purposes and the influence of privacy protection is thus researched in both theory and practice, to contribute to creating more knowledge that is valuable for both science and society.

1.3 Defining assumptions, goals and scope

In this paragraph, the research content will be further defined and visualised. Important choices made will be explained and thereby the scope of this research becomes clear.

1.3.1 Assumptions

Based on literature, one major assumption is made: problems in the field of privacy will occur when data is being used for smart city purposes, due to (European) privacy protection. Furthermore, it is assumed that data is needed in smart city programs. This assumption is based on literature, but also logically, data is needed in ICT and therefore also in smart city programs. In the model, both of these assumptions are represented by arrows and lines (see figure 3, p.10).

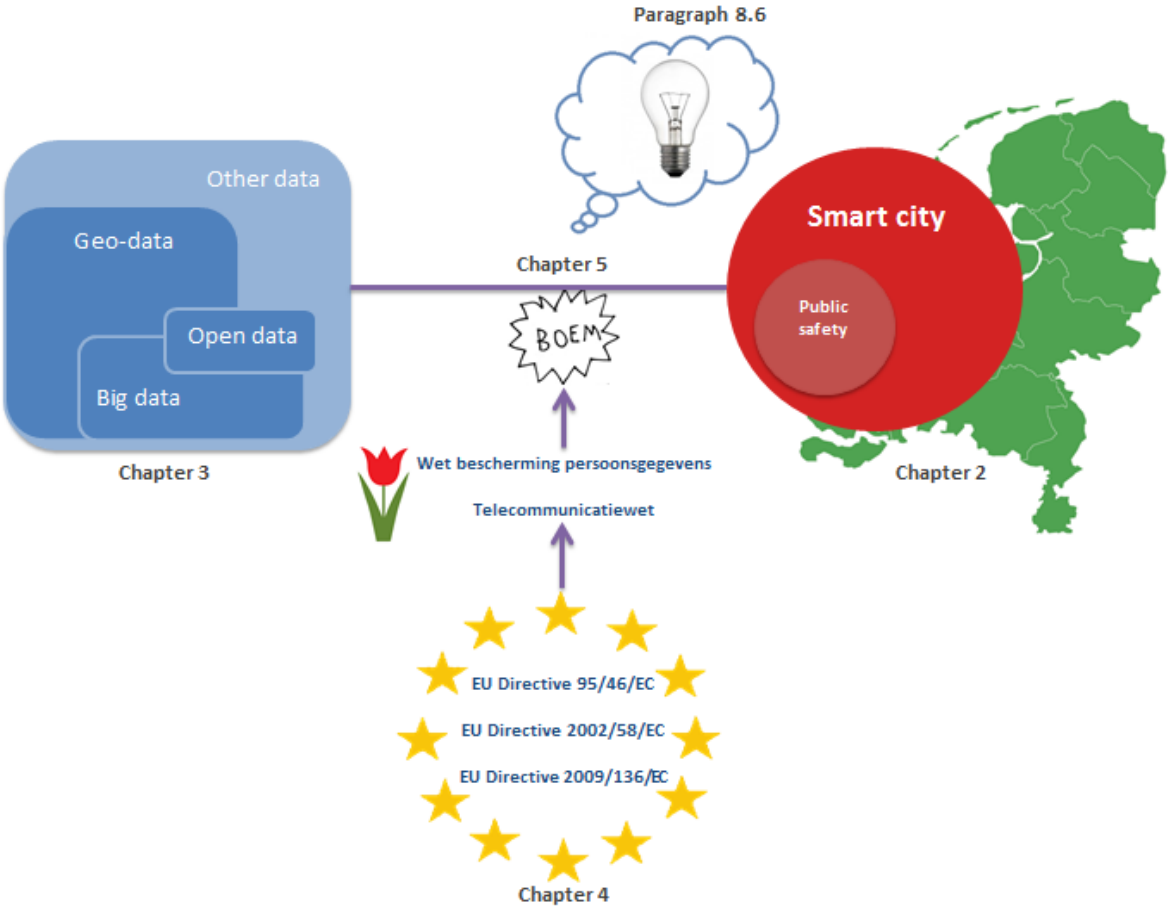


Figure 3: Overview of the research components and chapters/paragraphs in which they are discussed.

1.3.2 Research goal

The goal of this research is *to study the influence of European privacy protection directives on the use of open, big and geo-data for Dutch public safety programs in a smart city context, and to suggest improvements for this situation* (see figure 3). This research goal is translated into a research question and sub questions (see paragraph 1.4, p.11).

Two main steps need to be taken to achieve this goal. First, literature research is needed to gain basic knowledge on the three main subjects, and to create an understanding of the issues in privacy protection arising from theory. Second, the situation in practice will be explored by making use of interviews. It will be interesting to find out to what extent the privacy issues arising from theory occur in practice. These two main steps are divided in sub goals or steps that need to be taken (see table 1, p.13).

1.3.3 Scope

The scope of this research will become clear by explaining the choices that were made. The main choice that was made is to focus on a limited part of the smart city concept: public safety. The smart city is a broad concept, covering many aspects of urban life. The diversity of the concept would lead to the fact that in the end, little can be said on the influence of privacy protection on smart city programs. Furthermore, for programs focussing on improving air quality, for example, it is expected that little issues in privacy protection occur, since people are not the direct subject of the program. In public safety, on the other hand, people are the subject of data processing and it is more likely that personal data is used. From a more practical point of view, safety is in many cases among the smart city objectives in the Netherlands. Furthermore, safety is a subject that many people find an

important one and it should be well taken care of.

It should also be kept in mind that this research focuses on public safety programs in a smart city context. Public safety is also monitored on other levels, for example on the national level, but this is not in the scope of this research.

Also, within the subject of safety, the choice was made to focus on a specific part of this concept. In this research, safety refers to monitoring and forecasting safety. By making this choice, other public safety components such as securing and controlling mass events, and optimizing the capacity and response time of emergency services are not in the scope of this research. Because of the diversity of public safety programs in a smart city context, safety is further specified.

It is important to note that this research will focus on the Dutch situation, in the European privacy protection context. European directives are taken into account and attention is paid to their implementation in the Netherlands. This translation from the European to the Dutch national level is studied in order to identify differences. In general, privacy protection is taken care of at European Union level, but there might be differences in implementation between member states. For other countries outcomes can thus be different.

Privacy is being seen as a basic human right. In article 12 of the Universal Declaration of Human Rights the protection of privacy is addressed (OHCHR, 2015). Also on the European Union level, there are multiple agencies in which privacy is being protected. Among others, there is article 8 in European Convention on Human Rights (ECHR), Convention 108, a draft General Data Protection Regulation and the current directives on data and privacy protection: Data Protection Directive 95/46/EC, ePrivacy Directive 2002/58/EC and ePrivacy Amendment Directive 2009/136/EC (Convention 108, 1981 ; ECHR, 2010 ; EU Directive 95/46/EC, 1995 ; Directive 2002/58/EC, 2002 ; EU Directive 2009/136/EC, 2009). The choice was made to focus on the Data Protection Directive and the ePrivacy (Amendment) Directive, since these directives are the heart of current European data protection. It was decided not to include the draft General Data Protection Regulation, because there are too many uncertainties at this moment in time to incorporate it usefully in this research. In this research the focus will thus be on the current directives.

It should also be noted that it is not in the scope of this research to come up with a proposal for changes in law, but the outcomes can be useful for future research on this topic. It was chosen not to finalize this research by only presenting the results of the analysis, but to take it one step further and come up with suggestions and thoughts on improvements to overcome the problems in the field of privacy protection that are likely to be identified for data use in smart cities. These suggestions can address legislative issues, but proposing an in-depth change in law is not in the scope of this research.

Last but not least, this research focuses on the influence of (EU) privacy protection as explanatory factor for issues in processing data for smart city purposes. Other factors that might influence this link are not taken into account, since literature indicates the relevance of researching the influence of privacy protection on data use for smart city purposes.

1.4 Research questions

Since the content of this research was explained in the previous paragraphs, the main research question can now be formulated:

To what extent does the European Data Protection Directive and the ePrivacy (Amendment) Directive influence data use, and in particular the use of big, open and geo-data, for monitoring and forecasting public safety in smart city programs in the Netherlands and how can difficulties be overcome?

In the conclusion of this research, an answer to the main research question is formulated (see chapter 9, p.64). To ease this process and to provide a structure for addressing all elements of the main research question, ten sub questions are formulated:

1. *How can smart city and public safety be defined?*
2. *How can big data, open data and geo-data be defined?*
3. *How is privacy protection taken care of in the European Union and in the Netherlands?*
4. *What are the European Union requirements for processing personal data?*
5. *What difficulties in processing personal big, open and geo-data for monitoring and forecasting public safety in a smart city context can be identified from the field of privacy in theory?*
6. *To what extent is the use of big, open and geo-data for monitoring and forecasting public safety in a smart city context adhering to the rules set in the European Data Protection Directive and the ePrivacy Directive in practice?*
7. *What difficulties in processing personal big, open and geo-data for monitoring and forecasting public safety in a smart city context can be identified in the field of privacy in practice?*
8. *What is the influence of the European Data Protection Directive and the ePrivacy Directive on achieving smart city objectives in practice?*
9. *How can monitoring and forecasting public safety in smart city programs adhere to the requirements of the Data Protection Directive and the ePrivacy Directive?*
10. *How can difficulties in processing personal big, open and geo-data for monitoring and forecasting public safety in a smart city context be addressed?*

The chapters in which these questions are answered are listed in table 1 (see paragraph 1.5.1, p.13). In the conclusion of this research, the questions will be addressed (see chapter 9, p.64).

1.5 Methodology

In this paragraph, attention will be paid to the methods that are being used in order to enable answering the research questions. This paragraph addresses the methodology in general. First, the major research steps will be discussed. Second, the methods will be introduced. The methodology focussing on the case studies and operationalization is discussed after the theoretical chapters are finished (see chapter 6, p.34).

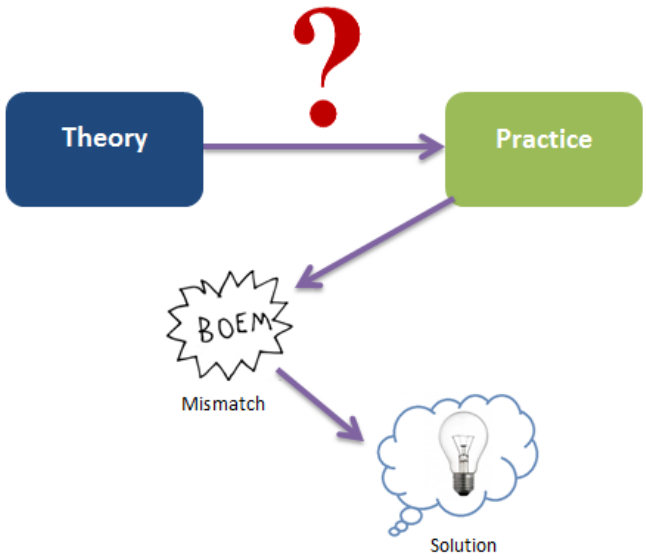


Figure 4: From theory to solution.

1.5.1 Research steps

The steps that need to be taken can be divided in a theoretical and a practical component (see paragraph 1.3.2, p.10). It is important to note that, for this research, scientific literature is not only used to serve as background information for the execution of the research itself, but is an important part of the methodology. The influence of data protection on data use in smart city programs focussing on monitoring and forecasting public safety, can only be studied in practice when a deeper level of knowledge on data and smart cities, and especially of privacy regulation, is reached. Both theory and practice should thus be seen as important parts of the steps and methodology. Thereafter, theory and practice can be bridged (see figure 4, p.12 and table 1).

Table 1: The research steps, related to sub questions and the location where they are discussed.

Category	Step nmbr	Step description	Question nmbr	Chapter / paragraph
Theory	1	Gaining in-depth knowledge on smart cities and public safety	1	2
	2	Gaining in-depth knowledge on big, open and geo-data	2	3
	3	Creating an understanding of European privacy protection, of the content of the privacy directives and of the Dutch implementation	3	4
	4	Creating an understanding of the requirements for processing personal data	4	4.3.1
	5	Identifying the theoretical difficulties in the field of privacy in processing personal (big, open and geo-) data for monitoring and forecasting public safety in smart cities	5	5
	6	Coming to a conclusion in theory	-	5.3.2
From theory to practice	7	Establishing the methodology for the practical part of this research: operationalization, cases (selection) and interview techniques	-	6
Practice	8	Gaining knowledge on data use and privacy protection in the four individual cases	-	7
	9	Researching the use of personal data for monitoring and forecasting public safety in smart cities	-	8.1
	10	Researching the level of specification of objectives for monitoring and forecasting public safety in smart cities	-	8.2
	11	Identifying to what extent big, open and geo-data for monitoring and forecasting public safety in smart cities is being process according to the rules set in the privacy directives	6	8.3
	12	Identifying the practical difficulties in the field of privacy in processing personal (big, open and geo-) data for monitoring and forecasting public safety in smart cities	7	8.3
	13	Researching the influence of privacy on achieving smart city objectives in practice	8	8.4
	14	Researching how personal data for monitoring and forecasting public safety in smart city programs can be processed according to the rules set in the privacy directives	9	8.4
	15	Coming to a conclusion in practice	-	8.5
Bridging theory and practice	16	Coming up with suggestions on addressing the difficulties in processing personal big, open and geo-data for monitoring and forecasting public safety in smart city programs	10	8.6
	17	Coming to an overall conclusion and answering the main research question	MRQ	9

Seventeen research steps need to be taken in order to answer the main research question (MRQ) (see table 1, p.13). Besides the steps, the table also lists which sub question will be answered in which step, as well as the number of the chapter or paragraph. Furthermore, a time schedule is established to keep the research on track (see appendix A, p.75).

1.5.2 Research approach

For the theoretical and practical part of this research, different methods are being used. Logically, the theoretical part will be based on literature and European directives as well as their Dutch implementation. For the practical part, four cases are being used. In-depth interviews are held with project leaders of smart city programs.

Theory

Research steps 1 to 6 will be completed by making use of scientific literature (see table 1, p.13). Logically, literature is needed to gain more in-depth knowledge on the main topics in this research. Knowledge on European data protection is needed for steps 3 to 6 and is not only gained from European directives and Dutch laws. Explanations, opinions and other related articles concerning these documents are also studied. This approach is also taken in many articles that focus, for example, on providing explanations and opinions on the directives (for example: Article 29 Working Party, 2007 ; Article 29 Working Party, 2011 ; Article 29 Working Party, 2013 ; Cuijpers & Koops, 2013 ; Cuijpers & Marcelis, 2012 ; Kulk & Van Loenen, 2012a ; Kulk & Van Loenen, 2012b ; Van Loenen et al., 2008 ; Wong, 2011).

Practice

For steps 8 to 15, the situation in practice, two other methods are used (see table 1, p.13). These methods are discussed into detail in step 7 (see chapter 6, p.34). First, (policy) documentation of these programs will be used. Analysys Mason Limited (2010) researched public safety user needs by analysing documentation from different organisations. The usability of this method depends on the availability and findability of policy documents. However, policy documents can be used to prepare the interviews and to gain basic information on the programs, and especially on its goals. Furthermore, it turned out that in the interviews held, documents are often shown and discussed to clarify statements or to ensure that enumerations (for example lists of data) are complete. The interviews are combined with the documentation analysis and altogether this will provide insights in data use and the influence of privacy protection.

Second, case studies of smart city programs in Dutch cities, focussing on public safety are selected. In-depth interviews are held with people involved in these programs (see chapter 6, p.34). This research method was chosen because of multiple reasons.

A qualitative research method was chosen over a quantitative one. A qualitative method, in this case an interview, enables to gain information on the way the interviewee approaches privacy issues. It provides inside in the thoughts and way of reasoning of the interviewee, as well as insight in experiences and perceptions. This is needed in the practical component of this research because it is important to gain an understanding of the way privacy issues are approached and being seen. Furthermore, this is an explorative research. Not much is known yet and it is thus not clear what aspects of dealing with privacy are likely to be influencing the situation in practice. In-depth interviews provide a method to find this out. All interviews are between 50 and 80 minutes long, so a lot of information was gained and this amount of time also indicates the relevance of the subject and the interest of the interviewees in it. Also, the selected cases each have their own character and are comparable on a few key characteristics (see paragraph 6.3, p. 34). Their individual characters cannot be taken into account well by using a quantitative methodology; flexibility in approach is needed. There is a limited amount of smart city programs that focus on public safety in the Netherlands. These programs differ from each other in many ways and only have a few core characteristics in

common. For these reasons it is a logical decision to adopt a qualitative approach instead of a quantitative one (Boeije et al., 2009).

The number of interviews needed is not set beforehand, but depends on the availability of interview candidates and on the point of saturation. Four interviews were held, one for each case. Since little was known on this subject in practice beforehand, it was decided to interview a leading case first to gain an idea of data use in smart city programs in practice. The Eindhoven Stratumseind 2.0 program serves as a well-known example of monitoring public safety in a smart city context. In the interview, it was tried to address all major elements of this research. The interview was semi-structured by using these elements, but only a simple version of the topic list was used. After this interview, the topic list was extended and this version was used for the other interviews (see appendix B, p.77). These other three interviews are also semi-structured, but a bit more structured compared to the first interview in Eindhoven. The semi-structured approach was taken since it is for each case important to address all elements, but at the same time it is important not to structure it too much because of the individual and distinguishing character of the cases (for more detailed information of the cases and the interviews, see chapter 6, p.34).

Bridging theory and practice

Steps 16 and 17 are completed by combining the results of the case studies and the theoretical knowledge (see table 1, p.13). For step 16, suggestions from the interviewees will be taken into account. This step is meant to be mainly inspiring and to give the interviewees a voice on suggestions to overcome difficulties. In step 17, in the conclusion of this research, the results of steps 1 to 16 are combined in the answer on the research questions.

1.6 Content

In the next chapters, chapters 2 to 5, the scientific context of the main research topics will be discussed. Chapter 2 focusses on the smart city and public safety, chapter 3 focuses on big, open and geo-data and in chapter 4 privacy protection (directives) in the European Union and the Netherlands is/are discussed. The theoretical knowledge gained from these chapters is combined in chapter 5. In chapter 6, the extended methodology for the practical part of this research is discussed. The results of the interviews and the policy documentation study are for each individual case described in chapter 7. In chapter 8, these results are combined and the answers on the practical sub questions are provided. Also, an advice for improvements is provided. Thereafter, in chapter 9, the main research question is answered in the conclusion of this research. Reflection and recommendations are given in chapter 10. The references used in this research are listed in chapter 11. In appendices A to F, the time schedule, topic list and the reports of the interviews and policy documentation can be found.

2. Theoretical context: smart cities

2.1 Smart city: a buzzword?

The term smart city is difficult to define well. There are two main reasons that contribute to the 'vagueness' of the term smart city.

First, the term smart city is often used for marketing purposes, since it is associated with technological innovations in the urban environment, making the exact meaning of the term unclear (Hollands, 2008 ; Peek & Toxler, 2014). The term is used for multiple kinds of marketing, of which city marketing is the most logical one: *"As no city wants to be a 'dumb' city, the Smart City concept is quickly adapted for benchmarking cities"* (Peek & Toxler, 2014, p.156). Cities protect and improve their status as attractive places for business, citizens, events and visitors and thereby work on their competitive positions (Boisen, 2007 ; Caragliu et al., 2011 ; Giffinger et al., 2007). Cities are working on programs, linking technological innovations to economic, political and social/cultural change and use the smart city term to market these programs (Hollands, 2008). Another type of marketing for which the term smart city is often used, is for selling innovative (ICT) products. Large (ICT) companies such as Cisco, IBM, Siemens, Philips and Hitachi use this concept to sell their innovative technologies which contribute to 'smarter' cities (Bélissent et al., 2010 ; Harrison & Donnelly, 2010 ; Peek & Toxler, 2014 ; Philips, 2014 ; Siemens, 2014). Examples of such products are the smart grids, security solutions and traffic control systems of Siemens and the smart lightning system of Philips (Philips, 2014 ; Siemens, 2014).

Second, There are multiple terms used to describe (almost) the same phenomenon (Batty, 2013 ; Hollands, 2008). Examples are the innovative city, wired city, digital city, virtual city, information city, creative city, cultural city and intelligent city (Batty, 2013 ; Hollands, 2008). These terms often have in common is that they link technological innovation with economic, political and socio-cultural change (Hollands, 2008).

Both the public sector and private sector are thus involved in smart cities and both use the term to associate with technological innovations in the urban environment. However, the exact meaning of the term is unclear. It could thus be stated that 'smart city' is a buzzword: "*a word or phrase, often sounding authoritative or technical, that is a vogue term in a particular profession, field of study, popular culture, etc.*" (Dictionary.com, 2014).

2.2 Drawing up a definition

Scientific literature cannot provide a clear definition for the smart city concept either. Multiple definitions are used and there seems to be no 'true' or widely shared definition. When comparing multiple definitions, it turns out that there is a large focus on the importance of ICT (for multiple definitions, see Chourabi et al., 2012 ; Gil-Garcia et al., 2013 ; Nam & Pardo, 2011). Besides this technological component (ICT-use), definitions often include a more societal/human capital component and in some there is also a policy and institutional component or environmental interest (Caragliu et al., 2011 ; Nam & Pardo, 2011). A definition that combines these elements well is the definition that a city may be called "*'smart' when investments in human and social capital and traditional (transport) and modern (ICT) communication infrastructure fuel sustainable economic growth and a high quality of life, with a wise management of natural resources, through participatory government."* (Caragliu et al., 2011, p.70 ; Schaffers et al., 2011, p.432). This definition also highlights the fact that the smart city is a city in which the citizens are 'smart': *…self-decisive, independent and aware"* (Giffinger et al., 2007, p.11).

Besides these definitions, the components of the smart city can contribute to creating an understanding of the concept. There are six characteristics of the smart city: smart economy, smart people, smart governance, smart mobility, smart environment and smart living (Caragliu et al., 2011 ; Cohen, 2012 ; Giffinger et al., 2007 ; Lombardi et al., 2012 ; Papa et al., 2013). These characteristics are clearly visualized by Cohen (2012) after researching multiple characterisations of the smart city concept (see figure 1, p.7). Cohen (2012) believes that definitions focussing on ICT and the use of

information are too narrow and tries to provide a more integrated approach with the smart city wheel.

2.2.1 Smart city objectives

Although the wheel seems to focus on making a division in components, Cohen (2012) argues that these components can also be seen as sub goals. For a city to be smart, a smart economy, a smart government, smart living etc. need to be established. For each of these sub goals, three key drivers have been identified.

Other thoughts on the objectives of smart cities focus on the three major developments addressed in the introduction (see paragraph 1.1, p.7): urbanisation, environmental sustainability and the rise of Information and Communication Technology (Bélissent et al., 2010 ; Bettencourt, 2013 ; Gil-Garcia et al., 2013 ; Naphade et al., 2011). If the overarching objective of the smart city is defined in literature, in most cases the goal is to improve the quality of living in the city for its residents (Bakici et al., 2013 ; Hall, 2000 ; Martinez-Ballesté et al., 2013 ; Papa et al., 2013). Others seem to use the objective of the smart city almost as a reason to implement ICT and seek the goal of the smart city in providing "*a framework for the implementation of information services for monitoring public areas and infrastructures*" (Filipponi et al., 2010, p. 281). However, so far, these definitions of smart city objectives remain very broad. Giffinger et al. (2007) discovered that German smart cities focus on different objectives and objectives are thus case-specific. Even within a case there can be problems with smart city goals: a "*lack of alignment of organizational goals and project*", as well as "*multiple or conflicting goals*" (Chourabi et al., 2012, p.2291) are among the main organisational challenges. It can thus be concluded that smart city goals or objectives are broad, can differ per case and, within cases, they can also cause ambiguities. This is an important issue, since it can cause problems in using data for smart city objectives (see paragraph 5.3, p.31).

2.2.2 Smart city: a fuzzy concept

It can be concluded that "*In general, the term 'smart city' is a fuzzy concept which is not used consistently in the literature*" (Steenbruggen et al., 2014, p.3). THE smart city does not seem to exist, since there are no agreements on when to label a city as smart and since every city would want to be labelled smart city. For this research, therefore, the focus will not be on the city itself, but on programs aiming to make the city 'smarter'. The focus is on achieving sub goals, as Cohen (2012) illustrated with his smart city wheel. In the remaining parts of this research, 'smart city' will refer to the following working definition: Programs taking place in the urban environment aiming to contribute to innovation in this urban environment by making use of ICT.

2.3 Smart public safety

In this research, the focus will be on smart public safety. Safety is one of the key drivers of the smart living sub goal in the smart city wheel (see figure 1, p.7) and is prominent in some smart city definitions (Bélissent et al., 2010 ; Chourabi et al., 2012 ; Gil-Garcia et al., 2013 ; Nam & Pardo, 2011).

Surprisingly, little scientific literature has been written about safety in the smart city context, while at the same time public safety improvement programs in many cases are on the priority list of governmental bodies (Bélissent et al., 2010). Public safety can be seen as the work of police, fire and ambulance services in case of an emergency, but also in the wider context of public protection such as in securing and controlling mass events, public administration transactions and workflows and in providing surveillance of public spaces (Analysys Mason Limited, 2010 ; Bélissent et al., 2010). Also, a division in natural disasters and disasters caused by humans can be made (Bartoli et al., 2013). To address these components of smart public safety, Bartoli et al. (2013, p.2), list tasks that need to be accomplished in a logical order:

"*Monitoring and forecasting*	*adopting specific monitoring activities and analysis of results arising from the monitoring campaigns to prevent natural or man-made disasters and crimes.*

Planning	*preparing action plans to be adopted in case of disaster.*
Emergency responding	*management and coordination of the operations of the first responders, which follow a natural disaster to limit damages and restore security.*
Recovering	*handling post-emergency activities with the aim of coordinating, designing and verifying the restoration works for a rapid return to normal life conditions."*

In this research, it is the first task that will be focused upon: monitoring and forecasting urban public safety. It addresses using data to prevent threats to public safety caused by human activity (so natural disasters are excluded), by monitoring and forecasting. This is an activity that is likely to take place in multiple cities in the Netherlands and has the largest focusses on data out of the four listed tasks. Monitoring and forecasting can be done by making use of big, open and geo-data and is thereby also likely to cause difficulties in the field of privacy.

3. Theoretical context: data

In smart cities, data is of great importance. Data is needed as input for ICT innovations, and ICT is an important component of the smart city (Gil-Garcia et al., 2013). There are two main recent developments in the field of data: big data and open data. Both kinds of developments offer opportunities for smart cities (Analysys Mason Limited, 2010 ; Batty, 2013 ; Batty et al., 2012 ; Schaffers et al., 2011). It is important to note that big and open data are not completely separate categories of data. Big data can, for example, also be open data. Furthermore, a large percentage of data can be considered geo-data and it is expected that in the near future the percentage will increase (Batty et al., 2012 ; Van Oortmarssen & De Vries, 2014). The overlap of big, open and geographic data is illustrated in the data component of figure 3 (see paragraph 1.3.1, p.9).

3.1 Big data

"Big Data is surely the Gold Rush of the Information Age." – Marshall, 2012, p.213.

"W*ithin the next twenty years, most of the data that we will use to understand cities will come from digital sensors of our transactions and will be available in various forms, with temporal tags as well as geotags in many instances"* (Batty et al., 2012, p.488). This citation indicates the importance of big data in cities in the near future (Batty, 2013 ; Batty et al., 2012 ; Marshall, 2012). There are many big data definitions (Kamp, 2014), but big data can easily be understood by the definition that it is "*any data that cannot fit into an Excel spreadsheet*" (Batty, 2013, p.274). Big data is data in large and complex datasets. It cannot be processed and analysed with traditional systems. New processors, software and algorithms are necessary in working with these datasets (European Commission, 2014a ; Kamp, 2014). Furthermore, big data is often produced automatically and routinely, for example by making use of sensors, chips or the internet (Batty et al., 2012, Kamp, 2014).

The Dutch national government states that connecting multiple sensors to the internet is an important development (Kamp, 2014). This is known as 'the Internet of Things' (IoT): objects (things) that are able to interact with other objects via the internet to achieve a certain goal (Jara et al., 2014). There are many possibilities, for example *"Radio-Frequency Identification (RFID), tags, sensors, actuators, mobile phones, etc."* (Atzori et al., 2010, p.2787). It connects electronic devices that can be located in private places, such as in houses, transportation (cars) but also in public space, such as in streets and public buildings (Komninos et al., 2011). The government expects that these developments will lead to high-quality data that is becoming cheaper to collect, store, transport and process in the near future, because of technological progress (OECD, 2012 ; Kamp, 2014).

To make use of these large data sets, data mining will become essential (Batty et al., 2012 ; Kamp, 2014). An easily understandable working definition for data mining is *"the analysis of (often large) observational data sets to find unsuspected relationships and to summarize the data in novel ways that are both understandable and useful to the data owner"* (Hand et al., 2001, p.6). An important part of data mining is the fact that there often is no predefined hypothesis, as is highlighted in the citation by the words 'to find unsuspected relationships' (Kamp, 2014).

It should be noted that big data is not only important in the ICT sector, but also for other sectors such as services, industry, healthcare and education. It can be used for improvements in efficiency in both private and public sectors, as well as in obtaining knowledge for both commercial and scientific use (Kamp, 2014). Many of these sectors are also related to cities, indicating the relevance of big data for the urban environment.

Since big data can be of use for many sectors, it is expected that the big data business will become booming in the near future (European Commission, 2014a ; Kamp, 2014 ; OECD, 2014). Big data in itself also contributes to economic activity, since, among others, data analysts, legal advisors and storage capacity are needed (Kamp, 2014). Many estimates on the economic contribution of big data have been made, all highlighting the benefits of big data use (European Commission, 2014a ;

Kamp, 2014 ; OECD, 2014). The European Commission (2014a), for example, expects the economic value of big data technology and services to grow annually at a rate of about 40.0%, to 16.9 billion dollar in 2015 and the organisation estimates that the number of big data specialists working in the United Kingdom will increase by 240.0% in the next five years. This example of numbers on the economic contribution of big data indicates the high expectations of benefits of big data, just like many other studies do (European Commission, 2014a ; Kamp, 2014 ; OECD, 2014). However, it might be more interesting to pay attention to what actually is being done and can be done with data, instead of on rough numerical estimates on economical contributions. As Zijlstra (2014a) points out: interesting metrics are about interaction; they are on what is being done with data, the connections made.

3.2 Open data

"*Open data is a goldmine*" – Neelie Kroes, in Zijlstra, 2014b

Besides big data, open data is also of interest for smart cities (Analysys Mason Limited, 2010 ; Batty et al., 2012 ; Schaffers et al., 2011). There are multiple definitions of open data. In short, "*open data and content can be freely used, modified, and shared, by anyone for any purpose*" (Open Knowledge, 2014a). Zijlstra (2014b) defines the major characteristics of open data clearly: Data that is gathered for a public task, pro-actively published, to be used by others and with no legal, technical and monetary barriers. In some cases, the term open government data is used, to refer to the fact that open data still is primarily provided by governments (Kulk & Van Loenen, 2012a). The Dutch government defines open government data by referring to five main characteristics (Rijksoverheid, 2014 ; Turksema et al., 2014) (see figure 5).

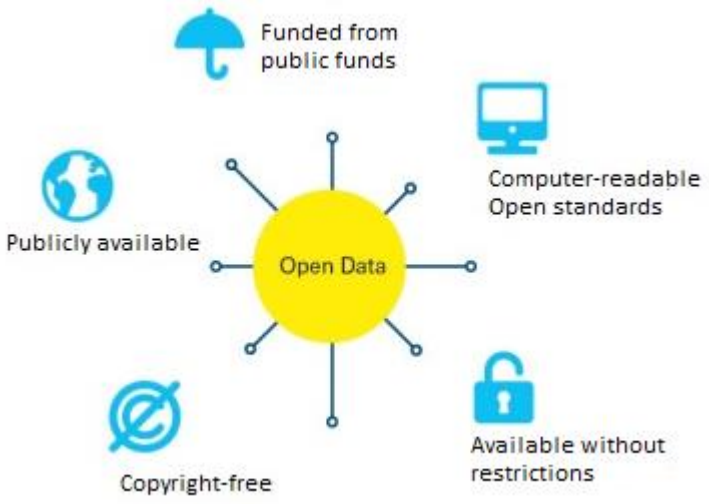


Figure 5: Open data criteria according to the Dutch national government
Source: Turksema et al., 2014 (texts translated from Dutch to English)

There are multiple reasons for creating open (government) data, which will shortly be addressed. First, there is enhanced transparency. By opening up data, the governmental transparency increases (European Commission, 2014b ; Kronenburg et al., 2012). In a democracy and with citizens as tax payers, citizens should be able to access information from the government as well as sharing this information (Open Knowledge, 2014b). Second, open data can enhance innovation in both the private and public sector (European Commission, 2014b ; Kronenburg et al., 2012 ; Open Knowledge, 2014b). Third, participation from citizens in the government can be improved. Opening up data connects citizens and the government, since it makes citizens more informed and enables them to contribute (European Commission, 2014b ; Open Knowledge, 2014b). Fourth, from a governmental

perspective, open data can contribute to a better service for companies and citizens. Open data can be both time-saving and money-saving and is thus an efficient way of working (European Commission, 2014b ; Kronenburg et al., 2012). The European Commission is also very positive in its estimates on financial benefits of open data: the Commission estimates the direct and indirect gains of open government data EU-wide at €140 billion a year (European Commission, 2011).

In making the most out of the benefits of open data, re-use of open data is important. Companies, governmental bodies and citizens can re-use datasets and create added value. The national government has therefore created its own data portal and established a network of knowledge on open data to identify and demolish barriers in (re-)using and sharing open data (Ministerie van Buitenlandse Zaken en Koninkrijksrelaties, 2013).

The European Union is also concerned with government-held data and re-use. In 2003, a directive on the re-use of public sector information (PSI) was established (EU Directive 2003/98/EC, 2003). More harmonisation in data-sharing within the EU is among the goals of this directive (European Commission, 2011). Also in the Infrastructure for Spatial Information in the European Community (INSPIRE) directive data sharing is central (EU Directive 2007/2/EC, 2007). INSPIRE is an important contribution to the re-use of PSI (Van Loenen & Grothe, 2014). As a result, data portals on multiple governmental levels are established and available for (re-)use (European Commission, 2011).

3.3 Geographic data

The majority of (open) data relates to locations on earth, so-called geographical data (geo-data) (Batty et al., 2012 ; Kulk & Van Loenen, 2012a ; Van Oortmarssen & De Vries, 2014). In the near future, it is expected that most data will have both geo-tags and temporal tags (Batty et al., 2012). It is estimated that about 60.0-80.0% of all data has a geo-component (Dempsey Morais, 2012). Just like Batty et al. (2012) estimated two years earlier, it is plausible that the percentage of geographic data will increase in the near future. The increasing use of sensors (see paragraph 3.1, p.19) is an important development herein, since sensors very often collect location data (Van Oortmarssen & De Vries, 2014). The fact that a large percentage of data is geo-data also illustrates that big data, open data and geo-data are not separate categories of data. Both big and open data are thus in many cases also geo-data.

Geo-data refers to "*the position of someone or something at a certain point in time and with certain accuracy. It links place, time, and attributes. Some attributes are physical or environmental in nature, while others are social or economic"* (Van Loenen et al., 2008, p.42). In this definition, the fact that geo-data is able to link place, time and attributes is important. Geo-data thus has the ability to link different kinds of data, by arranging on location. This is, according to Van Oortmarssen and De Vries (2014) an efficient system, since geographic coordinates are very precise and this storage system can relatively easily be used worldwide.

4. Theoretical context: privacy and data protection

4.1 Privacy versus data protection

First of all, it is important to pay attention to privacy in general. There is an important difference in privacy and data protection: privacy is a broader concept (Cuijpers & Koops, 2013 ; Van Loenen et al., 2008). After studying privacy definitions and typologies in literature, Van Loenen et al. (2008) identify four types of privacy. Privacy is divided in privacy of the body, psychological privacy, territorial privacy and behavourial privacy. Privacy in the field of data protection addresses territorial privacy to some extent, but mainly focusses on behavourial privacy. Behavourial privacy can be categorised in physical privacy, informational privacy and privacy of communications (see figure 6) (Van Loenen et al., 2008).

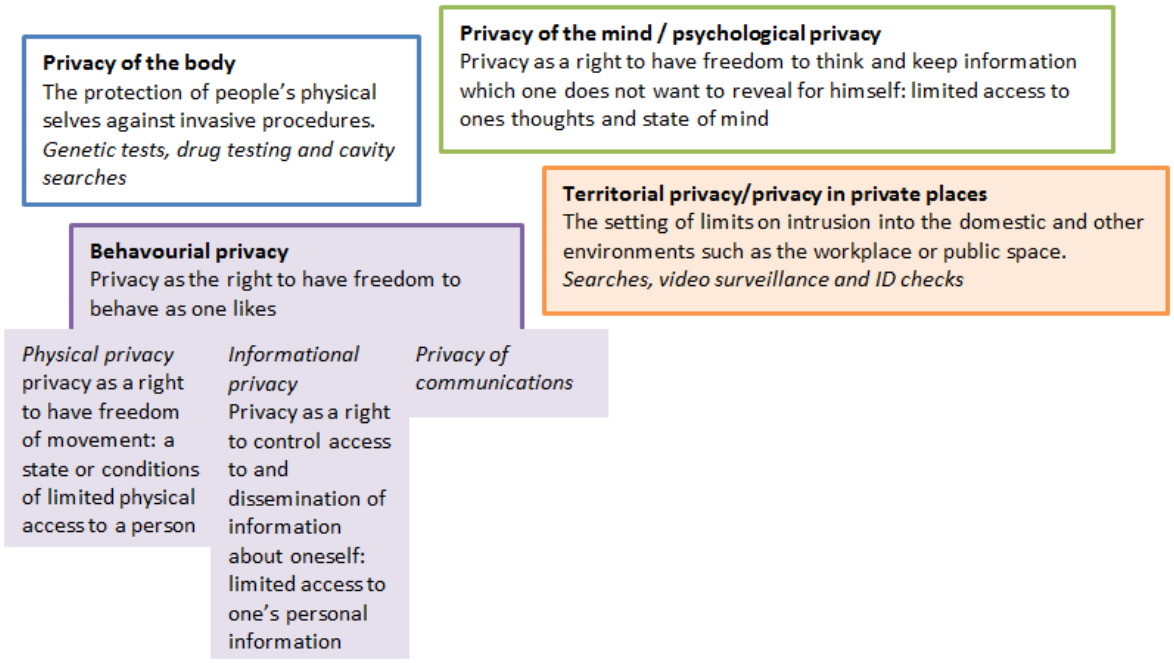


Figure 6: Different types of privacy.
Source: Van Loenen et al., 2008, p.19

Cuijpers and Koops (2013) also make a division in privacy aspects; in a territorial component, a relational component and an information component (Cuijpers & Koops, 2013). This research focuses on this last component of privacy: data protection. In both privacy definitions, data protection is one of the components. So, it should be kept in mind that privacy in this research refers to data protection, which is also often the case when people speak about privacy in today's society (Cuijpers & Koops, 2013).

Since this research focuses for a large part on geo-data, it is also interesting to pay attention to the privacy of location. Van Loenen et al. (2008) provide a definition for this type of privacy: "*the ability to prevent other parties from learning one's current or past location" (Beresford et al. 2003). It may also be defined as the ability to control the extent to which personal location information is being used by others*" (Van Loenen et al., 2008, p.42).

4.2 An introduction to European data and privacy protection

In the European Union, privacy is being seen as a fundamental right. In the European Convention on Human Rights (ECHR), in articles 5 and 8, fundamental privacy rights are taken into account: "*Everyone has the right to respect for his private and family life, his home and his correspondence"* and *"Everyone has the right to liberty and security of person"* (ECHR, 2010, pp.7-8 & 11). The ECHR

dates from 1950, and was amended by protocols since then (ECHR, 2010). In 1981 Convention 108 of the Council of Europe was established, in which "*the protection of individuals with regard to automatic processing of personal data*" is taken into account (Convention 108, 1981, p.1). It was established to ensure data protection in view of the increasing use made of computers, which can process data automatically instead of manually (Convention 108, 1981). Privacy protection on the European level is thus not a recent development.

Besides the articles in the ECHR and Convention 108, data protection and privacy is taken care of in multiple directives. In 1995, Data Protection Directive 95/46/EC "*on the protection of individuals with regard to the processing of personal data and on the free movement of such data"* was established (EU Directive 95/46/EC, 1995, p.31) (see paragraph 4.3). This Directive is today still the most important European directive on privacy protection.

In addition to the 1995 directive, in 2002 the ePrivacy Directive 2002/58/EC was established and this directive was amended in 2009 by the ePrivacy Amendment Directive 2009/136/EC (EU Directive 2002/58/EC, 2002 ; EU Directive 2009/136/EC, 2009) (see paragraph 4.4, p.25). The directive concerns *"the processing of personal data and the protection of privacy in the electronic communications sector (directive on privacy and electronic communications)*" (EU Directive 2002/58/EC, 2002, p.37). The increasing use of mobile phones and the internet led to the need for an addition to the 1995 Data Protection Directive.

In the field of data protection, these directives are the most important ones. They are implemented by all member states; in the Netherlands the Data Protection Directive is implemented in the Wet bescherming persoonsgegevens (Wbp) and the ePrivacy Directive is implemented in the Telecommunicatiewet (Telecommunicatiewet, 1998 ; Wet bescherming persoonsgegevens, 2000). Recently, on EU-level, a plan for a General Data Protection Regulation (GDPR) has been proposed, replacing the Data Protection Directive of 1995 (European Commission, 2014c). However, it is yet unclear whether or not the GDPR will be implemented and if so, when this will be and what the final contents will be (Costa & Poullet, 2012 ; Cuijpers & Marcelis, 2012 ; De Hert & Papakonstantinou, 2012 ; De Hert et al., 2013 ; European Data Protection Supervisor, 2012). For these reasons it is decided not to focus on this draft. However, it is an important development in European data and privacy protection and should therefore be briefly mentioned in this paragraph.

4.3 Data Protection Directive

The Data Protection Directive is the most important directive on data protection and pays attention to processing personal data. In this paragraph, the content of the directive that is most relevant for data use in a smart city context will be outlined.

4.3.1 Personal data

The Data Protection Directives main aim is the protection of personal data. Personal data is defined in this directive as *"any information relating to an identified or identifiable natural person ('data subject'); an identifiable person is one who can be identified, directly or indirectly, in particular by reference to an identification number or to one or more factors specific to his physical, psychological, mental, economic, cultural or social identity"* (EU Directive 95/46/EC, 1995, p.38). This definition is very general (on purpose), so all information concerning an identifiable individual would be included (Article 29 Working Party, 2007).

In recital 26 of the Data Protection Directive it is stated that anonymising data is not enough to change personal data into non-personal data: *"account should be taken of all the means likely reasonably to be used either by the controller or by any other person to identify the said person"* (EU Directive 95/46/EC, 1995, p.33). This recital ensures that even presumably anonymised data can become personal data again by making use of data mining and modern techniques, and should still be considered personal data (for examples see Kulk & Van Loenen, 2012a).

With this broad definition of personal data, recital 26, technological progress, data mining, and the increasing amount of location-specific data, more and more data will become personal data (Article 29 Working Party, 2011 ; Cuijpers & Marcelis, 2012 ; De Hert et al., 2013 ; European Commission, 2014c ; Van Loenen et al., 2008).

If data is personal data, this does not mean that it is impossible to process personal data. There are rules set in article 6 of the Data Protection Directive for processing personal data (EU Directive 95/46/EC, 1995, p.40):

"1. Member States shall provide that personal data must be:

a. *Processed fairly and lawfully;*
b. *Collected for specified, explicit and legitimate purposes and not further processed in a way incompatible with those purposes. Further processing of data for historical, statistical or scientific purposes shall not be considered as incompatible provided that Member States provide appropriate safeguards;*
c. *Adequate, relevant and not excessive in relation to the purposes for which they are collected and/or further processed;*
d. *Accurate and, where necessary, kept up to date; every reasonable step must be taken to ensure that data which are inaccurate or incomplete , having regard to the purposes for which they were collected or for which they are further processed , are erased or rectified;*
e. *Kept in a form which permits identification of data subjects for no longer than is necessary for the purposes for which the data were collected or for which they are further processed. Member States shall lay down appropriate safeguards for personal data stored for longer periods for historical, statistical or scientific use.*

2. It shall be for the controller to ensure that paragraph 1 is complied with."

When personal data is being processed, it must be done according to these rules. Otherwise processing becomes illegitimate and there might be sanctions (specified in the national implementations of the directive). For this research, it is important to pay attention to the fact that there must be a (predefined) specified, explicit and legitimate purpose to process personal data (article 6.1.b).

4.3.2 Content of the Data Protection Directive

Besides the articles on personal data and processing personal data, there are other articles that are useful to address briefly. In many of them, the term 'data subject' is used, which refers to "*an identified or identifiable natural person*" (EU Directive 95/46/EC, 1995, p.38). The controller is "*the natural or legal person, public authority, agency or any other body which alone or jointly with others determines the purposes and means of the processing of personal data"* (EU Directive 95/46/EC, 1995, p.38).

In article 7 of the Data Protection Directive, criteria are set for data processing. The data subject needs to be aware of the fact that data is being collected and processed. The subject can give direct permission, or one of the other criteria for processing is satisfied, for example to protect his or her interests.

In article 10 of the Data Protection Directive, information that needs to be provided to the data subject is listed. At least, the data controller must reveal its identity, the purpose of data collection and processing and further information, depending on more specific circumstances.

In article 12 of the Data Protection Directive, there are rules set for the rights of the data subject. The data subject has for example the right to obtain information regarding whether or not his or her data is being processed and on the way this is being done. The data subject then also has the right to object (article 14 of the Data Protection Directive).

In general it is not allowed to process personal data about ethnicity, political opinion, religion, trade-union membership, health data and data about sex life. These rules are set in article 8

of the Data Protection Directive, but in this article also exceptions are listed. Among others, exceptions include direct permission from the data subject or processing to protect his or her interests.

It is important to note that there is room for national differences in implementation. Article 13 of the Data Protection Directive is important herein, because it states that member states can narrow the scope of multiple articles of the Directive (6(1), 10, 11(1) & 21) when this is necessary because of, among others, defence, national security and public security. This last one is important for this research, because of its focus on public safety programs in a smart city context. The Dutch implementation will be discussed in paragraph 4.5 (see p.26).

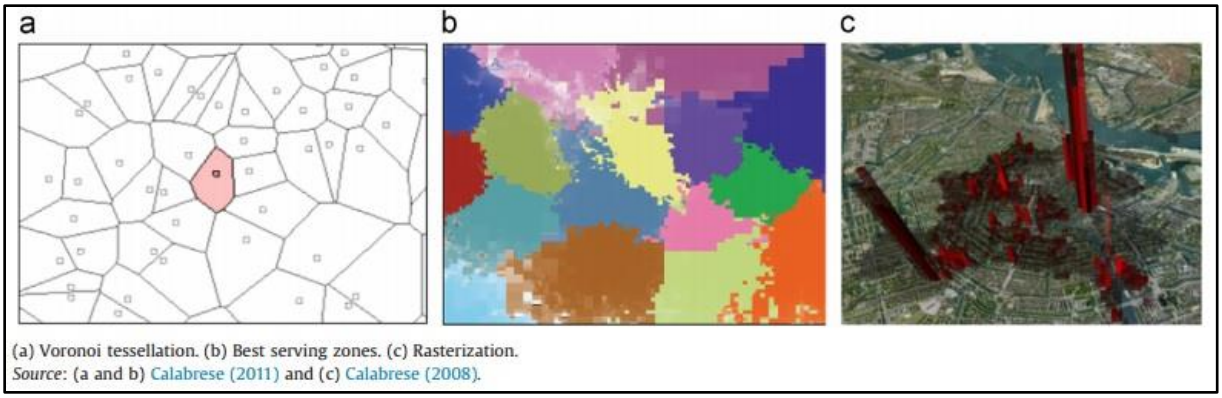
Each member state is responsible for ensuring that there is a public authority concerned with the implementation and supervision of (living up to) these rules (EU Directive 95/46/EC, 1995). For the Netherlands, this is the College Bescherming Persoonsgegevens (CBP). Furthermore, in article 29 of the Data Protection Directive, it is stated that a Working Party will be established, with advisory status on EU-level (EU Directive 95/46/EC, 1995). The Working Party is often named the Article 29 Working Party, after the number of the article that established the organisation. The Article 29 Working Party works independent and advises through working documents and opinions on European data protection (Kulk & Van Loenen, 2012 ; Van Oortmarssen & De Vries, 2014). The Article 29 Working Party consists of the national public authorities concerned with privacy protection (for the Netherlands this is the CBP). Besides providing advice on living up to the 'rules' of the directive, the Working Party also checks the national privacy protection authorities (Van Oortmarssen & De Vries, 2014).

4.4 The ePrivacy (Amendment) Directive

In 2002, the ePrivacy Directive 2002/58/EC was established and this directive was adjusted in 2009 with the ePrivacy Amendment Directive 2009/136/EC. The directive concerns "*the processing of personal data and the protection of privacy in the electronic communications sector (directive on privacy and electronic communications)"* (EU Directive 2002/58/EC, 2002, p.37). It extends the processing of personal data of the Data Protection Directive to personal data processed by publicly available communications services of the electronic communications sector (Wong, 2011). Technological development leads to the quest for this directive, especially because of the widespread and intensive use of digital mobile networks and (communication over) the internet. This development also has consequences for processing personal data, since these networks offer possibilities for processing such data (EU Directive 2002/58/EC, 2002). For this research, this directive is mostly interesting because of its framework for processing location data.

In the ePrivacy Directive, there is more emphasis on location data, compared to the Data Protection Directive. Location data means "*any data processed in an electronic communications network, indicating the geographic position of the terminal equipment of a user of a publicly available electronic communications service*" (EU Directive 2002/58/EC, 2002, p.43). But there is another kind of location data in the directive; so-called 'traffic data' and this can be location data too: "*any data processed for the purpose of the conveyance of a communication on an electronic communications network or for the billing thereof"* (EU Directive 2002/58/EC, 2002, p.43). Traffic data include data of a communication, among which are cell-phone locations; locations at the beginning and end of the conversation (Van Loenen et al., 2008). Van Loenen et al. (2008) explain that these data are in the scope of traffic data instead of location data, because these data are needed to enable the transmission of communications.

It is for mobile phone operators in most cases not difficult to collect data, since many data can be considered traffic data. With this information, an estimation of location at different spatial and temporal scales can be made (Steenbruggen et al., 2014). This can be visualized in three ways (see figure 7, p.26).



(a) Voronoi tessellation. (b) Best serving zones. (c) Rasterization.
Source: (a and b) Calabrese (2011) and (c) Calabrese (2008).

Figure 7: Different visualisations of mobile phone location data.
Source: Steenbruggen et al., 2014, p.4

There is also a downside to working with traffic data. Traffic data are only collected when a service is being used. Furthermore, the number of cell towers differs across regions and can result in location data that is not detailed (Steenbruggen et al., 2014).

In article 9 of the ePrivacy Directive it is stated that location data can only be used when it is anonymised or when processing location data is necessary for the provision of a service and the data subject has given direct permission for data collection (EU Directive 2002/58/EC, 2002). The data subject needs to be enabled to stop or prevent collection of his or her location data (EU Directive 2002/58/EC, 2002). Also, information on the type, the purpose and the duration of the location data collection and processing should be provided (EU Directive 2002/58/EC, 2002). Furthermore, the objectives of data processing need to be clearly defined and data processing needs to be done with supervision of public authority. There are exceptions on these rules for emergency calls for services such as the ambulance, fire brigade and police. Location data (non-anonymised) can be used to respond to these calls (EU Directive 2002/58/EC, 2002). Just like the Data Protection Directive, it is up to member states to decide on narrowing the scope of the Directive when this is necessary because of, for example, defence, national security and public security (article 15 of the ePrivacy Directive).

4.5 Data protection in the Netherlands

In the Netherlands, the Data Protection Directive 96/46/EC and the ePrivacy (Amendment) Directives 2002/58/EC and 2009/136/EC are translated to national laws in the Wbp and the Telecommunicatiewet (Telecommunicatiewet, 1998 ; Wet bescherming persoonsgegevens, 2000).

In article 6 of the Data Protection Directive there are rules set for collecting and processing personal data (see paragraph 4.3.1, p.23). These rules are fully included in the Dutch Wbp. All criteria for data processing of article 7 of the Data Protection Directive implemented in the Netherlands[1]. It is stated that data processing is only allowed when it is needed based on at least one of the reasons listed in the Wbp[2]. Legitimate reasons are, among others, processing with explicit permission from the data subject, processing because of legal obligation or contract with the data subject and processing to protect vital interests of the data subject. There are major exceptions to (parts of) articles[3], for example for national safety and prevention, detection and prosecution of crime. The supervision of these rules is also taken care of in the Wbp[4]. The major Dutch institution concerned with this task is the College bescherming persoonsgegevens (CBP).

Only a small part of the Telecommunicatiewet is relevant for this research: the articles focussing on the protection of personal data in the private environment[5] and more specifically on

[1] In article 8 of the Wbp
[2] In article 8 of the Wbp
[3] In article 43 of the Wbp
[4] In articles 51-64 of the Wbp
[5] In chapter 11 of the Telecommunicatiewet

processing geo-data. In the Dutch implementation of the ePrivacy Directive, the distinction in traffic data and location data is also made. The Telecommunicatiewet states that traffic data should not be stored longer than strictly necessary by providers of electronic communications, that traffic data should be anonymised and that the data subject should be notified about data storage[6]. Location data processing (other than traffic data) is only allowed if data is anonymised or the data subject has given permission. Before the data subject can give permission, it is required that the provider gives information on the kind of location data, the goals and the period of time of data processing and whether or not data will be provided to a third party. Again, for purposes as national security and the prevention, investigation and prosecution of crime, there are exceptions to these rules. The supervision of the Telecommunicatiewet is the task of the Autoriteit Commerciële Markt (ACM), who works closely together with the CBP on processing personal data (Telecommunicatiewet, 1998 ; Van Oortmarssen & De Vries, 2014).

Overall, there are no large differences between data protection on European level and the Dutch implementation. In general, the main articles of the Data Protection Directive discussed in this chapter, are literally copied to the Dutch law (Korff, 2002). Small differences are mainly in the level of detail; the Dutch implementation is, logically, more detailed than the European framework set in the directive. This is especially the case in the Telecommunicatiewet, which has a broader scope than the ePrivacy Directives.

[6] In article 11.5 of the Telecommunicatiewet

5. Theoretical context: Data and privacy in a smart city context

In chapters 2, 3 and 4, the three major concepts of this research are discussed individually. In this final theoretical chapter, they are combined and thereby the theoretical difficulties in the field of privacy for processing personal big, open and geo-data for monitoring and forecasting public safety in smart cities will become evident. Thereafter, the theoretical conclusion will be formulated.

5.1 Data in the smart city

The value of data for smart cities is clarified in multiple researches, most of them focussing on big data. Services and infrastructure in the near future will become more engineered, invisible, predictable and automatic and therein big data can be used to improve these systems (Bettencourt, 2013). But besides improving systems, there are opportunities for big data in social interaction and more informed decision-making (Batty, 2013). This confirms the statement of the Dutch national government that there are many applications of big data, for both the public and private sector and also in non-ICT-related sectors (Kamp, 2014).

So far the use of big data for smart cities remains very general, since big data can be used in many sectors (European Commission, 2014a ; Kamp, 2014). And even in public safety, the range of data that is relevant is diverse (Analysys Mason Limited, 2010). However, there are more specific developments in big data use for the urban environment.

First, big data enables real-time systems monitoring (Jara et al., 2014). Sensors, for example, collect data to be able to directly respond to a situation at a particular moment in time (Calabrese et al., 2009). Calabrese et al. (2009) and Komninos et al. (2011) define four key components of a real-time control system. These are: an "*entity to be controlled in an environment characterized by uncertainty; sensors able to acquire information about the entity's state in real-time; intelligence capable of evaluating system performance against desired outcomes; physical actuators able to act upon the system to realize the control strategy*" (Komninos et al., 2011, p.3). There are many applications for real-time systems monitoring, among which is managing emergency response, part of public safety (Jara et al., 2014) (see paragraph 2.3, p.17).

Second, another use of big data in smart cities is location data from mobile phones (Jara et al., 2014). Location-aware mobile applications can be used to gain spatial (big) data (Jara et al., 2014 ; Komninos et al., 2011 ; Tene & Polonetsky, 2013). "*Mobile devices–always on, location aware, and with multiple sensors including cameras, microphones, movement sensors, GPS, and Wi-Fi capabilities–have revolutionized the collection of data in the public sphere and enabled innovative data harvesting and use*" (Tene & Polonetsky, 2013, p.247). This citation highlights the possibilities of data collection through mobile phones and the possibilities for this in the public sphere. GPS and Wi-Fi enable location data collection, and there are numerous ways for obtaining data on indoor locations, for example by making use of infrared, wireless networks, sensor networks or Bluetooth (Larkou et al., 2014 ; Tene & Polonetsky, 2013). Mobile phone operators collect location data. These data can, for example, be used to track mobility patterns (Steenbruggen et al., 2014).

Steenbruggen et al. (2014) also studied the potential of using mobile phone data in public safety programs. Mobile phone location data turns out to be very useful in these programs, because it enables to predict the number of people attending an event or in an emergency situation and enables to detect patterns in movement. It thus enables to obtain real-time information. This is useful in urban decision making, especially when it can be combined with qualitative data. "*Also in the field of crisis management, there is a growing interest in using telecom data for crowd management and anomaly detection. For instance, responses to a crisis require a high level of preparedness, and the precise knowledge of how many people could be exposed in the incident area*" (Steenbruggen et al., 2014). Data from mobile phones can thus be of use in smart city programs and therefore the ePrivacy Directive is relevant for this research.

Besides big data use for real-time systems monitoring and mobile phone location data, other interesting technological developments for smart cities are smart sensors which can be connected to

each other and the internet (the Internet of Things) and smart devices. Sensors (networks, Internet of Things) are an interesting development for monitoring and forecasting in public urban space (Komninos et al., 2011). Batty et al. (2012) even state that sensor use in cities will become increasingly important in the near future. Furthermore, open data and cloud computing can be considered interesting developments for smart city purposes (Komninos et al., 2011).

The relevance of big data for smart city (public safety) programs has now become evident. Furthermore, specific attention has been paid to the value of mobile phone location data, an implementation of big data and geographic data in the smart city. The large majority of data is geo-data (Batty et al., 2012 ; Kulk & Van Loenen, 2012a ; Van Oortmarssen & De Vries, 2014). However, there is less literature available on the linkage of open data and the smart city. It should be kept in mind that open data can be big data too, and can thus also be gained, for example, from sensors. Open data, like big data, can also contribute to urban decision making. *"Open data from various sources, government, sensors, citizens and businesses, offer opportunities for advanced analytics and intelligence to detect patterns, generate alerts, visualise information and predict trends*" (Komninos et al., 2011, p.5).

Open data can be seen as data that is essentially linked to the urban citizen, by the openness of Dutch government. In words of Tene and Polonetsky (2013, p.255): "*if you're not paying for it, you're not the customer; you're the product*". One of the reasons for creating open data is governmental transparency. Open data is financed by the government, and thus indirectly by citizens (European Commission, 2014b ; Kronenburg et al., 2012). Also, open data aims to connect the government and citizens, since it informs and involves them (European Commission, 2014b ; Open Knowledge, 2014b). This clarifies the existential link of open data in cities and its citizens. An example of open data use for smart city purposes comes from Bakici et al. (2013). They agree with the statement that open data and citizens belong together: "*Recently, the Barcelona City Hall became involved in the Open Data movement with the Open Data project, whose objective is the opening up of government information to public access. These data involve territory, population, management and procedure indicators, urban environment and documental data. It is society's right to use this data, whether to brief themselves or for creating new services, increasing social value and perhaps also commercial value.*" (Bakici et al., 2013, p.144).

In the above statement, Bakici et al. (2013), provide examples on open data that is available for use in smart cities, for the case of Barcelona, but there are many possibilities for open data use in smart city programs. The Dutch national government facilitates (spatial) open data, for example in data portal 'Publieke Dienstverlening op de Kaart' (PDOK) and in the national georegister (PDOK, 2014a ; PDOK, 2014b). On the European level there is also attention for open spatial data (INSPIRE Directive, see paragraph 3.2, p.20). These developments lead, among others, to more accessibility of data and possibilities for smart city programs.

This paragraph clarified the role of data in the smart city. Especially the role of big data becomes evident from literature. Big data enables (real-time) analysis of urban processes and monitoring locations. These data are relevant for monitoring public safety. Also, the role of location data has become clear. The large majority of data is geo-data, and sensor data and mobile data often involve locations. On open data in the smart city, less literature is available. The relevance of open data for the smart city can mostly be found in the fact that open data links citizens and government. Open data contributes to smarter and more involved citizens, which is an essential part of the smart city concept.

5.2 Data and the protection of privacy

The broad definition of personal data, together with recital 26 and technological progress leads to the fact that many data comes under the scope of the Data Protection Directive (Article 29 Working Party, 2011 ; Cuijpers & Marcelis, 2012 ; De Hert et al., 2013 ; European Commission, 2014c ; Van

Loenen et al., 2008). In order to process personal data lawfully, it should be processed according to the rules of article 6 of the Data Protection Directive (see paragraph 4.3.1, p.23). In this chapter, three of these rules are addressed: article 6.1.b on purpose limitation, article 6.1.c on the demand for data to be adequate, relevant and not excessive and 6.1.e on the amount of time the data can be stored for (EU Directive 95/46/EC, 1995.

When the definition of open data (see paragraph 3.2, p.20) is compared to the rules for processing personal data in article 6 of the Data Protection Directive, this leads to the conclusion that, by definition, open data cannot be personal data. In particular, open data cannot satisfy two article 6 rules. First, open data is by definition not collected for a specified, explicit purpose (article 6.1.b DPD) (EU Directive 95/46/EC, 1995 ; Kulk & Van Loenen, 2012a). Data re-use is key to open data (Ministerie van Buitenlandse Zaken en Koninkrijksrelaties, 2013). In re-use, a specified and explicit purpose that has been defined beforehand is missing, and data is likely to be further processed in a way that is incompatible with the original purpose. Second, in open data, there is no predefined amount of time to keep the data (article 6.1.e DPD) (EU Directive 95/46/EC, 1995). The Article 29 Working Party even states that "*Making data available for reuse under an open license should be avoided unless it can be clearly demonstrated that compliance with data protection law can be effectively ensured"* (Article 29 Working Party, 2013, p.50). Before data becomes open data and thus available for re-use, an impact assessment should be done to determine the sufficient level of anonymisation (Article 29 Working Party, 2013). Since so many data is personal data, this might lead to problems with processing open data according to the rules, making it unlawful.

Just like open data, big data raises privacy issues. In large datasets, data mining is needed to make use of it (Batty et al., 2012 ; Kamp, 2014). Thereby, categorization of data is likely to take place and these categories can be applied on a person, which is often referred to as profiling. This way, by combining data, information can become personal information (Kamp, 2014 ; Kulk & Van Loenen, 2012a). The Article 29 Working Party sees, among others, privacy concerns in "*the sheer scale of data collection, tracking and profiling, also taking into account the variety and detail of the data collected and the fact that data are often combined from many different sources*" (Article 29 Working Party, 2013, p.45). Big data is easily being used for other purposes than the original purpose. It is tempting to use big data for different purposes (article 6.1.b DPD), as well as to keep data longer (article 6.1.e DPD) and to process more data than is strictly necessary (article 6.1.c DPD) (EU Directive 95/46/EC, 1995 ; Kamp, 2014). This is not according to the rules set in article 6 of the Data Protection Directive (EU Directive 95/46/EC, 1995). Furthermore, the data subject should be informed about data collection, but it is not easy to do this in an easily-understandable manner (Kamp, 2014). This undermines transparency: "*unless they are provided with sufficient information, individuals will be subject to decisions that they do not understand and have no control over*" (Article 29 Working Party, 2013, p.45). However, there are exceptions to these rules listed. Data processing can in some cases occur without explicit permission from the data subject (EU Directive 95/46/EC, 1995 ; Kamp, 2014). For example, when this is in the vital interest of the data subject (article 8 Wbp). National safety and prevention, detection and prosecution of crime are other exceptions to the rules (article 43 Wbp) (Wet bescherming persoonsgegevens, 2000). Especially in the field of public safety, this is an interesting exception.

Geo-data are in many cases personal data (Van Oortmarssen & De Vries, 2014). The extent to which geo-data can be considered personal data depends on "*the type of information, the level of detail of the location information, the timeliness of the information, and the context to which it is linked*" (Van Loenen et al., 2008, p.52). Whether or not spatial traffic data are personal data depends, for example, on the spatial scale of the data and on its context: "*In a general sense, the use of highly detailed (e.g., scale 1:500), real-time location data linked to a sensitive context, such as a church, can generally be expected to be at a higher 'privacy level' than less detailed data (e.g., scale 1:25,000) of a decade ago without a link to a specific sensitive context"* (Van Loenen et al., 2008, p.52). Geo-data can thus be personal data and in that case need to be processed according to the rules. In itself, geo-data is not in violation with the law. But, a geographic component makes it easier to relate data to

identifiable persons. As became evident in this chapter, this can lead to legal challenges. Furthermore, it is expected that in the future, more and more data will become location-specific and more data will become personal data (Article 29 Working Party, 2011 ; Cuijpers & Marcelis, 2012 ; De Hert et al., 2013 ; European Commission, 2014c ; Van Loenen et al., 2008).

5.3 Data, privacy and the smart city

5.3.1 Purpose limitation

The relevance of the combination of data and privacy in a smart city context becomes evident in the 'rules' of article 6 of the Data Protection Directive. One of these rules is that personal data can only be "*collected for specified, explicit and legitimate purposes and not further processed in a way incompatible with those purposes"* (article 6.1.b DPD) (EU Directive 95/46/EC, 1995, p.40). It is this rule that is of importance for this research, because it combines the subjects of data and data protection in a purpose: smart cities.

The Article 29 Working Party wrote an opinion on the subject of deciding to what extent a purpose for personal data processing should be defined; in *"Opinion 03/2013 on purpose limitation"* (Article 29 Working Party, 2013, p.1). Purpose limitation is a complex understanding, but also a very important one, because it aims to "*protect the data subject by setting limits on how controllers are able to use their data and reinforce the fairness of the processing*" (Article 29 Working Party, 2013, p.11). Purpose limitation deals with a dilemma: "*The limitation should, for example, prevent the use of individuals' personal data in a way (or for further purposes) that they might find unexpected, inappropriate or otherwise objectionable. At the same time, the notion of compatible use also offers some degree of flexibility for data controllers*" (Article 29 Working Party, 2013, p.11). The content of the Working Party 29 opinion will in this paragraph be discussed into detail, because of its major relevance for this research since it addresses the expected problems in data use for smart city purposes.

Article 6.1.b of the Data Protection Directive can be divided in two parts, of which the first part is: "*collected for specified, explicit and legitimate purposes"* (EU Directive 95/46/EC, 1995, p.40). This text is literally translated in the Dutch Wbp[7]. The three demands listed in this citation are clarified in an Article 29 Working Party opinion (Article 29 Working Party, 2013).

First, the purpose must be specified: "*sufficiently defined to enable the implementation of any necessary data protection safeguards, and to delimit the scope of the processing operation"* (Article 29 Working Party, 2013, p.12). This must be done before data collection takes place. It should be clear what is and is not in the scope of the processing of personal data. The level of detail needed for description, depends on the context. It is stated that more detailed information is not in all cases useful. A 'layered notice' is a good approach: "*key information is provided to data subjects in a very concise and user-friendly manner, while additional information (…) is provided for the benefit of those who require further clarification*" (Article 29 Working Party, 2013, p.16). If there are multiple purposes, they should each be specified (Article 29 Working Party, 2013).

Second, the purpose must be explicit: "*the purpose must be sufficiently unambiguous and clearly expressed*" (Article 29 Working Party, 2013, p.12). This must also be done before data collection takes place. It is important that the purpose of data collection can be easily and unambiguously understood by everyone involved, from data subject to data controller (Article 29 Working Party, 2013).

Third, the purpose must be legitimate. It is important to note that legitimacy "*goes beyond the requirement to have a legal ground for the processing under Article 7 of the Directive and also extends to other areas of law"* (Article 29 Working Party, 2013, p.12). At least one of the criteria of

[7] In article 7 of the Wbp

article 7 of the Data Protection Directive should be satisfied, but the purpose must be in accordance with all laws available in a country, in the broadest sense. It is stated that determining legitimacy even goes beyond law and, for example, also ethics and customs should be taken into account (Article 29 Working Party, 2013).

The second part article 6.1.b can be divided in, is: "*and not further processed in a way incompatible with those purposes"* (EU Directive 95/46/EC, 1995, p.40). This text is literally translated in the Dutch Wbp[8]. A new purpose is not necessarily incompatible with the original one. To what extent a new purpose is compatible can be determined by a compatibility test (Article 29 Working Party, 2013). There are two kinds of assessments to test compatibility: "*A formal assessment will compare the purposes that were initially provided, usually in writing, by the data controller with any further uses to find out whether these uses were covered (explicitly or implicitly). A substantive assessment will go beyond formal statements to identify both the new and the original purpose, taking into account the way they are (or should be) understood, depending on the context and other factors*" (Article 29 Working Party, 2013, p.21). The Working Party advises the substantive assessment, since this method is more flexible. There are criteria listed that help to define whether or not a new purpose is compatible:

"a) *the relationship between the purposes for which the data have been collected and the purposes of further processing"* (Article 29 Working Party, 2013, p.23).
"b) *the context in which the data have been collected and the reasonable expectations of the data subjects as to their further use"* (Article 29 Working Party, 2013, p.24).
"c) *the nature of the data and the impact of the further processing on the data subjects"* (Article 29 Working Party, 2013, p.25).
"d) *the safeguards applied by the controller to ensure fair processing and to prevent any undue impact on the data subjects"* (Article 29 Working Party, 2013, p.26).

These criteria to determine compatibility are all implemented in the Dutch Wbp[9]. It could thus be stated that the Dutch law supports the flexible approach for determining compatibility.

Importantly, member states are allowed to narrow the scope of the rules set for processing personal data in article 6 of the Data Protection Directive (Article 29 Working Party, 2013 ; EU Directive 95/46/EC, 1995). In the Dutch Wbp there only is an exception on the further processing of personal data[10] (second part of article 6.1.b DPD). It is stated that further processing is allowed when it is in the interest of national security, prevention, detection and prosecution of crime, significant national or public economic/financial interests, protection of the data subject and the rights and freedoms of others[11]. Especially the allowance of further, incompatible, processing for prevention, detection and prosecution of crime is interesting for this research, because of its focus on public safety programs. Further processing of data "*in a way incompatible with those purposes"* (EU Directive 95/46/EC, 1995, p.40) is thus in the Netherlands only allowed in case of significant public interests[12].

The demands for purpose limitation set by the Data Protection Directive, are well summarized in the following citation of the Working Party 29 opinion on purpose limitation: "*In practice, it is not sufficient for such a law to only mention the final objectives of the legislative measure and designate the controller of the processing. It should, at least, also specifically describe the objectives of the relevant data processing, the categories of personal data to be processed, the specific purposes and means of processing, the categories of persons authorised to process the data, the procedure to be followed for the processing, and the safeguards against any arbitrary interference by public*

[8] In article 9.1 of the Wbp
[9] In article 9.2 of the Wbp
[10] For article 9.1 of the Wbp
[11] In article 43 of the Wbp
[12] Listed in article 43 of the Wbp

authorities" (Article 29 Working Party, 2013, p.38).

Two main lessons can be learned from this paragraph. First, as long as the purpose is described according to the rules explained in this paragraph, there should be no problem with purpose limitation. As long as all article 6 rules are taken into account, of which purpose limitation is an important one, personal data can be processed lawfully.

Second, although it is possible to process personal data, the smart city context might lead to problems in purpose limitation, due to the fact that it is a vague concept. If the smart city purpose is not further specified, made explicit and legitimate according to the rules for processing personal data, personal data cannot be lawfully processed in smart city programs.

5.3.2 Conclusion theoretical context

It became evident that personal data can be processed lawfully if all rules listed in article 6 of the Data Protection Directive are taken into account. For smart cities, difficulties are expected to arise in purpose limitation, one of the demands of article 6. Since the smart city is a vague concept, it might be difficult to sufficiently specify purposes, ensure that they are explicit and legitimate, and personal data is not further processed in a way incompatible with the original purpose. Issues in data processing and (European) privacy protection for smart city purposes are expected to arise in two ways.

First, the use of big data and open data can conflict by nature and definition with privacy protection regulation (see chapter 5.2, p.29). Big data, open data and geo-data are in many cases also personal data and therefore need to be processed according to the rules set in the Data Protection Directive. However, processing personal data according to the rules is difficult, and for open data this seems to be impossible by its definition.

Second, the context for this research, smart city programs, raises issues with privacy in the field of purpose limitation (see paragraph 5.3.1, p.31). It is purpose limitation that links all three major subjects of this research: data, privacy protection and the smart city. Thereby, the relevance of this research becomes evident: there is little information on these issues yet, while in today's society, smart city programs become increasingly popular, and developments in (open and big) data occur at a rapid pace.

To conclude, it can be stated that as long as the article 6 of the Data Protection Directive rules are met, personal data can be used for smart city purposes. Out of all article 6 rules, purpose limitation is expected to be the most difficult rule to satisfy, because of the fact that the smart city is a vague concept. The demands for purpose limitation are clarified in an Article 29 Working Party opinion (Article 29 Working Party, 2013). Purpose limitation should be taken care of in formal documentation in which, among others, processing personal data for a well-specified, explicit and legitimate purpose is justified and personal data protection measures are taken. Smart cities thus need to work on purpose limitation in formal documentation in order to process personal data lawfully.

It will be most interesting to research how these issues in processing personal data, and especially in purpose limitation, are being experienced in practice.

6. From theory to practice

In order to bridge theory and practice, this chapter describes the way the outcomes of the theory can be studied for the situation in practice. The choice for a qualitative methodology for the practical part of this research was explained in the introduction (see paragraph 1.5, p.12).

6.1 Operationalization

In theory, two major problems with privacy regulation were identified in working with personal data in smart city programs (see paragraph 5.3.2, p.33).

First, both personal big data and personal open data conflict with data protection regulation. For open data, problems mainly arise in purpose limitation and in an undefined amount of time of storing data. In processing personal big data, problems arise in the fact that data is easily being used for other purposes than the original one, which is also part of purpose limitation. For big data too, there is often an undefined amount of time of data being kept for processing. Informing data subjects about personal data collection is another difficulty. So, for both personal big and open data, purpose limitation is an issue, and difficulties also arise in other demands of the Data Protection Directive.

Second, the smart city context in which personal data is being used also raises issues in purpose limitation. The smart city is a concept that is difficult to define and is, if purposes are not further defined, too vague to lawfully process personal data.

These two issues arising from literature need to be addressed in the interviews, to research to what extent they arise in practice. In the topic list, example questions are listed to ensure that these issues are covered (see appendix B, p.77). So, in this way, theory and practice are bridged.

Since the cases involved in this research are diverse, it is important that each case is explored well. The identification of the programs objectives is of importance, because of the problems that are expected to arise in purpose limitation. These problems will only occur if personal data is being used, so this is an important subject to address in the interviews. The use of open, big and geo-data needs to be researched in all interviews as well. Furthermore, questions about the experience of issues in the field of privacy protection in general need to be asked.

6.2 Interview techniques

All semi-structured interviews are held in Dutch, since this benefits the quality of the interviews since it is in all cases the native language of both the interviewer and the interviewee. The interviews themselves are voice recorded. This way, the interviewer can pay full attention to the conversation itself during the interview, which benefits the quality of the interviews. Thereafter, the voice recordings are transcribed to structured summaries of the interview and they are reviewed by the interviewee for approval. The information that is gained from policy documents is added to these summaries (see appendices C-F, pp.79-95).

For each subject in the interviews, at first very general and open questions are asked, to avoid steering answers. Thereafter, more focussed questions are asked. In general, the interviews have the character of a conversation, whereby it is tried to let the interviewee speak freely and unbiased and at the same time address all topics.

6.3 Case selection

The selection of cases is based on several characteristics. The most important one is that the case can be considered a smart city program: programs taking place in the urban environment aiming to contribute to innovation in this urban environment by making use of ICT (see paragraph 2.2.2, p.17). Data and ICT solutions are being used to achieve a certain goal. Thereby, public safety must be

among the goals or sub goals. The programs must take place in Dutch cities. Cities with a large population (above 100,000 inhabitants) are selected, since they are likely to have comparable smart city programs. Furthermore, in these cities only municipalities working on smart city programs and policy are selected. Companies are also involved in smart cities, but mostly in the field of selling smart products. Governmental organisations are concerned with the general wellbeing of society and have an important role in public safety, together with the police and justice. Since policy and smart city programs are the focus of this research, municipalities of Dutch cities are the most logical organisations to be selected. Furthermore, the cases can be easier compared when only municipal policy is taken into account. Last but not least, cases in which privacy issues are expected to occur are selected. This is difficult to determine beforehand, but if, for example, only sensors are being used to measure air quality, this data does not relate to people (and public safety, in this case) and privacy issues are not expected to occur.

It turned out that it is not easy to find smart city programs with these characteristics. The case of Eindhoven is well-known and could be found easily by searching the internet. Almost all cases are found by reviewing the programs of smart city symposia and conferences. Also documentation on privacy issues by Geonovum (Van Oortmarssen & De Vries, 2014) was helpful in finding respondents. Furthermore, interviewees and people in this field of study helped to find interesting cases. A snowball method was thus used to select cases and respondents. The four selected cases are Eindhoven, Almere, Den Haag and Zwolle. The respondents are project leaders of smart city programs, since it is expected that they have knowledge on different aspects of the smart city program.

7. Results

In this chapter, the outcomes of the interviews and policy documentation study will be discussed for each case individually. In the next chapter, these results are combined in an analysis, in order to provide answers on the research questions. For each case, the main elements found in theory are discussed. It is important to take into account that there are differences between the cases, even though they are all smart city programs focusing (among others) on public safety, and the result is that not all topics apply to each case. In appendices C-F (pp.79-95), the extended version of the case reports can be read.

7.1 Eindhoven

7.1.1 Introduction of the case

Stratumseind is a street full of bars and pubs in the centre of Eindhoven. A couple of years ago, the street faced economic and social challenges. It was estimated that bars would need to close in the upcoming years and there were social problems. Therefore, Stratumseind 2.0 was established: an overarching project towards a new Stratumseind. In this program, multiple smaller projects will take place. Recently the smart light project started, in which different light colours and intensities are being used to influence the ambiance in the street. In the next months, also projects on smell, gaming and changing the scene of the street will start.

All projects are monitored in the Living Lab, located on Stratumseind, which can be seen as a thermometer that measures the effects. It is "*an 'instrument' to measure influences of interactions (light, smell, design). Smart sensors, smart interfaces, smart actors, smart lights, smart data, smart design, gaming"* (Kanters, 2015, p.13).

It is important to keep in mind that all findings are about the Stratumseind 2.0 program, not about the municipality in general.

7.1.2 Objectives

The overarching objective of Stratumseind 2.0 is: "*Together with all partners, as entrepreneurs, breweries, property owners, police, City Council, we will structurally improve and increase the economic and social functioning and activities on the Stratumseind. This structural improvement will have three main-themes: Safety, Liveability and Attractivity"* (Kanters, 2015, p.12). Safety is thus a major objective in Stratumseind 2.0.

For the smart light project, the goals are to turn Eindhoven into a vibrant city and a sustainable city. These objectives are not further specified, and would thus not be enough to enable lawfully processing personal data.

7.1.3 Data

All data is being monitored in the Living Lab. There is a clear overview available of the datasets that are being used. Since safety is among the three main objectives, all data used to monitor the street will be viewed as data being processed for this purpose. The label 'other' is given if data cannot be included in at least one of the categories of open data, big data or geo-data. The following data is being used in Stratumseind 2.0:

1. Temperature, sun, rain and wind. This is measured with sensors, so it can be classified as big data, as well as geo-data.
2. Sounds (3D). This is measured with sensors, and can be classified as big data, as well as geo-data.
3. Bluetooth and Mac addresses. This is classified as 'other' data.
4. Cell phone counting. This is geo-data as well as big data.
5. Video people counting. This is classified as big data and geo-data.
6. Light amount (Lux) and colour (Kelvin). This is classified as 'other' data.
7. Events calendar of Stratumseind. This is classified as 'other' data.
8. Police recordings. This is classified as 'other' data.

9. Social sensors (media watching and interactive). This is classified as big data.
10. Brewery registrations. This is classified as 'other' data.
11. Waste/energy. This is classified as 'other' data.
12. Open data from the municipality. This is, off course, classified as open data.
13. Survey results of residents. This is classified as 'other' data.
14. Car and bike parking information. This can be considered geo-data.

At this moment in time, open data is not being used in Stratumseind 2.0 and data from the program is not opened up either. The municipality of Eindhoven in general is working on making data openly available for citizens, but this is not part of Stratumseind 2.0. Currently, a portal is being made. Data that would be relevant for the Stratumseind 2.0 program would, for example, be data on the amount of bicycles parked near Stratumseind. It is expected that these data are included in the system by the end of February.

On all five entrances/exits of Stratumseind there are lampposts with sensors. All sensor data is big data and can also be considered geo-data in this case. All data is combined in the Living Lab. There is an interface available, in which all data is shown in easily readable way (see figure 8).

The most important geo-data being used is Cell phone counting; this is Vodafone mobile location data that is used to determine from which municipality the visitors of Stratumseind originate from.

There are possibilities for a more accurate determination of locations. The newest technology in this field is small cells. There are small Wi-Fi connectors on the lampposts and if someone connects with them, the connection will be good, and the accuracy of location determination is about 25 meters. Another method is iBeacons for iPhone. This is an application that sends text messages about the direct environment. This application can be used in Eindhoven, since its marketing organization Eindhoven 365 developed it.

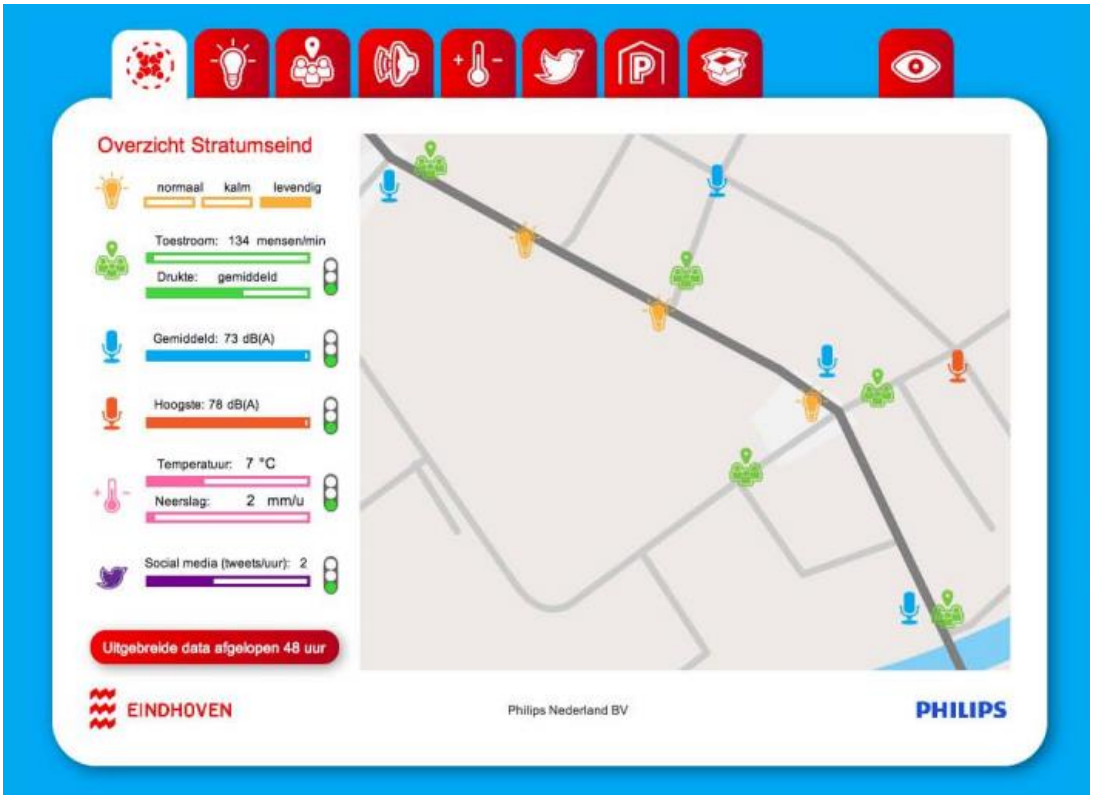


Figure 8: The Stratumseind Living Lab Interface.
Source: Kanters, 2013, p.16.

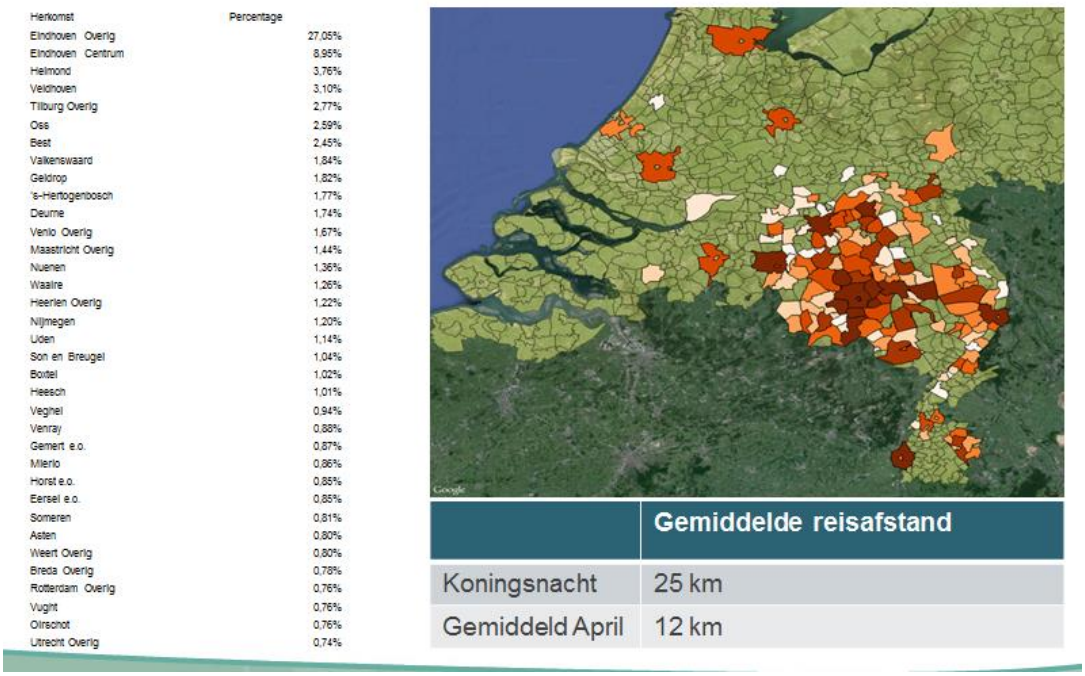
7.1.4 Data and privacy

In Stratumseind 2.0 it is tried not to use personal data. This way, there will be no problems in privacy. This does not mean that there are no privacy issues that need to be addressed. Issues arise in determining to what extent data should be considered personal or non-personal data. The boundary is in some cases fuzzy. Sometimes, the boundaries of personal data are being reached and perhaps crossed. In the past there were issues with sounds (3D) and video people counting.

Current techniques enable to detect not only the amount of sound, but also recording the sound itself. If this is recorded, sound data becomes personal data and privacy problems would occur. In the past, sound data of what is being said was saved, but this is no longer being done. It was useful to detect the origins of sound, because then the cause of differences in the amount of sound can be detected. Nowadays, only the amount of sound and the direction where noise comes from are registered.

In video people counting, cameras used to save a blurred image of what is going on in the street. However, in the blurred image it was possible to recognize people. Therefore the camera images are no longer used, and nowadays only the number of people is registered. On the cameras there is software installed that is able to detect people and to count them. Furthermore, their direction can be registered.

Currently, there is also data being used that balances on the boundary of personal and non-personal data. This is the case for cell phone counting (see figure 9). Vodafone mobile location data is being used. The data is already anonymised when it arrives in Eindhoven, but it is stated that people often find this data type a bit scary in terms of privacy. Vodafone made an agreement with the CBP on the aggregation of numbers. Anonymity is also protected by the fact that place of origin is only registered when there are at least fifteen telephones that originate from the same municipality. Every unique telephone is given a code and it is determined where this telephone was most of the time in the past weeks. This location is being seen as the place of origin. So, no addresses or conversation data are being used. To determine the location, the location of cell towers is being used and this system has an accuracy of about 200 meters. The data is not real-time; it takes two or three days before it arrives in Eindhoven. Because anonymisation is well taken care of and the spatial accuracy is not very detailed (see EU Directive 2002/58/EC, 2002 ; Van Loenen, 2008), this mobile location data is not considered personal data.



Herkomst	Percentage
Eindhoven Overig	27,05%
Eindhoven Centrum	8,95%
Helmond	3,76%
Veldhoven	3,10%
Tilburg Overig	2,77%
Oss	2,59%
Best	2,45%
Valkenswaard	1,84%
Geldrop	1,82%
's-Hertogenbosch	1,77%
Deurne	1,74%
Venlo Overig	1,67%
Maastricht Overig	1,44%
Nuenen	1,36%
Waalre	1,26%
Heerlen Overig	1,22%
Nijmegen	1,20%
Uden	1,14%
Son en Breugel	1,04%
Boxtel	1,02%
Heesch	1,01%
Veghel	0,94%
Venray	0,88%
Gemert e.o.	0,87%
Mierlo	0,86%
Horst e.o.	0,85%
Eersel e.o.	0,85%
Someren	0,81%
Asten	0,80%
Weert Overig	0,80%
Breda Overig	0,78%
Rotterdam Overig	0,76%
Vught	0,76%
Oirschot	0,76%
Utrecht Overig	0,74%

	Gemiddelde reisafstand
Koningsnacht	25 km
Gemiddeld April	12 km

Figure 9: the origin of visitors per municipality in percentages.
Source: Kanters, 2015, p.18.

However, one dataset that is currently being used could be considered personal data: police recordings. The numbers for different categories of crime are registered. In some categories, there are small numbers involved, and these categories can be seen as personal data (see figure 10). Recital 26 of the Data Protection Directive indicates that anonymizing data is not enough and data that is likely to become de-anonymised by combining it with other data should still be considered personal data. The police data is available in a time span of two hours and is thus detailed. By combining it with for example news items, data could lead to individuals. However, Stratumseind 2.0 focuses on the categories with the highest numbers of recordings because the incidents of these categories are likely to be influenced by different light colours and intensities. The other categories could thus be left out, which would lead to this dataset becoming non-personal data.

Tabel 1: Alle uitgaansgebieden opgesplitst naar incident voor de periode 2009-(15-12)2011 tijdens uitgaansavonden (absolute aantallen)

kode	Klasse-omschrijving	2009	2010	2011 (tot 15-12-2011)	2011 t.o.v. 2010 (in %)
F11	Openlijke geweldpleging tegen goederen	3	2	2	0
F12	Openlijke geweldpleging tegen personen	56	47	24	-49
F16	Lokaalvredebreuk	3	0	0	0
F17	Wederspannigheid (verzet)	16	40	36	-10
F18	Niet voldoen aan bevel/vordering	63	56	54	-4
F40	Bezit hard-drugs (lijst 1)	34	19	21	11
F41	Bezit softdrugs (lijst 2)	4	3	3	0
F42	Handel e.d. hard-drugs (lijst 1)	7	1	2	100
F46	Aantreffen drugs (geen verdachte)	3	5	2	-60
F47	Overige drugsdelicten	4	0	0	0
F50	Discriminatie	1	0	1	100
F51	Belediging	116	72	142	97
F530	Bedreiging	22	16	40	150
F531	Overige misdrijven tegen de persoonlijke vrijheid	1	0	0	0
F550	Eenvoudige mishandeling	187	217	225	4
F551	Zware mishandeling	17	34	39	15
F552	Overige mishandeling	15	10	0	-100

Figure 10: Police recordings 2009-2011.
Source: Kanters, 2015, p.19.

The solution that is often being used to turn personal data into non-personal data, is not to save data. This is being done with sounds (3D), video people counting, Bluetooth and Mac addresses and social sensors.

Bluetooth and Mac addresses are processed, but immediately after the collection deleted, since they cannot be aggregated or made anonymous yet. Social media is also being monitored, but messages are not saved. It is only being registered if these messages were positive, negative or neutral. It is not possible yet to include semantics in the software, leading to many messages becoming classified as 'neutral'. If these data would be saved, it would become possible, for example, to find out how many times someone wrote something negative. On the other hand, it is stated that if someone posts a message on social media, it is this persons own choice to publish it. As soon as these texts would be combined with information that this person did not provide, such as an address, problems in privacy would occur. This debate often is more ethical than juridical.

In general, Eindhoven tries to stay away from using personal data, but it is difficult to decide on the border between personal and non-personal data. There are multiple organisations involved in the field of privacy, such as the CBP, TILT, Tilburg University, Ministry of Internal Affairs and the Ministry of Justice (Kanters, 2015). Privacy is a difficult subject to deal with, since it cannot be directly read from the law. Since Stratumseind 2.0 does things that have not been done before in the Netherlands, help is needed. According to Eindhoven, the law does not cover all things being done in Stratumseind 2.0. Technological developments take place at rapid pace and laws are not able of keeping up. Eindhoven did not receive official warnings from the CBP on using personal data. The problem is that even for these people it is hard to decide on how to deal with privacy. If someone

start collecting names, addresses and pictures it is obvious that it is not allowed to do that. The problem is that there is a large area in which it is unclear if something is allowed. It would be great to know for sure if something is allowed, especially in the field of combining datasets.

7.1.5 The influence of privacy on the smart city program

Besides these difficulties, there is data that would be valuable for the Stratumseind 2.0 program, but is not included because of privacy legislation. The amount of people is being counted, but it not known where they are in the street. It would be valuable for the smart city program to gain more information on the location of people within the street, to see if there are parts of Stratumseind that get crowded. It is states that this could be useful information when combined with for example, sound data, police recordings and the colour of the lights that were turned on.

Furthermore, there is data that is being seen as relevant for the smart city program, but due to the fact that this data cannot be anonymised or aggregated, it is not being saved. This is another way of privacy having an influence on the smart city program, since, apparently these data are considered relevant for the program. Another example is provided by sound data. By saving not only the amount of sound, but also the sounds itself, this data can be used to determine the source of sounds in order to explain the causes. Sounds are not always caused by humans, but for example also by church bells or cleaning machines. Saving the data would lead to the fact that the data becomes personal data and this is unwanted in Eindhoven, while this data can be very useful for explaining sound levels.

However, according to the respondent, privacy protection does not stand in the way of the project, since the focus is on the mass and not on individuals. The ambiance on Stratumseind needs to be improved and this goal focusses on the collective. The police are subject to different rules for privacy which allow them to focus on the individual.

7.1.6 Suggestions for solving the problems

Changing the law is not being seen as an option on short term, because this would take too long. Since the start of the Living Lab, a lot has changed in this period of time. If the law was changed, it would soon be outdated again. In order to keep data non-personal, it is important to take care of data storage and ensure that it remains impossible to de-anonymize data by combining datasets. But if someone wants to misuse data, this will always be possible somehow. Developments are occurring at such a rapid pace that it is difficult to protect data.

7.1.7 Conclusion

In the Stratumseind 2.0 smart city program in Eindhoven, personal data processing is avoided as much as possible. However, the data on police recordings could be seen as personal data. In this case, the rules in article 6 of the Data Protection Directive, including purpose limitation, should be taken into account in order to process these data lawfully. Currently, this is not being done, which can easily be explained by the fact that Eindhoven tries not to use personal data. However, since it is stated that the program focusses on the categories with large amounts of incidents, personal data could also be avoided in this case. Eindhoven proves that it is possible to avoid using personal data. On the other hand, avoiding personal data leads to not using data that is relevant for the program. Examples are data on the location of people in the street and data on saving sounds. The way privacy issues currently are addressed is by asking help from experts and if data is considered personal data it is not being saved. In the smart city program, mostly big data is being used (sensor data). The sensor data is also geo-data and geo-data is also important in this smart city program, especially mobile phone location data (cell phone counting). Open data is not being used or produced. The main privacy difficulties arise in determining the border between personal and non-personal data. This cannot directly be read in the law. Technological developments take place at rapid pace and laws are not able of keeping up.

7.2 Almere

7.2.1 Introduction of the case

In Almere, a new information management system was developed: the 'StraatKubus'. It is a geographic information system (GIS) that serves as an early warning tool in signalling liveability issues at an early stage (RRAAM, 2014). The system is useful for monitoring the city and preventing problems in liveability. The StraatKubus is a tool that enables to link data to developments in the city and thereby monitor even small changes. The tool can be used on a daily basis to answer questions and is useful in conversations and discussions with colleagues about developments in the city. It is important to keep in mind that all findings are about the StraatKubus, not about the municipality in general.

7.2.2 Objectives

Monitoring the liveability and ultimately preventing liveability issues in Almere is the main goal of the StraatKubus. It serves as a tool to use in conversations, whereby it can be used to check presumptions with facts. Liveability is a large concept that contains a number of sub-concepts. Safety is being seen as an important part of liveability, but others might see safety as a separate concept next to liveability. Safety is also a large concept, containing different elements.

In the Gedragsrichtlijn StraatKubus, the objective of the StraatKubus is defined. "*The goal of the StraatKubus is to provide insight in the development of liveability. In order to do so, the StraatKubus combines data from the physical, social and safety domain and presents these data on 6ppc-level. By keeping an eye on trends and developments that can put liveability under pressure and by having conversations about it, local partners want to detect problems in an early stage. The StraatKubus is a tool in detecting liveability problems in an early stage by local partners. It is a communication tool that helps local partners to have conversations on trends and developments in areas. What gets attention and how should be worked together to address the problems identified?*" (translated from Dutch to English) (Gemeente Almere, 2014, p.8).

7.2.3 Data

In Article 13 of the Gedragsrichtlijn StraatKubus it is stated that data can be included in the StraatKubus if they concern one of six (main) themes: "*1) population and housing characteristics 2) socio-economic situation and poverty (prevention) 3) Safety and bonding 4) social support and welfare 5) environmental incidents/reports of nuisance in public space 6) developments in housing values"* (Gemeente Almere, 2014, p.10). In articles 14, 15 and 16, subthemes and data are further specified. Subthemes relating to safety are: reports of nuisance in subsidised housing complexes, reports of nuisance in public space, burglaries and neighbourhood mediation/conflicts between neighbours. Thematic spatial data that can be included in the StraatKubus, relating to safety, are social cohesion and intercultural intercourse, reports of public space, reports of nuisance in subsidised housing complexes and reports on physical nuisance (dirt dumping, graffiti, vandalism, odours, dust, noise, parking spots, trees) (Gemeente Almere, 2014).

Big data is not included in the StraatKubus. The privacy risks concerned with big data are highlighted. Data becomes interesting if it is possible to research patterns in a protected way and to find out what data is relevant in a certain pattern and what is not relevant. Thereafter, about 90.0% of data can be removed. The risk of big data is being seen in the possibility that it might get in the wrong hands, while the majority of data is not needed.

Open data is not being produced with the StraatKubus, also because of privacy. It would be great for citizens to access data, but this is impossible because personal data is included. Almeres opinion is that with the StraatKubus, a greater purpose can be served than by opening up generalised data. Furthermore, it is believed that people tend to speak too easily about opening up data and it is important to keep in mind how open data actually is allowed to be. On the other hand, the StraatKubus does make use of open data. In the near future, newly opened up data from Kadaster will be included. At this moment in time there is open data from the Centraal Bureau voor de

Statistiek (CBS, Statistics Netherlands) in the StraatKubus. Open data are thus integrated in the StraatKubus, but the StraatKubus does not publish open data because of privacy issues.

All data in the StraatKubus is geo-data, since the StraatKubus is a geographic information system. The data are available on 6ppc-level (6 digit zip-code level, also called street level). This level provides detailed information and this was an important choice made by the municipality. Data on neighbourhood level is too little detailed: if something is going wrong in a small part of the neighbourhood and if the rest of the neighbourhood is doing fine, this cannot be seen. The 6ppc level enables field workers to do their work focused. This amount of detail has the consequence that data leads too easily to individuals. Even though the StraatKubus data is personal data, a threshold of five is used. This is an extra security guarantee.

7.2.4 Data and privacy

By now it has become clear that both goals and the data included in the StraatKubus are clearly defined. Almere is aware of the fact that the broad concepts of liveability and safety, combined with processing personal data, lead to problems in purpose limitation. To address these issues, the Gedragsrichtlijn StraatKubus is established. In the Gedragsrichtlijn, all concepts are described individually and the value of their connections are described as well. The objective of the StraatKubus is defined, and it is stated that everything thereafter needs to relate to this objective, such as authorisation and publication. As soon as there is doubt, it is not allowed to use the StraatKubus.

In the Gedragsrichtlijn, the purpose is specified, made explicit and legitimate. All elements of the StraatKubus are defined and there is a limitative enumeration of data that can be included in the StraatKubus. Furthermore, safeguards are well taken care of, both in the field of authorisation and cybercrime. The aim of the StraatKubus is easily understandable and unambiguous and explains the purpose for which the StraatKubus is established. Also one of the article 7 criteria in the Data Protection Directive is met and further processing for different purposes is taken care of. For example, using images for presentations is only allowed when the rules on this topic in the Gedragsrichtlijn are satisfied (Gemeente Almere, 2014) (for the analysis, see paragraph 8.2, p.51). It can be tempting to upload data in the StraatKubus that is gathered for a different purpose. In many situations, it is not allowed to use data for another purpose than the original one. If more data is needed, it is important to ask the data subject for permission. The Gedragsrichtlijn is part of a contract that needs to be signed by the StraatKubus users. If data are uploaded into the StraatKubus, all Wbp demands are taken care of.

According to Almere, coping with purpose limitation mainly is a matter of starting to write down what the aims and objectives are and why personal data is needed: start peeling down the main objective. This was a difficult process and establishing the Gedragsrichtlijn took a long time. It is not a fun job to do, but it is being seen as very important that this is well taken care of.

Privacy protection and especially purpose limitation needs to be taken care of since the StraatKubus contains spatially detailed data on 6ppc-level. At this level, information leads for too many people to individuals and it becomes personal data. However, this amount of detail is needed for professionals to do their job well and focused. This is explained in the Gedragsrichtlijn. The data in the StraatKubus belong to privacy category 2. Category 1 should be seen as data that does not lead to privacy issues. Category 3 data is very sensitive data and is not included in the StraatKubus.

For Almere, it turned out to be difficult to find a correct way to enable processing personal data lawfully. Boundaries, criteria and a way of reasoning were unclear. Help was asked, but even professionals could not come up with a workable answer. Eventually, the CBP advised Almere to start writing down the purpose for processing personal data and this was the beginning of the Gedragsrichtlijn.

A lot of effort was needed to find out what the municipality needed to do. Almere felt discouraged in this process, while all the municipality wanted is to process data in accordance with the law. The municipality simply wanted to work on liveability issues and suddenly needed to cope

with unclear regulations, while little help was offered. Two separate worlds were coming together and answers were missing. People often shy and overlook why Almere wants to work with the StraatKubus: to work on creating a better Almere, not to match data that is not needed or irrelevant for the StraatKubus' objectives. The CBP never warned Almere officially, but they did ask questions. The cause of the fact that many people could not help solving the privacy issue with the StraatKubus, lies, according to Almere, in the simple fact that people do not always know what they are talking about. Specialists often say every case needs a different approach. Furthermore, people find it difficult to think about this issue from a domain that is not their own.

7.2.5 The influence of privacy on the smart city program

Privacy does not influence achieving the objectives of the StraatKubus anymore, because it is well taken care of in the Gedragsrichtlijn. However, there are important limitations that the Gedragsrichtlijn imposes. The main one is that only professionals are allowed to work with the StraatKubus. An important question of the Wbp is: why does information need to be combined and who needs this information in order to do his or her job well? Therefore, the municipality had to decide who should get access to the StraatKubus data. It was decided that professionals of the municipality of Almere and housing corporations should be allowed access if they need the data in order to do their job well. This was a difficult and long debate. It is stated that it would be great to share information with more people, because of the function of the municipality in society. This way, data could be used to start conversations with, for example, inhabitants. Lawyers decided that this is simply not allowed and only municipal professionals and housing corporations can work with the data. Access is denied to all other people. So, the StraatKubus is not only unavailable for citizens, but also for professionals and private parties. Furthermore, in order to add new data to the StraatKubus, this can only be done when the Gedragsrichtlijn is revisited. Flexibility can thus be seen as another limitation of the solution found in Almere.

7.2.6 Suggestions for solving the problems

According to Almere, the law itself could be improved, but that is not the most important thing to do. The most important thing is to keep in mind why the Wbp was established. Purpose limitation should be taken care of if personal data is being processed. It is needed to ask yourself a number of questions and a law will not be capable of doing this. It mostly is a matter of starting to work on writing down the purposes and justification for personal data processing. The law could use some updating. It could be more focussed on the network society we live in nowadays and it could provide a checklist, for example, to enable people to adapt the law easier. In that way, the long and difficult process that Almere experienced in finding answers can be addressed.

In general, according to Almere, privacy protection should not be seen as an obstacle. It is about taking good care of personal data and about justifying personal data use. It is no more than decent to do so. There is enough space to do great things with data, but some effort needs to be put into it.

7.2.7 Conclusion

In the StraatKubus, all data is personal data. This is a choice Almere made because spatially detailed data is needed in order to achieve the objective on monitoring liveability. Since the data is available at 6ppc-level, data relates too easily to individuals and thereby becomes personal data. The StraatKubus is a geographical information system. Geo-data plays an important role in this case. Big data is not being used, and open data is only being used, not produced. This is due to privacy protection. The Gedragsrichtlijn StraatKubus was established in order to take care of purpose limitation and all other rules of article 6 of the Data Protection Directive. This is a good solution to enable lawfully processing personal data. On the downside, this also leads to a very limited group of people with access to the StraatKubus and little flexibility. It was a difficult process to establish the Gedragsrichtlijn and help and clarity were missing. In the future this needs to be improved. The most important lesson to learn from Almere is that it is possible to process personal data lawfully, as long as the rules of the Data Protection Directive are taken into account. The Gedragsrichtlijn is thus an

example of a way to enable processing personal data according to the rules set in the Data Protection Directive.

7.3 The Hague

7.3.1 Introduction of the case

The Hague recently published the Smart City Road Map, in which the smart city goals for 2014-2018 are defined (Gemeente Den Haag, 2014). In the Road Map, working together with public and private institutions is important. This program used to be called 'ICT voor de Stad'. With the Smart City Road Map, the ICT voor de Stad program did not become more 'smart city' than it already was. Safety is one of the two priority themes of the Smart City Road Map. Opening up data and sharing data in events are central in The Hagues smart city program in the field of safety.

7.3.2 Objectives

For the period of 2014-2018, the goal of Smart City The Hague is "*to boost the competitiveness, the quality of life and the sustainability of The Hague, by working together with representatives of Triple Helix by working with pilots and projects, making innovative use of technology*" (translated from Dutch to English) (Gemeente Den Haag, 2014, p.2).

In the Smart City Road Map there are eight themes identified for The Hague Smart City. Two of them are marked as priority themes: Quality of Life and Safety. For 2018, two main goals are identified in the field of safety: "*The Hague is leading on ICT use and open data use for safety (worldwide) and The Hague is the Silicon Valley of safety*" (translated from Dutch to English) (Gemeente Den Haag, 2014, p.2). ICT and open data are thus being seen as important for safety. Therefore, The Hague Smart City contributes to projects working on these aspects. The safety cluster of The Hague is taken care of in the Hague Security Delta (HSD). There are four safety programs identified: event safety, protection of vital infrastructure, cyber security and protection of the international zone. For multiple parties there is an interest in events and their safety, such as the municipality, police and event organization, but also for bars and restaurants.

In the field of ICT, it is stated that it is important to share data cross-sectorial. The Smart City Road Map tries to achieve this by working on opening up and linking open data (cross-sectorial), by enabling digital communication in the city and by working on the infrastructure of knowledge, in order to make the most out of facilities and services (Gemeente Den Haag, 2014).

7.3.3 Data

In the field of safety there is a focus on event safety. In events it is necessary to share data between partners. The police are an important partner. They often like to gain data from the municipality and event organisation, but do not share much data themselves. This is due to privacy. Data is combined in an event cloud, for a limited amount of time. These data are only opened up for the event partners. Crowd control is an important feature in event management. On this topic, much information is shared between the partners. There is, for example, data on the use of mobile phones and the amount of people. Mobile phone data is made available by providers to the event organisation. It is, for example, known when squares need to be closed and people are informed about routes they can take. It is stated that in these situations, privacy is not a big issue. Mobile phone location data can be seen as big and geo-data and is used real-time in The Hague.

The data in event clouds is only shared for a limited amount of time, but currently people are working on a network to share data structurally. After an event is over, data is analysed and thereafter the data is being seen as not interesting any more. Data might be kept to get back to it if it turns out to be interesting after all, but the main problem is that the system for sharing is removed after an event. For the organisation of events, maps are mostly used. In the smart city not much data is shared between partners yet, both private and public organisations, with the exception of maps.

Most data on safety originates from the police. But there is also data available that is gathered by city guards and the neighbourhood watch. Police data are not publicly available. In The Hague there is a focus on event safety, but there is also more general data on safety available. There is data openly available on safety, on www.hoeveiligismijnwijk.nl (an open platform of the police in which crime rates are made available) and on the municipal open data portal 'Den Haag in Cijfers' (Gemeente Den Haag, 2015 ; Politie Den Haag, 2014). Liveability and safety is one of the data categories available in the portal.

Open data sub categories that relate to safety are crime and nuisance, opinion on neighbourhood (safety monitor), common nuisance (safety monitor), sense of security (safety monitor) and social cohesion (safety monitor). In this open database, the majority of data is (also available as) geo-data. However, data is in most cases only available on the spatial level of The Hague (so not on neighbourhood levels, for example) and displays the overall number or percentage for the whole city. The data is thus not detailed (Gemeente Den Haag, 2015).

Data is also published on Columby, a platform for distribution of open data (Columby, 2015). Eventually 8,000 data sets will be published on this platform, but until now only a few are made available. Compared to the past years, data is now opened up at rapid pace. Standardisation is an important issue that needs to be taken care of in order to make use of data.

7.3.4 Data and privacy

In the field of privacy, The Hague mainly focuses on opening up data. For opening up data, a method for a Privacy Impact Assessment (PIA) was developed: a decision tree. The decision tree is being used to determine if it is possible to open up data (Penninga & Siebert-Han, 2014). Lawyers are involved in this process. In general, lawyers tend to say that it is better not to open up data as soon as there is doubt about it. Opening up is an interesting process since this is a new development and it is yet unknown how society will react on opening up data. It is stated that it will be interesting to see if there will be lawsuits, and how judges will decide in them. Opening up data also has to do with the courage of people responsible, according to The Hague. Some will decide to just go for it and others need to know for sure if opening up is allowed. The municipality recently started working on the demands of the Wbp. It is expected that in many cases it will lead to not opening up data, due to privacy protection. It is stated that more data could lead to issues than would be expected at first sight. There is also data that is not privacy sensitive at all, such as the location of public toilets and trees. For municipalities, it can be scary to open up data, because citizens can easily check what is being done in the municipality. An example is maintenance data on playgrounds.

For every dataset that might be opened up, the decision tree is used, a model to decide if opening up data is safe. A lawyer helps to decide in this process. The ministries of Economic Affairs and of the Interior and Kingdom Relations are contacted about how to deal with this issue, but it turned out that these ministries had also just started working on it. As a municipality, The Hague finds it difficult to find out for itself what needs to be done. It is important to take care of anonymisation, because personal data cannot become publicly available.

In The Hague, opening up data and sharing data between event partners thus leads to privacy issues, but in the portal there is no personal data opened up. The Hague is aware of the fact that many data is or can become personal data, however, which data is personal data is not specified. The data that is opened up is generalized and highly anonymised data with little (spatial) detail. The goals of The Hague are not specified enough in order to process personal data lawfully.

The police own most data on safety but do not want to share. According to The Hague, this indicates that privacy is an issue that matters. For the establishment of a system in which safety data is shared among event partners, the national police would be an important partner. However, because of privacy legislation they do not share much data. For safety it would be interesting to know, for example, which people could be a danger for society. The police often know, but because of privacy legislation they are not able to do something with their information. It is stated that in events, sometimes it would be great to share more information and forget about privacy for a moment, and

to agree with the partners on removing all data immediately after an event is over. Especially if things go wrong in an event, privacy would be something you would like to forget about for a moment, according to The Hague.

7.3.5 The influence of privacy on the smart city program

Privacy issues in The Hague can be found in opening up data and sharing data between event partners. Since opening up data and sharing data in events are important parts of the smart city and safety objectives of The Hague, it could thus be stated that privacy influences achieving the smart city goals of The Hague.

Privacy issues are mostly addressed in the field of opening up data, by the establishment of the decision tree and by involving lawyers in the decision process for opening up data. Before data is opened up, a PIA is done. No personal data is opened up, but on the other hand, the data in the open data portal is not detailed, and because of that, has little value.

In the field of sharing data between event partners, privacy mostly influences sharing data with event partners, among which is the police. The police are an important partner of the municipality in event safety. The fact that the police are not sharing (personal) data is in The Hague being seen as a limitation for the shared goal of a safe city. On sharing mobile phone data in events for crowd management, it is surprisingly stated that privacy is not an issue.

7.3.6 Suggestions for solving the problems

The Hague experiences issues in opening up and sharing data. In opening up, there were difficulties in discovering what needs to be done in order to open up data in a responsible way. Help was asked for by the national government, but the national government was also still in the beginning of this process. It is important to take care of privacy because personal data cannot become publicly available. Also, anonymising data should be taken good care of. Furthermore, more flexibility for sharing data among event partners is desirable. In The Hague this is needed in order to share data between event partners, so event safety can be better taken care of.

7.3.7 Conclusion

In The Hague opening up data and sharing data are important in public safety in the smart city program. The privacy risks in opening up data are evident for The Hague and personal data is not being opened up by making use of the decision tree and of lawyers. This results in open data that is well anonymised but also very general and less valuable. Privacy thus does influence opening up data, but not in the field of purpose limitation. Also in sharing data, privacy problems are experienced. It would be desirable to share more data between partners in an event, especially data from the police. Privacy influences achieving the smart city objectives, since opening up and sharing data are among the main objectives (in safety). In The Hague, besides open data, also geo-data and big data are processed, such as mobile phone location data. Personal data is being used, but which data this is, is not specified. Thereby purpose limitation and the other rules of article 6 of the Data Protection Directive should be taken care of, but the goals of The Hague are not specified enough in order to process personal data lawfully. The Hague does not seem to be well aware of privacy yet and sees privacy problems in its smart city program mainly in opening up and sharing data.

7.4 Zwolle

7.4.1 Introduction of the case

The focus in Zwolle is not on being a smart city, but on finding useful (smart) solutions. The approach of Zwolle is that a few smart solutions do not make the city smart yet. Internally, a data platform was established in which the possibilities of using data are discussed. This involves a wide range of topics such as applications, consequences, regulation, policy and communication. Since municipalities gained more freedom and also more (financial) responsibility from the national government, it became more important for municipalities to check the effects of their policies and to monitor

society. There is lots of data available, but the municipality is still working on creating the right culture, infrastructure and competencies to make use of data. This process is rapidly developing. Zwolle is being seen by other municipalities as leading in both geo-information as well as the policy side. In the municipality, geo-information, research, statistics and policy started working together, using each other's strengths. Initiatives and potential of the local community is being used; an important part of the smart city concept.

7.4.2 Objectives

Zwolle does not market itself as smart city and there are no official objectives formulated. Data is being used in the municipality for two main reasons. The first one is prevention. It is important to know what indicates a negative change in society and to take action. The second one is measuring the effectiveness of policy. In the new municipal role, it is important to support choices with facts. Therefore it is needed to invest in data. Monitoring and forecasting is thus important in Zwolle. Safety is an important aspect of monitoring society. In protecting safety there are roles for the municipality as well as for lawyers and police. The Gebiedsscan serves as the documentation that is used in deliberation between these three parties and is a report on crime in Zwolle. In this Gebiedsscan, going out safely, and safety in events and soccer is listed as one of the five priorities (Team Zwolle, 2014).

7.4.3 Data

In Zwolle, safety is being protected in a number of ways. In large events such as soccer matches or Liberation day, data is being gathered to control the mass. Sensors are being used in the city, but only to measure air quality and traffic movement. Sensors are thus not being used in public safety.

In the field of safety, monitoring data are being used. Most of these data originates from a survey among citizens, a neighbourhood survey and data from the police. Most of the data on safety originates from the police, and the Gebiedsscan is a report of police registrations.

Zwolle also has open data. There is a central (geo-)database and some data is publicly available. The municipality nowadays does not need to collect all data; much data is already available. This is increasingly becoming the role of the department. In the open data portal, there is little open data on safety available. Datasets 'Buurt voor Buurt 2014' (Neighbourhood monitor 2014) and 'Opgaven Zwolle' (Societal challenges) are the only open datasets containing safety data (Gemeente Zwolle, 2015). In general, the open data available is not privacy sensitive (data that do not relate to people in any way).

All open data is geo-data. It is stated that about 80.0% to 90.0% of all data is geo-data and the percentage of geo-data in Zwolle is expected to be similar to these estimates. In the municipality there is no distinction is made in geo-data and other types of data. Data is being seen as needed to do useful things, whether or not it is geo-data is not important.

In the municipality, big data (however not for monitoring safety), open data and geo-data are thus being used.

7.4.4 Data and privacy

In one of the neighbourhoods of Zwolle, Holtenbroek III, there were issues in safety. The municipal council decided to take extra measures. CCTV was being used, as well as preventive searching. The council decided on doing this for as long as specific monitoring data remained negative. The competency of supervisory authorities was thus linked to the values of data. This can only be done for excessive cases and there are strict conditions defined. Finding a balance between safety risks and privacy is important. If data are personal data, benefits in the field of safety should be evident. In the case of severe criminality or heavy nuisance, which was the case in Holtenbroek III, and is also in vandalism in soccer games, for example, extra measures can be taken.

In protecting safety, there are roles for the municipality as well as for lawyers and police. Thereby, the police have other rights concerning privacy. It is not easy for the municipality to use police data, because of a difference in focus. This is being explained by the fact that the police are focussing on the prosecution of individuals, while the municipality is interested in the preconditions

of a situation. As soon as it is tried to get more detailed information, the police are becoming more protective of their data. It is being seen as important to find the right 'Hygiëneregels' in the interaction between police and municipality.

The municipality owns a lot of personal data, for example data on social security payments and social assistance (social services). Even within the municipality, this can cause difficulties in using data for different purposes. It is thus needed to take care of a good information system within the municipality, without violating privacy unnecessarily.

Opportunities for linking data have developed at rapid pace. There is a debate going on in society about data use. It is often a matter of balancing on the border between what is allowed and appropriate and what is not. These borders are still unclear and they are the subject of a debate on ethics.

In Zwolle, privacy is becoming more and more of a complicating factor. The alderman therefore decided that the Hygiëneregels should be established. In the Hygiëneregels, privacy is taken care of. These rules are still being worked on and the exact content is not clear. Until they are finished, privacy in Zwolle mainly is a matter of using good judgement. Everyone working with personal data is aware of the importance of taking care of privacy, but to make sure that everyone does this in the same and correct way, the Hygiëneregels are established.

Privacy determines the boundaries of the playing field of data. But, according to the municipality of Zwolle, it is a dynamic concept that remains difficult to capture in general rules. Privacy is seen as one of the risks people need to cope with, just like with financial or image risks. It is stated that if a lawyer is approached to help, this often results in the advice not to do something, since than there will be no risks either. It is therefore being seen as important to gain experience in estimating privacy risks. The Hygiëneregels are only being seen as a starting point and cannot go further than that.

7.4.5 The influence of privacy on the smart city program

Since Zwolle has no specified smart city program, privacy cannot influence the objectives either. But, since Zwolle is working with personal data for municipal tasks, privacy is an important issue. Currently, using good judgement is the way privacy protection is being coped with and the Hygiëneregels are being established.

7.4.6 Suggestions for solving the problems

Zwolle has clear ideas on changes needed. For Zwolle, it would help to create a safe space in which it is possible to experience how to deal with privacy. In municipalities, information at the individual level is needed in order for people to do their work well, but it is important to use personal data wisely. It is stated that many people tend to stay away from personal data, but in that way no innovation will occur. Space is needed to discover whether data combinations are useful or not. A safe environment, wherein it can be assured that no harm is done, would be ideal. In this way, data linkages can be explored and tested in a responsible way. The creation of such a space will be part of the Hygiëneregels too.

A way to overcome problems is not being seen in changing the law. A juridical discussion follows the ethical one and currently we are still in the beginning of the ethical one. Thereby, regulation is always lagging behind, according to Zwolle. Laws will never be able to fully take care of privacy protection, because technological developments occur at rapid pace.

7.4.7 Conclusion

In Zwolle personal data is being used, but there are no smart city objectives defined. In order to take care of privacy, the Hygiëneregels are being established. These are yet unfinished and privacy is currently taken care of by using good judgement. It is unclear what should be expected from these rules. At this moment in time, purpose limitation and the other rules of article 6 of the Data Protection Directive are not taken into account. Personal data is thus not lawfully processed. Zwolle seems to take a more flexible approach and tries to do what should be done in a responsible way. For

public safety monitoring and forecasting, open data is being used and produced and almost all data being processed in the municipality is geo-data. Big data is not being used for safety purposes. The ethical debate about data use and the borders between what is and is not allowed and appropriate, together with the fact that privacy is a dynamic concept that is difficult to capture in rules, lead to uncertainty and a lack of clarity. Zwolle sees no solution in changes in law. Flexibility is needed in order to innovate and learn about the use of data combinations. The establishment of a safe test environment would be ideal according to Zwolle.

8. Analysis

This chapter combines the results of chapter 7 with the outcomes of the theoretical chapters 2 to 5. The sub questions on the situation in practice will be addressed and thereby the main quests arising from literature will also be answered for the situation in practice.

8.1 Personal data in the smart city

In theory it became evident that many data is personal data and therefore rules should be taken into account in order to process personal data lawfully for smart city purposes. Personal data use is thus key in this research and for this reason, this chapter starts with personal data use in a smart city context.

Table 2: Personal data use in the four smart city programs for public safety.

	Eindhoven	Almere	The Hague	Zwolle
Personal data being used?	✓	✓	✓	✓
Which data is personal?	Police recordings (other)	All data in the StraatKubus (geo)	Not specified	Not specified

yes: ✓ no: X does not apply: — remains unclear: ?

In all four case studies personal data is being used, although the (type of) data that is personal is different per case (see table 2). In Eindhoven it is 'other' data (not big data, open data or geo-data) that is personal data, while in Almere all personal data is geo-data.

For the case of Eindhoven, this might be a surprising result, since, in general, it is tried not to use personal data. However, the data on police recordings that is being used can be considered personal data (see figure 10, p.39). Recital 26 of the Data Protection Directive indicates that anonymizing data is not enough, and data that is likely to become de-anonymised by combining it with other data should still be considered personal data. Police data is available in a time span of two hours and is thus very detailed. By combining it with for example news items, data could lead to individuals. However, Stratumseind 2.0 focuses on the categories with larger amounts of incidents and it tries to influence them by using, among others, lights. If the categories with only a few incidents would be left out, no personal data would be used in this smart city program.

In Almere, working with personal data in this smart city program is a choice made by the municipality, because of the amount of (spatial) detail that is wanted. The 6ppc-level and the fact that the StraatKubus' aim is to combine datasets on data subjects that are category 2 sensitive data, lead to the fact that all data in the StraatKubus is personal data. Almere is aware of this and decided to undertake action in order to do this lawfully by establishing the Gedragsrichtlijn StraatKubus. The StraatKubus' aim is to monitor and improve liveability and since this concept focuses on actions and feelings of people, this type of data is likely to become personal data. The fact that the StraatKubus shows data at a spatially detailed scale is the most decisive factor in data becoming personal data.

In The Hague and Zwolle personal data was not further specified. In The Hague, the focus of its smart city program on opening up data and sharing data between event partners leads to privacy issues. In both The Hague and Zwolle people are well aware of the fact that many data is or can become personal data. Zwolle emphasizes the fact that the municipality needs to be close to the citizen and therefore personal data is needed, which is agreed upon by Almere.

The fact that in all four smart city programs personal data is being used, as well as the fact that there are multiple types of data being personal data, makes it interesting to see how this is being dealt with, since privacy issues inevitably occur and need to be addressed.

8.2 Smart city objectives and purpose limitation

In order to lawfully process personal data, a number of measures need to be taken, listed in article 6 of the Data Protection Directive. The fact that personal data must be "*collected for specified, explicit and legitimate purposes and not further processed in a way incompatible with those purposes*" is a rule that is of major interest for using personal data for smart city programs, since the smart city is an unclear understanding (EU Directive 95/46/EC, 1995, p.40). A clearly defined purpose for processing personal data should be defined in all four cases, since they all work with personal data.

The expected difficulties in purpose limitation become visible in all four case studies (see table 3). In Eindhoven and Almere there are clearly defined and concrete smart city programs: the StraatKubus and Stratumseind 2.0. In The Hague and Zwolle the smart city programs are more general. In these cases the goals are not specified or not specified enough to enable processing personal data lawfully. These cities are in the process of becoming aware of privacy issues and are starting to address them. Both more concrete cases of Eindhoven and Almere are further progressed in this process.

Table 3: Specification of the objectives for the four smart city programs.

	Eindhoven	Almere	The Hague	Zwolle
Goals officially specified?	✓	✓	✓	X
Goals-well specified enough to use personal data?	X	✓	X	X

yes: ✓ no: X does not apply: — remains unclear: ?

Only Almere has officially specified its objectives and did this well enough in order to lawfully process personal data (see table 3). The specification of objectives is included in the Gedragsrichtlijn StraatKubus. The Gedragsrichtlijn is part of a contract that needs to be signed by the StraatKubus users. The purpose for processing personal data must be specified, explicit and legitimate (Article 29 Working Party, 2013 ; EU Directive 95/46/EC, 1995) (see paragraph 4.3.1, p.23 & 5.3.1, p.31). The purpose of Almere (see paragraph 7.2.2, p.41) can be considered sufficiently *specified*, since in the Gedragsrichtlijn it is clearly expressed what is and is not in the scope of data processing and protection safeguards are established. All elements of the StraatKubus are defined and there is a limitative enumeration of data that can be included in the StraatKubus. Furthermore, safeguards are well taken care of, both in the field of authorisation and cybercrime. The purpose can also be considered *explicit*. The aim of the StraatKubus is easily understandable and unambiguous and explains the purpose for which the StraatKubus is established. Due to the limitative enumeration of data themes, it becomes clear what data is included in the StraatKubus. The purpose also needs to be *legitimate*. A legitimate purpose means that at least one of the article 7 criteria in the Data Protection Directive should be met. Article 7.e of the Data Protection Directive is applicable to the StraatKubus: "*Processing is necessary for the performance of a task carried out in the public interest or in the exercise of official authority vested in the controller or in a third party to whom the data are disclosed*" (EU Directive 95/46/EC, 1995, p.40). Legitimacy is not only about meeting at least one of these criteria, but also all other Dutch laws, ethics and customs should be taken into account. It is difficult to decide if the StraatKubus is in accordance with other laws, ethics and customs, since this is another field of research which is not in the scope of this one. Because of the fact that at least one of the article 7 criteria is met, it is plausible that the Gedragsrichtlijn StraatKubus can also be considered legitimate.

Furthermore, the data must not be used for purposes that are incompatible with the original purpose. In the StraatKubus this is also taken care of. For example, use of StraatKubus images for presentations is only allowed when the rules on this topic in the Gedragsrichtlijn are met. An example is that it is not allowed to do publications on 6ppc level.

In the Article 29 Working Party opinion on purpose limitation (Article 29 Working Party, 2013), a practical advice was provided on what should be included in a justification document such as

the Gedragsrichtlijn (see paragraph 5.3.1, p.31). The Gedragsrichtlijn StraatKubus includes all demands, but more attention to the justification of the themes and the subthemes of data could be paid. The themes are a limitative enumeration and it is therefore very clear what is and is not included in the StraatKubus, but it is not explained why these themes are needed to monitor liveability. However, it is evident that Almere is aware of the fact that purpose limitation needs to be addressed in order to lawfully process personal data and the municipality has well taken care of this.

Both Eindhoven and The Hague specified their smart city goals, but not sufficiently enough to enable lawfully processing personal data. For Eindhoven this is easy to explain, since Eindhoven tries to stay away from personal data. According to the municipality no personal data is used, and in this case it is not needed to work on purpose limitation. In The Hague, smart city objectives have been defined, but these objectives are broad policy objectives. An example is provided by one of the two goals on safety and is for The Hague to become "*the Silicon Valley of safety*" (Gemeente Den Haag, 2014, p.2).

Zwolle has no explicitly formulated smart city program. The municipality works on improvements in the city and uses smart solutions if they are suitable. In Zwolle, the smart city is being seen as a holistic concept and it is argued that some smart solutions do not make the city smart yet.

Both Zwolle and The Hague involve no very specific or concrete program, compared to the StraatKubus in Almere and Stratumseind 2.0 in Eindhoven. A major difference in approach between The Hague and Zwolle is that The Hague markets itself as smart city whereas there is no large difference in the ICT program compared to the previous years in which the smart city term was not yet used, while Zwolle use smart solutions but does not profile itself as a smart city.

An interesting result is that personal data is involved in all cases, but in only in Almere the purposes are specified well enough in order to process personal data lawfully. At the same time, in all cases people are aware of the fact that privacy is an important issue that needs to be addressed. Almere put a lot of effort into arranging all necessary measures. This was a long, difficult and uncertain process. In the remaining of this chapter, it will become evident that even though people in all cases are aware of privacy issues, the fact that it is difficult to find out how to deal with privacy is a major issue.

8.3 Big, open and geo-data and privacy in practice

In this paragraph, the outcomes on data use in practice and the influence of privacy protection are discussed. It should be kept in mind that the outcomes are on data use for public safety, and not in general and/or for other purposes.

8.3.1 Big data and privacy

In two out of the four case studies, big data is being used for public safety purposes (see table 4, p.53). Eindhoven makes use of many sensors. They are being used to register temperature, sun, rain and wind as well as sounds (3D) and social sensors (media watching and interactive). Also video people counting data that is collected by using software that is installed on cameras, can be seen as big data. The 3D sound data can be used real-time during an event. If a predefined level of sound is reached, a text message can be sent to the event organization so they can undertake action. The sensor data is linked to other data and can be displayed in an information system (see figure 8, p.37).

Besides real-time systems monitoring, another application for big data in the smart city was defined by Jara et al. (2014) and Steenbruggen et al. (2014): the use of location data from mobile phones. This type of data is also being used in Eindhoven. Even though people might find this data a bit 'scary', it can be considered non-personal big and geo-data. The data is anonymised by Vodafone in accordance to the rules they agreed upon with the CBP and then, via another company, the data is transferred to Eindhoven. The data is registered as the percentage of visitors from a certain municipality (using an accuracy of about 200 meters), and is thus not very spatially detailed. Furthermore, there must be at least fifteen people originating from one municipality for the

percentage to be registered. Taking the ePrivacy Directive (EU Directive 2002/58/EC, 2002) and the clarification on personal location information of Van Loenen et al. (2008) into account, this data can be classified as non-personal data (see paragraph 5.2, p.29).

In Eindhoven location data is not being used real-time, but in The Hague it is. The smart city program in the field of safety of The Hague focuses on event safety. Crowd control is an important element herein and data on the use of mobile phones is being used to gain knowledge on the amount of people. It is known, for example, when squares need to be closed and people are informed about routes they can take, making it real-time information. Especially for The Hague, Steenbruggen et al. (2014) were correct about the potential of mobile phone location data for detecting patterns in movement, using information for emergency situations and estimating the number of people present. In The Hague it is stated that privacy is not an issue in working with location data from mobile phones. This statement indicates that the data used is at least not being seen as personal data. The reason for this does not become evident in the interview. It might have to do with the fact that the police are mainly concerned with protecting public safety and in this case other privacy regulations are valid. Personal data would then not be under the scope of the Data Protection Directive. Another possibility is a lack of awareness of personal data or no personal data is involved.

In the past, sounds (3D) and video people counting were personal big data that was used in Eindhoven. Sound data did then not only include registering the level of sound, but also what was being said. In video people counting, the counting software used to be installed on computers instead of directly on the camera and the (blurred) video images were used. Nowadays, only the level of sound is registered, as well as only the number of people passing a camera, to ensure that personal data is not being used in Stratumseind 2.0. Eindhoven is thus exploring the borders between personal and non-personal data and tries to stay on the non-personal data side.

Table 4: The use of big data in the four smart city programs for public safety purposes.

	Eindhoven	Almere	The Hague	Zwolle
Big data being used?	✓	X	✓	X
Which data being used is big data?	- Temp., sun, rain & wind[13] - Sounds (3D) - Cellphone counting (mobile phone location data) - Video people counting - Social sensors	—	- Mobile phone location data - Possible other big data that is not specified	—
Personal big data being used?	X	—	?	—
Purpose limitation personal big data?	—	—	—	—
Specified amount of time personal big data kept?	—	—	—	—

[13] Not all big data in Eindhoven seems to focus on public safety, at first sight. However, this data is marked as relevant for public safety since in Stratumseind 2.0 safety is one of the main objectives and these measures are used to gain insight into the circumstances in which an incidents takes place. The same accounts for the data of Eindhoven in table 6, p.56.

Personal big data being updated?	–	–	–	–
People informed about big data collecting?	–	–	–	–

yes: ✓ no: X does not apply: – remains unclear: ?

In Eindhoven, big data for public safety is thus being used for both (real-time) systems monitoring and for gaining location data from mobile phones. In The Hague big data for public safety is being used for location data in events. The other two cases, Almere and Zwolle, do not use sensor data for monitoring and forecasting public safety in the city. In Zwolle sensor data is being used, but not in the field of safety. In the StraatKubus from the municipality of Almere no use is made of sensors.

Since no personal big data is being used in the cases, possibly with the exception of The Hague, the purpose limitation demands are not relevant. From literature, problems in processing personal big data were expected to occur in purpose limitation, an unspecified amount of time of data being kept and in informing people. In the four case studies these issues do not become evident, since no personal big data is being used, possibly with exception of The Hague.

8.3.2 Open data and privacy

In open data use in smart city programs, a division between on the one side the two concrete smart city programs of Eindhoven and Almere and on the other hand the more general smart city programs of The Hague and Zwolle can be made. In the more concrete programs, open data is not produced and only in the StraatKubus of Almere open data is being used. In the more general programs of The Hague and Zwolle, on the other hand, open data is in both cases produced and in Zwolle it is also used (for The Hague this was not explicitly mentioned in the interview) (see table 5, p.55). This can be explained by the fact that in the municipality in general, open data is very likely to be produced as a consequence of INSPIRE (EU Directive 2007/2/EC, 2007).

The cases of The Hague and Zwolle have an open data portal. In the portal of The Hague, data is included on public safety, while in Zwolle there is little information available on this topic. In Zwolle there is spatial open data available on public safety from the neighbourhood monitor and there is data on the societal challenges the municipality works on. The role of the municipality becomes increasingly to find data that is already available. In The Hague there is more open data on public safety available (see table 5, p.55). However, these open datasets are not detailed and in most cases only available on the spatial level of the city. Furthermore, if data is not being shown on a map, the overall number or percentage for the whole city is displayed. The data is thus not detailed and thereby in most cases not very interesting, since it does not allow comparing within the city.

The fact that data remains very general also has a positive side, since in both Zwolle and The Hague no personal data is opened up. The very general open data published by The Hague is surprising, keeping in mind the smart city purposes of The Hague and the focus in the interview on open data, sharing data and privacy issues herein. Not much data has been opened up yet. On the other hand, the process of opening up data in The Hague is speeding up and the Smart City Road Map is scheduled from 2014 to 1018, so there is time left to work on achieving the objectives.

In Almere, open data is being used in the StraatKubus. StraatKubus data is not made publicly available and only a small group of people is allowed to use the data. The aim of the StraatKubus is to support conversations on liveability with facts, but the StraatKubus data cannot be shared with the subjects of these conversations: the inhabitants of Almere. According to the municipality, data can only be opened up if it is highly anonymised and generalised, because open data cannot be personal data due to privacy legislation. The highly anonymised and generalised open data can be seen in the cases of Zwolle and The Hague. For the case of the StraatKubus in Almere, privacy protection stands in the way of opening up the data of the StraatKubus. The data in the StraatKubus is personal data and cannot be opened up. For this reason, only a limited amount of people is allowed to access the StraatKubus; authorisation is taken care of in the Gedragsrichtlijn. It is argued that with the

StraatKubus a greater purpose can be served than by opening up data, because it addresses societal challenges.

In the other more concrete program, Stratumseind 2.0 in Eindhoven, open data is not being used nor opened up. The municipality is working on opening up data in a portal, but the data available in the Living Lab is not opened up. In the near future, data from the portal will be included in the system if it is considered relevant. An example is data on the amount of bicycles parked on or near Stratumseind.

Table 5: The use of open data in the four smart city programs for public safety.

	Eindhoven	Almere	The Hague	Zwolle
Open data being used?	X	✓	?	✓
Which data being used is open data?	—	Open data from Kadaster and CBS	Not specified	Not specified
Open data being produced?	X	X	✓	✓
Which data being produced is open data?	—	—	On themes: - Crime & nuisance - Opinion neighbourhood - Common nuisance - Sense of security - Social cohesion	- Neighbourhood monitor 2014 - Societal challenges
Personal open data being produced?	—	—	X	X
Purpose limitation personal open data?	—	—	—	—
Specified amount of time personal open data kept?	—	—	—	—
Personal open data being updated?	—	—	—	—
Impact of personal open data identified?	—	—	—	—
yes: ✓ no: X does not apply: — remains unclear: ?				

In theory, problems in purpose limitation and an undefined amount of time were identified. These issues are not visible in the four cases, since no personal open data is produced. The data that is opened up is very general and anonymous municipal data, that is not (spatially) detailed. Furthermore, in the two concrete smart city programs, the role of open data is limited. For Almere, this is fully due to privacy protection.

To conclude, the two theoretical problems are thus not being seen in the four cases, but there are other problems experienced in privacy and open data. In The Hague there is a large focus in the smart city program on opening up and sharing data, but this turns out to be a difficult process because of privacy protection. The decision tree is used in order to decide if opening up is allowed and lawyers are involved in the final decision. In Almere, opening up StraatKubus data is not possible since the data in it is personal data and can thus not be shared.

In literature, the link between open data and the smart city was mainly found in the fact that opening up data enhances the transparency of the government and enables the citizen to participate and innovate. In all involved cases, municipalities are working on opening up data, whether or not it

is part of the involved smart city program. Especially in The Hague it becomes evident that this process of opening up has just started. The problems expected in purpose limitation did not become visible in these four cases, because of the fact that this process of opening up started only recently in most of the cases and because of the fact that in opening up data the privacy issues are clear. Especially in Almere and The Hague the influence of privacy is evident. These cases are trying not to open up data that can be considered personal data, which results in data portals with general and highly anonymised data. Privacy thus does influence opening up data, but not in the field of purpose limitation.

8.3.3 Geo-data and privacy

Geo-data is being used in all cases (see table 6). The capability of geo-data to link place, time and attributes becomes evident in both of the more concrete cases. In Eindhoven, all data is combined in the Living Lab interface (see figure 8, p.37). In the interface, a map is being used as background in which the locations of the measurements are shown. The most important measurement locations are the five entrances/exits of Stratumseind, at which the sensors and cameras are located in a lampposts. Since the location of sensors is known, sensor data is also geo-data, just like Van Oortmarssen and De Vries (2014) concluded. For this reason the list of big data and the list of geo-data of Eindhoven are largely the same. In Eindhoven, there is no personal geo-data being used. Only police recordings can be considered personal data and this data is not geo-referenced (see paragraph 8.1, p.50). In Almere, the StraatKubus is a geographic information system (GIS), so all data in it is spatially referenced. The 6ppc-level is one of the most important characteristics of the system. This is also the main cause of privacy difficulties, which are addressed in the Gedragsrichtlijn StraatKubus (see paragraph 8.2, p.51 & 8.4, p.58), since this scale contributes to data easily becoming personal data. The statement that geographical data contributes to data easily becoming personal data, is thus true for this case.

Table 6: The use of geo-data in the four smart city programs for public safety.

	Eindhoven	Almere	The Hague	Zwolle
Geo-data used?	✓	✓	✓	✓
Which data being used is geo-data?	- Temp., sun, rain & wind - Sounds (3D) - Cell phone counting - Video people counting - Car and bike parking information	All StraatKubus data	- Mobile phone location data - Most open data - Not further defined	- All open data - Not further defined
Personal geo-data being used?	X	✓	?	?
Which personal data being used is geo-data?	–	All StraatKubus data	–	–
Purpose limitation personal geo-data?	–	✓	–	–
Personal geo-data processed according to the rules set in article 6 of the Data Protection Directive?	–	✓	–	–
yes: ✓ no: X does not apply: – remains unclear: ?				

Mobile phone location data is an important example of geo-data use in the smart city (see paragraph 8.3.1, p.52). In the interview about Stratumseind 2.0 in Eindhoven, other methods to gain location information are mentioned, but not yet implemented in the system. Examples are the Wi-Fi-connectors and iBeacons. Their spatial accuracy is better, but also has the consequence that this data would become personal data, keeping the quote on personal geo-data of Van Loenen et al. (2008) in mind (see paragraph 5.2, p.29). In The Hague, mobile phone location data is being used real-time in crowd management. It is stated that privacy issues in using this data do not occur (an explanation for this is provided in paragraph 8.3.1, p.52). Furthermore, in The Hague, it is stated that geo-information is the frontrunner in the process of data sharing.

In open data portals, geo-data plays an important role. In the open data portal of Zwolle all data is location-based and in the open data portal of The Hague, the majority of data is (also available as) geo-data. However, in The Hague open data is not spatially detailed and mostly displayed on the level of the city. In Zwolle, most data is available on the level of neighbourhoods. In both cases, there are no problems with privacy regulation, since data is anonymised and generalised. In Zwolle, it is estimated that about 80.0% to 90.0% of all data is geo-data. This indicates that it is likely that the personal data that Zwolle processes is also geo-data, but this is not further specified in the interview.

Geo-data is thus being used in all four case studies, and in Almere it is spatial scale that leads to data becoming personal data. This is well taken care of in the Gedragsrichtlijn StraatKubus.

8.3.4. Data and privacy in practice: a conclusion

Personal data is being used in all four cases, but this is not personal big data nor personal open data. Big data mainly is being used for public safety purposes in Eindhoven and in The Hague. In The Hague open data plays an important role in its smart city program, and in Zwolle data is opened up too. However, data remains very generalized and anonymised. The problems in purpose limitation that were defined in the theoretical part of this research are thus not being seen in the practical part of this research.

A large percentage of all data is geo-data. For Almere it becomes evident that the detailed geographical scale of 6ppc level, largely contributes to data becoming personal data. Almere is aware of this and decided to develop the Gedragsrichtlijn to ensure that personal data is processed according to the rules set in the Data Protection Directive and the Dutch implementation in the Wbp.

In the field of public safety, data from the police is often important. In Eindhoven and Zwolle it was mentioned that the police have different rules for privacy, which allow them to focus on the individual. The difference in role between the police and the municipality was highlighted. The police focus on the individual and are allowed to do that, while the municipality focuses on the context in which an incident occurs. In The Hague too, it is emphasized that the police and municipality are working together in the field of safety. The fact that the police are not sharing (personal) data is in The Hague being seen as a limitation for the shared goal of a safe city. In general, the cases are aware of the fact that the police have different rules in privacy protection compared to the municipality and that the municipality has a different role since it focusses on the community instead of on individuals.

A division in the cases can be discovered when the outcomes of the interviews and policy documentation study are compared. Eindhoven and Almere are more defined and concrete cases. In the field of privacy these cases are further progressed. However, two opposite strategies were taken: Almere chooses to use personal data and covers this in the Gedragsrichtlijn, while Eindhoven tries to stay away from personal data because of privacy legislation (see paragraph 8.4, p.58). The other group of cases is formed by The Hague and Zwolle. These cases have a broad smart city program and objectives are not specified or not specified enough to enable processing personal data. In these case studies privacy issues occur, and they are still in the process of addressing these issues.

Since in all case studies there are privacy issues, but not in personal open data and personal big data, it is interesting to see what these issues are and how they are being dealt with.

8.4 The influence of privacy on the smart city

In this paragraph, the difficulties experienced by the four cases in the field of privacy will be addressed. Furthermore, the influence of privacy issues on achieving the smart city objectives is clarified.

The more concrete cases of Eindhoven and Almere are further progressed in dealing with privacy, compared to The Hague and Zwolle. However, the problems they are experiencing are different (see table 7). In Eindhoven, problems are mainly found in a lack of clarity on the border between personal and non-personal data. Legislation does not provide clear answers on this subject and therefore lawyers and research institutes are involved to provide advice on the subject of privacy. Examples of partners are the CBP, TILT, Tilburg University, and the Ministries of Justice and Internal Affairs. Thereby it is tried to stay on the side of non-personal data. This has the consequence that not all data can be saved. For sound data and video people counting, this problem is now solved (see paragraph 8.3.1, p.52), but also Bluetooth and Mac addresses as well as social media messages (social sensors) are not being saved. Even though this data is relevant for the smart city program, it is not being used because of privacy. The same problem occurs in location data. The amount of people in the street is registered, but it is not known where people are in the street. It is stated that this type of data would be interesting for monitoring safety to see if parts of Stratumseind get crowded. In Eindhoven, privacy difficulties are thus experienced and they influence achieving the smart city objectives.

In Almere these problems do not occur. While Eindhoven tries to stay away from personal data, the aim of Almere is to work with personal data in order to monitor liveability and to do this lawfully. In the Gedragsrichtlijn, all rules of the Data Protection Directive are taken into account (see paragraph 8.2, p.51). This is thus a useful way to work with personal data and at the same time to overcome privacy problems. But also in this approach there is a 'downside', in the field of the many limitations the Gedragsrichtlijn imposes. The data can, for example, only be used by a limited amount of people and only if they need the data in order to do their job well. Because of the focus on privacy protection, the attitude towards opening up data and big data is negative because of privacy risks and the fact that open data must then be very general data. Thereby, the StraatKubus as a smart city program does not contribute to smarter and involved citizens, which is an element that is often used in smart city definitions (Giffinger et al., 2007). However, Eindhoven is not opening up data either. Altogether, Almere does provide a good example of using personal data lawfully. It proves that working with personal data is not impossible, as long as personal data use can be justified, and processing personal data is allowed as long as the rules of the Data Protection Directive are satisfied.

Table 7: An overview of the experience of privacy difficulties and the influence on smart city objectives in the four smart city programs for public safety.

	Eindhoven	Almere	The Hague	Zwolle
Experience privacy difficulties?	✓	✓/✗	✓	✓
Wherein are privacy difficulties experienced?	In a lack of clarity on the border between personal and non-personal data	In the past they were experienced in using personal data. It was unclear how to deal with this	Opening up and sharing data	Working with personal data for municipal tasks
How are difficulties addressed?	Personal data is avoided. Lawyers and research institutes are asked for help and personal data is not saved	Help was asked and eventually the Gedragsrichtlijn StraatKubus was established	The decision tree was established and lawyers are asked for help	Using good judgement and Hygiëneregels are being established

Does privacy influence achieving the objectives?	✓	X	✓	X
In what way does privacy influence objectives or why does it not influence them?	A more precise location of people on Stratumseind is useful for monitoring safety but is not used for privacy reasons. Also, data that would be interesting for the project cannot be saved because of privacy	By establishing the Gedragsrichtlijn, the objectives can be achieved since personal data can be used lawfully	Opening up data and sharing data between partners is an objective but open data cannot be personal data	There are no specified objectives. The municipality does what is necessary and mainly uses good judgement. To ensure this and achieve alignment, the Hygiëneregels are established
yes: ✓ no: X does not apply: — remains unclear: ?				

Establishing the Gedragsrichtlijn was not an easy process. Since the law is difficult to interpret, help was needed. But even professionals could not provide a concrete and workable advice on how to process personal data lawfully. Two different worlds, of lawyers and of municipalities trying to monitor and improve the city by making use of data, needed to be bridged.

In the broader smart city programs of The Hague and Zwolle, there is a focus on open data. Both municipalities are aware of the fact that many data are or can become personal data. This results in opening up data on subjects that do not relate to people in any way, which is the case for almost all open data in Zwolle, or in opening up data that is very well anonymised and generalised, which is especially the case in The Hague. Because of avoiding opening up personal data, the data that is opened up is not (spatially) detailed and thus also less valuable. For The Hague, privacy difficulties in sharing and opening up data are the main issues. Open data is a central element in safety in the Smart City Road Map. For opening up data the decision tree is being used and lawyers are involved. This way, opening up can be decided for each dataset. In both The Hague and Zwolle, it is stated that lawyers tend to be against opening up data; if data is not opened up, there will be no problems either. Furthermore, it is still uncertain how citizens will react to opening up data. Privacy and the 'fear' of opening up personal data thus influences the quality of open data. In general, it is a positive result that personal data is not opened up in these cases, but, on the downside, the quality of open data is suffering from the fear of opening up personal data.

The smart city program of The Hague mostly works on safety in events. In events it is useful to share data among the event partners, including the police. Privacy limits the possibilities for data sharing among event partners. Since opening up data and sharing data in events are important parts of the smart city and safety objectives of The Hague, it could thus be stated that privacy influences achieving the smart city goals of The Hague.

Municipalities need to work with personal data, in order to be close to citizens and in order to focus their policies, according to Almere and Zwolle. In Zwolle, personal data use is not yet taken care of in a formal document as is the case in Almere, but such a document is currently being established (the Hygiëneregels). At the same time, privacy is being seen as a concept that is difficult to cover in rules. The purpose for personal data use should remain central and thereby attention must be paid to privacy risks. Currently, this is mainly a matter of using good judgement, but privacy is becoming more and more of a complicating factor. The alderman therefore decided that the Hygiëneregels should be established. Everyone working with personal data knows that privacy should be taken into account, but to make sure that everyone knows about it and deals with it in a correct way, the Hygiëneregels are established.

The approach of dealing with privacy and personal data of Zwolle is thus more flexible, compared to Almere. Doing what needs to be done and to do it wisely seems the be the motto in

Zwolle. Compared to Almere, taking care of privacy is less controlled and, depending on the exact content of the Hygiëneregels, the rules for processing personal data are likely not to be fully complied with.

In the four case studies, different privacy issues are experienced (see table 7, p.58). All difficulties focus on personal data (unclear borders, unclear how to process lawfully, and unclear how to work with it for municipal tasks) or on open data in the case of The Hague. Different approaches were taken (see table 7, p.58).

Altogether, there is one major problem that all four cases share: a lack of clarity. It is not clear what is allowed and what is not, where the borders are between personal and non-personal data, how personal data can be used and in what way privacy issues can be addressed. In short: there are many questions that need to be answered to address the lack of clarity.

8.5 Conclusion situation in practice

In theory, the importance of purpose limitation in using big, open and geo-data for monitoring and forecasting public safety in smart cities became evident. If personal data is being used, the rules listed in article 6 of the Data Protection Directive, including purpose limitation, need to be taken care of. It was expected that personal data would be used in all cases, which indeed is being done. However, only in Almere, purpose limitation and all other demands for processing personal data in the Data Protection Directive and Wbp are fully taken care of.

Furthermore, it was expected that especially the use of big and open data would conflict by their nature and definition with privacy protection. In practice, personal big data or personal open data is not being used in any of these four cases. In opening up data, part of the smart city programs of The Hague and Zwolle, privacy is well taken care of and this results in opening up data that do not relate to people in any way, or in opening up highly anonymised and generalised data. Big data is being used in Eindhoven and The Hague. Eindhoven thereby balances on the border between non-personal data and personal data and chooses to stay on the non-personal side, whereas for The Hague it is not clear to what extent mobile phone location data is personal data (for an explanation see paragraph 8.3.1, p.52). For geo-data, on the contrary, personal data is being used in at least Almere, and possibly also in The Hague and Zwolle. In Almere this is well taken care of in the Gedragsrichtlijn StraatKubus.

All cases experience privacy issues. A lack of clarity is the main issue in all cases. The law is difficult to interpret and clear answers are missing. Almere is further progressed on this subject, and Eindhoven works closely together with experts in the field of privacy to make sure that no personal data is being used. Zwolle is in the process of working on addressing privacy issues. The establishment of the Hygiëneregels is an important step. The Hague seems to be the least aware of privacy issues, and mostly focuses on addressing privacy issues in opening up data. All cases experience comparable issues, since in all of them personal data is being used and a lack of clarity in how to deal with it is widely shared. Therefore, in the next paragraph, suggestions for solutions and improvements will be done.

8.6 Bridging the smart city and privacy: suggestions on solving the issues

There are two main ways of addressing the problems. The first option can be found in a change in law. The second option is to work on a practical solution, by addressing the needs of the cases (see table 8, p.60).

It is evident that the cases do not see changing the law as an option for improvements (possibly with exception of The Hague, where this topic was not addressed). A change in law is expected to take a lot of time, and by the time the changes are implemented, the law will be outdated because of technological developments. In Zwolle, the law is being seen as the official registering of the

outcomes of an ethical discussion, whereby the ethical discussion is more interesting. A more practical approach in solving the issues and not focussing on changes in law is thus being viewed as a better way to address issues.

Table 8: the changes needed according to the four cases.

	Eindhoven	Almere	The Hague	Zwolle
Opinion on change in law on personal data processing needed	A change in law is not needed, since laws will always be behind. Establishing changes takes too long and the law will soon be outdated again due to technological developments	A change in law is necessary since it is lagging behind, but this is not most important. Could be more focused on our network society and could maybe provide a checklist to enable adopting the law easier	?	Eventually the law will be changed, but the ethical discussion just started and the outcomes of this discussion will eventually be registered in a law. Laws will always be behind and never will be able to fully take care of privacy protection.
Opinion on practical changes needed	More clarity on the border between personal data/non personal data	The Gedragsrichtlijn is a good solution but this was a difficult process. More help and clarity is needed in the future	More flexibility for sharing data among event partners	More space for learning on data combinations and privacy. Establishing a safe test environment
yes: ✓ no: X does not apply: — remains unclear: ?				

In the suggestions for practical changes a distinction can, again, be made in the concrete cases of Eindhoven and Almere and the broader smart city programs of The Hague and Zwolle. The Hague and Zwolle both state that more flexibility in coping with privacy is needed. In The Hague this is needed in order to share data between event partners, so event safety can better be taken care of. In Zwolle, flexibility is needed because there must be space to discover possibilities and risks in order to innovate. Especially in combining datasets for monitoring and forecasting public safety it is needed to experiment on making data combinations in order to see if combinations are relevant.

In Eindhoven and Almere, there mainly is (or was, in case of Almere) a quest for more clarity and help. Zwolle also experiences these issues and it is expected that The Hague will experience these issues too, as soon as the city is more aware of difficulties in privacy. Issues arise in the fact that the law is lagging behind on the current situation in which technological innovations occur at rapid pace, resulting in difficulties in interpreting regulation. In all cases it became evident that people find it difficult to start from scratch on the issue of privacy. It is thus interesting to see how this issue can be addressed and possibly solved.

It is important to notice that using personal data for smart city purposes is certainly not impossible. Even the heaviest category of personal data (3) can be processed if there is an explicit permission from the data subject, according to article 23 of the Wbp. As was concluded from theory, there are many possibilities for working with personal data, as long as the rules set in the Data Protection Directive and in the Wbp are taken into account. Almere proves this concept: by working with the Gedragsrichtlijn in order to protect the personal data of the StraatKubus, all demands are taken into account and personal data can be processed lawfully.

There is no other way than a formal way to enable lawfully processing personal data. In Almere, and to some extent also in Zwolle, this approach is already taken. Purpose limitation is the main demand for using personal data, and this needs to be addressed by specifying the purpose, making it explicit and ensuring its legitimacy. Furthermore, further processing of data is only allowed if the new purpose is not considered incompatible with the original one. There is an important exception to the rule for further processing. If this is needed for the prevention, detection and prosecution of crime, further incompatible processing is allowed. However, in practice, it turned out that the focus is on the mass and the circumstances in which an incident occurs, instead of on

individuals. This is being seen as the task of the police. Municipalities thus need to work on purpose limitation. The Article 29 Working Party listed more concrete demands in its opinion on purpose limitation (Article 29 Working Party, 2013) (see paragraph 5.3.1, p.31). To put it simply, these elements need to be written down for the smart city program according to the demands listed in the Article 29 Working Party opinion on purpose limitation and in the Data Protection Directive.

This was also the advice Almere was given by the CBP. Eventually it resulted in the Gedragsrichtlijn. Since municipalities are very likely to work with personal data, as was highlighted by Almere and Zwolle, Almere and the Gedragsrichtlijn can serve as an example for other municipalities. The process of getting help and getting a clear advice on what needs to be done in order to lawfully process personal data was difficult and long. Almere can help the other cases in achieving more clarity and certainty.

On the down side, the Gedragsrichtlijn also imposed a number of restrictions in the field of authorisation, flexibility and innovation (see paragraph 8.4, p.58). In Zwolle these limitations were taken into account, and at the same time a formal approach was adopted by establishing the Hygiëneregels. Since these rules are unfinished, it is difficult to decide to what extent the rules of the Data Protection Directive and the Wbp on personal data are sufficiently taken into account to lawfully process personal data. It is stated that space for experimenting on data combinations is also part of the Hygiëneregels. Zwolle thus tries to work responsibly with personal data and at the same time leaves space for innovation and learning about data combinations and privacy risks. This is a more flexible system than Almere has, and from the point of learning and flexibility, this system might be more attractive. However, this approach does not seem to be lawful when keeping in mind the rules of the Data Protection Directive. When data is combined just to see if it makes sense, the demands of purpose limitation cannot be met.

The ideal situation would be to combine the formal approach of Almere, with the flexibility of Zwolle. This way, personal data can be processed according to the rules and at the same time there will be space to stimulate innovation. A balance between the correct but restrictive approach of Almere and the incorrect but flexible and innovation-allowing approach of Zwolle needs to be found. However, this is not an easy combination to make. The quest for more flexibility in processing personal data is difficult to address, due to the strict rules of purpose limitation and the other demands of article 6 of the Data Protection Directive. The following can, for example, be done:

- ***Establishing a panel of stakeholders***

People from different backgrounds can be brought together to discuss and exchange ideas on combining the formal and correct approach of Almere and the flexibility of Zwolle. Important stakeholders are municipalities that experience issues and are searching for solutions (especially Almere and Zwolle), specialists with a background in law as well as scientists on data and law. The Vereniging van Nederlandse Gemeenten (VNG, Association of Dutch Municipalities) could also be involved in order to represent municipalities.

Almere can serve as an example for other municipalities, since this case proves that it is possible to process personal data lawfully. Knowledge can be shared with other municipalities and together with all stakeholders, ideas can be shared on different approaches and a way to implement more flexibility.

- ***Communicating the outcomes in clear and easily understandable documentation***

These meetings will result in knowledge on the most important aspects of processing personal data lawfully, which can be written down in concise and easily understandable documentation. The approval of the CBP is important herein and the CBP has an interest herein because this solution will stimulate municipalities to make arrangements for processing personal data. This way, the lack of clarity on dealing with privacy and personal data can be addressed. Because of the function of the municipality in society, being close to citizens, personal data will in many cases be used. Municipalities will benefit largely from clear guidelines and best practices on privacy issues. The step to start processing personal data responsibly, and in accordance with the Data Protection Directive

and the Wbp, will become smaller.

These solutions could improve the situation of the case studies that were included in this research. Eindhoven will, for example, gain more information on how to use personal data lawfully in Stratumseind 2.0. The fact that personal data is avoided can be seen as a chance that is missed, because as long as the rules are complied with, personal data can be used. Almere might benefit from the position as leading city in this field of interest, by sharing their approach with others. Furthermore, the meetings can be interesting to see if more space and flexibility can be created in their system. The Hague is still in the beginning of the process of becoming aware of privacy, and this platform would be an interesting opportunity to learn more and start working with the results. Zwolle would highly benefit from the debate about the possibility of adopting flexibility into the formal justification for processing personal data that needs to be established in order to lawfully process personal data.

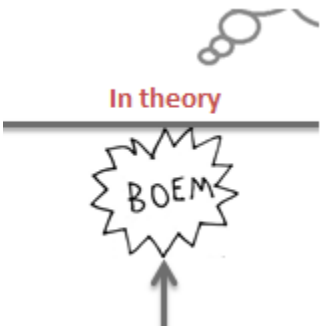
Currently, it is unclear how to deal with privacy and personal data in data processing for smart city purposes. A helping hand is missing and municipalities have to find answers themselves. These solutions address the main demands for solving the identified issues in using personal data for smart city purposes: achieving more clarity and certainty. It enables cities and municipalities to move from a passive attitude due to a lack of clarity, towards applicable knowledge and processing personal data lawfully and responsibly.

9. Conclusion

In this chapter, all research questions will be answered. Thereby, elements of the research model (see figure 3, p.10) will be used. The seventeen research steps and ten research sub questions identified in table 1 (see paragraph 1.4, p.11 & paragraph 1.5.1, p.13) will provide the structure of paragraph 9.1. As a research question was answered, this is indicated with (RQx) in the text. Questions 1, 2 and 3 will not be explicitly answered, since these questions focus on gaining the basic knowledge on the mean components of this research needed to answer the other questions. By using this method, a summary of the main findings will thereby automatically be established. Thereafter, a concise answer on the main research question is provided in paragraph 9.2.

9.1 Summary of the research findings

The smart city does not exist. It is a vague concept, with no widely shared definition. In the Netherlands, there are cities implementing elements of the smart city concept. In this research, four examples of such smart city programs, focusing on public safety, are included: Eindhoven (Stratumseind 2.0), Almere (StraatKubus), The Hague and Zwolle. Data use is central in smart city programs and especially big data, open data and geo-data are data types that are often used in smart city programs. However, data is often personal data; data that, possibly by combining it with other data, can lead to individuals. Personal data is in the European Union protected in the Data Protection Directive, which is implemented in the Netherlands in the Wet bescherming persoonsgegevens. Article 6 of the Data Protection Directive lists rules for processing personal data. Personal data can only be processed lawfully if these demands are satisfied. One of them is on purpose limitation: the purpose for data processing needs to be specified, explicit and legitimate and data cannot be further processed for purposes that can be considered incompatible (RQ4). But, since the smart city is a vague concept, it is expected that difficulties arise in following the rules of the Data Protection Directive and especially in purpose limitation. These problems are identified as follows:

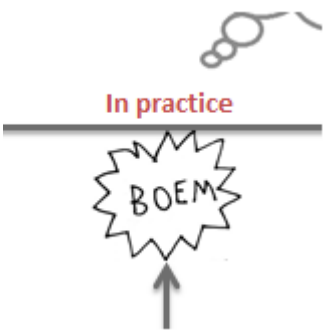


1. Data and purpose: The use of big data and open data can conflict by their nature and definition with privacy protection regulation:
- Big data: further processing is tempting (*purpose limitation*), and in relation to purpose: data is easily kept longer than strictly necessary, updating is difficult, and it is tempting to process more data than strictly necessary.
- Open data: by definition not collected for specified and explicit purpose (*purpose limitation*), and in relation to purpose: data is easily kept longer than strictly necessary, and keeping the data up to date is difficult.

2. Smart city and purpose: If personal data is being used in a smart city context, purpose limitation should be taken care of. Since the smart city is a vague concept, difficulties in *purpose limitation* are expected (RQ5).

Thereafter, in interviews on four cases, it was researched if these problems occurred in practice:

1. Data and purpose: No personal big or open data is used in the four cases, so the rules of article 6 of the Data Protection Directive, among which is purpose limitation, do not apply. For these four cases, privacy issues do not occur in open and big data.
2. Smart city and purpose: In all four cases personal data is used, either geo-data or 'other' data. The rules of article 6 including purpose limitation should thus be taken care of, but in only one case personal data is being processed according to the law: Almere (RQ6). In



explaining why only one case is processing according to the rules, a distinction can be made between the two concrete smart city programs of Eindhoven and Almere and the broader programs of The Hague and Zwolle. The concrete programs of Stratumseind 2.0 and the StraatKubus are, both in their own way, further progressed in taking care of privacy. In the broader programs, it is more difficult to define objectives, to justify all personal data use and to do this well enough to enable lawfully processing personal data (RQ7).

In The Hague and Zwolle open data is included in the smart city program, but data is highly generalised and anonymised and it is not personal data. Big data is used in Eindhoven and The Hague. Only in The Hague there is a chance that this data is personal, but this is not likely, since it is stated that privacy problems do not occur in processing this data. In geo-data, on the contrary, personal data is used. In Almere, the spatially detailed scale of data use, 6ppc-level, contributes to data becoming personal data. In The Hague and Zwolle, it remains unclear what data exactly is personal data, but this could be geo-data. The personal data that is being used in Eindhoven is data on police recordings and this can be categorised as 'other' data. Out of the three data types included in this research, it is geo-data that is most often personal data. Open data never is personal data in these four cases, which is a positive outcome (RQ6+7).

In Eindhoven and The Hague, privacy has an influence on achieving smart city goals. In Eindhoven the approach is taken to not include personal data in the program. Thereby, data that can be valuable in monitoring and forecasting safety on Stratumseind, such as data on the location of people in the street, is excluded. In The Hague, opening up data is one of the goals in the field of safety, which is central in their smart city program. In opening up data, as well as in sharing data between event partners privacy issues occur (RQ8).

Also in Almere and Zwolle privacy is an issue that influences the smart city program, but for these cases it does not influence the objectives (anymore). In Zwolle there were no clear objectives identified. In Almere, privacy used to be an issue in achieving the objective (RQ8). However, in this case, a solution was found by addressing all the demands of article 6 of the Data Protection Directive, and mainly purpose limitation, in the Gedragsrichtlijn StraatKubus (RQ9). Zwolle currently works on a comparable solution, by establishing Hygiëneregels. Eindhoven tries to avoid personal data and The Hague is still in the beginning of the process of identifying privacy issues.

Four different cases, four different problems and four different approaches. However, there is one major shared problem in privacy: a lack of clarity. All cases, with exception of Almere, still have many questions on privacy issues relating to personal data. It is not evident where the border between personal and non-personal data is, it is not clear what is and is not allowed with personal data and there is no clear guideline on how to process personal data lawfully (RQ7). Almere experienced these issues in the past, and successfully developed a solution: formal documentation to justify personal data use, addressing purpose limitation and all demands of article 6 of the Data Protection Directive. And this is the only solution that works (RQ9).

On the downside, Almere also deals with major limitations. Only a small number of people is allowed to work with the StraatKubus data. Furthermore, there is no space for a learning process on data combinations and flexibility. In Zwolle, also a formal approach is taken, but these elements are also taken into account (RQ10).

Altogether, the ideal situation would be to combine the methods of Almere and Zwolle. Almere, because there is no other way to address purpose limitation and all other rules of article 6 of the Data Protection Directive than by doing this in a formal way in which all choices made are justified. But, it is desirable to include more space for innovation and flexibility, and avoid getting stuck in rules (RQ10).

It is expected that many municipalities will work with personal data, since the function of the municipality is to be close to Dutch citizens and focus their policies well. Especially

by Almere and Zwolle this idea was highlighted. Thereby, from finding very easily understandable, concrete and applicable solutions for personal data use, municipalities and citizens would benefit. By the cases it is agreed upon that solutions should not be searched in changes in law. Two suggestions to work on achieving a good solution and establishing more clarity in the field of personal data use are done (RQ10):

- Establishing a panel of stakeholders. In order to discuss and brainstorm on possibilities and best practices for a mix-up of the formal way to enable processing personal data lawfully and at the same time include possibilities for innovation and flexibility.
- Communicating the outcomes in clear and easily understandable documentation. To help municipalities processing personal data according to the rules and address the lack of clarity

9.2 Final conclusion

To what extent does the European Data Protection Directive and the ePrivacy (Amendment) Directive influence data use, and in particular the use of big, open and geo-data, for monitoring and forecasting public safety in smart city programs in the Netherlands and how can difficulties be overcome?

To conclude, European data protection and privacy directives influence data use for monitoring and forecasting public safety in smart city programs mainly in the field of purpose limitation. It is difficult to define smart city purposes into enough detail to enable processing personal data and to justify personal data use. Since many data is personal data, and personal data indeed is being used in all four cases, purpose limitation and the other demands of article 6 of the Data Protection Directive should be met. There is no other way to do this other than the formal way. This was being done in Almere, but there are also downsides on the approach. More flexibility and space for learning and innovation are missing in this solution. The main privacy problems experienced by the four cases deal with a lack of clarity. There are many questions and difficulties in finding the answers. Difficulties cannot be overcome by changing the law, but by trying to find a mixture of addressing all demands of the law and including more flexibility and space, and fore mostly by communicating easily understandable, concrete and applicable solutions. This can, for example, be done by establishing a panel of stakeholders to think through the possibilities of a mix-up and by communicating the outcomes in clear and easily understandable documentation.

10. Reflection and scientific recommendations

Every research inevitably has its limitations. This research is an important contribution to both theoretical and practical knowledge on the role of privacy protection in data use for smart city purposes, but in scientific research it is very important to be critical. Therefore, in this chapter, a critical reflection will be provided. Thereafter, scientific recommendations will be provided. Coming up with practical recommendations was part of the main question and for this reason are these recommendations discussed in paragraph 8.6 (p.60).

10.1 Reflection

The main limitations of this research can be found in its methodology. In a methodology many important choices are made and this automatically also leads to limitations.

First of all, remarks on the case study selection can be made. Between the case studies, there are large differences. Beforehand, it was expected that there would be differences between smart city programs, since the smart city concept covers a wide range of subthemes. To overcome this problem, the subject of safety was selected. Public safety is in many cities among the main smart city objectives (although for Zwolle this was less evident). In all four case studies it was part of a broader program, in which also other smart city themes are included. This resulted in difficulties in interviews to find out if, for example, data is being used for public safety or for other purposes. In practice, this was a difficult distinction to make, especially if there were no overviews in policy documentation available, which was the case for The Hague and Zwolle.

The large differences in case studies can also be explained by the fact that a snowball method was used to select them, because by searching the internet, case studies were difficult to find. A difficulty arised in the fact that THE smart city does not exist and therefore the focus was on smart city projects. Even though cities above 100,000 inhabitants were selected, there still were large differences in the way smart city programs are being seen and implemented. The largest difference in organisation is between The Hague and Zwolle, which is also logical because The Hague is four times larger than Zwolle, in terms of the number of inhabitants. In The Hague, the organisation is fragmented and this led to an interview with little depth and details compared to the other interviews. The Hague profiles itself as smart city, while Zwolle has a more practical approach and does not market itself as smart city. These (organisational) difference thus became evident in the interviews.

Second, more general remarks on the methodology can be made. The choice for a qualitative research method is a logical decision, because, among other reasons, it is explorative in nature. But for the case studies, it turned out that also data on facts was needed. In interviews, it is difficult to, for example, find out which data is being used and how the smart city objectives are exactly formulated. It was tried to address this downside of the qualitative research method by studying policy documents. Especially for the concrete cases of Eindhoven and Almere, this was available and it turned out to be very useful. Therefore, it was easier to come up with clear statements for these cases, compared to the more vague smart city programs of The Hague and Zwolle. Including the policy documentation analysis proved thus to be a valuable addition to the interviews. The interviews in the broader smart city programs were difficult to structure, since there was no detailed information available and it became difficult to keep focussing on data and privacy protection for monitoring and forecasting public safety.

For each case study, one interview was held with a project leader (with exception of Zwolle, where there was no project leader identifiable). In this research, the opinions of the interviewees are important, but because there was only one interview held for each case study, opinions become attached to the cases themselves and possible errors in information provided do not become easily evident. In practice it was not possible to interview more people for each case, and the project leader is a logical choice since this person has knowledge on the range of topics that are addressed in this

research. In Zwolle, an interview was held with two interviewees at the same time, and this might be a way to overcome this problem and gain information on multiple thoughts within one case. On the other hand, on the main outcomes of the interview that are used for this research, both interviewees agreed upon each other. It is thus expected that for the outcomes of this research, the fact that only one respondent was interviewed does not have a large influence.

The last methodological remark that should be addressed is on the processing of interview data. The interviews were voice recorded, which can affect the honesty and openness of the interviewees. However, the voice recordings were only used to make a report from the interview results. Voice recordings enabled the interviewer to pay full attention to the conversation, and since there are no citations used in the text, it is expected that the negative effects are compensated for. Furthermore, the fact that the interview and policy documentation reports were first written in Dutch and then translated to English, might have the consequence that minor details in data are lost. It is expected that this does not influence the outcomes of this research, since the main findings will not be lost in this translation.

The main problem in the field of privacy that was found in the case studies is the lack of clarity. The problems in a lack of clarity cannot only be found in these cases, but also by the researcher. It became evident that interpreting the law and translating the interpretation to easily understandable and workable outcomes is not easy, not even for professionals. It remains difficult to make statements on this subject, but it is tried to explain all choices well and to provide nuances if possible. Examples of important but difficult choices in which the law was interpreted are the conclusion that the StraatKubus processes personal data according to all rules set in the Data Protection Directive. In the field of purpose limitation, questions could arise in legitimacy, since this is a difficult concept to decide upon. Another example is the decision that the police recordings that are used in Eindhoven are personal data. The municipality does not seem to see these data as personal data, but by taking into account the definition of personal data, it can be stated that the police recordings that are currently being used are personal data.

With the outcomes of this research, suggestions for improvements were done in order to bridge the world of the smart city and municipal policies, and the world of law. These suggestions are no more than ideas that are based on the findings from the case studies. It should thus be kept in mind that there might be more and other options to address problems, but a first step towards addressing the privacy issues identified in this research was taken.

Furthermore, in the choices made in what is and is not in the scope of this research, limitations can also be found (see paragraph 1.3.3, p.10).

It is expected that there are more critical remarks that can be made on this research, but the most important ones are addressed in this paragraph. Every research has limitations and it is important to be aware of them. But overall, this research is an important contribution to knowledge in the influence of European privacy protection for using data in the smart city, on both a scientific and practical level. It hopefully contributes to addressing the issues that were identified.

10.2 Scientific recommendations

This research contributes to new knowledge in the field of the influence of privacy protection on data use in the smart city. It can serve as a base for other research in this field, especially because of its extensive theoretical component. Further research can be done on the possibilities in law, in order to establish a balanced solution between meeting the formal requirements for processing personal data and including flexibility herein, as was suggested as the perfect option to overcome problems. It would be useful to address this topic from the background of law. Furthermore, since privacy protection is expected to be fully taken care of on European Union level in the future, it is interesting

to research the influence of this change on personal data use for smart city purposes. At this moment in time there are too many uncertainties in the Draft General Data Protection Regulation, but in the future this can become an interesting subject for further research. Furthermore, research can, off course, be done on the influence of privacy protection for other smart city themes than public safety. Last but not least, further research in the field of ethics can be conducted. This research focuses on the influence of law, but it would be interesting to research this topic from the angle of ethics. As was explicitly stated by Zwolle: we currently are in the beginning of the ethical discussion, and the juridical discussion always follows the ethical one. Last but not least, it was not in the scope of this research to come up with suggestions for changes in law, but this could be studied in further research. Even though the four included case studies do not see helpful short-term options in changes in law, they agree upon the fact that the law is lagging behind.

11. References

Analysys Mason Limited (2010), Public safety mobile broadband and spectrum needs – Report for the TETRA Association. UK: London.

Article 29 Working Party (2007), Opinion 4/2007 on the concept of personal data. WP 136.

Article 29 Working Party (2011), Opinion 13/2011 on geolocation services on smart mobile devices. WP 185.

Article 29 Working Party (2013), Opinion 03/2013 on purpose limitation. WP 203.

Atzori, L., A. Iera & G. Morabito (2010), The internet of things: a survey. Computer Networks vol. 54, pp. 2787-2805.

Bakici, T., E. Almirall & J. Wareham (2013), A smart city initiative: the case of Barcelona. Journal of the Knowledge Economy, vol. 4, pp. 135-138.

Batty, M. (2013), Big data, smart cities and city planning. Dialogues in Human Geography, vol. 3, no. 3, pp. 274-279.

Batty, M., K. Axhausen, F, Giannotti, A. Pozdnoukhov, A. Bazzani, M. Wachowicz, G. Ouzounis & Y. Portugali (2012), Smart cities of the future. The European Physical Journal-Special topics, vol. 214, pp. 481-518.

Bartoli, G., R. Fantacci, F. Gei, D. Marabissi & L. Micciullo (2013), A novel emergency management platform for smart public safety. International Journal of Communication Systems, doi: 10.1002/dac.2716.

Bélissent, J., C. Mines, E. Radcliffe & Y. Darashevich (2010), Getting clever about smart cities: new opportunities require new business models. Cambridge: Forrester research.

Bettencourt, L. (2013), The uses of big data in cities. Sante Fe Institute working paper.

Boeije, H., H. 't Hart & J. Hox (2009), Onderzoeksmethoden. Den Haag: Boom Lemma.

Boisen, M. (2007), The role of city marketing in contemporary urban governance. Paper presented at the conference: Future of Cities: Impacts – Indicators – Implications.

Calabrese, F., C. Ratti, C. &K. Kloeckl (2009), WikiCity: Real-Time Location-Sensitive Tools for the City. Handbook of Research on Urban Informatics: The Practice and Promise of the Real-Time City, pp. 390-413.

Caragliu, A., C. Del Bo & P. Nijkamp (2011), Smart cities in Europe. Journal of Urban Technology, vol. 18, no. 2, pp. 65-82.

Chourabi, H., T. Nam, S. Walker, J. Gil-Garcia, S. Mellouli, K. Nahon, T. Pardo & H. Scholl (2012), Understanding smart cities: an integrative framework. Proceedings of the 45th Hawaii International Conference on System Sciences (HICSS).

Cohen, B. (2012), What exactly is a smart city? http://www.fastcoexist.com/1680538/what-exactly-is-a-smart-city [cited October 15, 2014].

Columby (2015), Open Data Portaal Gemeente Den Haag. https://beta.columby.com/explore/organisation/gemeente-den-haag [cited February 5, 2015].

Convention 108 (1981), Convention for the Protection of Individuals with regard to Automatic Processing of Personal Data. http://conventions.coe.int/Treaty/EN/Reports/Html/108.htm [cited November 4, 2014].

Costa, L. & Y. Poullet (2012), Privacy and the regulation of 2012. Computer Law & Security Review, vol. 28, pp. 254-262.

Cuijpers, C. & B.-J. Koops (2013), Smart metering and privacy in Europe: Lessons from the Dutch case. In: Guthwirth, S., R. Leenes, P. de Hert & Y. Poullet (2013), European Data Protection: Coming of Age. Dordrecht: Springer Science and Business Media. Chapter 12, pp. 269–293.

Cuijpers, C. & P. Marcelis (2012), Oprekking van het concept persoonsgegevens beperking van privacybescherming? Computerrecht, vol. 13, pp. 339-351.

De Hert, P. & V. Papakonstantinou (2012), The proposed data protection regulation replacing directive 95/46/EC: A sound system for the protection of individuals. Computer Law & Security Review, vol. 28, pp. 130-142.

De Hert, P., V. Papakonstantinou, D. Wright & S. Gutwirth (2013), The proposed Regulation and the construction of a principles-driven system for individual data protection. Innovation: The European Journal of Social Science Research, vol. 26, no. 1-2, pp. 133-144.

Dempsey Morais, C. (2012), Where is the Phrase "80% of Data is Geographic" From? http://www.gislounge.com/80-percent-data-is-geographic/ [cited February 3, 2015].

Dictionary.com (2014), Buzzword. http://dictionary.reference.com/browse/buzzword [cited November 11, 2014].

ECHR (2010), Convention for the Protection of Human Rights and Fundamental Freedoms. http://www.echr.coe.int/Documents/Convention_ENG.pdf [cited November 4, 2014].

EU Directive 2002/58/EC (2002), Directive 2002/58/EC of the European Parliament and of the council of 12 July 2002 concerning the processing of personal data and the protection of privacy in the electronic communications sector (Directive on privacy and electronic communications). Official Journal of the European Communities, L 201, pp. 37-47.

EU Directive 2003/98/EC (2003), Directive 2003/98/EC of the European Parliament and of the council of 17 November 2003 on the re-use of public sector information. Official Journal of the European Union, L 345, pp. 90-96.

EU Directive 2007/2/EC (2007), Directive 2007/2/EC of the European Parliament and of the council of 14 March 2007 establishing an infrastructure for spatial information in the European Community (INSPIRE). Official Journal of the European Union, L 108, pp. 1-14.

EU Directive 2009/136/EC (2009), Directive 2009/136/EC of the European Parliament and of the council of 25 November 2009 amending directive 2002/22/EC on universal service and users' rights relating to electronical communications networks and services, Directive 2002/58/EC concerning the processing of personal data and the protection of privacy in the electronic communications sector and Regulation (EC) No 2006/2004 on cooperation between national authorities responsible for the enforcement of consumer protection laws. Official Journal of the European Communities, L 337, pp. 11-36.

EU Directive 95/46/EC (1995), Directive 95/46/EC of the European Parliament and of the council of 24 October 1995 on the protection of individuals with regard to the processing of personal data and on the free movement of such data. Official Journal of the European Communities, L281, pp. 31-51.

European Commission (2011), Communication to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions. Open data: An engine for innovation, growth and transparent governance. December 12 2011. Brussels: COM(2011) 882 final.

European Commission (2014a), Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee of the Regions. Towards a thriving data-driven economy. Brussels: COM(2014) 442 final.

European Commission (2014b), Digital agenda for Europe–Open data. http://ec.europa.eu/digital-agenda/en/open-data-0 [cited November 19, 2014].

European Commission (2014c), Justice - Data protection - Commission proposes a comprehensive reform of the data protection rules. http://ec.europa.eu/justice/newsroom/data-protection/news/120125_en.htm [cited October 17, 2014].

European Data Protection Supervisor (2012), Opinion of the European Data Protection Supervisor on the data protection reform package. Brussels: March 7, 2012.

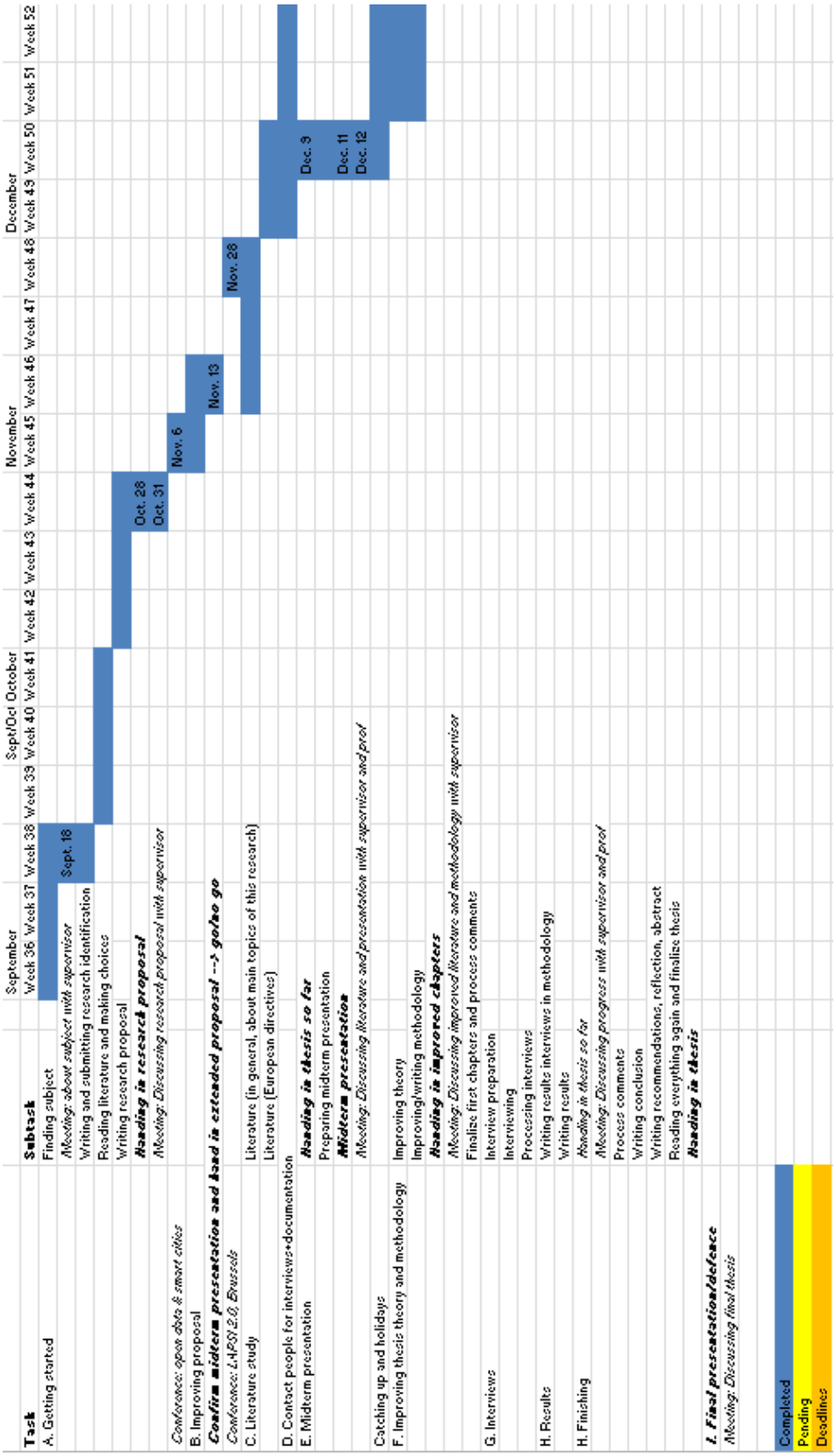
Filipponi, L., A. Vitaletti, G. Landi, V. Memeo, G. Laura & P. Pucci (2010), Smart City: An Event Driven architecture for Monitoring Public Spaces with Heterogeneous Sensors. Proceedings of the 2010 Fourth International Conference on Sensor Technologies and Applications, Venice, Italy, 18–25 July 2010, pp. 281–286.

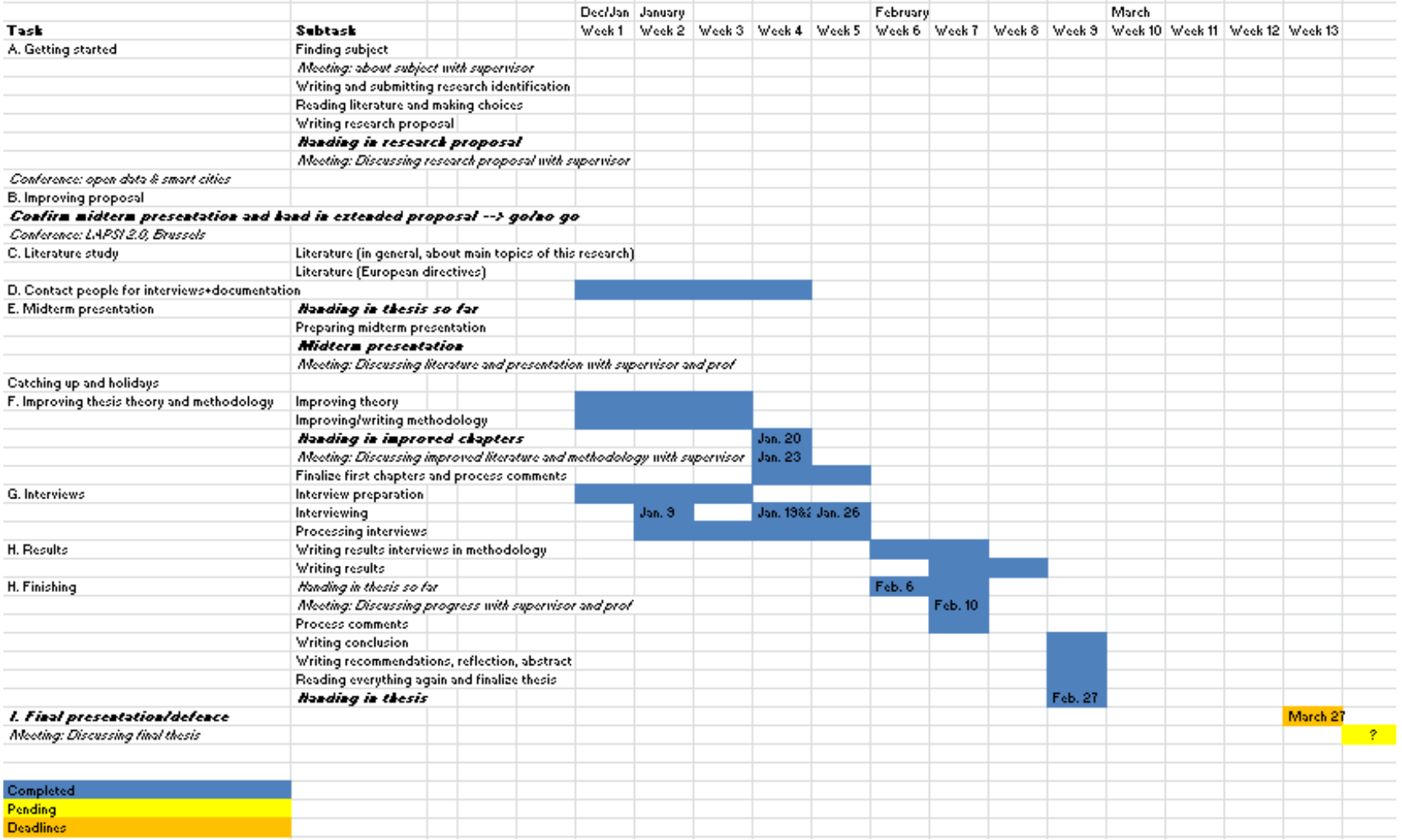
Gemeente Almere (2014), Gedragsrichtlijn StraatKubus, horende bij de samenwerkingsovereenkomst StraatKubus. Almere, October 2014.
Gemeente Den Haag (2014), Road Map Smart City Den Haag: Samen naar een Slimme Stad. The Hague, March 2014.
Gemeente Den Haag (2015), Den Haag in Cijfers. http://www.denhaag.buurtmonitor.nl/ [cited February 5, 2015].
Gemeente Zwolle (2015), Geo informatie portaal. http://www.geopoort.nl/geo-online/ [cited February 24, 2015].
Giffinger, R., C. Fertner, H. Kramar, R. Kalasek, N. Pichler-Milanović & E. Meijers (2007), Smart cities: ranking of European medium-sized cities. Vienna: Centre of Regional Science (SRF).
Gil-Garcia, J., T. Pardo & A. Aldama-Nalda (2013), Smart cities and smart governments: using information technologies to address urban challenges. Proceedings of the 14th Annual International Conference on Digital Government Research.
Gutwirth, S., R. Leenes, P. De Hert & Y. Poullet (2013), European data protection: coming of age. Dordrecht: Springer Science and Business Media.
Hall, E. (2000), The vision of a smart city. Proceedings of the 2nd International Life Extension Technology Workshop, Paris, France, September 28, 2000.
Hand, D., H. Mannila & P. Smyth (2001), Principles of data mining. United States of America, Massachusetts: MIT Press.
Harrison, C. & I. Donnelly (2010), A theory of smart cities. Proceedings of the 55th Annual Meeting of the ISSS.
Hoevenaars, R. (2013), Blog: Boyd Cohen: the smart city wheel. May 14, 2013. http://www.smartcityevent.com/boyd-cohen-the-smart-city-wheel/ [cited October 22, 2014].
Hollands, R. (2008), Will the real smart city please stand up? Intelligent, progressive or entrepreneurial? City, vol. 12, no. 3, pp. 303-320.
Jara, A., D. Genoud & Y. Bocchi (2014), Big data for smart cities with KNIME a real experience in the SmartSantander testbed. Softw: Practice and Experience, DOI: 10.1002/spe.2274.
Kamp, H. (2014), Kamerbrief over big data en profilering in de private sector. Den Haag: Ministerie van Economische Zaken.
Kanters, T. (2013), Living Lab, onderdeel van Stratumseind 2.0. Smart sensors, smart interfaces, smart actors, smart lights, smart data, smart design, augmented reality, gaming. PowerPoint Gemeente Eindhoven.
Kanters, T. (2015), Living lab and smart light project. PowerPoint Gemeente Eindhoven.
Komninos, N., H. Schaffers & M. Pallot (2011), Developing a policy roadmap for smart cities and the future internet. Proceedings of the eChallenges e-2011 Conference.
Korff, D. (2002), EC study on implementation of Data Protection Directive. Annex 3: Report on the findings of the study. United Kingdom: Cambridge University.
Kronenburg, T., T. Monasso, E. Boschker & M. Thaens (2012), De waarde van open data: Keuzes en effecten van open-datastrategieën voor publieke organisaties. Den Haag: Zenc
Kulk, S. & B. Van Loenen (2012a), Brave new open data world? International Journal of Spatial Data Infrastructures, vol. 7, pp. 196-206.
Kulk, S. & B. Van Loenen (2012b), Open data and beyond: Exploring existing open data projects to prepare a successful open data strategy – deelrapport privacy. Delft: Onderzoeksinstituut OTB.
Larkou, G., M. Mintzis, P. Andreou, A. Konstantinidis & D. Zeinalipour-Yazti (2014), Managing big data experiments on smartphones. In: Distributed and Parallel Databases. New York: Springer Science and Business Media.
Lombardi, P., S. Giordano, H. Farouh & W. Yousef (2012), Modelling the smart city performance. Innovation – The European Journal of Social Science Research, vol. 25, no. 2, pp. 137-149.
Marshall, C. (2012), Big data, the crowd and me. Information Sciences & Use vol. 32, pp. 213-224.
Ministerie van Buitenlandse Zaken en Koninkrijksrelaties (2013), Visie open overheid. September 2013. Den Haag: Ministerie van Buitenlandse Zaken en Koninkrijksrelaties.

Nam, T. & T. Pardo (2011), Conceptualizing smart city with dimensions of technology, people, and institutions. Proceedings of the 12th Annual International Conference on Digital Government Research.
Naphade, M., G. Banavar, C, Harrison, J. Paraszczak & R. Morris (2011), Smarter cities and their innovation challenges. Computer, vol. 44, no. 6, pp. 32-39.
OECD (2012), OECD technology foresight forum 2012 – Harnessing data as a new source of growth: big data analytics and policies. http://www.oecd.org/sti/ieconomy/iccptechnologyforesightforum-harnessingdataasanewsourceofgrowthbigdataanalyticsandpolicies.htm [cited November 20, 2014].
OECD (2014) Data-driven innovation for growth and well-being. Interim Synthesis report. Paris: OECD Publications.
OHCHR (2015), United Nations Human Rights, Office of the High Commissioner for Human Rights. Universal Declaration of Human Rights. http://www.ohchr.org/en/udhr/pages/Language.aspx?LangID=dut [cited February 2, 2015].
Open Knowledge (2014a), The open definition. http://opendefinition.org/ [cited November 19, 2014].
Open Knowledge (2014b), What is open? https://okfn.org/opendata/ [cited November 19, 2014].
Papa, R., C.Gargiulo, A. Galderisi (2013), Towards an urban planners' perspective on Smart City. TeMA Journal of Land Use, Mobility and Environment, vol. 6, n. 1, pp. 5-17.
PDOK (2014a), Publieke Dienstverlening Op de Kaart – Nationaal Georegister. https://www.pdok.nl/nl/producten/nationaal-georegister [Cited December 7, 2014].
PDOK (2014b), Publieke Dienstverlening Op de Kaart. https://www.pdok.nl/nl [Cited December 7, 2014].
Peek, G.-J. & P. Toxler (2014), City in transition: urban open innovation environments as a radical innovation. Vienna: Real Corp 2014.
Penninga, F. & L. Siebert-Han (2014), Eindrapport Professionalisering Open Data Infrastructruur. Gemeente Den Haag, oktober 2014.
Philips (2014), Smart cities – Switch on to the latest thinking in city lightning. http://www.lighting.philips.com/main/smartcities/index.wpd [cited October 23, 2014].
Politie Den Haag (2014), Hoe veilig is mijn wijk. www.hoeveiligismijnwijk.nl [cited February 24, 2015].
RRAAM (2014), Staat van de Stad 2013/2014. Gemeente Almere, maart 2014.
Rijksoverheid (2014), Het opendataportaal van de Nederlandse overheid. https://data.overheid.nl/ [Cited October 15, 2014].
Schaffers, H., N. Komninos, M. Pallot, B. Trousse, M. Nilsson & A. Oliveira (2011), Smart cities and the future internet: towards cooperation frameworks for open innovation. In: J. Domingue et al. (Eds.): Future Internet Assembley, LNCS 6656, pp. 431-446.
Siemens (2014), Smart city – through intelligent automated infrastructure. http://www.siemens.com/information-technology/smart-city.html [cited October 23, 2014].
Special Eurobarometer 359 (2011), Attitudes on data protection and electronic identity in the European Union. Conducted by TNS Opinion & Social at the request of Directorate-General Justice, Information Society & Media and Joint Research Centre.
Steenbruggen, J., E. Tranos & P. Nijkamp (2014), Data from mobile phone operators: a tool for smarter cities? Telecommunications Policy, in press.
Team Zwolle (2014), Gebiedsscan 2013 Gemeente Zwolle. Politie Oost Nederland, District IJsselland, Team Zwolle. Zwolle, May 2014.
Telecommunicatiewet (1998), Wet van 19 oktober 1998, houdende regels inzake de telecommunicatie (Telecommunicatiewet). Law as was valid on November 13, 2014.
Tene, O. & J. Polonetsky (2013), Big data for all: privacy and user control in the age of analytics. Northwestern Journal of Technology and Intellectual Property, vol. 11, no. 5, pp. 239-273.
Turksema, R., P. Boers, M. Kingma & M. Schaefers (2014), Algemene Rekenkamer: Trendrapport open data. Den Haag: Algemene Rekenkamer.

Van Loenen, B. & M. Grothe (2014), INSPIRE Empowers Re-Use of Public Sector Information. International Journal of Spatial Data Infrastructures Research, vol. 9, pp. 86-106.
Van Loenen, B., J. De Jong & J. Zevenbergen (2008), Locating mobile devices, balancing privacy and national security. NWO Research report.
Van Oortmarssen, A. & M. De Vries (2014), Privacy op zijn plaats: tussen willen weten en wetten. Witboek over de spanning tussen privacyregels en het realiseren van het locatie-informatie potentieel. Amersfoort: Geonovum.
Wet bescherming persoonsgegevens (2000), Wet van 6 juli 2000, houdende regels inzake de bescherming van persoonsgegevens (Wet bescherming persoonsgegevens). Law as was valid on November 13, 2014.
Wong, R. (2011), Data protection: the future of privacy. Computer & Law Security Review, vol. 27, no. 1, pp. 53-57.
Zijlstra, T. (2014a), A visionary view on the future of PSI: the big picture on open data. Or: some patterns, thoughts and ideas. Presentation at LAPSI Conference Brussels, November 28, 2014.
Zijlstra, T. (2014b), Open (gov) data. The what, why and how in 50 examples or less. Guest lecture open data, October 30, 2014.

Appendix A: Time schedule




Dec/Jan
January
February
March
Task
Subtask
Week 1
Week 2
Week 3
Week 4
Week 5
Week 6
Week 7
Week 8
Week 9
Week 10
Week 11
Week 12
Week 13
A. Getting started
Finding subject
Meeting: about subject with supervisor
Writing and submitting research identification
Reading literature and making choices
Writing research proposal
Handing in research proposal
Meeting: Discussing research proposal with supervisor
Conference: open data & smart cities
B. Improving proposal
Confirm midterm presentation and hand in extended proposal --> go/no go
Conference: LAPSI 2.0, Brussels
C. Literature study
Literature (in general, about main topics of this research)
Literature (European directives)
D. Contact people for interviews+documentation
E. Midterm presentation
Handing in thesis so far
Preparing midterm presentation
Midterm presentation
Meeting: Discussing literature and presentation with supervisor and prof
Catching up and holidays
F. Improving thesis theory and methodology
Improving theory
Improving/writing methodology
Handing in improved chapters
Jan. 20
Meeting: Discussing improved literature and methodology with supervisor
Jan. 23
Finalize first chapters and process comments
G. Interviews
Interview preparation
Interviewing
Jan. 9
Jan. 19&2
Jan. 26
Processing interviews
H. Results
Writing results interviews in methodology
Writing results
H. Finishing
Handing in thesis so far
Feb. 6
Meeting: Discussing progress with supervisor and prof
Feb. 10
Process comments
Writing conclusion
Writing recommendations, reflection, abstract
Reading everything again and finalize thesis
Handing in thesis
Feb. 27
I. Final presentation/defence
March 27
Meeting: Discussing final thesis
?
Completed
Pending
Deadlines

Appendix B: Topic list

Topic list (translated from Dutch to English)

Introducing the research and reason for interview:
Research on data use for smart city purposes and the influence of privacy regulation. Specific: programs focusing on using data to improve public safety: monitoring and forecasting.

General info:

- On the smart city program and the role of the interviewee.

Goal:

- Smart city program?
- Goals smart city program?
- Sub goals?
- Officially formulated? → Documentation?
- The role of public safety in the smart city goals
- Difficulties in formulating a goal?
- What are the difficulties?
- *Providing information:* in order to work with personal data, a clearly formulated goal must be provided. The smart city is a vague concept and is insufficiently specified. Are you aware of this? Do you think that the goals are specific enough to work with personal data? Are there problems? How did you try to tackle this issue?

! To keep in mind: focus on public safety and monitoring & forecasting.

Open data:

- Is open data being used?:
  - produced
  - used
- Personal data?
- Is there a specific goal of open data within the smart city program? Purpose limitation?
- For how long is data being kept?
- Is data being updated?
- Is a risk analysis being performed/has the impact of opening up data been identified?

Big data:

- Is big data being used?
  (real-time, sensors, internet linkages, chips, mobile phone (location) data?
- Personal data?
- Is there a specific goal of data within the smart city program? Purpose limitation?
  Is data being used for a different purpose than the original one?
- For how long is data being kept?
- Is data being updated?
- Are people informed about data collecting? How?

Geo-data:

- Geo-data (which data, proportion)?
- What does location add up to a dataset?
- Does geo-data make data easier personal data?

Exceptions:

- Are there exceptions to the rules? Mazes in the law? Ways to get around these rules?
- *Providing information:* In the law it is stated that data use in order to prevent, detect and prosecute criminal activities can be an exception to the rules. Does this apply to the smart city program? Is it being used?

Privacy in general:

- How is privacy being coped with?
- Did you ever experience problems in privacy?
- How do you check whether or not something is allowed?
- Is there any control? Do you have experience with it?
- How would you judge the influence of privacy regulation on data use for your smart city program?
- Do you experience privacy regulation as an obstacle? Wherein?
- What would you change in legislation? What would you like to see differently?
- Is there enough space in current legislation in order to achieve the programs objectives/to do your job well?

Future:

- What do you expect from the future? (general, data use for this kind of programs, privacy)?
- What is the role of privacy therein?

Practical:

- Is there documentation available on the goals?
- Is there an overview of data used available?
- Do you have contacts with people working on smart city programs in other cities that I can contact for interviews?
- Can I contact you if I have more questions or if I have forgotten to ask about something?
- I am making a report out of this interview. Could you check if I understood everything well?
- Are you interested in the final version of my thesis?

! To keep in mind: In essence there are two major problems:

1. Open data & big data being personal data, cannot go with data protection:

   Problem PERSONAL Open data:

   - Not gathered for specific purpose (by definition, aimed at re-use – OPEN data).
   - Unclear amount of time the data is being kept for.
   - Is an impact assessment being done before data becomes open data?

   Problem PERSONAL Big data:

   - Tempting to use for another purpose than the original one.
   - Unclear amount of time: tempting to use longer than strictly necessary.
   - Difficult to inform people about their data being collected.

2. Smart city: vague purpose. Purpose limitation is a demand for working with personal data.

Appendix C: Interview Eindhoven

Interview Eindhoven

Friday, January 9, 2015, 2.00-3.15 PM, Living Lab De Oude Rechtbank, Eindhoven

The first interview was held with Tinus Kanters on January 9, 2015. Additional documentation will in some cases also be used (Kanters, 2013 ; Kanters, 2015). The text in this appendix is the full report of the interview, and only contains the opinions and way of seeing things of the interviewee (with exception of additional documentation that was included).

Introduction of the respondent

Tinus Kanters is working at the municipality of Eindhoven and is the project leader of Stratumseind 2.0 and the Living Lab. He is working on the integration of new technologies and social media in 24/7 measurements in Stratumseind, Eindhoven.

Introduction of the case

Stratumseind is a street full of bars and pubs in the centre of Eindhoven. A couple of years ago, it was estimated that ten to fifteen bars would need to close. Nowadays, four bars have closed and ten bars do not pay rent. Rent often goes to the brewery and they would rather pay for the rent themselves and have the bar still sell their drinks, then the bar would close and not sell their drinks anymore. This means that altogether, one million euros a year are missing. This situation needs to be changed.

The overarching project towards a new Stratumseind is called Stratumseind 2.0. Recently the smart light project was started in which light is being used to influence the ambiance in the street. The lights were turned on for the first time the day before the interview. In the next months, also projects on smell, gaming and changing the scene of the street take place.

All these projects are monitored in the Living Lab, located on Stratumseind, which can be seen as a thermometer that measures the effects. It is "*an 'instrument' to measure influences of interactions (light, smell, design). Smart sensors, smart interfaces, smart actors, smart lights, smart data, smart design, gaming"* (Kanters, 2015, p.13).

Objectives and the role of safety

Objectives

The overarching goal of Stratumseind 2.0 is: "*Together with all partners, as entrepreneurs, breweries, property owners, police, City Council, we will structurally improve and increase the economic and social functioning and activities on the Stratumseind. This structural improvement will have three main-themes: Safety, Liveability and Attractivity"* (Kanters, 2015, p.12).

For the smart light project, the goals are to turn Eindhoven in a vibrant city and a sustainable city, each divided in three components. The creation of a sustainable city also focuses on sustainable social contacts and taking care of each other. In the smart light project it is tried to influence the ambiance on Stratumseind, by making use of lights. If people are very excited, aroused and astonished, this can easily turn into an alarmed, tense or angry state of mind. People should then become more relaxed. But people can get sad, miserable, frustrated and bored and then a shift towards a happier state of mind is necessary. In laboratory, tests showed that people's mood can be changed by lights. Not only the subject itself is most deciding for someone's state of mind, but its surrounding is determining as well. By working with light, the surroundings can be influenced. The lights are attached to all eighteen streetlights in Stratumseind. The colours can easily be changed. It is known that people get angry from white lights. For the police, this colour is handy, since they have a good view on the street. But with the light project it is tried to find out if other colours have a better influence on people and decrease, for example, the amount of incidents.

Eventually the light system should become self-regulating. But, an operator will always be needed. It will become possible to make an automatic system, but then important questions arise

such as: who is making the final decision and who is responsible?

Supply based or demand based

The smart city is a vague term. It is often interpreted in what technologies can best be used for maintenance, increasing safety and (health)care in a city. Stratumseind 2.0 and the smart light project fit into this smart city story. The citizen is central in this project. Technique helps in taking care of them. Eventually, what is being done with the help of techniques should be central. Often, developments in this field are supply based. People then want to work with new techniques, simply to try them out. This used to be the case in Eindhoven too. But nowadays it is agreed upon that new techniques are not the reason to do things; they must be useful. So, both supply and demand should meet each other, which is the case for the smart light project.

Data

A lot of data is being collected in the Living Lab (Kanters, 2015):

1. Temperature, sun, rain and wind.

The temperature is being measured with sensors near Stratumseind. It is not open data. In the near future, three sensors will be placed on Stratumseind, to see if there are differences. If there turn out to be little or no differences, these sensors will be removed again.

2. Sounds (3D).

Sensors detect the loudness of sounds and nowadays also the direction of where the sounds come from are known. Current techniques enable to detect not only the amount of sound, but also what is being said, but then problems with privacy occur. Only the level of sounds is being measured and this can be shown in graphs and can then be interpreted. In the past, sound data of what is being said was saved, but this is no longer allowed. It was useful to detect the origins of sound. Not only people make noise, but for example the church bells and road sweepers make noise too.

3. Bluetooth and Mac addresses.

These can be measured, but are not saved since they cannot be made anonymous or aggregated yet.

4. Cellphone counting.

Data from Vodafone is being bought. About 30.0% of all people uses Vodafone and this is enough to be representative. This data is not real-time and takes two or three days before it is available. The data is already anonymised. Tinus says this is data that people often find a bit scary in terms of privacy. Vodafone made an agreement with the CBP on the aggregation of numbers. Problems with data leading to individuals only occur when different datasets can be combined. Anonymity is also protected by the fact that place of origin is only registered when there are at least fifteen telephones from the same municipality.

5. Video people counting.

Camera's used to save a blurred image of what's going on in the street. But even in the blurred image it was still possible to recognize people. Therefore the images are no longer used, and since half a year only the amount of people is registered. On the camera's there is software installed that is able to recognize people and to count them. Furthermore, their direction can be registered. It is not needed to know who it is, only the numbers are relevant. If there is a fight, for example, this is a case for the police. The police has different rules for privacy and is allowed to focus on the individual. All five entrances and exits of Stratumseind have cameras with video counting software. Furthermore, cameras can count the number of people per m2. Stratumseind is not the only location where people are counted using cameras. In the hallway of the train station, for example, people are being counted as well.

6. Light amount (Lux) and colour (Kelvin).
7. Events calendar of Stratumseind.
8. Police recordings.

Only the numbers for different categories of crime are registered (see figure 10, p.39). If there are small numbers of people involved, this can be sensitive data. The registrations are anonymised, but are available for a time span of two hours. For Stratumseind numbers on insults, threats and assaults

are most interesting. These are the categories with the highest numbers of recordings and are likely to be influenced by using lights. By working with lights the number of stolen bikes or drug dealing will not be limited and these categories are not so interesting because of the limited number of incidents.

9. Social sensors (media watching and interactive).

Social media messages are being gathered. The messages are not saved, but it is registered if messages were positive, negative or neutral. This is being done automatically, by software that checks words that are known to be negative or positive. Until now, no real interpretation is available and thereby about 75.0% of all messages are neutral. Atos and Intel are working on improving this process and try to add semantics to the software.

10. Brewery registrations.

Register how much is being brought to Stratumseind in general, so not for each pub or bar.

11. Waste/energy.

The amount of waste collected, entered manually into the system.

12. Open data from the municipality.

This is a new development and not much is being done yet.

13. Survey results of residents.
14. Car and bike parking information.

On all five entrances/exits of Stratumseind there are lampposts with sensors. All data is brought together in the Living Lab. There is an interface available, in which all data is shown in easily readable way (see figure 8, p.37). All data is being communicated with open standards. If a sensor is not interesting anymore, it can easily be replaced.

Data is only interesting if it leads to intelligence. Therefore, many partners are involved in Stratumseind 2.0., including both academic institutions (Tilburg, Eindhoven and Amsterdam) as well as companies and institutions (Philips, Munisense, Icen, Geodan etc.).

Location data

By the Vodafone data it is known where someone originates from. Every unique telephone is given a code and it is determined where this telephone was most of the time in the past weeks. This location is being seen as the place of origin. So, no addresses or conversation data are being used. To determine the location, the location of cell towers is being used and this system has an accuracy of about 200 meters (see figure 9, p.38).

There are possibilities for a more accurate determination of locations. The newest technology in this field is small cells. There are small Wi-Fi connectors on the lampposts and if someone connects with them the connection will be good, and the accuracy of location determination becomes about 25 meters. Another method is iBeacons for iPhone. This is an application that sends text messages about the direct environment, such as: do you see this shop? You should check it out! In Eindhoven this application can be used, since its marketing organization Eindhoven 365 developed it.

Until now, the amount of people is being counted, not where they are. Tinus states he would like to have more information on their locations as well. He would like to see if there are parts of Stratumseind that get crowded. This could be useful data when combined to for example, noise data, and a fight when the red lights were turned on. If the week thereafter, the same situation occurs but then with green lights and no fights, this could mean that the light experiment works.

Furthermore, the sensors are location based and are displayed in a map in the interface (see figure 8, p.37).

Open data

At this moment in time, open data is not being used yet. Eindhoven is working on making data openly available for citizens. This is data that is being gathered to improve the functioning of the government, that has been aggregated. Examples are data on population growth and the amount of trees, for example. Currently, a portal is being made.

Data that would be relevant for the Stratumseind 2.0 project would, for example, be data on the amount of bicycles parked near Stratumseind. It is expected that these data are included in the system at the end of February.

In the near future, there is a meeting to discuss how data can remain anonymous if it is opened up. Also, questions on safety and the location of servers need to be answered. These are new issues that need to be addressed.

Personal data

In general, there is no personal data being used in this project. Sometimes, the boundaries of personal data are being reached and perhaps crossed. Camera images of people are an example. At first it was said that it would be enough to blur the camera images. But then people were still recognizable. If someone is wearing a red jacket, for example, you could go out and find this person easily. For this reason it was decided to place the people counting software directly on the camera and not on a computer. This way, only the amount of people is being registered and no video images can be seen anymore.

Social media is another example of touching the boundaries. You can easily see what people write about Stratumseind. But if you would save these data, it would become possible, for example, to find out how many times someone said something negative. But, on the other hand, if someone posts a message on social media, it is this persons own choice to publish it. As soon as these texts would be combined with information that this person did not provide, such as an address, it would become a different story. It is often more of an ethical question than a juridical one.

Tinus states he tries to stay away from personal data as much as possible, to make sure there will be no problems with privacy. It does not stand in the way of the project, since the focus is on the mass and not on individuals. The ambiance on Stratumseind needs to be improved and this goal focusses on the collective.

Privacy

There are multiple organisations involved in the field of privacy, such as the CBP, TILT, Tilburg University, Ministry of Internal Affairs and the Ministry of Justice (Kanters, 2015). This is a difficult subject to deal with, since it cannot be directly read from the law. Since Stratumseind 2.0 does things that haven't been done before in the Netherlands, other organisations are helping with the privacy issue. The law does not cover all things being done in Stratumseind. Technological developments take place at rapid pace and laws are not able of keeping up.

Tinus spoke to companies that were willing to sell their personal data. There are possibilities to gain data that is so specific that it is possible to determine which bars someone is visiting. But especially as a governmental organization, this cannot be done. The main problem in privacy is caused by commercial companies. Governmental organizations will try to stick to the rules and will not have the intention to misuse information. Large companies don't care. It is difficult to go against this and maybe the only method is for consumers and clients to take action and not accept this. Companies are often afraid that if they do something bad with personal data and this goes public, clients will go to competitors.

Tinus has not been warned about privacy, neither by the CBP nor by municipal lawyers. The problem is that even they find it hard how to deal with privacy. If you start collecting names, addresses and pictures it is obvious that it's not allowed to do that. The problem is that there is a large area in which it is unclear if something is allowed. In the case of the camera images, it was probably not allowed. But in fact it is only not allowed if the images would be saved. It would be great to know for sure if something is allowed. There is also a difficulty in combining datasets.

The future and improvements

Changing the law is not an option, since this takes too long. And if the law was changed, it would soon be outdated again. Tinus says that since the start of the Living Lab, about a year ago, a lot has changed in this one year.

It is important that people take care of data storage and ensure that it remains impossible to de-anonymize data by combining datasets. But if someone wants to misuse data, this will always be possible somehow. In electronics, it is often the case that people in data management and protection are lacking knowledge compared to other specialists such as hackers. Developments are so new and occurring at rapid pace, that it is difficult to protect data.

Appendix D: Interview Almere

Interview Almere

Monday, January 19, 2015, 0.05-1.25 PM, Stadhuis Almere

The second interview was held with Thijs van der Steeg on January 19, 2015. This is the first interview that makes use of the topic list (see appendix B, p.77). Additional documentation will in some cases be used (Gemeente Almere, 2014 ; RRAAM, 2014). The text in this appendix is the full report of the interview, and only contains the opinions and way of seeing things of the interviewee (with exception of additional documentation that was included).

Introduction of the respondent

Thijs van der Steeg is strategical advisor in the municipality of Almere, in 'Programmabureau Stad'. He is project manager of Almere 2.0. In the municipality, the needs of the inhabitants of the municipality are central; a bottom-up approach is adopted. Many colleagues are field workers in neighbourhoods of the municipality of Almere. There are also tasks that need to take place on an overarching level and Thijs, among others, works on these tasks. One of these tasks is to work on a new information management system that enables to monitor the city on a detailed level.

Introduction of the case

A new information management system was developed: the 'StraatKubus'. It is a geographic information system that serves as an early warning tool in signalling liveability issues at an early stage (RRAAM, 2014). The system is useful to monitor the city and work on prevention of problems in liveability. The StraatKubus is a tool that enables to link data to developments in the city and thereby monitor even small changes. It is easy to work with the tool and by using it, statistics/data can be linked to the observations from municipality (field) workers. The tool can be used on a daily basis to answer questions and can be used in conversations and discussions with colleagues about developments in the city.

The StraatKubus is part of an overarching program: Almere 2.0. The number of inhabitants in Almere still is rapidly growing and to keep an eye on the growth of the city it was decided in a meeting with people working for the national government, the municipality and local partners, that the information management system needed to be improved.

Objectives and the role of safety

Monitoring the liveability and ultimately preventing liveability issues in Almere is the main goal of the StraatKubus. The StraatKubus should answer questions on liveability. It serves as a tool to use in conversations. It can be used to check presumptions with facts. Sometimes presumptions are confirmed, but it is even more interesting to see that many presumptions turn out not to be true when compared with facts. The StraatKubus is thus not only about using technique, but it is also a way of working in the municipality of Almere.

Liveability is a broad concept that contains a number of sub-concepts. Examples are the physical environment, quality of houses and data of citizens. Safety is part of liveability, but others might see safety as a separate concept next to liveability. Safety is also a large concept, containing different elements. Sub-subjects in safety are, among others, criminality, an (un)safe feeling or perhaps also subjects as unclean streets.

The fact that liveability (and safety) are large concepts, leads to problems in purpose limitation. Therefore, a behavioural guideline 'Gedragsrichtlijn StraatKubus' was established.

In the Gedragsrichtlijn StraatKubus, the goal of the StraatKubus is defined. "*The goal of the StraatKubus is to provide insight in the development of liveability. In order to do so, the StraatKubus combines data from the physical, social and safety domain and presents these data on 6ppc-level. By keeping an eye on trends and developments that can put liveability under pressure and by having*

conversations about it, local partners want to detect problems in an early stage. The StraatKubus is a tool in detecting liveability problems in an early stage by local partners. It is a communication tool that helps local partners to have conversations on trends and developments in areas. What gets attention and how should be worked together to address the problems identified?" (translated from Dutch to English) (Gemeente Almere, 2014, p.8).

In Article 13 of the Gedragsrichtlijn StraatKubus it is stated that data can be included in the StraatKubus if they concern one of six (main) themes: "*1) population and housing characteristics 2) socio-economic situation and poverty (prevention) 3) Safety and bonding 4) social support and welfare 5) environmental incidents/reports of nuisance in public space 6) developments in housing values* (Gemeente Almere, 2014, p.10). In articles 14, 15 and 16, subthemes and data are further specified. Subthemes relating to safety are: reports of nuisance in subsidised housing complexes, reports of nuisance in public space, burglaries and neighbourhood mediation/conflicts between neighbours. Thematic spatial data that can be included in the StraatKubus, relating to safety, are social cohesion and intercultural intercourse, reports of public space, reports of nuisance in subsidised housing complexes and reports on physical nuisance (dirt dumping, graffiti, vandalism, odours, dust, noise, parking spots, trees) (Gemeente Almere, 2014).

Data

Data in the StraatKubus thus concerns the themes and subthemes listed in articles 13-16 of the Gedragsrichtlijn StraatKubus (Gemeente Almere, 2014).

Open data

The StraatKubus does not contain data that are openly accessible. It would be great for citizens to access data themselves, but this is impossible. People tend to speak too easily about opening up data. It is important to keep in mind how open data actually is allowed to be. Data can only be opened up if it is anonymised and generalised, because open data cannot be personal data because of privacy legislation. In the StraatKubus, data is available on a detailed level and it helps the municipality to address societal challenges. With the StraatKubus, a greater purpose can be served then by opening up generalised data. Data from the StraatKubus is thus not publicly available.

On the other hand, the StraatKubus does make use of open data. The municipality starts working on integrating open data into the StraatKubus. One of the open data sources is data from Kadaster. Kadaster will start publishing more data online. At this moment in time there is open data from the Centraal Bureau voor de Statistiek (CBS, Statistics Netherlands) in the StraatKubus. Integrated CBS data is data that was published online by the CBS about four years ago, but was removed soon afterwards.

Open data are thus integrated in the StraatKubus, but the StraatKubus does not publish open data.

Geo-data

All data in the StraatKubus are geo-data. The StraatKubus is a geographic information system. All data are mapped on 6ppc-level (6 digit zip-code level). This level provides detailed information. If data would be aggregated to the level of neighbourhoods or districts, data would become too generalized for people to do their job focused. If something is going wrong in a small part of the neighbourhood and if the rest of the neighbourhood is doing fine, this cannot be seen in the data on neighbourhood level. The 6ppc level enables field workers to do their work focused. This amount of detail has the consequence that data becomes too close to becoming personal data. For this reason, StraatKubus data can only be used by professionals that need these data in order to do their job well. Within a geographic area in the StraatKubus, there is a threshold value of five needed, otherwise data is not shown. This is an extra security guarantee. The CBS uses the same system. If there is for example only one person that receives social security payments, it is not difficult to find out who it is. In this case, data leads too easy to individuals and this is not allowed in the StraatKubus.

Big data

In big data, patterns are sought in an enormous amount of data. Data becomes interesting if it is possible to research patterns in a protected way and to find out what data is relevant in a certain pattern and what is not relevant. Thereafter, about 90.0% of data can be removed. The risk with big data is that it might get in the wrong hands, while the majority of data is not even needed. There is no big data in the StraatKubus.

Privacy

Target group

Only professionals are allowed to work with the StraatKubus. This is because of purpose limitation and the self-imposed rules of the Gedragsrichtlijn StraatKubus. An important question of the Wbp is: why does information need to be combined and who needs this information in order to do his or her job well? Therefore, the municipality had to decide who should get access to the StraatKubus data. It was decided that professionals of the municipality of Almere and housing corporations should be allowed. This was a difficult and long debate. It would be great to share information with more people, because of the function of the municipality in society. This way data could be used to start conversations with, for example, inhabitants. Lawyers decided that this is simply not allowed and only municipal professionals and housing corporations can work with the data. Access is denied to all other people. For example, people working for employment agencies, banks or bailiffs are denied access. So, the StraatKubus is not available for citizens, but it is also unavailable for professionals and private parties. The following question needs to be answered: does someone need the data in the StraatKubus in order to do his or her job well? For these groups, the answer is: no.

Privacy risks

The StraatKubus needs to deal with privacy regulation because of the spatially detailed scale of 6ppc-level. At the level, information leads for too many people to individuals. On the other hand, this amount of detail is needed for professionals to do their work well.

The data in the StraatKubus belong to privacy category 2. Category 1 should be seen as data that doesn't lead to privacy issues. Category 3 data is very sensitive data and is not included in the StraatKubus.

Process

It turned out to be extremely difficult to find out a correct way to cope with these privacy difficulties. Help was needed, but it was missing. Boundaries, criteria and a way of reasoning are unclear. Lawyers could explain what the law states, but couldn't clarify its boundaries. Also technicians were asked to help, especially to answer questions about classifications such as ISO standards and NEN-norms. For this case too, it turned out to be very difficult for these experts to come up with a workable answer. It was then decided to start working on the Gedragsrichtlijn StraatKubus. A lawyer was approached to do this job, but the amounts of money needed were unacceptable. Thijs therefore started to work on it himself, with help from a lawyer.

It is a pity that so much effort was needed to keep working according to rules set in the law. The municipality simply wanted to work on liveability issues and then needed to cope with unclear regulations and little help is being offered. Especially since the municipality is part of the Dutch government, it cannot process data unlawfully. The municipality of Almere even went to the Dutch parliament to talk about the issues and went to the CBP to discuss what Almere tries to achieve. People often shy and overlook why Almere wants to work with the StraatKubus: to work on creating a better Almere, not to match data that is not needed or relevant for the StraatKubus' goal. The CBP never warned Almere officially, but they asked some questions after a certain news item popped up.

The cause of the fact that many people could not help solving the privacy issue with the StraatKubus, lies in the fact that apparently people don't know what they are talking about. On congresses and symposia people tend to speak superficially about it, but as soon as you try to find real answers, people get quiet. Specialists often say every case needs a different approach.

Furthermore, people find it difficult to think about this issue from a domain that is not their own.

Gedragsrichtlijn

Purpose limitation is included in the Gedragsrichtlijn StraatKubus. All concepts are described individually and the value of their connections is described as well. This seems to be a good approach. If there are developments in the city in the field of liveability, the StraatKubus can provide insight in data this development relates to and thereafter people can start working on it. This is written in the Gedragsrichtlijn. Everything thereafter needs to relate to the StraatKubus' goal, such as who is allowed to use data, authorisation, publication. As soon as there is doubt, it is not allowed to use the StraatKubus.

A CBP advisor suggested starting with writing down the purpose of combining data. In the Gedragsrichtlijn, it is described what the StraatKubus is, what its purpose is and what is needed. It is a limitative enumeration of data. The data that can be included in the StraatKubus is written down and if another dataset would be added, it is necessary to first change the Gedragsrichtlijn.

The Gedragsrichtlijn looks like a contract. Important terms are clarified, general terms of use are explained and the process of authorisation is clarified (Gemeente Almere, 2014). If data are uploaded into the StraatKubus, all Wbp demands are taken care of.

Coping with purpose limitation mainly is a matter of starting to write down what you wish to do and why: start peeling down the main goal. Thijs started to work on this in the Gedragsrichtlijn. It takes some time and is not very exciting work, but it is very important that this basis is well taken care of.

Re-use

If more data is needed, people can be asked for permission to use their data for other purposes. One can inform people about the other purposes and explain why their data is needed. It is normal to work responsibly with personal data, but someone needs to start working on dealing with purpose limitation.

It can be tempting to upload data in the StraatKubus that is gathered for a different purpose. In many situations, it is not allowed to use data for another purpose then the original one. If you would like to include data into the StraatKubus, permission is needed to do so.

The future and improvements

The Wbp should be read a bit better. It is important to realize why this law was established. The law itself could be improved, but that's not the most important thing to do. It mostly is a matter of starting to work on writing down your goals and needs. You need to ask yourself a number of questions. A law will never be capable of doing this, but there is a reason for the establishment of a law. The Wbp tries to protect personal data. What does the law try to protect and why? Purpose limitation is the thing you need to start working on. A broad or vague law can provide possibilities, but at the same time you are a citizen of this country too and protection is important. Once you start working on privacy issues, you realize how much data is gathered by companies and institutions. Companies with commercial interests are likely to work differently with data and privacy compared to governments, but this only is a presumption. A vague or broad Wbp is not necessary, but it can use some updating. It should be more focussed on the network society we live in nowadays and it could provide a checklist, for example, to enable people to adapt the law easier.

Improving the situation in the future is not in changing laws, but in the fact that people simply don't want to do time-consuming and boring jobs such as making a Gedragsrichtlijn. It is important to start doing new things, but this must be done in a careful manner. Everyone that tries to combine data should go back to the core: what is necessary and what is not.

In general, privacy protection will not be an obstacle. It is about taking good care of personal data and being able to explain what it is being done. It is no more than decent to do so. There is enough space to do nice things with data, but you will have to put some effort in it.

Appendix E: Interview The Hague

Interview The Hague

Thursday, January 22, 2015, 1.20-2.10 PM, Gemeentehuis Den Haag

The third interview was held with Hedwig Miessen on January 22, 2015. Additional documentation will in some cases be used (Columby, 2015 ; Gemeente Den Haag, 2014; Gemeente Den Haag, 2015). The text in this appendix is the full report of the interview, and only contains the opinions and way of seeing things of the interviewee (with exception of additional documentation that was included).

Introduction of the respondent

Hedwig works as program manager at the Bestuursdienst department of the municipality of The Hague. She has been working on the development of the Smart City Road Map, mostly on providing an ICT fundament. This is called 'ICT voor de Stad', of which Hedwig is program manager. The ICT voor de Stad program already exists for about ten to fifteen years and nowadays contributes to the smart city vision and roadmap. Furthermore, Hedwig has a national role in the Digitale Steden Agenda (DSA) and works on the Safe City theme (Veilige Stad) in this overarching program.

Introduction of the case

Recently, the ICT voor de Stad program was adapted to the smart city concept. By working with both public and private institutions, the Smart City Road Map was developed. In the Road Map, smart city goals for the period 2014-2018 are described (Gemeente Den Haag, 2014). The department of Economics (part of Urban Development) of The Hague municipality is also involved in this project, following the example of Amsterdam and Eindhoven.

The ICT voor de Stad program did not become more 'smart city' than it already was. However, it now focusses more on cross-sectorial and cross-departmental sharing of data and ICT. This increases efficiency and opportunities for data use. Creating a more efficient system is not easy, since the municipality of The Hague is a large and organisation and works very compartmentalised and fragmented. Standardisation is very important. Data currently is being shared mostly within a department, not between them. If data is more widely available, this creates more opportunities, especially when it is combined with open and big data and the internet.

To address this issue, basic ICT facilities are needed and this is where Hedwig works on in the ICT voor de Stad program.

Objectives and the role of safety

For the period 2014-2018, the goal of Smart City The Hague is "*to boost the competitiveness, the quality of life and the sustainability of The Hague, by working together with representatives of Triple Helix by working with pilots and projects, making innovative use of technology*" (translated from Dutch to English) (Gemeente Den Haag, 2014, p.2). The Hague Smart City is needed because of the strategic importance of ICT and technology and this is of economic interest. Furthermore, more efficiency should be achieved: 'doing more with less'. Also, high technology and telecom offer opportunities (Gemeente Den Haag, 2014).

In the Smart City Road Map there are eight themes identified for The Hague Smart City. Two of them are marked as priority themes: Quality of Life and Safety. It is stated that The Hague wants to be leading on these two themes and facilitates innovation. For 2018, two main goals are identified in the field of safety: "*The Hague is leading on ICT use and open data use for safety (worldwide) and The Hague is the Silicon Valley of safety*" (translated from Dutch to English) (Gemeente Den Haag, 2014, p.2). The Hague is the city in which the Dutch national government is located, as well as the international court of justice, and hosts many international organisations and meetings (Gemeente Den Haag, 2014).

ICT and open data are being seen as important for safety. Therefore, The Hague Smart City

contributes to projects working on these aspects. The safety cluster of The Hague is taken care of in the Hague Security Delta (HSD). The Hague Smart City works closely together with the HSD. There are four projects identified: event safety, protection of vital infrastructure, cyber security and protection of the international zone.

In the field of ICT, it is stated that it is important to share data cross-sectorial. The Smart City Road Map tries to achieve this by working on opening up and linking open data (cross-sectorial), by enabling digital communication in the city and by working on the infrastructure of knowledge, in order to make the most out of facilities and services (Gemeente Den Haag, 2014).

The municipality of The Hague focuses on event safety. The focus of organisations involved in safety is for a large part on ICT. Safety is a broad concept and it depends on the person what is included in this concept and what is not.

Event safety

Multiple organisations are working together in the safety domain, such as the municipality, event organisations and emergency services. In events it is necessary to share data between partners and they all have an interest in sharing. This is the reason the municipality focuses on events.

The police often like to gain data from the municipality and event organisation, but do not share much data themselves. This is due to privacy.

Data is brought together in an event cloud, for a limited amount of time. This is being done for multiple events, such as Queens day/Kingsday, SAIL and the gay parade. Data in the cloud is not opened up for everyone, but only for the event partners. Crowd control is an important feature in event management. On this topic, much information is shared. There is, for example, data on the use of mobile phones and the amount of people. Mobile phone data is made available by providers to the event organisation. It is known when squares need to be closed and people are informed about routes they can take. In these situations, privacy is not a big issue.

Data in event clouds is only shared for a limited amount of time. People are currently working on a network to share data on a structural level. Future Events is developing a public-private platform to share data structurally. This network can be used in events to share information from multiple sources. For years there have been experiments in sharing data for a small period of time. Data is thereafter analysed and as soon as the event is over, the data is not interesting any more. Data might be kept to get back to it if it turns out to be interesting after all. The main problem is that the system for sharing is removed after an event.

The risks in events are high, since there are many people involved. But citizens value events highly positive and important for the attractiveness of a city. For multiple parties there is an interest in events and their safety, such as the municipality, police and event organization, but also for bars and restaurants. Events are thus a nice platform for innovations in the field of safety and sharing data.

Data

Most data on safety originates from the police. But there also is data available that is gathered by city guards and the neighbourhood watch. The municipal department of Openbare Orde en Veiligheid (Public Order and Safety, part of Bestuursdienst) owns most safety data. A large percentage of data comes from the police. These data are not publicly available. In The Hague there is a focus on event safety, but there is also more general data on safety available. For the organisations of events, maps are mostly used. In the smart city not much data is shared between partners yet, both private and public organisations, with the exception of maps.

Open data

There is data openly available on safety. Hedwig refers to the website www.hoeveiligismijnwijk.nl. It is an open platform of the The Hague police in which crime rates are made available. The municipality developed an open data portal ‘Den Haag in Cijfers’ (Gemeente Den Haag, 2015). Liveability and safety is one of the categories.

Open data sub categories that relate to safety are crime and nuisance, opinion on neighbourhood (safety monitor), common nuisance (safety monitor), sense of security (safety monitor) and social cohesion (safety monitor). However, data is in most cases only available on the spatial level of The Hague (so not on neighbourhood levels, for example) and displays the overall number or percentage for the whole city. The data is thus not detailed (Gemeente Den Haag, 2015).

Besides this viewer, data is also published on Columby, a platform for distribution of open data (Columby, 2015). Eventually 8,000 data sets will be published on this platform, but until now only a few are made available. The process of opening up data has recently started speeding up. This mainly is because of the fact that some key persons, such as statisticians and researchers, opened up their data. These people also collect data from other sources, and if there are no privacy issues, these data are opened up as well.

The process of opening up data has been speeding up recently. What would really help is if there would be someone that stood up for it, or there would be more money available. If citizens would start showing a lot of (economic) interest, such as software developers, that would be very helpful. One of the problems herein is that the Netherlands is a small country, and if data is only available on a local level, it is from a commercial point of view not interesting to start developing applications or software. Standardization is thus very important. This way, data on local level can become regionally or nationally available. There are some agreements on standardisations, but more effort is needed. Geographic data is leading in this process.

Privacy

A decision tree is being used to determine if it is possible to open up data. Lawyers are involved in this process. In general, lawyers tend to say that it is better not to open up data as soon as there is doubt about it. This is an interesting process since this is a new development and it is yet unknown how society will react on opening up data. It will be interesting to see if there will be lawsuits, and how judges will decide in them. Opening up data also has to do with the courage of people responsible. Some will decide to just go for it and others need to know for sure if opening up is allowed. Until now this is a subject that needs more attention. The municipality recently started working on the demands of the Wbp. It is likely that in many cases it will lead to not opening up data, due to privacy protection.

There is more data that leads to issues than would be expected at first sight. Some colleagues find it scary to open up data, because citizens can easily check what is being done.

For every dataset that might be opened up, the decision tree is used. Hedwig is concerned with the preparations for a system to do this in; a structural facility. A lawyer helps to decide in this process. The ministries of Economic Affairs and of the Interior and Kingdom Relations are contacted about how to deal with this issue. But it turned out that these ministries had also just started working on it. As a municipality it is difficult to find out what needs to be done all by yourself, but you will have to.

In order to do this, structural facilities need to be established and this will be done in the next four years. The process thus just started and until now there was little money available to start working on it. It was not on the priority list of the municipality. Eventually it will be established, but the amount of time needed is unclear.

The police have the most data on safety but do not want to share. This indicates that privacy is an issue that matters. It is important to take care of privacy because personal data cannot become publicly available. Also anonymising data should be taken good care of.

The future and improvements

Eventually, a solid network of safety data should be established. This takes time. Hedwig compares it to the Dutch water network: on multiple spots people started creating dikes and polders and creating water ways and eventually a national network started to come into existence. Especially the national police would be an important partner in creating such a system, but because of privacy they don't share much data. For safety it would be interesting to know, for example, which people could be a

danger for society. The police often know, but because of privacy legislation they are not able to do something with their information. In events, sometimes it would be great to share more information and forget about privacy for a moment. It would be great to let go of all the privacy rules for a moment and to agree with all partners on removing all data immediately after an event is over. Especially if things go wrong in an event, privacy would be something you would like to forget about for a moment.

In the future it would be valuable to start working together with other event organisations, such as the organisations of TT Assen and the Nijmeegse Vierdaagse. The municipality of The Hague does not have the capacity to monitor everything well.

Appendix F: Interview Zwolle

Interview Zwolle

Monday, January 26, 2015, 3.10-4.20 PM, Hanzelaan Zwolle

The fourth interview was held with Marcel Broekhaar and Jaap Pleeging on January 26, 2015. Additional documentation will in some cases be used (Team Zwolle, 2014). The text in this appendix is the full report of the interview, and only contains the opinions and way of seeing things of the interviewee (with exception of additional documentation that was included).

Introduction of the respondents

Marcel is geo-information advisor at the department of Research and Information at the municipality of Zwolle. On this department, data is being used to answer questions. He is also involved in the Living Lab for the Internet of Everything of Geonovum. Herein he researches the relevance of big data for society.

Jaap is policy advisor on the department of Societal Development at the municipality of Zwolle. He focuses on monitoring. He states it is important to be asking the right questions and he then thinks of how data could help answering them.

Both Marcel and Jaap are involved in an internal data platform that is concerned with all kinds of data and also of regulations. The alderman requested to write down rules on how to deal with privacy: 'Hygiëneregels'. In the data platform the possibilities of using data are discussed. This involves a wide range of topics such as applications, consequences, policy and communication.

Introduction of the case

Municipalities used to execute policies coming top down from the national government. This changed recently and nowadays municipalities are freer in their policies and have more responsibility and (financial) risks. That's why it has become more important to check the efficiency of policies. Data can be used to check this. There is lots of data available, but the municipality is still working on creating the right culture, infrastructure and competencies to make use of data. But this process is rapidly developing. There are two main developments. The first one is on improving the data itself. The amount and quality of data, but also editing options and the speed of registration are developing. The second one is adapting political and policy processes to technological developments and to the changing task of the municipality. This results in many projects and pilots.

A product is often needed to answer a question. The Research and Information department is also responsible for data gathering and distribution. There is a central (geo-)database and some data is publicly available. It is not needed to collect all data yourself; much is already available. This is increasingly becoming the role of the department.

Zwolle is not a 'hard core' smart city. The municipality is doing well, but much more can be done. That's one of the reasons to start the data platform. Other municipalities view Zwolle as leading in both geo-information as well as the process side. In the municipality, people from different departments and domains started working together. Geo-information, research, statistics and policy were thereby linked together and this leads to utilization of each other's strengths.

Zwolle used to be ahead of developments. By developments in technique and knowledge it became possible to develop new products. There was no quest for these products yet, but Zwolle started to develop. As soon as there was a quest, Zwolle was able to come up with solutions. The development of the map table is an example. You always start with the question and then you start peeling it down into smaller parts and see how data and technique can contribute to answering it. Finally, you need to think about how to present it. Technique is only useful if you're able to connect it to policy and politics.

The focus in Zwolle is not on being a smart city, but on using smart solutions if they are useful for answering a question. Zwolle makes use of sensors, for example. But sensors don't make Zwolle a

smart city yet. Only Eindhoven can be considered a real smart city in the Netherlands. However, the Stratumseind 2.0 project only focusses on a small part of the city. Zwolle doesn't focus on technique itself, as Eindhoven does, but does what is needed and then decides what techniques can contribute to it. Thereby also initiatives and potential of the local community is being used; an important part of the smart city concept.

Objectives and the role of safety

Data is being used to check the effects of municipal policies and to monitor society. With the new role of the municipality it has become important to gain knowledge on the effects and on developments in the municipality. There are multiple ways to do this. First, there is prevention. It is important to know what indicates a negative change and to take action. Second, the effectiveness of policy can be measured. This is an example of the connection of technique and policy. Geo-information is important herein. In the new municipal role it is important to support choices with facts. Therefore it is needed to invest in data. This is not an easy task and many questions will arise. Examples are: How do you guarantee data quality? How to build a structure for data access? How to deal with privacy? These questions will be answered in the Hygiëneregels.

The municipality is not the only organisation working on these issues. It is important to keep an eye on others and sometimes to work together. The municipality has become a facilitator. It is the conversation where it is all about. Data can be input therein and help to create a shared reality.

In one of the neighbourhoods of Zwolle, Holtenbroek III, there were issues in safety. By the municipality council it was decided to take extra measures. CCTV was being used, as well as preventive searching. The council decided on doing this for as long as specific monitoring data remained negative. The competency of supervisory authorities was thus linked to the values of data. To take action like this is only being done in excessive cases. Holtenbroek III started to score on national ranking lists of murders and therefore it was decided to take these measures. But this only happens under strict conditions. It is needed to find a balance between safety risks and privacy. If data are personal data, benefits in the field of safety should be evident. In the case of severe criminality or heavy nuisance, which was the case in Holtenbroek III, and is also in vandalism in soccer games, for example, extra measures can be taken.

In general, Zwolle is a safe city. In large events such as soccer matches or Liberation day data is being gathered to control the mass. Sometimes measures can be used for multiple purposes. For example, in the city there are Wi-Fi access points. They could also be used for measuring the amount of people for crowd management, but this is not being done.

In protecting safety there are roles for the municipality as well as for justice and police. They work together on safety protection. Eventually, the mayor is responsible for public order.

Data

Big data

There are sensors used in the city, to measure air quality and traffic movement. The availability of parking spots in garages and the data from traffic nodes are connected. Traffic is being navigated by making use of data. There is an external party involved herein. In the field of public safety no use of sensor data is being made.

Monitoring

In the field of safety, monitoring data are being used. Most of them originate from a survey among citizens, a neighbourhood survey and data from the police. In the case of Holtenbroek, a few indicators were selected. If their values remained negative, additional safety measures were taken.

Most of the data on safety originates from the police. But it is not easy for the municipality to use their data, because of a difference in focus. The police are focussing on the prosecution of individuals, while the municipality is interested in the preconditions of a situation. In the 'Gebiedsscan' there is data on the police registrations (Team Zwolle, 2014). This document is at the

base of meetings of the three involved parties. As soon as it is tried to get more detailed information, the police are becoming more protective. It is important to find the right Hygiëneregels in the interaction between police and municipality.

In the Gebiedsscan there is data on numbers of crimes committed. In some cases the number of crimes committed per category is low. Locations are not mentioned, besides in hotspot maps. In the Gebiedsscan, going out safely, events and soccer is listed as one of the five priorities (Team Zwolle, 2014).

Geo-data

In discussions it is estimated that about 80.0% to 90.0% of all data is geo-data. The percentage of geo-data in the municipality of Zwolle is unknown but will be similar to the estimates. In the municipality no distinction is being made in geo-data and other types of data. Data is being seen as needed to do useful things, whether or not it is geo-data is not important.

Privacy

Personal data

The municipality owns a lot of individual data, for example data on social security payments and social assistance (social services). Sometimes there are problems with someone and by speaking to people you can gain an understanding of what's going on. This information in the hands of someone else, for example an inspector, who uses this information to hand out fines, can be used for the wrong reasons. It is thus needed to take care of a good information system within the municipality, without violating privacy unnecessarily. How can goal and means be combined in a decent manner?

Opportunities for matching data have developed at rapid pace. If something is effectual for someone, it will not be for another. There is a big discussion about data use. If you combine data on the probability that someone will commit a crime, you cannot go there and take measures. But what is possible, is to start the conversation with someone. It often is a matter of balancing on the lines between what is allowed and appropriate and what is not. How far can you go using technique? When is something becoming unethical? If, for example, a child dies in a family that is known to have issues in child abuse, the municipality would 'know' that something like this could happen, but is not able to take measures. If this would happen, major discussions on privacy would be held. Nowadays childcare is taken care of in municipalities, and the focus is on prevention. But how can this be done? You have detailed data but what is allowed and what is not and what parts of data exactly are relevant? In a few years, it will be possible to develop algorithms to predict developments. Large companies already are working with them. But it will always be only a part of the full story that can be taken into account by using data.

Hygiëneregels

Privacy is becoming more and more of a complicating factor. The alderman therefore decided that the Hygiëneregels should be established. Every civil servant knows that privacy should be taken care of, but to make sure that everyone knows about it and deals with it in a correct way, the Hygiëneregels are established. They are currently being worked on, and until they are finished, it mainly is a matter of using good judgement.

Privacy should not be seen as an obstacle. It is important to keep an eye on privacy. Privacy determines the boundaries of the playing field of data. But it is a dynamic concept that remains difficult to capture in general rules. Privacy should be seen as one of the risks you need to work with, such as financial or image risks are too. You need to gain experience in estimating privacy risks. The Hygiëneregels are only a starting point and cannot go further than that.

At this moment in time, dealing with privacy in the municipality of Zwolle mainly is a matter of using good judgement. This is because of the technological developments that occurred rapidly and you're being confronted with. Laws will always be behind this process and will never be able to fully take care of privacy protection. The privacy discussion will go on for years. The question is if it is possible to find a balance between societal gains and the violation of privacy. What will be

acceptable by society is an important question.

The Hygiëneregels are thus mainly a reassurance for the municipal board. If something goes wrong, they will be blamed. It is important to do useful things, without causing trouble. The Hygiëneregels serve as a base and will not change much in practice, since people are already working on privacy in a good way. At first sight privacy seems to be an obstacle, but it could perhaps also be viewed as prove that the municipal board also is aware of the possibilities of working with data and new techniques.

Juridical or ethical discussion

The juridical discussion follows the ethical one. The outcomes of an ethical discussion are fixed in laws. The techniques have developed rapidly and we're now at the beginning of the ethical debate. Rules are always behind. That's also why you often end up in the ethical field, because the juridical one is outdated. If a lawyer is approached to help, it often is the advices not to do something, since then there will be no risk.

Almere versus Zwolle

The approach in Almere and Zwolle is not very different. The main difference is the context. In Zwolle everything is going fine and there are very little problems. In Almere there is a more tensed situation and if something happens, this influences for example the housing market and liveability, and economic problems might occur. In Zwolle there are little social problems, low unemployment, a relatively young population and lots of educational possibilities. In terms of technique the differences are not very large.

The future and improvements

It would help to create a space in which it would be possible to experience how to cope with privacy. The municipality gained more freedom in policy, since it is closer to the citizen. But it is difficult to use this feature if you are told to keep your eyes and ears closed. Information at the individual level is needed in order to do your work well. However, it is very important to do this in a good way. It would be fine if an independent party would watch what is being done and check that it is being done in a good way. What you don't want to happen is that you need to do it secretly. Space to experiment what is and is not useful is needed. For a large part the municipality will be able to decide on the outcomes, and in some cases it would be great to get help.

Many people tend to stay away from privacy-sensitive data. But in that way nothing will happen. Room is needed to discover whether something is useful or not. That way it can determine what data is needed to investigate a particular phenomenon, and to what extent this is allowed.

A safe environment, wherein you're assured that no harm is being done, would be ideal. That is part of the hygiene rules too. It would be nice to be able to find out whether a data link is useful. Often you will find disappointing linkages, but in other cases, interesting result will emerge. In this way, you have had the chance to explore them and to test them in a responsible manner.