GIS departments as potential promoters of big data driven innovation in policymaking

A case study of a GIS department of a regional governmental organisation in the Netherlands

Master Thesis

Elmer Boulonois
Master Economic Geography
Faculty of Geosciences
Supervised by: Egbert van der Zee

University of Utrecht
November 2018

Preface

I would like to thank Egbert van der Zee for his supervision, guidance, and support throughout this thesis project. Due to the novelty of the subject studied, finding a research focus and thesis structure that works proved to be challenging. However, thanks to the optimistic guidance provided by Egbert I managed to overcome these challenges. Furthermore, my gratitude goes out to all the members of Team GIS of the Province of Utrecht who made me feel part of the team, even though I did not directly cooperate with most of them. In particular I would like to thank Peter Hooiveld for his support during the internship. Next to being an optimistic and constructive discussion partner, Peter helped me navigate through the organisation, get in touch with potential respondents, and overall, proved instrumental in providing me with the tools necessary to gather the desired empirical material. When looking back on how the project evolved, simply saying I learned a lot would not do justice, as not only did I gain more (practical) knowledge on subjects such as big data, governmental organisations and policymaking, but also did I find a new passion which I hope to pursue further in my upcoming career.

Table of contents

Abstract

Driven by the fact that governments are falling behind in terms of big data use compared to the private sector – despite the great potential it holds for them, this study set out to explore whether GIS departments of governmental organisations are capable of promoting the adoption of big data driven innovations into the policymaking process. On the basis of literature derived from a variety of disciplines, four roles are proposed that GIS departments likely could fulfil within an organisation to foster the adoption process: an encouraging role, a supporting role, a directing role and a developing role. Drawing on a case study of a GIS department of a regional governmental organisation in the Netherlands, the study shows that GIS departments which form part of an enterprise GIS organisational structure, in their current state, are capable of successfully fulfilling two of the four roles proposed: the encouraging role and the supporting role. However, the study also shows that GIS departments are likely to still face several challenges before their promoting efforts will actually result in successful adoption of big data driven innovations in policymaking, such as gaining support from the top-management, creating a data culture within the governmental organisation, becoming more integrated into the policymaking processes, and acquiring more big data experience and expertise. To help them overcome these challenges, four suggestions are provided.

1 Thesis introduction

In today's information age, organisations are increasingly becoming aware of the value of data (Gandomi & Haider, 2015). At a continuous rate, an overwhelming amount of data is being collected, processed and analysed with the objective to discover new knowledge. Collection takes place through a multitude of new applications and devices that have been introduced to society and the economy in the last few decades, such as social media, websites, business software, mobile phones, smart house appliances, smart cars and industrial machines (Klievink et al, 2017). Since these devices and applications are often equipped with sensors and software, they capture information about their users and their surrounding environment. As a result, an unprecedented variety of voluminous and timely data is increasingly becoming available for organisations to extract insights from. This together with the corresponding advances in data storage, processing, analysis and visualisation techniques, led to the emergence of a new concept called: big data.

Even though for many scholars the phrase 'big data' raises mixed feelings due to the hype that surrounds it, the lack of a consistent definition and the way it is being used as a buzz phrase, most of them do agree that the underlying concept holds great potential (Giardullo, 2015). For both private and public organisations, big data offers several new opportunities with regard to marketing, research, health care, business processes and governance. Some even go as far as to say that big data is a revolution; one that holds the potential to significantly transform the economy and ultimately society for the better (Hashem et al, 2015).

Currently, private organisations are at the forefront when it comes to big data usage and development (Mayer-Schönberger & Cukier, 2013; Kim et al, 2014; Gamage, 2016). Ranging from large hard- and software developers like Google and Facebook, who created their entire business model around it, to financial institutions like Morgan Stanley, who use big data to better understand the market and thereby improve investment performance (Davenport and Dyché 2013; Groenfeldt, 2012). Through a process of what is often called 'big data analytics', these companies extract insights or 'business intelligence' from these new forms of data (Chen

et al, 2012). Insights as such are of great value to organisations since it allows them to make better-informed decisions.

While the private sector is making big strides in terms of big data development, the public sector seems to be falling behind (Mullich, 2013; Klievink et al, 2017). This is despite the great number of possibilities big data use is believed to have for the public sector. According to scholars like Kim et al (2014) and Bertot et al (2014), big data holds the potential to significantly improve a government's governance capabilities. Often mentioned values in this regard are improved efficiency, effectiveness, transparency, accountability and engagement (Jetzek et al, 2014). Moreover, big data use "can foster collaboration; create real-time solutions to challenges in agriculture, health, transportation, and more; promote greater openness; and usher in a new era of policy- and decision-making" (Bertot et al, 2014: 5). According to Clarke (2016) and Klievink et al (2017) however, it is dangerous to think that governments are likely to only benefit from big data. Because of big data's complex nature, its usage by governments could potentially have unforeseen negative consequences as well: i.e., by breaching the privacy and security of the public due to a data leak, by making bad policy decisions due incorrect data handling, and by making large IT-investments that may never pay off.

Although big data comes with a host of challenges and risks, most scholars do agree that because of its great potential, its use by governments should be encouraged (Bertot et al, 2014; Kim et al, 2014; Mullich, 2013). To properly do this, certain questions still have to be answered in more depth, such as why the use of big data by governments in most countries is very limited? And what then could governments do to increase their big data use at different governmental levels? The ultimate goal of this study is to provide more insight into the issues raised by these questions. However, the scope of these issues is still too broad and big for the extent of a single study alone. Therefore, the focus is narrowed down further.

First off, the study adopts an organisational perspective. It thereby focuses on the organisational challenges governments face when trying to adopt big data driven innovations. In doing so, it builds upon the foundations laid by a special issue of the scientific journal 'Information Systems Frontier' (ISF, 2017), which includes nine research papers with a focus on big and open linked data (BOLD) innovation in the public sector. Based on these studies, Janssen et al (2017) identified a multitude of factors that are of influence on the adoption of data-driven innovations by government organisations. When zooming in on these factors, several organisational necessities become apparent, such as the need for a data infrastructure that supports the use of big data, forward-thinking leadership that is aware of the added-value of big data, more public-private collaboration, and data innovations that are in line with a government's activities and main statutory tasks.

Second, the study focuses on the organisational challenges of adopting big data driven innovations in one of the core activities of governments: policymaking. In the context of evidence-based policymaking, multiple academics believe that the use of big data in the policymaking process would lead to better informed and more effective policies. Big data would offer new ways to make sense of the policy playing field, predict policy outcomes and prescribe policy decisions. Authors like Höchtl et al (2016) take it even a step further and argue that big data has the potential to significantly transform the policymaking process by offering new ways to evaluate and monitor previously taken policy decisions. To illustrate their ideas, Höchtl et al have proposed the big data-revised policy cycle, which will be discussed in more detail later on in this paper. What is important to understand now, however, is that with these new forms of policymaking also come new challenges in terms of organisation and management. In the literature, attention is paid to challenges such as the threat of increasing institutional complexity, the risk of breaching data security and privacy, and the achieving of a so-called data culture (Giest, 2017).

Last, the study pays specifically attention to GIS departments of governments and how they could potentially foster the adoption of big data driven innovations into the policymaking process. Optimistic estimates report that up to 80% of data stored in government databases contain a spatial component (Worral, 1990; Alrwais et al, 2015). Although one could question the legitimacy of such an estimate (no empirical evidence provided), it does show the significance of geospatial data for governments. As Alrwais et al (p: 1) point out, "a large share of a government's work revolves around geography; e.g. planning, facility management, taxation, property maintenance, crime analysis, environmental monitoring, etc.". It is therefore not surprising that the majority of governments around the world have a GIS infrastructure in place to make sense of their geospatial data and thereby improve their decision- and policymaking capabilities (Longley et al, 2010; Carton; 2007). Now, since on top of that GIS technologies are increasingly able to process big data thanks to new developments in the field (Li et al, 2016), the idea is that the departments responsible for the organisation's GIS might have the potential to function as promoters of big data driven innovation in the policymaking process. Especially since GIS departments often already form an integrated part of the organisation's operations (Somers, 1998), which might make them relatively more capable to foster big data driven innovation that is actually in line with the organisation's policymaking goals, needs and organisational structure. Klievink et al (2017) identified this necessity through the use of IT organisational alignment models, which pose that for IT-innovations to be successfully adopted, they must be aligned with the organisation's strategy, structure and activities. They learned that such organisational alignment is especially important for governments, because they have predefined goals and needs, and thus do not have the freedom like private sector organisations do to unlock new possibilities and enable new goals through big data use.

The possibility of GIS departments as promoters of big data driven innovation within governments seems not to have been noticed yet by the academic world however; especially with the additional focus on policymaking, it received little to no attention yet. The aim of present study is therefore to further explore whether GIS departments are indeed capable of fostering the adoption of big data driven innovations into the policymaking process. It does so by focusing specifically on governmental organisations in the Netherlands. The study seeks to answer the following research question: *To what extent could GIS departments of governmental organisations in the Netherlands foster the adoption of big data driven innovations into the policymaking process?*

What differentiates this study from those in ISF (2017), and in particular the one by Klievink et al (2017) who developed a framework to measure a government's big data readiness, is that it takes on a less generic approach to big data use in government organisations. By focusing on GIS departments of governmental organisations as possible promoters of big data driven innovation in policymaking, new avenues are explored, and thereby, a more detailed and concrete understanding is gained of the challenges governments face to increase their big data use.

To answer the research question, a case study was conducted of a GIS department of a regional governmental organisation in the Netherlands. The researcher was interning at this GIS department of the regional government for a period of roughly five months. Before the start of the internship, the management of the department knew about big data, but had not paid any serious attention to it yet. The start of the research internship corresponds to the first attempts to seriously explore the possibilities of big data use for policymaking, as from then on, the subject was put on the agenda and gained in attention from both management and personnel. This allowed the researcher to not only observe the early phases of the innovation adoption process, but also become part of it himself.

The remainder of this paper is divided into 7 chapters. In chapter 2, a foundation for the research is laid by discussing relevant literature. Attention is paid to literature on big data and the potential link with GIS, the usage of big data in policymaking and the adoption process of data-driven innovations in government organisations. In chapter 3, the methodological approach used is described. In chapter 4, attention is paid to the contextual backdrop of Team GIS of the Province of Utrecht, which forms the case of this study. In chapter 5, the current state of the Province of Utrecht in terms of big data use is described. In chapter 6, the findings on the efforts of Team GIS to explore roles to foster the adoption of big data driven innovations into the policymaking process are presented. In chapter 7, a discussion is held by reflecting on the findings based on the theory discussed in the literature review. At last, in chapter 8, the study is wrapped up with conclusions.

2 Literature review

2.1 The big data buzz; what is it about?

"What is big data exactly?" and, "could all large datasets be considered big data?". Questions as such, which are commonly asked by people not directly involved with the topic, are a good representation of the ambiguity that surrounds the phrase. As could be derived from the phrase itself, big data concerns large datasets. However, as the remainder of this section will show, such an explanation does not suffice. In order to fully understand the power of big data and why it is currently penetrating almost every realm of the economy, in this section a more detailed explanation will be provided.

2.1.1 Definition

Although big data as phrase is often interpreted in differing ways throughout the economy (e.g. see SAP, 2012), there is somewhat of a common understanding in the academic world. As coined by Laney (2001), most scholars make use of the so called 'three v's' as a framework to make sense of the concept, which stands for: volume, variety and velocity (Chen et al, 2012; Kwon et al, 2014; Gandomi & Haider, 2015). Laney originally drafted this framework to pinpoint the challenges in data management. Nowadays it is seen as one of the most concise and effective frameworks to describe what is meant by big data.

The first v: volume is used to describe the magnitude of data. As the word 'big' in big data suggests, it often concerns datasets that are large in size. What is considered large is relative and varies by multiple factors such as data type and time. In the case of data type, take for example the difference between large amounts of video and textual data. When expressed in bytes, video data is significantly larger in size than textual data. Furthermore, in the case of time, what may now be perceived as large, may not be in the future. The reason for this is found in the fact that capacity thresholds are still increasing thanks to developments in compression and data storage techniques.

The second v: variety is used to describe the structural heterogeneity in available data. Data comes in three forms: structured, semi-structured and unstructured. Nowadays, the fast majority of data gathered falls within the categories of semi-structured and unstructured. The less structured data is, the harder it is to analyse it. However, due to the technological advances in fields such as big data analytics and machine learning, organisations are increasingly able to successfully analyse data no matter its original structure. They often do this by combining datasets with varying structures or by systematically converting data into different easier to analyse formats.

The last v: velocity is used to describe the rate at which data are generated and the speed it is analysed and acted upon. Technological advances such as the internet, sensors and smartphones have ushered in an era in which large streams of data are being generated at an unprecedented rate. At the same time, the speed at which these large streams of data can be analysed and visualised experienced a substantial increase as well, thanks to advances in computing frameworks and processing technologies (e.g. parallel computing). As a result, new possibilities in terms of data analysis and visualisation opened up. One possibility that particularly gained attention is that of 'real-time', which allows analysts to make sense of data while it is being generated. This possibility to analyse data in real-time is one of the reasons why big data has become so popular, because it offers a whole new range of opportunities for multiple fields.

In addition, one other v is frequently mentioned in the literature as well: veracity (Gandomi & Haider, 2015). The term veracity points at the uncertain nature that characterises

certain data. Data derived from social media is often used as example to explain this. Since users of social media are (almost) entirely free in what they upload and under which conditions, data derived from social media is inherently uncertain. However, as several examples in the past have shown, such data might still prove to be useful. The challenge of veracity is thus to make sound assumptions, even though the data used is uncertain in nature.

2.1.2 Geospatial big data

Recently, several car manufactures have been implementing a new functionality to the GPS system of their cars. This new functionality called 'eco routing' calculates the most energy efficient route to a destination based on big data. It does so by analysing a combination of multiple datasets that contain information about factors that vary spatially (Lee & Kang, 2015). Now, one could wonder why this section starts off with such an example; what differentiates this case from other big data use cases? Well, the reason for this is found in the fact that eco routing is a great example to show the power of a particular kind of big data: that of geospatial big data. What differentiates geospatial big data from traditional big data is that it includes locational information with each data point, allowing each data point to be linked to a specific geographical location. This opens up a wide array of possibilities with regard to analysis, mapping and visualisation. In the case of eco routing, GPS data of the car is combined with geospatial data about roads, the location of traffic lights, elevation, speed limits and other relevant factors that vary spatially and affect fuel consumption. An analysis as such is possible thanks to the geographical location that is linked to each data point; by identifying the route that offers the most beneficial variation over space for each factor in terms of fuel consumption, the most energy efficient route can be calculated.

In the last decades, the collection of geospatial big data has increased substantially. Whereas collecting it used to be a relatively expensive activity, currently, with the advent of advanced technological devices such as smartphones, drones and other sensor equipped devices, this has changed for the better and takes place at an unparalleled rate (Li et al, 2016). Take for instance the smartphone; this device carries multiple sensors that together with its GPS functionality generates a myriad of geospatial data at a relatively low cost. Since the smartphone is an immensely popular device and used nearly all over the world, the amount of geospatial data generated is unprecedented. Furthermore, another example that is often mentioned when it comes to the generation of geospatial data is 'the internet of things' (IoT) (Gubbi et al, 2013). Increasingly, devices found throughout the physical environment are equipped with sensors and software that allows them to be interlinked and connected to the internet infrastructure. This way, devices of which the geographical location is known or that are equipped with GPS receivers generate a large quantity of geospatial data, which thanks to its internet connectivity can be stored and processed remotely.

2.1.3 Big data analytics

For organisations to successfully reap the benefits of big data, simply acquiring and possessing big datasets is insufficient. Before value can be obtained, several processes of extraction are required in which high volume, diverse and fast-moving data are turned into meaningful insights (Maltby, 2011; Gandomi & Haider, 2015). The activity of extracting insights from big data through a multitude of extraction techniques are often summarized under the umbrella term 'big data analytics'. Since organisations are increasingly trying to harness the power of big data, big data analytics as field managed to grow into a fully developed area of expertise. Organisations are increasingly willing to pay large sums of money to obtain expertise in related topics such as data mining, statistics, machine learning, and other relevant analysis techniques. Since the number of analysis techniques used are far too numerous for the scope of this study, only the broad idea behind it will be discussed here. This will be done according to the five

steps that organisations are required to take to successfully extract insights from data, as distinguished by Labrinidis and Jagadish (2012). This is but one of such explanations found throughout the literature, all describing a somewhat similar process in a different manner (e.g. see Sagiroglu & Sinanc, 2013). In sake of clarity, this study remains to the explanation as by Labrinidis and Jagadish. Furthermore, since two steps show strong overlap, the number of steps have been reduced to four.

The first step that organisations take is data acquisition. When data is acquired, it often takes the form of raw data (e.g. data generated by sensors, social media or other sources). Much of this data is irrelevant for the ultimate goal and therefore needs to be filtered and compressed by orders of magnitude. The main challenge is then to ensure that the filters do not filter out useful information. This is done by first determining what specific information is required. Subsequently, the right filters can be identified. Once the right filters are determined, the actual filter process can commence, resulting in a dataset comprising only the information that is relevant for the analysis.

The second step concerns the format of the datasets used. After obtaining the required datasets and filtering out irrelevant information, it is necessary to format it in such a way that it is suitable for analysis. In most cases this means that the different data structures and semantics found in the datasets have to be transformed into a machine readable and resolvable form. This allows data to be processed and computed, which is a requirement for most big data analysis techniques. Database design plays a crucial role here, because the way in which data is stored can be of influence the final outcomes of the analysis.

Once the data is structured and formatted properly, it is ready for analysis, which is the third step organisations take. The actual analysis, or as it is often called; the process of 'data mining', comes with a host of challenges and decisions that have to be made concerning query and mining interfaces, scalable algorithms and big data computing environments. To demonstrate, one of the challenges organisations run into when working with big data is finding the best analysis method for the goal that has been set. Which method is best is dependent on factors such as data structure, the size of datasets used, compatibility, efficiency and so on.

The last step is to make sense of the information resulting from the analysis. What is important here is that the person who is in charge of interpreting the results is "data-competent" and has access to supplementary information on how these results were derived. Having knowledge of the assumptions made in the analysis and the precise inputs used (e.g. is the data used reliable? how is the data formatted?) is vital to be able to draw valid conclusions, because these determine how the results should be interpreted.

2.1.4 Big data analytics and GIS

Now, what about big data sets that have a geographical element to it? In other words, what methods are available for analysing geospatial big data sets? Currently, there are multiple tools available that allow the user to visualise geospatial data for both descriptive and analytical purposes. These tools are often referred to as geographic information systems (GIS). In their early years, these systems were almost exclusively being used by larger organisations due to their expensive and resource intensive requirements (Jardine & Teodorescu, 2013). However, in the 1990's this changed, as two relatively inexpensive and easy to use desktop packages were released, called ArcView and MapInfo. Mainly Esri's ArcView has risen to prominence as tool to analyse geospatial data. Especially in the current 'data age', this tool is experiencing a significant increase in popularity because of the high rate at which geospatial data is being generated and the increasing awareness of organisations of the power of such data. In the case of geospatial big data, some say that currently a strong process of 'coupling' is taking place with GIS (Li et al, 2016). GIS's mature toolbox that allows both spatial and non-spatial data to be processed through computational as well as visual means, in combination with a parallel

computing framework that allows processing of large volume datasets (e.g. Hadoop), offers an efficient and relatively inexpensive system to extract insights from geospatial big data (e.g. see "Location intelligence is transforming how government thinks about big data", by Carlson, 2017).

2.1.5 Privacy and data security

However, as the recent controversy surrounding Facebook's data leak shows, working with big data is not all about roses. On the 17th of March, 2018, The Guardian and The New York Times together reported that 50 million (later revised to 87 million) Facebook profiles were harvested for a company named Cambridge Analytica. As a result, Facebook has broken its promise to its users to not share any personal information without consent (CNBC, 2018). Consequentially, a debate got ignited questioning the legitimacy and trustworthiness of these large multinational corporations that are in possession of sensitive personal data. Questions like "who or what will control these organisations?" and "does the possession of vast amounts of data by these organisations make them too powerful?" predominate the debate. For this study, the example of Facebook raises an important question that cannot be left unasked: What are the responsibilities that organisations have when working with big data?

Many of today's datasets are on a personal level of measurement; containing information about customers, clients, patients, and other types of users (Maltby, 2011). When these datasets are used without taking any additional measures to anonymise it or protect it from falling in the wrong hands, this could have severe consequences for the concerning individual (i.e. when a security company leaks information about a person's house security system). Therefore, when dealing with sensitive data, organisations have an obligation to ensure it is irreducible and well protected from external influences. In the literature, the measures that organisations are required to take to ensure this fall under the header of 'data security' (Maltby, 2011; Tene & Polonetsky, 2013). Currently, due to the increasing amount of privacy laws and regulations that are put in place, data security has become a hot topic. To prevent data breaches and thereby breaking the law, organisations are investing heavily to create strong security systems.

2.2 Big data based policymaking; a new era of governance?

Whereas the previous section focused on how big data is defined and how value is obtained from it by organisations in general, in this section the focus is narrowed down to the public sector. As already became clear in the introduction, it is widely believed that the public sector has a lot to gain when it comes to big data. In the context of research streams such as e-government and e-policy, many scholars stress that big data holds the potential to herald a new era of governance; one that is characterised by better public services and, as of interest for this research, more efficient, sound, responsive and transparent policies (Giest, 2017; Kim et al, 2014; Bertot et al, 2014). In the remainder of this section, the above-mentioned research streams in relation to big data will be reviewed. Subsequently, to gain a solid understanding of how policies are formulated and what role big data could play in this, the concepts of evidence-based policymaking and the policy cycle will be discussed. At last, an overview will be provided of the challenges governments face when using big data in the policymaking process by discussing the concepts of Institutional complexity, data culture and data security.

2.2.1 E-government and e-policy

In the literature, the use of big data by the public sector is often discussed in the context of e-government. Originally derived from e-commerce, e-government refers to the natural evolution

of governments due to advances in information and communication technology (ICT) (Howard, 2001). Since the 1970's, economies around the world have been shifting from a predominantly 'industrial' orientation towards a predominantly 'electronic' orientation. Along the same lines, the public sector is increasingly using ICT related tools and techniques to improve their governance capabilities. Until recently, the concept of e-government was mainly used to describe the improvement of delivery systems for public services, such as the use of electronic 'channels' (e.g. the internet) for the provision of information to citizens and promoting citizen participation. However, according to Dunleavy et al (2006), the concept of e-government has moved beyond mere policy delivery; increasingly, the concept is used to contextualise processes and structures resulting from ICT related developments in the policy formation process as well.

In terms of ICT, it is generally known that the public sector lags behind when compared with the private sector (Dunleavy et al, 2006). Low levels of literacy in new technologies and in some cases even in computers in general result in governments encountering high costs (e.g. training personnel, bad bargains due to lack of knowledge, etc.) when trying to adopt todays newest ICT innovations. When this ICT illiteracy is overcome, scholars such as Höchtl et al (2016) believe that ICT advances not only hold the potential to improve the standard procedures of bureaucratic organisations (e.g. administrative tasks), but also influence the internal logic and structure of how decisions are made. More specifically put, ICT advances would hold the potential to influence which problems are given priority and restructure the decision-making process.

Now, what does this mean for the evolution of policy formation? According to Höchtl et al (2016), just like the evolution of government due to advances in ICT are termed 'e-government', the evolution of the policy formation process can be termed 'e-policy'. The concept of e-policy shares many features with concepts such as 'policymaking 2.0' and 'policy informatics' (see respectively: Ferro et al, 2013; Johnston, 2015), but takes on a broader scope. It goes beyond the idea that new technologies are just mere tools to improve the policymaking process. Instead, it poses that the introduction of today's newest technologies asks for a new way of thinking in which policymaking is perceived and executed differently. As Höchtl et al (p: 148) point out, "although policy formation is still essentially a political activity, technological advances reduce the timeframe and increase
the evidence base for policy decisions". Big data is a good example of such a technological advancement that could potentially reduce the time required to formulate well-substantiated policies and that asks for a restructuring of the way policies are made. To better understand this, in the literature the usage of big data in the policymaking process is often linked to the concept of evidence-based policy making.

2.2.2 Evidence-based policy making

The underlying idea of evidence-based policymaking is not necessarily new; examples of policies conducted by governments based on empirical evidence could be traced back centuries (Banks, 2009). The term itself, however, got only popularised relatively recently. Under the adage of 'what matters is what works', the Blair government of the UK in 1999 aimed to 'modernise' and 'rationalise' the way in which policies are made (Blair and Cunningham, 1999). Ever since, the term received considerable attention in the social sciences, mainly questioning the relative value of research and other kinds of evidence inputs into the policymaking process (Marston & Watts, 2003). For instance, the term is often criticized for putting too much emphasis on evidence as decisive factor, while in reality policymaking is an inherently political activity driven by the forces of expediency, ideology and finance (Sanderson, 2002). As Banks (p: 4) argues, it is "values, interests, personalities, timing, circumstance, and happenstance – in short, democracy – that determines what actually

happens". Nonetheless, it is widely agreed upon that evidence and analysis can be of great value for policymakers to make correct judgements and to set the conditions of the political environment in which the decisions have to be made. Therefore, authors like Nutley et al (2002) prefer to speak of 'evidence-influenced' or 'evidence-aware' policy making instead.

At the time of writing, the question of what weight evidence should have in the policymaking process remains a topic of debate. However, in the current data era in which unprecedented volumes and varieties of information are available at unparalleled speeds (big data), several scholars believe that the relative value of evidence inputs into the policymaking process could potentially increase (Höchtl et al, 2016). Since data forms the basis of evidence, the more fast and diverse data is available, the larger, quicker and more diverse the evidence base is for policymakers to work with.

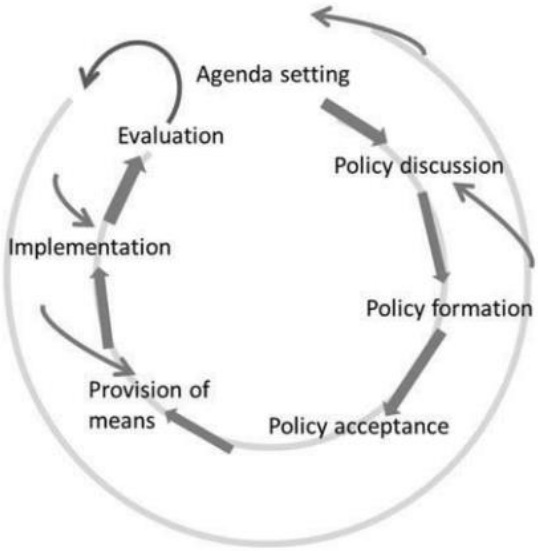
It is important to note here that using big data analyses as a means of evidence does not necessarily result in higher quality policy and decision making. As Giest (2017) points out, working with big data in the policymaking process comes with a host of challenges that need to be overcome. Before going into detail about this, it is important to first have a solid understanding of the policymaking process itself and what role big data could play in it.

2.2.3 The policy cycle

One of the most popular conceptual models to make sense of the policymaking process is the policy cycle (see figure 1). Originally proposed by a political scientist named Laswell in 1951, the goal of the model is to illustrate the different stages of the policy decision lifecycle and open them up for investigation. The model consists of seven inseparable stages (Höchtl et al, 2016). The first stage is agenda setting, where issues are identified that require government intervention. In the stage that follows, the identified problems are then discussed in order to find the best solution. This stage often goes hand in hand with more public awareness for the problem in question, which helps in deciding what policy options form the best match with the needs of the public. Once it is clear what policy offers the best solution and is most satisfactory to the stakeholders involved, the actual policy is then formulated in the policy acceptance stage. In this stage the policy is formulated in a legislative and executive language to prepare them for adoption. Before the policy can be adopted, the provision of necessary means has to be sorted out first, which is done in the subsequent stage. Only then is the proposed policy ready for actual implementation. Since most policies are implemented with a certain goal in mind, policymakers are required to keep a close eye on them to monitor whether they perform accordingly. This last stage, often called evaluation, is crucial if effective public policies are desired. The evaluation stage enables policymakers to evaluate the performance of previously taken policy decisions and improve them where necessary. For instance, it is common practice to perform an outcome evaluation soon after implementation to investigate whether the implementation was successful (e.g. were the granted means sufficient?). In addition, a long-term evaluation is often executed where each stage of the entire policymaking process is examined once more to see what parts could be improved in the future.

Even though the policy cycle is praised by many for offering a clear depiction of the policymaking process, it has been subject to serious criticism as well. The distinction between clear-cut functional activities is considered to be somewhat arbitrary; in practice, the activity of policymaking would be less straightforward than the model suggests (Nachmias and Felbinger, 1982; Everett, 2003). Furthermore, it would put too much emphasis on process, while in reality it is quality and performance that matter most. Regardless of such criticism, the model still offers a clear theoretical representation of the undertakings involved. In particular, as will become clear in next section, the model proves to be useful as a means to elucidate how big data could improve the policymaking process.

Figure 1: the policy cycle.



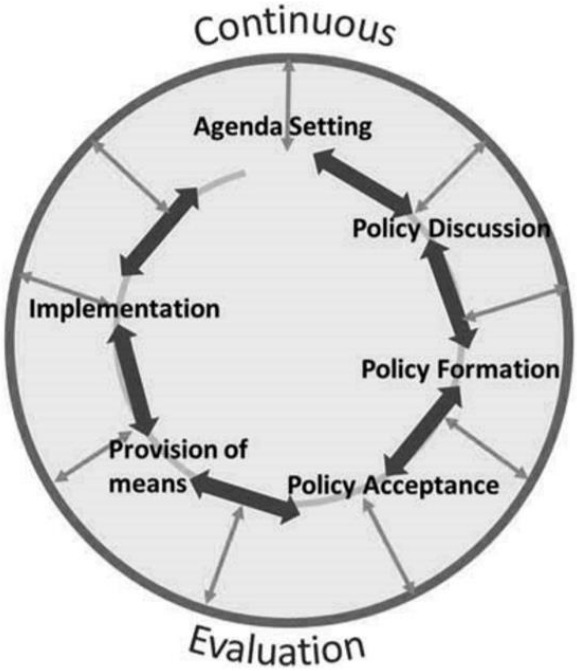
Source: Höchtl et al, 2016.

2.2.4 The big data altered policy cycle

In their article on policymaking in the digital era, Höchtl et al (2016) propose the big data-revised policy cycle (see Figure 2). The model is based on the idea that advantages resulting from the defining characteristics of big data (volume, variety and velocity), together with advances in the field of big data analytics, could potentially change how policies are made. To explain this, they discuss what the opportunities of big data are for each stage of the policy cycle. For example, during the stages of agenda setting and policy discussion, policymakers would now have the possibility to gain insights into policy preferences of the public by gathering data on political discussions held on social media platforms. The speed and continuity at which such data can be gathered allows them to continuously monitor the public opinion and adjust their policy decisions to it. What is more, during the implementation stage, policymakers would now have the opportunity to pinpoint the areas that are in most need of policy by analysing a combination of multiple (big) datasets; ultimately enabling them to adjust the intensity of their policies accordingly.

In each of the examples above big data is used as a tool to gain insights and increase the evidence base. However, Höchtl et al (2016) argue that big data could be more than just a mere tool; since big data enables policymakers to early react to the adverse effects of policy decisions, it holds the potential significantly alter the policymaking process as a whole. To explain this in more detail, they have proposed the big data-revised policy cycle (see Figure 2). What differentiates the big data-revised policy cycle from the traditional one is its focus on evaluation. Whereas in the traditional model evaluation forms the last step of the process, in the big data-revised model the activity of evaluation is inherent to every stage. As Höchtl et al point out, big data's feature to process data in 'real-time' allows policymakers to evaluate policy decisions in a near instantaneous manner. This ability to reassess and reconsider policy decisions right after they are made asks for a new way of policymaking, where each stage of the process is subject to continuous evaluation.

Figure 2: the big data altered policy cycle.



Source: Höchtl et al, 2016.

2.2.5 Institutional complexity, data culture and data security

Until now, mainly the potential benefits of big data for the policymaking process have been discussed. In reality however, working with big data comes with several challenges that need to be overcome if successful application is desired. One such challenge that is often highlighted in this context is institutional complexity. According to Giest (2017), successful integration of big data driven innovations into the existing institutional context of governments is dependent on the organisation's capacity to be able to find and utilize data-based information. This capacity to collect, manage and effectively utilize data into the decision-making process is defined as 'political analytical capacity' (Pawson, 2006; Nutley et al. 2007; Howlett, 2015). When a government's political analytical capacity is limited, it can lead to the involvement of additional actors (e.g. outsourcing). Even though the involvement of additional actors helps with obtaining a larger evidence base for policymakers to work with, it also adds more complexity to an organisation (institutional complexity). As complexity grows, coordination across departments becomes increasingly difficult, slowing down the policymaking process and possibly decreasing the organisation's ability to effectively utilize the gathered evidence (Best & Holmes, 2010).

Institutional complexity within governments is therefore considered as something that should be kept to a minimum. One way to achieve this is by promoting a so called 'data culture' (Giest, 2017). Similar to the concept of political analytical capacity, data culture "emphasises the importance of civil servants and policymakers understanding how to find, analyse and utilize data and the institutional structure to support this through, for example, training or sharing of data among government departments" (P: 372). It asks for an attitude in which data is not only perceived as IT issue, but as integral to the operations of the organisation. Furthermore, it requires support from organisation-wide structures and capabilities; ensuring the free flow of data throughout the organisation. Since many governments possess a siloed data structure (isolated data only accessible to designated departments), effectuating this could prove to be particularly challenging. Especially when specifically looking at big data; due to its velocity characteristic, it is not just the openness in which data is being shared throughout

the organisation that matters, but also the speed at which it happens (Höchtl et al, 2016). Yet, governments are generally not considered to be speedy organisations.

In the context of data culture within governments, another topic that requires attention is that of data security. Governments collect, process and disseminate a wide range of sensitive information on personal, financial and medical aspects (Ebrahim & Irani, 2006). To prevent leakage of such information and thereby harming individuals or organisations, governments have an obligation to take adequate security measures. Scholars such as Gefen et al (2002) point out that this requirement has proven to be a significant barrier to implementing e-government applications (e.g. the use of big data in policymaking). For instance, in the case of big data applications in policymaking, it is detailed knowledge about citizens that offers the most potential to forecast public behaviour with high precision (Höchtl et al, 2016). However, detailed knowledge about citizens also happens to be the kind of data that comes with the highest risk of breaching someone's privacy – demanding policymakers to treat it with utmost caution.

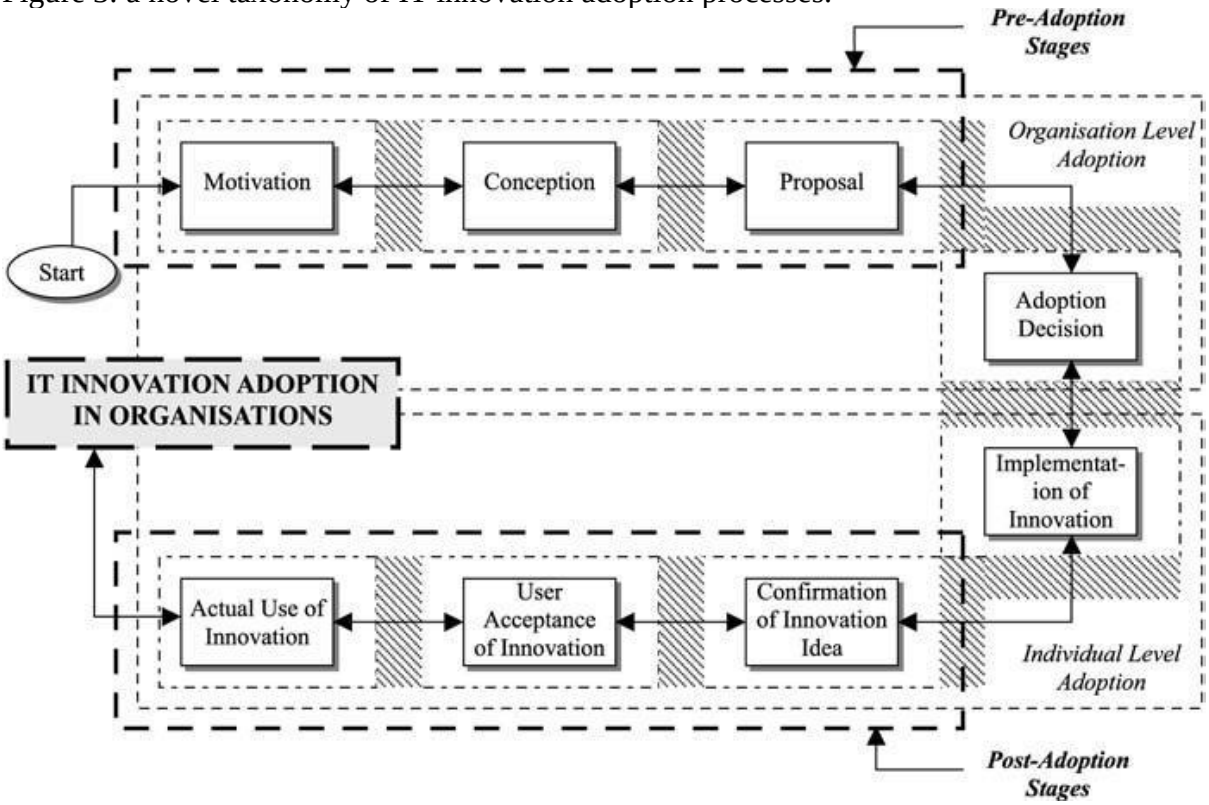
2.3 Adopting big data; data driven innovations in governments

With a clear understanding of big data and its potential for policymaking, the next step is to look at what processes governments undergo to actually adopt such innovations. Within the academic world, the subject of innovation adoption has received much attention from a broad variety of disciplines. Especially after the advent of Rogers' book 'diffusion of innovations' in 1962, which laid the foundations for many of today's innovation adoption theories, the subject increased in popularity. As a result, an extensive amount of literature is available; some of which will be discussed here. The remainder of this section consists of two parts. First, to get a basic understanding of the course of the innovation adoption process, a multistage model will be discussed proposed by Kamal (2005), who focused specifically on information technology (IT) innovations. Second, attention will be paid to a list by Janssen et al (2017), which provides insight into the underlying factors driving and influencing data-driven innovation adoption in the public sector.

2.3.1 The adoption process of IT innovations in organisations

According to Rogers (2003: 12), "an innovation is an idea, practice, or object that is perceived as new by an individual or other unit of adoption". In the current information era, in which IT is significantly restructuring almost every realm of the economy, many of such new ideas and practices are being adopted to increase quality, to find new solutions to problems and improve profitability (i.e. in office automation, telecommunications, data storage) (Knol & Stroeken, 2001). To gain a better understanding of these adoption processes, Kamal (2006) set out to develop an all-encompassing theoretical model that depicts the multiple stages organisations go through (see Figure 3). To develop his model, Kamal has consulted the wide range of literature that is available on the subject (e.g. Rogers, 2003; Agarwal and Prasad, 1998; Gallivan, 2001; etc.). The model consists of eight different stages; from the initial motivation to adopt an IT innovation, to the actual use of it. As the arrows in de model indicate, the innovation adoption process is not simply sequential – rather, the stages overlap and backtrack. Furthermore, the model differentiates between pre-adoption stages and post-adoption stages, as well as between stages that take place on the organisational level and the individual level.

Figure 3: a novel taxonomy of IT innovation adoption processes.



Source: Kamal, 2006.

The first stage of the model is motivation. During this stage, an organisation becomes aware of a technological innovation, tries to gain knowledge of it and subsequently forms an attitude towards it (Kamal, 2006). When it is believed that the innovation is beneficial to the operations of the organisation, the initiators will try to convince the decision-makers that adoption of it should be pursued. This is done during the conception stage; where a plan of action is drafted and presented to the decision-makers. Once the decision-makers are convinced, a formal innovation adoption proposition is made to the rest of the organisation, which is vital for the remainder of the adoption process, since it requires the concerning departments within the organisation to assess whether they possess the needed capacities. After proposition, the final decision can be made on whether the innovation will be implemented. If there is green light, the adoption process shifts from pre-adoption towards post-adoption. During the post-adoption stages the innovation is in use by the organisation. While in use, an assessment takes place of whether the adoption of the innovation was the right decision. If so, it often gets further diffused throughout the organisation.

2.3.2 Driving factors of data-driven innovation by governments

The model of Kamal (2006) provides a clear illustration of the process organisations go through when adopting IT innovations. However, what is more interesting for this study are the underlying forces driving and influencing these processes. Having knowledge of the underlying factors enables one to pinpoint areas where change could potentially benefit the adoption of an innovation. In an attempt to create a clear overview of these factors specifically for data-driven innovations in the public sector, Janssen et al (2017) have drafted the following list (see Table 1). The factors listed have been derived from a special issue of the scientific journal 'Information Systems Frontier' (ISF, 2017), in which nine papers have been published with a focus on big and open linked data (BOLD) innovation in the public sector. The factors

identified in these papers are predominantly on an organisational level of abstraction. However, according to the innovation adoption literature, also individual level factors affect the adoption process (see i.e. Damanpour & Schneider, 2008; Aarons et al, 2011). Therefore, Table 1 shows an enriched version of Janssen et al's list to which the category of 'individual/staff', derived from a narrative synthesis of Wisdom et al (2014), is added. Although this narrative synthesis mainly reviewed research on innovation in the healthcare sector, the category borrowed still proves to be useful thanks to its generic nature.

In the remainder of this section, the different categories listed in Table 1 will be discussed briefly. What is important to note here is that the listed factors are relevant for different stages of the adoption process. Since the precise nature of the relationship between most of the listed factors and data-driven innovation in the public sector is still unclear, Table 1 does not differentiate between these stages.

Strategic and political

The category of 'strategic and political' encapsulates factors related to strategic decisions which in the case of governments are often made on a political level (Janssen et al, 2017). Consequentially, the organisation itself does not always have complete control over these.

Organisational

As the name suggests, the category of 'organisational' concerns factors that are controlled by the organisation itself. Although many different themes are addressed by this category, one in particular requires further attention. According to Janssen et al (2017), to successfully adopt data-driven innovations, governments have to get rid of old government structures and replace them with new organisational forms in which collaboration with private parties and citizens is key. The underlying idea behind it is to mobilise society to share their data and help solve the societal problems governments deal with. Often mentioned examples are: hackathons; where a group of people are provided with a problem to which they have to find an innovative solution in a limited amount of time, living labs; where public and private organisations meet to explore innovative ideas cooperatively, data collaboratives; where public and private organisations share their data to solve societal problems together, and last; agile innovation management, where temporary multidisciplinary teams consisting of a mixture of people internal and external to the organisation are set up to create innovative applications in a short period of time (Gascó, 2017; Susha et al, 2017; Mergel, 2016).

Data governance

Along the same lines, governments are increasingly opening up their data to the public through so called open-data or open government initiatives (Attard et al, 2015). The idea behind these initiatives is that governments are in possession of a large amount of high-quality data that is of great value to other organisations or individuals. By creating an architecture through which data can flow freely not only internally to the organisation, but also externally, governments can foster data-driven public innovation. For instance, it allows private organisations to find solutions to societal problems through big data analyses, which often involves combining several datasets stemming from a multitude of sources. When they in turn sell these solutions to the government, both parties profit; the private organisation in monetary and reputation terms and the government in terms of successfully executing their statutory tasks. It thus shows that data-driven innovation does not necessarily takes place internally, but also externally to the organisation. Therefore, Janssen et al (2017: 191) stress that if governments want to improve values such as transparency, effectiveness, and accountability through data-driven innovations, "a shift from inward looking towards outwards-looking is necessary".

Technical

Next to having the required organisational structure and human resources, data-driven innovations need a technical infrastructure (Janssen et al, 2017). The factors listed in the category of 'technical' are therefore related to the hard- and software needed for data storage, processing, analysis, presenting, etc.

Individual/staff

The category of 'individual/staff' is the lower-level category added for this study to the list. The innovation adoption literature poses that the staff who uses the innovation cannot be left out of an analysis, since their attitudes, competence, social network, readiness for change and so on, could all either positively or negatively influence the innovation adoption process (Wisdom et al, 2014). To illustrate this, take for instance attitude; when an employee is not convinced of the advantages of a data-driven innovation, he or she could form a negative attitude towards it and become a barrier to further adoption.

Table 1: Factors possibly driving or influencing data-driven innovation.

Categories	**Factors**
Strategic and political	Type of public values targeted (efficiency, effectiveness, transparency, accountability or engagement)
	Type of societal problem addressed (different domains including smart cities)
	Data sharing licenses (enabling or constraining the reuse of data)
	Culture enabling or blocking innovation (risk-averse behaviour, enabling collaboration with others)
	Resources and budget (making available human resources, money and other means for innovation)
	Incentives for stimulating data-driven innovation
Organizational	Organizational form of innovation (hackathons, data collaborative, living labs)
	Public-private collaboration (parties involved)
	Division of costs and benefits
	Potential and actual risks (misuse, privacy violation, racial profiling, sensitive data sharing etc.)
	Trust among stakeholders (influencing willingness to collaborate)
	Capabilities and knowledge of staff (for being able to innovate using data)
	Project management (ability to manage the innovation process, to involve other parties etc.)
Data governance	Access to data (open, semi-open or closed access)
	Data quality (accuracy, veracity, completeness etc.)
	Ability to reuse data
	Ability to process and share data
	Local and country data privacy and access regulations
Technical	Technology readiness
	Systems
	Availability of supporting infrastructure (programming environment, data analytics software, cloud, etc.)
	Additional lower-level factors
Individual/staff	Affiliation with organizational culture
	Attitudes, motivations, and readiness towards data innovations
	Feedback on execution and fidelity
	Individual characteristics (awareness, competence, current practice, demographic factors, etc.)
	Social network (individual's personal network)
	Readiness for change/capacity to adopt

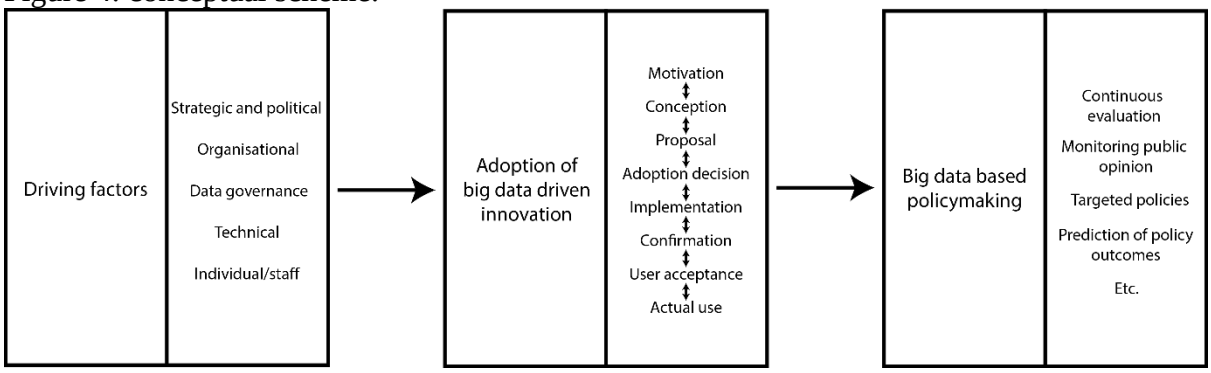
Sources: Janssen et al, 2017; Wisdom et al, 2014.

2.4 GIS departments as possible promoters; an overview

The purpose of this literature review was to present the theoretical foundation upon which the study is built. By clearly defining what is meant by big data, how its use can alter the policymaking process and what factors could drive or influence the adoption of data driven innovations, the goal was to get rid of the fuzziness that surrounds the topic and provide the reader with the knowledge necessary to understand the remainder of the study.

The literature discussed can be conceptualised into the following scheme (see Figure 4).

Figure 4: conceptual scheme.



The scheme illustrates how the concepts that have been discussed relate to each other. It shows that when a government tries to reap the benefits of big data in the policymaking process, an iterative adoption process will take place that is driven and influenced by a multitude of factors. However, as the attentive reader might have noticed, one important part of the research focus is still missing: the GIS department. As became clear in the introduction, this research seeks to find out whether GIS departments could foster the above illustrated adoption process. To further explain where this idea of GIS departments as possible promoters of big data driven innovation in policymaking is derived from, it is important to first look at what roles GIS departments generally fulfil within organisations.

2.4.1 The roles of GIS departments within organisations

The role GIS departments fulfil within organisations is dependent on the organisational GIS structure used. Somers (1998) distinguishes between two organisational structures in which a centrally organised GIS department is present.

First, an enterprise GIS structure, which "provides an information and operations framework for a major portion of the activities and applications within the organisation" (p: 158). In this form, GIS is integrally used by many departments throughout the organisation, but its coordination and control are executed centrally by a dedicated GIS department. Typically, a dedicated GIS department has multiple responsibilities, such as GIS design, implementation, expansion, standards, and the management of the core system software and geospatial databases.

Second, a data and service resource structure, in which GIS is used as a data and service resource throughout the organisation, but is not integral to the operations of its different departments. Through a centrally organised unit (e.g. GIS department), geospatial data and GIS services are occasionally provided to other departments of the organisation when needed. In doing so, this unit "may run a GIS site, build and provide access to databases, provide operational and applications services (do jobs), and offer guidance on standards" (p: 159).

The main difference between these two roles fulfilled by GIS departments is found in their integration. Whereas in an enterprise GIS structure the GIS department forms an integral part of the organisation's operations (e.g. by being directly involved in the policymaking process), in a service and data resource structure, a GIS department only functions as an occasional GIS service provider whenever there is need for it.

2.4.2 GIS departments as promoters of big data driven innovation adoption in policymaking

As has been mentioned earlier, a large share of a government's work, including policymaking, revolves around geography (Alrwais et al, 2015; Carton; 2007). When considering this fact in combination with the strong process of coupling that is currently taking place between GIS technologies and big data, in which GIS is not only used as a visualisation tool of big data analysis outcomes, but also increasingly as a processing tool thanks to advances in the field (Li et al, 2016; Carlson, 2017), it could then be assumed that the combination of GIS and big data holds great potential for policymaking.

It is this assumption that led to the idea that GIS departments of governments might now find themselves in the position to potentially foster the adoption of big data driven innovations into the policymaking process. Especially since the roles of GIS departments described above show that they already fulfil a major role for some of the factors that are believed to drive and influence the adoption of big data driven innovations, as listed in Table 1. For example, the factors listed under 'data governance' concern the data infrastructure of a government, which in the case of geospatial data is often taken care of by GIS departments. While the data infrastructure managed by GIS departments is usually not on par with the requirements of big data, it does show that GIS departments already possess somewhat of a basis when it comes to experience, expertise and resources necessary to work with data in an organisation, and thus may only require relatively little additional training and investment to foster the adoption of big data driven innovations into the policymaking process.

On top of that, when GIS departments form part of an enterprise GIS structure, their activities are already integrated into the policymaking process, which might make it relatively easier for them to foster big data driven innovation in policymaking that is in line with the government's goals, needs, and organisational structure. This idea that IT innovations like big data should be in line with the organisation's structure, strategy and activities to work as desired, is based on IT organisational alignment models like the one by Henderson and Venkatraman (1993). Klievink et al (2017) used this model to explain that organisational alignment of big data driven innovations is fundamentally more important for public sector organisations compared to private sector organisations. Whereas private sector organisations can unlock new possibilities and enable new goals through the use of big data (to then adapt their strategies and activities to it), public sector organisations cannot, as their goals and needs are typically set by laws and regulations imposed by politics.

2.4.3 Promoting the adoption of big data driven innovations through new roles

Even though previous subsection explained in more depth why GIS departments might be able to foster the adoption of big data driven innovations in policymaking, it remains rather vague how exactly. To fill in this void, in this subsection four roles are proposed that GIS departments possibly could fulfil to foster the adoption of big data driven innovations into the policymaking process. The roles have been drafted based on the literature discussed so far, by making a careful consideration between the roles GIS departments currently fulfil on the one hand, and the organisational necessities to successfully work with big data in the policymaking process on the other.

An encouraging role

As the model of Kamal (2006) in section 2.3.1 showed, in the first stages of the innovation adoption process, an individual or a group of individuals within the organisation takes initiative and subsequently tries to convince decisionmakers and the rest of the organisation that adoption is desired. The idea is that GIS departments could be those ones that take initiative and fulfil an encouraging role, in which they use their data expertise and reputation within the organisation to convince policymakers and decision-makers that adoption of big data driven innovation is desired. In doing so, they could emphasise the need for a data culture as defined in section 2.2.5.

A supporting role

Based on the premise that (big) data driven innovation for public issues can be promoted by ensuring that government data flows freely both inside and outside the organisation (Janssen et al, 2017), the idea is that GIS departments can fulfil a supporting role, where they make sure that the geospatial data of governments is easily accessible for the organisation itself and other organisations or individuals.

A directing role

For governments to reap the benefits of innovations for which they do not possess the required expertise, knowledge and tools, it is common practice to outsource it (Giest, 2017). However, in the case of big data related innovations, outsourcing comes with several risks, such as increasing institutional complexity, breaching the privacy and security of the public due to relinquish of control, and misinterpretation of big data analysis outcomes due to lack of expertise (see section 2.1.3). The idea is that GIS departments can mitigate these risks by fulfilling a directing role, in which they use their data-competence to direct outsourcing projects integral to the policymaking process.

A developing role

As described in section 2.1.4, GIS technologies are increasingly able to process big data. Since GIS is widely used in the policymaking process, the idea is that GIS departments of governmental organisations might fulfil a developing role, in which they execute big data analyses integral to the policymaking processes.

As a final remark, what should be mentioned here, is that these roles do not exclude one another. For example, a GIS department could hypothetically speaking fulfil all roles at the same time.

3 Methodology

Building on previous section, which laid out the theoretical foundations for the study, in this section the used strategy and methods are described. Attention is paid to what considerations led to the research strategy chosen, what research design is used, how data was collected and analysed, and at last, what ethical considerations have been taken.

3.1 Research strategy and design

With its focus on whether GIS departments could facilitate the adoption of big data driven innovations into the policymaking process, this study is exploring a new avenue for governments to increase their big data use. Therefore, the decision is made to adopt a qualitative research strategy that is explorative in nature. A qualitative strategy suits the research problem well, since it allows for a more unstructured and open approach, which is especially useful when little is known about the subject (Bryman, 2012).

The research design chosen is that of a case study, in which Team GIS of the Province of Utrecht, a GIS department of a regional governmental organisation in the Netherlands, forms the case. The choice for this particular case was made because the researcher got the opportunity to do a fulltime internship here. Preceding the internship, the researcher and internship supervisors of Team GIS negotiated about what form the internship would take. During these negotiations it was decided that the researcher would have to fulfil two interconnected tasks. First, to execute a thesis research on what roles Team GIS could fulfil in the adoption of big data driven innovations into the policymaking process, and second, to give advice on how the Province of Utrecht could promote big data use in policymaking. To successfully perform the assigned tasks, the researcher would be doing research independently within the team and organisation, and actively participate in all activities that may be of relevance.

3.2 Research methods and data collection

Since the researcher was interning at Team GIS of the Province of Utrecht, he got the opportunity to make use of a combination of research methods to gather data. In the remainder of this section, each research method will be discussed briefly.

3.2.1 Semi- and unstructured interviews

In total, 18 semi- or unstructured interviews were conducted during the research period (see Table 2 for an overview; for a more detailed list, see Appendix I). The questions asked during the interviews are based on the material discussed in the literature review. In particular, Table 1 proved to be instrumental in this. Through an iterative process of literature study and doing research at location, the researcher gradually learned more about the subject of interest and often adapted his interview questions likewise. The interview questions were adjusted based on the respondent's team and position. For example, whereas the questions asked to public officials of Team GIS were more focused on the technical aspects of big data and the role that the team could possibly fulfil to foster its adoption in policymaking, the questions asked to policymakers were more focused on the substantive aspects of policymaking and how big data could add to it. Furthermore, in both cases, questions were asked about the relationship with the other team to better understand the position that Team GIS currently holds within the organisation and what this could mean for their ability to foster the adoption of big data driven innovations.

Table 2: overview of respondents.

Category	No. of respondents
Team GIS	10
Policymakers	5
Open data	1
Province of North-Holland	1
Province of South-Holland	1

As GIS departments form the main focus of the research, most interviews were conducted with public officials of Team GIS. Out of the ten respondents interviewed, two respondents hold a management position: the team leader and one of the strategists. The eight remaining respondents were sampled based on factors such as age and position. The aim was to get an as representative sample as possible.

The other respondents were all sampled through a process of snowball sampling. Since the study focuses specifically on big data driven innovations in policymaking, the researcher wanted to get in touch with policymakers as well. Doing so proved to be relatively simple, as the public officials of Team GIS work in close cooperation with policymakers and could thus introduce the researcher to them. Consequentially, the researcher conducted five interviews with policymakers in total: three policymakers of the domain Urban Development, one policymaker of the domain Soil and Environment, and one policymaker of the domain Traffic and Transport.

In addition, upon recommendation of one of the policymakers interviewed, an external employee who got hired by the Province of Utrecht to take care of the open data housekeeping was interviewed as well. Furthermore, to put the developments at the Province of Utrecht in perspective, two other provinces were visited. In both cases, an in-depth interview was conducted with a representative who could tell more about the organisation's big data use and the role fulfilled in this by their GIS department.

3.2.2 Participant observations

As mentioned above, one of the tasks of the researcher was to actively participate in activities of Team GIS that are of relevance to his research. As a result, the researcher was actively participating in several projects initiated by Team GIS with regards to fostering the adoption of big data driven innovations into the policymaking process. By giving presentations, joining in on a big data analysis experiment, and forming part of negotiations with a company about a big data related outsourcing project, the researcher was able to get in-depth insights into the current efforts of Team GIS and the organisation it is part of.

Moreover, since the researcher formed part of the team and performed his working activities at the same workplace, he got the opportunity to not only observe the inner-workings of the team, but also of the organisation it is part of. In doing so, the researcher was having multiple conversations with public officials, took part in meetings and attended presentations.

3.2.3 Focus groups

Next to participant observations and interviews, the researcher was given the opportunity to conduct two different focus group discussions. The focus groups were attended by policymakers, public officials of Team GIS, managers and other interested. The first focus group was attended by around 8 public officials, the second around 13. During the first focus group, the question of "should the Province of Utrecht focus on adopting big data driven innovations in policymaking?" was central. During the second focus group the focus was put on the question of "how then should the Province of Utrecht approach the adoption of big data

driven innovations in policymaking?". Data was gathered by making notes and using interactive presentation software which allowed the focus group participants to give their opinion by voting, presenting statements and creating word-webs. The exact course of events will be elaborated upon further later on in this paper in the results section, as it forms part of the results. What is important now however, is that these focus groups allowed the researcher to get a broader sense of how the subject of big data is currently perceived within the Province of Utrecht, as they were attended by public officials from various corners of the organisation. Such pieces of information are vital, because the research showed that to fully understand the developments that are currently taking place in Team GIS, they have to be seen in their wider organisational context.

3.2.4 Desk research

To supplement the empirical material gathered, additional data is gathered through desk research. Due to the Province of Utrecht being a public organisation, much contextual information and information about the organisation's activities could be found on their own website.

3.3 Data analysis

The empirical material gathered as a result of the various research methods used came in a variety of forms, such as transcribed interviews, unstructured text, notes, documents, diagrams and so on. To analyse this material, a software package called ATLAS.ti was used which is specifically developed to deal with substantial amounts of unstructured data. The software allows one to load in documents of differing kinds and offers the analyst a set of tools to analyse it. By using this software, the unstructured material was 'coded' into categories similar to those found in the literature review (see Appendix II for a list of the codes used). However, the researcher tried not to limit himself to these categories alone, as the explorative nature of the research demanded of the researcher to be open for unforeseen insights. Once the data was coded, existing links in the data were identified and insights into the research objective were gained.

3.4 Ethical considerations

During the full extent of the research process, the empirical material was gathered and processed with confidentiality in mind. The findings presented here should in no case harm or damage the individuals or organisations who cooperated with this study. To ensure that such confidentiality is maintained, sensitive information such as names and other personal details have been left out of the analysis.

4 The province of Utrecht and Team GIS; context

Before getting into detail about the current state of the Province of Utrecht regarding big data use in the policymaking process, it is important to first have a thorough understanding of the organisation and team itself. This section is therefore dedicated to provide the reader with the contextual backdrop necessary to fully comprehend the findings described in the remainder of the study. For both the Province of Utrecht and Team GIS, attention is paid to topics such as the role they fulfil in their broader context, the assigned tasks they execute, and the way they are organised.

4.1 The Province of Utrecht

The Province of Utrecht is one of the twelve provinces of the Netherlands which together represent the middle administrative layer between the national government and the municipalities (Provincie Utrecht, ND). With an area spanning 144,915 hectares, 28 municipalities and a population of around 1.2 million, the province of Utrecht is the smallest and one of the most densely populated provinces in the Netherlands. Within the Dutch state apparatus, provinces like Utrecht execute a wide range of tasks often summarised under the term of 'regional director'. By connecting municipalities and other stakeholders, provinces direct and mediate developments that transcend municipality borders. In addition, they monitor the activities of municipalities and provide support when necessary.

When taking a closer look at the responsibilities of provinces, seven core tasks can be distinguished (IPO, ND):

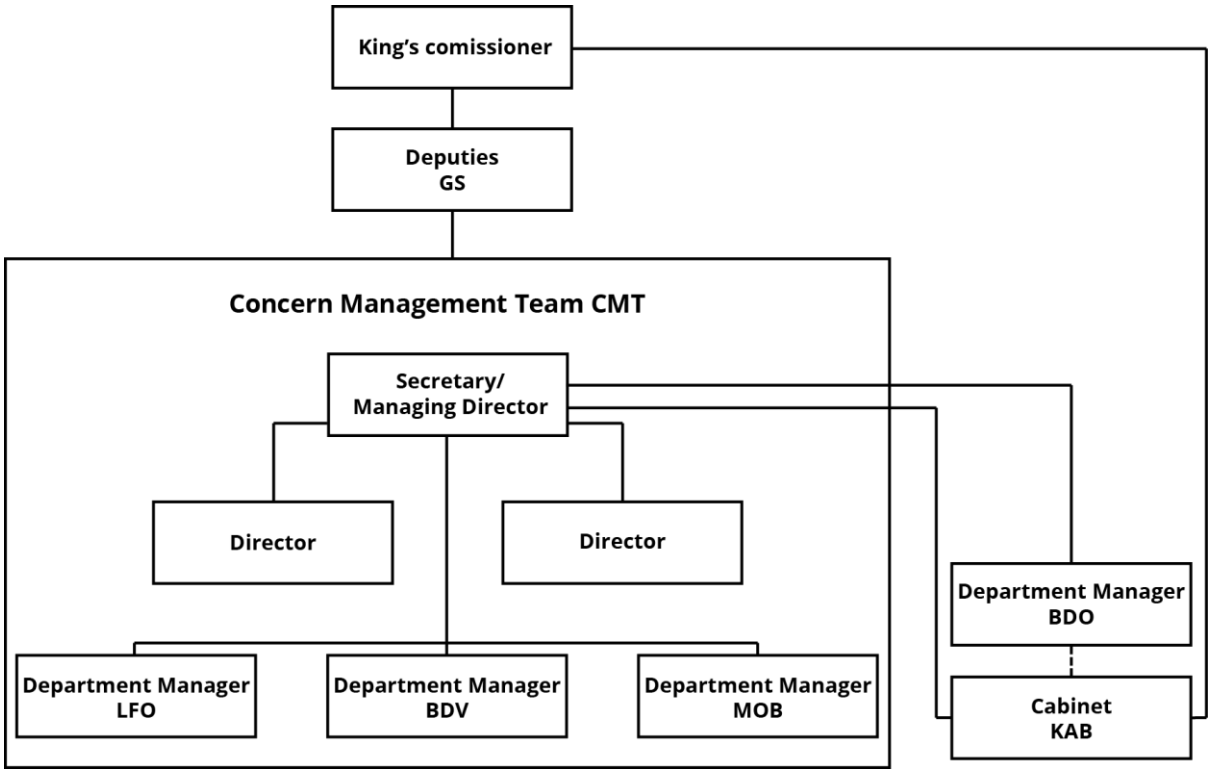
1. Sustainable spatial development, including water management
2. Ensuring a clean, sustainable and secure living environment
3. Directing developments of nature reserves, recreational areas and agricultural land.
4. Ensuring regional accessibility
5. Promoting a strong regional economy
6. Providing a cultural infrastructure and taking care of historical monuments
7. Supervising municipalities

To execute these tasks, provinces are authorized to implement policies, collaborate with stakeholders, and perform management activities (e.g. management of roads and traffic).

4.1.1 Organisation

In total, the Province of Utrecht employs roughly 800 public officials, which taken together form the institution that ensures the above described tasks are executed (Provincie Utrecht, ND). Since the Province of Utrecht is a governmental organisation, its management is chosen through the Dutch democratic system. Every four years a new provincial parliament is elected by the public. Next to determining the broad lines of the policy conducted, this parliament decides who leads the organisation, resulting in the following top-management structure (see Figure 5 for an organisational chart).

Figure 5: top-management structure Province of Utrecht.



Source: Provincie Utrecht, 2017.

At the head of the organisation is the King's Commissioner, who functions as the figurehead and leading representative of the organisation. While the King's Commissioner has a variety of tasks, his or her main task is to chair both the Provincial council, which is the parliament of the organisation, as well as the Provincial Executive, which is the organisation's daily administration.

The Provincial executive consists of six deputies, each with their own areas of attention for which they are accountable. Every four years, the deputies are appointed by the provincial council. The main tasks of deputies are to make proposals and carry out decisions related to the challenges the organisation faces. To ensure these proposals and decisions are executed, the deputies give instructions to the management board.

The job of the management board is to lead the organisation in executing the instructions of the deputies by translating them into concrete tasks. The board consists of three members of which one occupies the role of managing director. Together, they give directions to three departments which combined make up the lion's share of the organisation: Living Environment, Business Management and Mobility.

Each department is represented by its own department manager and is again subdivided into multiple domains. See Table 3 for a clear overview. As a side-note, it should be mentioned that these domains differ over time due to continuously changing organisational goals. The purpose of Table 3 is therefore solely to provide the reader with a general idea of the domains that make up the organisation. Whether it provides a completely valid and up-to-date representation, is not guaranteed.

Table 3: the three main departments and their subordinate domains.

Living Environment (LFO)	Business Management (BDV)	Mobility (MOB)
Soil and Environment	Administration	Management and Monitoring
Culture, Heritage and Recreation	Advising	Exploitation
Economy	ICT management (I&A)	Mobility Infrastructure and Public Transport
Municipal Spatial Development	Facility management	OV Asset Management
Geographical Information Systems (GIS)	Framework, Projects	OV Projects
IPO Affairs and Subsidies	Office Support	Project and Program management
Nature and Agriculture	Execution	Project Management
Commissioning RUD		Traffic and Use
Planning & Control Cycle		Road Safety and Traffic
Program		Information
Program, Process and Project Management		
Regular Duties and Management		
Strategic Spatial Development		
Licensing Nature and Landscape		
Water		

Source: the intranet of the Province of Utrecht, ND.

4.2 Team GIS

Despite being grouped under the department Living environment only, Team GIS manages the geospatial data of the whole Province of Utrecht. Similar to most governments, the Province of Utrecht collects, creates and possesses a large quantity of high-quality geospatial data that being is used in many of the organisation's tasks. The responsibility of Team GIS is to make this geospatial data accessible and usable both for the public and the organisation itself. In their efforts to do this for the organisation itself, the team tries to fulfil more than a mere supporting or facilitating role. As the head of the team explained (respondent 10), Team GIS and its activities form an integral part of the organisation's operations. In the case of policymaking for example, they are directly involved in the process by cooperating with policymakers and providing them with advice on issues related to the use of geospatial data. In addition, Team GIS develops several applications such as interactive maps, dashboards and other tools to help policymakers independently extract insights from geospatial data. Based on this, it could be stated that Team GIS finds itself somewhere between a resource and an enterprise GIS structure, but due to its integral activities, inclining more towards an enterprise GIS structure. Therefore, in the remainder of this study Team GIS is considered to be forming part of an enterprise GIS organisational structure.

4.2.1 Organisation

At the time of writing, Team GIS consists of 23 employees, each with their own role and corresponding tasks (See Figure 6).

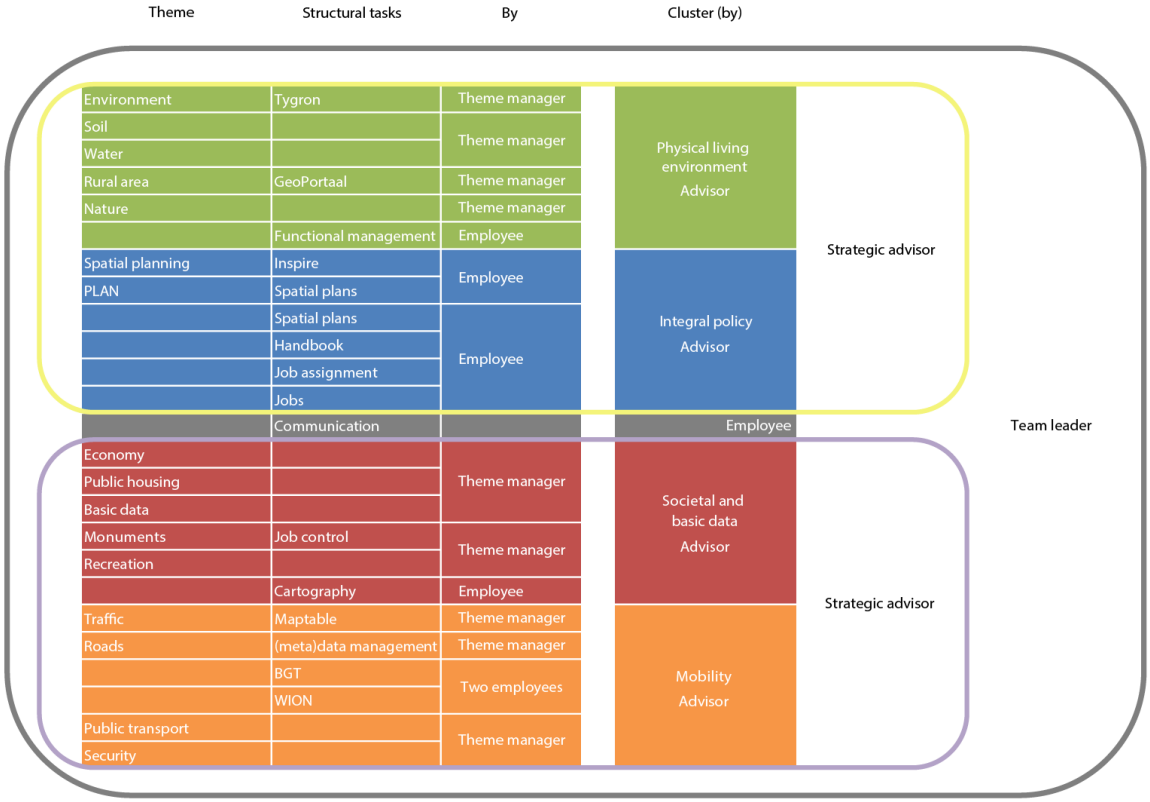
At the top of the team's hierarchy is the team leader, who bears the final responsibility and whose main job is to make decisions on the size, composition and working methods of the team. Furthermore, the team leader provides guidance, sets goals and determines which direction the team is heading. While doing this, the team leader stays in close contact with members of the top management of the organisation, to inform them on new developments that have been taking place and to ensure the team's efforts are in line with the broader organisational goals.

To successfully execute these tasks, the team leader is assisted by two strategic advisors, whose main jobs is to keep track of relevant developments that take place both in- and outside the organisation, ultimately to identify innovations that might have added-value for the team. Subsequently, it is also their job to then translate these into concrete goals for the team. In doing so, each strategic advisor is responsible for their own themes, which is divided according to what they have called 'clusters'.

The working activities of Team GIS have been divided into four different clusters: physical living environment, integral policy, societal and basic data, and mobility. Each cluster consists of a variety of themes that largely reflect the main subjects where the Province of Utrecht as a whole is concerned with, such as economy, environment, public housing, water, spatial planning and public transport. Moreover, each cluster has its own advisor, who next to having their own executive tasks, keeps track of the developments within the cluster and communicates this to the strategic advisors and team leader.

The remainder and largest share of the team consists of employees that have specific executive tasks (e.g. cartography and functional management), and so called 'theme managers'. This latter group consists of employees that each have their own policy theme for which they manage the geospatial data. They do this in close cooperation with the policymakers of that particular theme; whereby the theme manager ensures that the relevant geospatial data is up-to-date and easily accessible for the policymakers.

Figure 6: organisational structure Team GIS.



Source: collected data.

5 Big data driven policymaking and the Province of Utrecht; current state

When trying to summarise the status of the Province of Utrecht with regard to the adoption of big data driven innovations into the policymaking process, one could state the organisation finds itself in an explorative phase. Increasingly, the topic of big data is being explored collectively during meetings, presentations of guest speakers, information markets and living labs. While for many public officials the term big data remains ambiguous, most of them do agree that the Province of Utrecht should not and cannot "miss out" on what today's information age has to offer, and therefore its adoption into the policymaking process should be pursued. As one of the policymakers during one of the focus groups put it "the question is no longer if we have to, but rather how?".

Currently, actual big data use in the policymaking process in the Province of Utrecht is limited to pilots and small experiments. For example, during the research period, policymakers of the policy domain concerned with urban development started a pilot to gain insight through big data analyses into the effectiveness of previously taken policy measures to promote flow through on the housing market. To execute the pilot, the policymakers consulted a company specialised in urban analytics. This company has multiple large datasets at its disposal containing a great variety of information on household level. The idea was that by analysing this data in combination with data supplied by the Province, the company would be able to provide insight into whether the policy measures taken resulted in the desired patterns in the moving chain. Whether the pilot was completed successfully is unknown, as by the end of the research period, negotiations between the company and the Province about the exact details of the analysis were still going.

In the remainder of this section, the current state of the Province of Utrecht when it comes to using big data in the policymaking process is explained in more detail. Since actual big data use in policymaking is limited, attention is mainly paid to the state of the Province of Utrecht concerning the factors that are believed to be positively associated with the adoption of big data driven innovations into the policymaking process, as discussed in the literature review. The following topics are elaborated upon: the organisation's strategic and political considerations taken regarding the subject, its organisational capabilities, its technical readiness and its data governance.

5.1 Strategic and political

Although the Province of Utrecht does not have a strategy aimed specifically towards the adoption of big data driven innovations into the policymaking process, there are strategic developments going on in terms of data use in policymaking in a more general sense.

5.1.1 Data strategy and vision

In the December 2017's publication of the Provincial Committee Living Environment (PCL) of the Province of Utrecht, the PCL advised the Provincial and Executive Council on the need for a "vision on digitalisation and data-driven decision-making" (PCL, 2017). The advice was drafted as a response to the high priority given to the subject by both management and the administrative organisation. "The effectiveness of an organisation is to a large degree dependent on the way in which data is dealt with", so is stated in the document. Therefore, in the context of the environment vision (an all-encompassing vision being drafted for long-term policy on the physical living environment), the PCL advised that extra attention should be paid

to theme of data, and in doing so, awareness should be created throughout the organisation about its strategic importance.

Central to the PCL's advice is the idea of a network society. The Province of Utrecht would increasingly form part of a network of equal partners, consisting of a combination of other public organisations, knowledge institutions and market operators, who all possess their own valuable data. By exchanging this data throughout the network, all partners benefit, creating a win-win situation. The PCL advices the Province of Utrecht to do the following:

1. To answer the following questions for all strategic issues of the environment vision: what data are there about this subject? Who is the owner of the data? Who maintains it? What is the quality of the data? How accessible is the data for other organisations and citizens? Are there gaps in the data or its accessibility? Is the data suitable for long-term storage and is the corresponding software up to date?
2. To ask for attention for the theme of data within the organisation, by for example: creating awareness of data's strategic importance, creating a method to centralise all the data from the different domains of the organisation, training public officials to gain basic data skills (with a focus on open data), and ensuring that the goals set concerning data become a structural part of each domain's working activities.
3. To reap the benefits of the relationships with partners through exchanging data, by for instance making clear agreements with partners about data sharing and security, or by stimulating others to open up their data.
4. To perform an active open-data policy, by ensuring that all data that could be shared with the public, is actually shared with the public.
5. To encourage regional and national cooperation, by for example sharing know-how and learning from other governments, or by using the same data standards for disclosing data.

Currently, there is one policy program in which this emphasis on collaboration with partners to achieve data-driven innovation is put into effect: smart mobility.

5.1.2 The smart mobility program

With its central location, the Province of Utrecht possesses the busiest road and train network of the Netherlands. Due to the ever-increasing number of users on this network, the province is experiencing more and more pressure on the accessibility of the region. Solving this problem asks for more than simply expanding the current network, as expansion possibilities are becoming increasingly limited. Therefore, the Province of Utrecht has decided to look for possible solutions elsewhere.

Joining in on a partnership of public organisations, knowledge institutions and market operators, the Province of Utrecht aims to mitigate the increasing pressure on the accessibility of the provincial region through a new program called: smart mobility. The idea underlying smart mobility is to improve the alignment between supply and demand by finding smarter ways of using the existing infrastructure. Key in this is exploring the possibilities of today's newest technological developments, such as autonomous cars, automation in traffic and mobility management, and new IT-driven mobility services. Currently this is being done in the form of several pilot projects.

Even though the smart mobility program encapsulates the exploration of solutions that differ in many ways, there is one factor that most of them share: their dependency on data. For example, the technological developments mentioned before are all fuelled by data, and not just traditional data, but often big data. In the case of automated mobility and traffic managements systems for instance, real-time data streams on traffic intensities, road conditions, traffic lights,

etc. are used as input, allowing certain parts of the management process to be executed automatically.

The Province of Utrecht is aware of this importance of data in smart mobility, and correspondingly fulfils the role of data-supplier. By opening up their high-quality data to partners of the program, the Province of Utrecht forms an indispensable part in the search for 'smart' solutions to pressing mobility related problems. In addition, the Province of Utrecht fulfils a directing role, bringing partners together and mediating when necessary.

Now, when viewing the smart mobility program in the light of this research, the attentive reader might have noticed that the smart mobility program is more focused on the cooperatively solving of societal problems through concrete data-driven applications, rather than using (big) data as evidence-base to improve the policymaking process. Even though for a large part this appears to be true, it would be wrong to state that with the smart mobility program the Province of Utrecht does not pay any attention to the potential of data for policymaking in the mobility domain. In the publication of the PCL on smart mobility for example, there is one paragraph emphasising the need for a more data-based approach. They thereby advocate the use of both qualitative and quantitative data to gain insights on the mobility needs of citizens, employees working in the province, and visitors. In addition, they advocate the collection of real-time data to make traffic safer and promote flow through.

5.2 Organisational capabilities

As the literature review has shown, successfully adopting big data driven innovations into the policymaking process is far from straightforward and requires the organisation to possess over a divergent set of capabilities. The purpose of this section is to discuss whether the Province of Utrecht possesses over these capabilities. In doing so, attention is paid to topics such as big data experience and expertise, perceptions of big data, and the organisation's political analytical capacity.

5.2.1 Big data experience and expertise

Due to the limited amount of empirical material gathered, one should be careful making statements about big data experience and expertise found in the organisation as a whole. However, from what became clear during the interviews and focus groups, it seems that most public officials of the Province of Utrecht do not have experience working with big data. When zooming in on policymakers for example; apart from the ones concerned with setting up pilots and small experiments, none of the policymakers spoken with indicated to have worked with big data before. The same goes for the IT teams of the organisation: I&A and GIS. While the public officials of these teams possess over more data technical and analytical skills, none of the respondents indicated to have experience in the activities specific to working with big data, such as data mining, machine learning, and working with big data processing frameworks (e.g. Hadoop). Furthermore, also the different managers spoken to made clear that they never managed a big data related project before.

It would be unjustified however, to conclude based on this that near to none of the public officials of the Province of Utrecht have experience working with big data. There is for instance one domain of the organisation in which data is being used that actually corresponds to the defining characteristics of big data: the regional traffic management centre. Through a contiguous network of measurement devices such as bluetooth, traffic lights and camera's, this centre gathers real-time data on traffic on the provincial roads and subsequently analyses it. Based on the results of the analyses, the traffic conductors working at the centre take measures to promote traffic flow through. For example, a common measure taken is to alter the duration of red or green light for traffic lights.

The traffic management centre of the Province of Utrecht is a great example of how the organisation has already proved to be capable of setting up an infrastructure to not only collect big data, but also make sense of it and thereby improve their operations.

5.2.2 Perceptions of big data

On the basis of the interviews and focus groups, it could be stated that the subject of big data is perceived in differing ways throughout the organisation. While some of the public officials spoken to remained to describing big data as datasets that are large in size, most of them were able to provide the interviewer with a more detailed description. For instance, one of the respondents of Team GIS (respondent 4; Theme Manager) described big data as next to being large in size, also being "fugitive" (similar to big data's velocity characteristic). Others, like two of the policymakers (respondent 13; Program Manager, and respondent 14; Coordinator Area Development), mainly pointed at the advent of new data sources like social media and smartphones and how these produce new streams of data on a personal level (similar to big data's variety characteristic). None of the respondents spoken to however, provided a complete definition like the three v's described in the literature review.

Due to these differing perceptions throughout the organisation, the use of the term big data has shown to often cause ambiguity. This became particularly clear during the interviews and focus groups when the topic of big data use in policymaking was discussed. Although the initial goal was to discuss the possibilities of big data for policymaking, most of the time the people involved ended up discussing the possibilities of data in general for policymaking instead. As the one of the Strategists of Team GIS (respondent 9) stated "what I noticed is that people started calling everything big data; all of a sudden a regular register which has been around for a long time already is called big data as well". Furthermore, during one of the focus groups, one of the managers also noticed this and wondered if it is really that important to make the distinction between traditional data and big data. Instead, she suggested, the organisation should start with promoting data use in general.

Another topic that requires attention when discussing how the organisation perceives big data is that of data security and privacy. While the research was taking place, a new EU-wide regulation got implemented to further regulate the way in which data is handled: the General Data Protection Regulation (GDPR, ND). In short, this regulation "harmonizes data privacy laws for all organisations and institutions across Europe, protects and empowers all EU citizens data privacy, and reshapes the way organizations across the region approach data privacy". What it comes down to is that organisations across Europe, including the Province of Utrecht, have to deal with new (stricter) rules and laws when collecting, storing, processing and sharing data, in particular data on a personal level. Due to the implementation of the GDPR, the topic of data security and privacy was a hot topic during the research, as public officials of the organisation became increasingly aware of the responsibilities they have when handling data. This clearly showed during the interviews and focus groups; several of the public officials indicated to be somewhat hesitant about the idea of collecting and using big data on a personal level in their working activities, even when it is anonymised. As the team leader of the policy domain of Soil & Environment (respondent 11) pointed out: "what if a person like Erdoğan (president of Turkey) ever comes to power and such data is available to him or her? I am not sure if we should want to take such risks". Furthermore, one of the attendants of the first focus group stated "yes we should focus on adopting big data, but with respect to privacy and without blindly trusting on numbers". The statement received three upvotes on the interactive app used during the meeting.

5.2.3 Political analytical capacity

As became clear in the literature review, successful integration of big data driven innovations into the existing institutional context of governments is partly dependent on the organisation's capacity to be able to find and utilize data-based information (Giest, 2017). Since collecting data on a wide range of subjects forms part of the statutory tasks of the Province of Utrecht, the organisation has been collecting data-based information for a considerable period of time already. As a result, the Province of Utrecht possess over a large quantity of high-quality data. In an effort to determine whether the organisation is also able to successfully utilize this data in the policymaking process, it became clear that in several cases the organisation is. Through dashboards, geographic maps, diagrams or other (interactive) applications either developed internally by a team like GIS or externally by a partner, several policy domains within the organisation have shown to successfully integrate the use of data in the policymaking process.

To illustrate, the domain of Urban Development developed a housing market monitor in collaboration with an external partner. This monitor uses a combination of maps and diagrams to provide insights into topics such as the current and possible future housing stock, migrations, housing needs, etc. One of the policymakers involved in the project made clear during an interview that so far, the monitor has been instrumental for determining housing market related policy.

Another example that should be mentioned is the use of an interactive application called Tygron. Team GIS and the policy domain Soil and Environment purchased this application to ultimately improve policy related decision making. The use of Tygron allows policymakers to generate a 3D replication of an area by loading in a combination of datasets on a variety of themes. Once the datasets are loaded in and the 3D environment is generated, policymakers can take measures and directly see what consequences these measures have on the surrounding environment. Despite that Tygron is currently mainly used in an experimental fashion, it provides a good example of how the Province of Utrecht is exploring new innovative ways to integrate data use into the policymaking process.

While more examples like the ones above could be provided, it also became clear during the interviews that the consultation of data does not always form part of the standard procedures of policymaking yet. When asking the public officials why they think this is the case, many pointed at the fact that certain policymakers are not sufficiently aware of the potential of data and often lack data related experience and knowledge. One of the public officials of Team GIS clearly noticed this during her work (respondent 1; Integral Policy Advisor), as she works in close cooperation with policymakers from different domains. According to her, consulting data does not form part of some of the policymaker's their system, or in other words: their mindset. She explained that "they (policymakers) do not come to us to ask about it. We (Team GIS) can provide them with more information, we have the capabilities to do that". She went on to explain that "but apparently there is still too much distance between us and the policymakers for us to be able to think along with them; to provide them with more data-based knowledge" Similarly, the public official of the policy domain Traffic and Transport (respondent 12; Information Manager) also pointed at this lack of "data mindedness" of certain policymakers when being asked why the domain made little use so far of the big data collected by the traffic management centre in the policymaking process. She explained: "I do not think we are 'data minded' yet. There are some specific people within the mobility department that are, but I also know several colleague's of which I think they are not."

5.3 Technical readiness and data governance

Next to having the right capabilities, governments ideally have a technical infrastructure that supports both the big data analytics workflow, as well as the free flow of data inside and outside

the organisation. In the remainder of this section, attention is first paid to whether the Province of Utrecht possesses over the hard- and software that are typical for working with big data. Subsequently, attention is paid to how the organisation governs their data.

5.3.1 Hard- and software

Since currently structural big data use only takes place at the regional traffic management centre of the Province of Utrecht, the hard- and software required to collect, process and make sense of big data only seems to be found here. However, because the systems used are specifically designed for real-time traffic management, they are not aligned with other activities of the Province of Utrecht such as policymaking.

5.3.2 Data governance

In an effort to get a general sense of how data is organised in the Province of Utrecht, it became clear that an organisation-wide overview is lacking. However, from the interviews conducted the following could be derived. First off, Team GIS collects and manages most geospatial data of the organisation in a centralised fashion. Consequentially, a large share of the organisation's data is directly accessible both internally and increasingly externally to the organisation. Furthermore, several domains possess over their own data which, according to the public officials spoken to, is in most cases indirectly accessible to the rest of the organisation, usually on a request basis.

To prevent leakage and misuse of data that is considered sensitive (e.g. the basic registration of personal data), the organisation has strict rules and guidelines in place as to whom is allowed to access certain data and under which circumstances. The team leader of the policy domain Soil & Environment (respondent 11) indicated that these rules and guidelines are sometimes perceived as barrier. As example, she described how an intern tried to get access to geological data provided by Team GIS, but was rejected because in order to reach the data, the intern needed full access, which was not allowed as that would mean he or she also got access to certain sensitive data.

Moreover, another topic that requires attention when it comes to the data governance of the Province of Utrecht is that of 'open data'. Currently, on a national level, a proposal is made to impose new legislation which goes under the name of 'the open government law' (Eerste Kamer, ND). The goal of this legislation is to make government organisations more transparent. If this proposal passes the senate, the Province of Utrecht will be obliged to 'open-up' certain categories of data to the outside world. In doing so, they will have to maintain an online register of the datasets and documents they possess. As preparation for this legislation and in the context of the current open government trend, the Province of Utrecht has already started opening-up some of their data. The employee co-responsible (respondent 16; Advisor Dataplatform) for this stated in an interview that so far, the opening up of data has proven to be challenging. As she put it "several public officials are not really aware of the importance of the open data issue yet". As a result, "it is hard to get all the different domains of the organisation on board". Furthermore, an organisation-wide technical infrastructure aimed at opening up data is missing. When for instance she convinced certain public officials to open up a dataset on an open data platform, she experienced several complications, as "the systems used were not aligned and proved not to be ready for it".

5.4 The adoption of big data driven innovations in policymaking; an overview

This section showed that the subject of big data use in policymaking is still being explored within the Province of Utrecht. While the document by the PCL shows that the top-management of the organisation is aware of the strategic importance of data in general, there seems to be no strategy or vision aimed specifically at encouraging the use of big data in policymaking. However, the traffic management centre and the smart mobility program indicate that there are developments going on in which big data has a role to play. Although these developments are not specifically aimed at big data use in policymaking, they do show that the organisation is open to big data driven innovations when it supports their operations.

However, when zooming in on the capabilities and capacities of the organisation regarding big data use in policymaking, it becomes clear that the organisation may not be ready yet. Despite the fact that several policy domains make use of data in the policymaking process, the findings show that its consultation does not always form part of the standard procedure of policymaking. When looking specifically at big data, it is therefore not surprising that currently, the organisation does not possess over any experience or expertise in working with big data in the policymaking process. What is more, the subject of big data is perceived in differing ways throughout the organisation, which has shown to lead to ambiguity and uncertainty whenever the subject is dealt with. At last, even though a large share of the data of the organisation is accessible both internally and externally to the organisation due to the efforts of Team GIS, the findings showed that one cannot speak of a ‘free flow’ of data within the organisation, as several departments still possess over their own delimited data and a clear overview of the data found within the organisation is lacking.

6 Team GIS as promoter of big data driven innovation; exploration of roles

With a clear understanding of the current state of the Province of Utrecht regarding big data use in policymaking, only one question remains: to what extent is Team GIS fulfilling or could they fulfil the roles as defined in section 2.4.3 to foster the adoption of big data driven innovations into the policymaking process?

As already became clear in the methodology section, at the start of the research, the strategists of Team GIS were aware of big data's potential for the policymaking process and decided to further explore its possibilities. The main reason underpinning this decision is that the Province of Utrecht currently faces issues which according to one of the strategists (respondent 9) "deviate from regular policy", such as the upcoming Environment and Planning Act, climate adaptation and the energy transition. Due to the complexity of these issues, "the Province of Utrecht is increasingly looking for new innovative ways to deal with them", as traditional policymaking approaches may be insufficient. The idea of the strategists of Team GIS is that the use of big data might offer such a new way. Furthermore, since Team GIS' core activities revolve around working with data and form an integrated part of the organisation's policymaking processes, the team might be able to foster its adoption.

To further explore whether Team GIS could indeed foster the adoption of big data driven innovations into the policymaking process, the team has been fulfilling or trying to fulfil roles similar to those defined in section 2.4.3. In the remainder of this section, this process of exploration is discussed for each role. In doing so, attention is paid to what developments have been taking place and what challenges the team has been facing.

6.1 An encouraging role

During the research, Team GIS has been trying to create awareness of and support for the possibilities of big data for the policymaking process among both the top management of the organisation, as well as policymakers. As one of the strategists (respondent 9) pointed out "we find ourselves in exciting times in which we have to find a way to put innovations like big data and smart cities on the agenda of the top-management". In doing so, "one has to figure out whether the top-management shows interest in the subject and understands its necessity". Usually, "we are not on our own in doing this, as we work in close cooperation with several policy domains, such as Economy and Urban Development". The respondent went on to explain that "such collaboration with policymakers is crucial because they have the substantive knowledge necessary to align innovations like big data with the issues the organisation currently faces". Only "when you are able to show that an innovation has added value for the organisation's operations, will you be able to convince the top management that adoption should be pursued".

However, for many policymakers the topic of big data is uncharted territory. To change this for the better, Team GIS has been organising information meetings. During the research, two of such meetings took place. The first meeting (focus group 1) took the form of an ambassadors meeting in which representatives of several policy domains, I&A and Team GIS gathered together to discuss the possibilities of big data for policymaking and whether the Province of Utrecht should pursue its adoption. The underlying idea of these ambassador meetings is to bring those public officials together that want to contribute to the development of innovative solutions. The meeting consisted of big data use examples presented by guest speakers (including the researcher himself) and a group discussion. During the second meeting (focus group 2), the researcher himself presented scenarios describing what measures the

Province of Utrecht could take to adopt big data driven innovations into the policymaking process. The goal of presenting these scenarios was to provide food for thought about what it takes for the Province of Utrecht to adopt big data driven innovations into the policymaking process (e.g. by promoting a data culture, by collaborating with private partners or knowledge institutions, etc.), and what consequences it could have. The presentation was held for managers from various corners of the organisation and other interested.

While both meetings succeeded in informing relevant public officials, creating awareness and ensuring that a discussion got ignited, it also showed that gaining organisation-wide support – in particular from the top-management – remains challenging. Part of the reason for this seems to be found in the fact that the big data use examples presented during the meetings are not directly in line with the structure, strategy and activities of the Province of Utrecht. Questions like "how does it fit into the current goals and operations of the Province of Utrecht?" were often left unanswered. Consequentially, the adoption of big data driven innovations into the policymaking process remains too ambiguous and uncertain to secure a priority spot on the agenda of the top-management.

6.2 A supporting role

In the light of the open data trend and as a response to the Open Government legislation that is in the offing, Team GIS is next to making geospatial data accessible internally, increasingly concerned with opening their data to the public. The employee co-responsible for the open data housekeeping of the Province of Utrecht (respondent 16; Advisor Dataplatform) indicated that Team GIS is currently one of the forerunners of the organisation. All geospatial data managed by Team GIS is openly accessible, except for data that cannot be shared due to privacy and security reasons. The data of Team GIS is found on multiple open data platforms, such as the National Geo Register (NGR), Dataplatform, and the data portal of the Dutch national government.

To facilitate the process of opening up data to these platforms, Team GIS makes use of a tool called GeoPublisher. This tool is specifically aimed at publishing geospatial data. As one of the public officials of Team GIS (respondent 3; Executive Employee) working with the tool pointed out, "we are increasingly becoming aware that the data that we serve should be as neat and high in quality as possible". She went on to explain that "including clear metadata (data that tells something about other data; often included with datasets to provide supplementary information) is of great importance to ensure the data is used correctly".

However, in their efforts to open up data, Team GIS has not paid any specific attention to geospatial big data. The explanation for this is simple, as Team GIS currently does not collect or process any geospatial data that could be considered big data. Most geospatial data collected and processed by Team GIS concerns structured or semi-structured data that is relatively small in size and provides a snapshot of a certain phenomenon. The team has thus not been working with large unstructured data streams (as is typical for big data) yet.

6.3 A directing role

As became clear earlier, it is common practice for governments to reap the benefits of an innovation by collaborating with external partners or by outsourcing. When such forms of innovation take place, the parties involved undergo a process in which first goals are set and agreements are made. Subsequently the aimed-for product is developed according to the agreed-upon guidelines. In the case of collaboration or outsourcing projects to achieve big data driven innovations into the policymaking process, the strategist of Team GIS spoken to

(respondent9) indicated that he thinks that Team GIS should fulfil a directing role in this process. The idea is that Team GIS could direct "what the Province of Utrecht demands and receives of the other parties involved". In doing so, the emphasis is on the more data-technical aspects, as this is where Team GIS' expertise lies in contrast to policymakers. By "directing what is happening, what data is being used, and how the final 'product' returns to and ends up being used by the Province of Utrecht", Team GIS is able to mitigate some of the risks that come with outsourcing and collaboration (e.g. institutional complexity, breaching privacy and security, misinterpretation of big data analysis outcomes; see literature review).

During the research period, Team GIS got an opportunity to try out the role of director in a big data related outsourcing project. This project has already been discussed briefly in the introduction of section 5. With the help of a company specialised in urban analytics, the domain of Urban Development wanted to try out the possibilities of big data analysis to evaluate previously taken policy measures to promote flow through on the housing market. Team GIS (including the researcher himself) was asked to join the negotiations between the policy domain and company to ensure that the Province of Utrecht was also sufficiently represented in terms of data-technical knowledge. The goal of Team GIS was to check whether the big data solution was reasonably priced, whether the company operated in accordance with data privacy and security laws, whether the big data analysis would be well documented in sake of interpretation, and to what extent the big data solution offered could be reused.

Even though the representatives of Team GIS managed to achieve some of the goals outlined above, it would be amiss to conclude they directed the process. The role fulfilled by Team GIS had more in common with that of a checker, rather than a director, as they did not lead the process. A reason that possibly underlies this is Team GIS' lack of big data experience. During the interviews with the public officials of Team GIS, all respondents indicated that they do not have experience working with big data. While Team GIS has more data-technical knowledge than most policymakers, they lack the experience to be able to for example precisely determine whether the big data solution offered is reasonably priced, or precisely indicate what of the big data analysis process should be documented for policymakers to correctly interpret it. Due to this lack of big data experience, the team was not capable yet to take control of the negotiations.

6.4 A developing role

At the start of the research period, certain public officials of Team GIS had the idea that the team might try to do big data analyses themselves. Several public officials of the team indicated to have data analysis skills which in their current working activities are left lying dormant. As one of the respondents (respondent 6) stated: "I was getting several courses during my studies in which we had to do statistical analyses in many different forms, not just through GIS". Hence, the idea arose that by exploring the possibilities of big data analytics through experimentation, the team would be able to provide insights into policy issues and, at the same time, make better use of its members' potential.

To explore what the possibilities are, three public officials of Team GIS interested in the subject (including the researcher himself) set out to experiment with big data analyses. The idea was that by executing an experiment first, requirements of working with big data would be identified and a foundation for future use would be put into place. At the start of the experiment, the three public officials realised they first needed policy issues into which big data analyses could provide insights. This required of the public officials to not only get in touch with policymakers, but also find out what big data is currently available.

It was at this point that the project stranded already; finding a policy issue for which there is relevant big data available proved to be particularly challenging. Even though most

policymakers spoken to did come up with policy issues for which they thought big data could be of use (e.g. what are the best regions to place windmills?), the three public officials were unable to then also find suitable big data.

During this search for relevant big data, two problems occurred. First off, as there is no clear overview of all the data that the Province of Utrecht possesses, it was unclear whether the required big data could be found within the organisation itself. Second, when looking for big data sources outside the organisation (e.g. social media, data portals, companies, knowledge institutions), it became clear that most high-quality big data comes with a substantial price tag or, in the case of big data that is openly available, often is incomplete, irrelevant or of insufficient quality. Since the experiment was small in design and initiated bottom-up, neither the required budget to purchase high-quality big data nor time to experiment with uncertain big data was available.

6.5 Team GIS' promoting capabilities; an overview

On the basis of the exploring efforts of Team GIS, it can be concluded that, in their current state, they can successfully fulfil two of the four proposed roles to foster the adoption of big data driven innovations into the policymaking process: the encouraging role and the supporting role. What is striking however, is that the promoting efforts by Team GIS seem to have had little effect so far, as concrete big data related developments in policymaking remain limited. As the previous sections have shown, in exploring the different roles, Team GIS ran into several problems that might reflect why this is the case. There is however, one problem to which no specific attention has been paid yet, but which could be an important underlying explanatory factor.

In section 5.2.3 the integral policy advisor of Team GIS (respondent 1) was quoted "apparently there is still too much distance between us (team GIS) and the policymakers for us to be able to think along with them; to provide them with more data-based knowledge". This quote provides a clear representation of an issue that the team is facing which the researcher also identified during the participant observations: the team's position within the policymaking process of the organisation. Even though the team leader and strategists of Team GIS want the team to be an integrated part of the policymaking process, in reality, this seems not always to be the case. When asked to describe their relationship with Team GIS, several of the policymakers spoken to indicated that they approach Team GIS whenever they are in need of certain geospatial data, maps or other applications the team offers. This would thus imply that the team is not always directly involved in the policymaking process itself, but rather supports it on a demand basis. One of the theme managers of Team GIS (respondent 7) expressed it clearly:

"The Province of Utrecht has opted for a centralised GIS infrastructure. Consequentially, all GIS employees work together in one team. While this is great for sharing knowledge between the GIS employees, it also creates distance with the policymakers, as you are do not actually form part of the policy theme. How I personally experience it; policymakers are just doing their thing. They want a map or an application. Sometimes at the start of the policymaking trajectory, but often at the end, since then it has to be put on a map so they can present it to a deputy. So, I would call it a 'customer-supplier' relationship. That is how I experience it; you do not have to think along, you only execute what they say."

While the explanation by the theme manager provides a clear representation of the issue the team currently faces, it requires to be nuanced. The example of Tygron provided in section 5.2.3 for example shows that the team in certain cases actually does form an integral part of

the policymaking process, as the team works in close cooperation with the policy domain Soil and Environment to solve policy issues through a data-driven application.

To conclude, the findings show that the team is still working on its positioning within the organisation. Both within the team itself and outside of it, the exact role the team fulfils within the organisation is perceived and acted upon differently. As a result, the team might not be able to sufficiently get a grip on the policymaking processes that take place within the organisation to successfully foster the adoption of big data driven innovations.

7 Discussion

In 2014, the municipality of San Francisco adopted a new policy program named: Vision Zero (Vison Zero SF, ND). The ultimate goal of Vision Zero is to bring the number of traffic fatalities that occur in the city each year back to zero. To do this, the program aims to "build better and safer streets, educate the public on traffic safety, enforce traffic laws, and adopt policy changes". The municipality of San Francisco felt the need to implement this policy program as the city was experiencing a continuous increase in traffic accidents. To better understand where traffic accidents are taking place and what is causing it, the municipality is performing (big) data analyses. By developing a map that keeps track of traffic accidents that are taking place within the city (in near real-time), the municipality has been able to identify 'hot zones' which are in most need of targeted policy measures. Once the policy measures have been taken, the map subsequently shows to what extent they perform as desired. In addition, this map also enabled the municipality to identify factors which appear to have a stimulating influence on the number of traffic accidents in a region. By gathering (big) data on these factors, the municipality is now able – by using predictive analytics – to predict areas that in the future may become hot zones. So far, the Vision Zero policy program has shown to be successful, as the number of traffic fatalities is showing a clear downward trend from the program's inception on.

This discussion starts off with this example because it clearly shows why academics should bother researching how governments can increase their big data use in policymaking. It provides a great illustration of how through big data analyses a foundation is put into place on which a policy program could then thrive. Even though working with big data comes with several risks for governments (possibly breaching privacy and data security of the public, taking bad policy decisions due to incorrect data handling, making IT investments that may never pay off), the example above shows that these are risks worth taking.

7.1 Reflection on the findings

In line with the above statement, this research set out to explore whether GIS departments of regional governments can foster the adoption of big data driven innovations in policymaking. On the basis of the findings, it can be concluded that GIS departments which form part of an enterprise GIS organisational structure are, in their current state, capable of successfully fulfilling two of the four roles proposed (see section 2.4.3). By organising meetings with guest speakers, presentations and group discussions, GIS departments can use their data related knowledge and network to provide information, raise awareness and thereby encourage policymakers and others in a governmental organisation to start exploring the possibilities of big data. Furthermore, by ensuring that most geospatial data of the organisation are well accessible both internally and externally, GIS departments are able to indirectly support (big) data driven innovations.

However, the case of Team GIS of the Province of Utrecht shows that despite the team's efforts of fostering the adoption of big data driven innovations into the policymaking process, the organisation made little progress. When contrasting the findings to the theory discussed in the literature review, the following reasons can be derived that might explain this.

Firstly, since GIS departments typically do not have the use of big data included in their working activities, its members often lack big data experience and expertise. In the case of Team GIS, this appeared to have consequences for the team's ability to foster the adoption of big data driven innovations into the policymaking process. To illustrate, in their efforts to direct an outsourcing project, the lack of big data experience prevented representatives of the team to

take control of the negotiations. Consequentially, the team did not direct the outsourcing process, but rather 'checked' it. The question that has to be asked then is whether simply 'checking' is sufficiently mitigating the risks of outsourcing big data analyses for policy issues.

Secondly, one of the arguments provided in the introduction was that GIS departments often form an integrated part of the policymaking process, which might improve their capabilities of fostering the adoption of big data driven innovations. To explain this, the introduction pointed at organisational alignment models which were also used by Klievink et al (2017) to assess a government's big data readiness (Henderson and Venkatraman, 1993). The case of Team GIS showed however, that this 'integration' might not always be the case for GIS departments. Even though the management tries to position the team within the organisation in such a way that it is directly involved in the policymaking process, the findings showed that both inside and outside the team, this position is perceived and acted upon differently. Consequentially, the team might not be able to sufficiently get a grip on the policymaking processes that take place within the organisation to successfully foster the adoption of big data driven innovations.

Thirdly, the subject of big data is known for often causing ambiguity and uncertainty within an organisation due to differing perceptions of what big data is and the risks that come with it (e.g. see SAP, 2012). The case of Team GIS showed that such ambiguity and uncertainty is particularly problematic when trying to gain support from the top-management of the organisation. To successfully adopt big data driven innovations into the policymaking process, support of the top-management is of vital importance because they decide what the organisation is aiming for and how resources are distributed throughout the organisation. Since the agenda of the top-management of a governmental organisation like the Province of Utrecht is formed by its statutory tasks and imposed on them by politics (the Provincial Council), innovations within the administrative organisation have to not only be in line with the organisation's structure to be able to convince the top-management that adoption is desired, but also with the organisation's strategy and activities (see Klievink et al, 2017). The Smart Mobility program is a great example illustrating this; the (big) data driven innovations used in this program gained support from the top-management because they help solve one of the issues that the Province of Utrecht faces: the increasing pressure on the accessibility of the provincial region. However, the examples presented during the information meetings organised by team GIS lacked such alignment with the current structure, strategy and activities of the Province of Utrecht. Therefore, the topic of big data use in policymaking remains perceived as too ambiguous and too risky to secure a priority spot on the agenda of the top-management, despite the encouraging efforts of Team GIS.

Finally, many governmental organisations lack a data culture (Giest, 2017). The findings indicate that this also appears to be the case for the Province of Utrecht. Even though the advice of the PCL shows that the top-management is aware of the strategic importance of data and tries to diffuse this idea throughout the rest of the organisation (see section 5.1.1), the findings indicate that several public officials lack "data-mindedness". Consequentially, (big) data use opportunities for policymaking issues that are lying dormant may not be identified by policymakers, simply because they lack awareness and data related knowledge. Furthermore, a data culture also means that an organisation possesses an infrastructure that facilitates data driven innovation by allowing data to flow freely throughout the organisation. In the case of the Province of Utrecht, Team GIS ensures that most geospatial data of the organisation are well accessible both internally and externally. However, the findings show that the organisation possesses over more data than just geospatial data (e.g. the big data collected by the traffic management centre). What data exactly is unclear, as an overview is missing. As a result, valuable data found within the organisation may be left unused because policymakers or other interested do not know about its existence or how to access it. To illustrate, in trying to fulfil a

developing role, Team GIS was facing a similar problem, as they did not manage to figure out what data is available within the organisation.

7.2 Suggestions for GIS departments of governmental organisations

Despite that the Province of Utrecht has made little progress so far, the case of Team GIS showed that GIS departments are capable of fostering the adoption of big data driven innovations into the policymaking process. However, the case also showed that GIS departments are likely to still face several challenges that have to be overcome if structural adoption is desired. Therefore, based on the findings, the following suggestions are made for GIS departments to improve their promoting capabilities.

First, since outsourcing and collaboration with external partners offers a relatively convenient and, for most governmental organisations, familiar way of reaping the benefits of an innovation, it should be promoted. To ensure that the risks that come with outsourcing and collaboration are kept to a minimum, GIS departments can fulfil a directing role. However, as the case of Team GIS showed, this can only be done successfully if the department possesses over the required capabilities. Therefore, this study suggests that GIS departments should train their current personnel to become more experienced and competent in working with big data, and, if possible, hire new personnel that already has experience.

Second, as became clear earlier, gaining support of the top-management of the governmental organisation is of vital importance to make progress in the adoption of big data driven innovations into the policymaking process. The case of Team GIS showed that GIS departments can fulfil a significant role in this. Therefore, this study suggests that GIS departments should encourage the use of big data in the policymaking process throughout the organisation to ultimately gain support of the top-management. In doing so, the department should present concrete big data use examples that clearly offer a solution to issues that the organisation faces. Furthermore, the department should emphasise the need for a data culture in which public officials (policymakers in particular) are aware of the possibilities of big data and have basic data related knowledge, and in which a data infrastructure is present that promotes (big) data driven innovation instead of hampers it. At last, the department should try to convince public officials of the organisation to start seeing the adoption of big data driven innovations in policymaking as an issue that should be dealt with on its own, instead of an issue that is only dealt with when it supports, or in other words, is in line with the organisation's goals and needs. Since successful adoption of big data driven innovations into policymaking requires measures to be taken that stretch organisation-wide (e.g. centralised all-encompassing data infrastructure, organisation-wide data awareness, etc), operating purely guided by the substantive issues that the organisation is facing is no longer sufficient.

Third, for big data driven innovations to be successfully adopted into the policymaking process, it has to be aligned with the organisational structure, strategy and activities. For GIS departments to foster this process, they should form an integrated part of the policymaking processes that take place within the organisation. With an example like Tygron, the case of Team GIS showed that GIS departments are capable of doing so; by closely cooperating with a policy domain, they managed to develop a data-driven application that is being used in the policymaking process. While this example mainly shows the importance of integration for the proposed developing role, other roles might also benefit. For instance, when a GIS department works in close cooperation with a policy domain to present concrete big data driven examples to the rest of the organisation that are in line with the organisation's policy goals and needs; thereby fulfilling an encouraging role. The study thus suggests that GIS departments should try to become more involved in the policymaking process in order to strengthen their position within the organisation.

At last, GIS departments generally ensure that the organisation's geospatial data is openly accessible both internally and externally. On the basis of the findings, this study suggests that GIS departments should remain doing so to support (big) data driven innovations. However, they should be open to a restructuring of the data housekeeping if the organisation decides to start organising data centrally to promote the free flow of data. Herein, the department should try to become an integrated part.

7.3 Limitations of the study

Due to the novelty of the subject studied, this study had to draw on literature stemming from a variety of disciplines. While the literature discussed is certainly relevant, in some cases it does not fully align with the specificity of the research subject. For instance, Table 1, which is derived from the article by Janssen et al (2017), describes the factors that are of influence on the adoption of data-driven innovations. In this study however, the factors were used in the context of big data driven innovations instead. To ensure that despite this the conclusions made are valid, the researcher tried not to stick to these factors alone and supplemented it with literature that is specifically aimed at big data driven innovations in policymaking (e.g. Giest, 2017; Höchtl et al, 2016).

Furthermore, section 5 deals with findings that concern the Province of Utrecht as a whole. Since the empirical material gathered on the organisation as a whole is sparse, the researcher had to be extra cautious when drawing conclusions. To ensure that the findings described are valid, the researcher tried to be as nuanced as possible. Moreover, the decision was made to only include those findings that through a process of triangulation were confirmed.

8 Conclusion

Driven by the fact that governments are falling behind in terms of big data use compared to the private sector – despite the great potential it holds for them, this study set out to explore whether GIS departments of governmental organisations are capable of promoting the adoption of big data driven innovations into the policymaking process. Drawing on literature from a variety of fields, dealing with topics such as big data in general, e-government, evidence-based policymaking, the big data revised policy cycle and the adoption of data-driven innovations, four roles were proposed that GIS departments could possibly fulfil to foster the adoption of big data driven innovations into the policymaking process. To explore whether GIS departments can indeed fulfil these roles, a case study of a GIS department of a regional governmental organisation in the Netherlands was conducted. The case study showed that GIS departments which form part of an enterprise GIS organisational structure can, in their current state, successfully fulfil two of the four roles proposed: the encouraging role and the supporting role. However, the study also showed that GIS departments still have several challenges to overcome before these efforts will actually lead to successful adoption of big data driven innovations in policymaking. To help them overcome these challenges, based on the findings, the study provided four suggestions.

Finally, to wrap up, future academic endeavours can use the roles proposed by this study to better understand the role GIS departments might fulfil in the adoption of big data driven innovations into the policymaking process. In doing so, they should in particular focus on the possibilities of GIS departments to develop big data analyses themselves, as in the current research this was done insufficiently (since the big data analysis project by Team GIS already stranded at an early stage). Even though it is likely that GIS departments will run into similar problems, such as the lack of integration into the policymaking processes, the lack of a

data culture, lack of support from the top-management, and lack of actual big data experience and expertise, the possibility of a developing role played by GIS departments received too little attention to be written off already. Furthermore, it might be interesting to start thinking about what position GIS departments should take within a governmental organisation when the use of big data actually forms a structural part of the policymaking process. Should they form an integrated part of a dedicated data team who manages an all-encompassing and centrally organised data infrastructure? Or should they remain a separated entity who takes care of the geospatial data of the organisation only?

References

Aarons, G. A., Hurlburt, M., Horwitz, S. (2011). Advancing a conceptual model of evidence-based practice implementation in public service sectors. Administration and Policy in Mental Health and Mental Health Services Research, 38(1), 4–23.

Agarwal, R. and Prasad, J. (1998). The antecedents and consequents of user perceptions in information technology adoption. Decision Support Systems, 22(1), 15-29.

Attard, J., Orlandi, F., Scerri, S., Auer, S. (2015). A systematic review of open government data initiatives. Government Information Quarterly, 32(4), 399–418.

Banks, G. (2009). Evidence-based policy making: What is it? How do we get it?. Commonwealth of Australia. Productivity Commission, Canberra.

Bertot, J.C., Gorham, U., Jaeger, P.T., Sarin, L.C., Choi, H. (2014). Big data, open government and e-government: Issues, policies and recommendations. Information Polity, 19, 5–16.

Best, A., Holmes, B. (2010). Systems thinking, knowledge and action: Towards better models and
methods. Evidence and Policy, 6(2), 145–159.

Blair, T., Cunningham, J. (1999). Modernising Government. Prime Minister and Minister for the Cabinet Office. London. UK.

Bloomberg Businessweek. (2011). The Current State of Business Analytics: Where Do We Go from Here? Bloomberg Business, 1186 MIS Quarterly, 36(4).

Bryman, A. (2012). Social research methods. Oxford: Oxford University Press.

Carlson, C. (2017). Location intelligence is transforming how government thinks about big data. Statescoop. [online] https://statescoop.com/location-intelligence-is-transforming-how-government-thinks-about-big-data (June 20, 2018).

Carton, L.J. (2007). Map making and map use in a multi-actor context: Spatial visualizations and frame conflicts in regional policymaking in the Netherlands. Dissertation. Technical University of Delft.

Chen, H., Chiang, R. H. L., & Storey, V. C. (2012). Business intelligence and analytics: From big data to big impact. MIS Quarterly, 36(4), 1165–1188.

Clarke, R. (2016). Big data, big risks. Information Systems Journal, 26(1), 77–90.

CNBC (2018). Facebook-Cambridge Analytica: A timeline of the data hijacking scandal. https://www.cnbc.com/2018/04/10/facebook-cambridge-analytica-a-timeline-of-the-data-hijacking-scandal.html.

Damanpour, F., Schneider, M. (2008). Characteristics of Innovation and Innovation Adoption in Public Organizations: Assessing the Role of Managers. JPART, 19, 495–522.

Davenport, T. H., Dyché, J. (2013). Big data in big companies. International Institute for Analytics. [Online] http://docs.media.bitpipe.com/io_10x/io_102267/item_725049/Big-Data-in-Big-Companies.pdf (July 7, 2018).

Downs, G.W., Mohr, L.B., (1976). Conceptual issues in the study of innovation. Administrative Science Quarterly, 21, 700–14.

Dunleavy, P., Margetts, H., Bastow, S., Tinkler, J. (2006). New public management is dead—Long live digital-Era Governance. Journal of Public Administration Research and Theory, 16(3), 467–494.

Ebrahim, Z., Irani, Z. (2006). E-government adoption: architecture and barriers. Business Process Management Journal, 11(5), 589–611.

Eerste Kamer (ND). Initiatiefvoorstel-Snels en Van Weyenberg Wet open overheid. [Online] https://www.eerstekamer.nl/wetsvoorstel/33328_initiatiefvoorstel_snels_en (November 1st, 2018).

Everett, S. (2003). The policy cycle: Democratic process or rational paradigm revisited? Australian Journal of Public Administration, 62(2), 65–70.

Ferro, E., Loukis, E.N., Charalabidis, Y., Osella, M. (2013) Policy making 2.0: From theory to practice. Government Information Quarterly, 30(4), 359–368.

Gascó, M. (2017). Living labs: Implementing open innovation in the public sector. Government Information Quarterly, 34 (1), 90 – 98.

Gallivan, M.J. (2001). Organisational adoption and assimilation of complex technological innovations: development and application of a new framework. The Database for Advances in Information Systems, 32(3), 51–85.

Gandomi, A., Haider, M. (2015). Beyond the hype: Big data concepts, methods, and analytics. International Journal of Information Management, 35, 137–144.

Gamage, P. (2016) New development: Leveraging ‘big data’ analytics in the public sector. Public Money & Management, 36(5), 385–390.

GDPR (ND). GDPR.org. [Online] https://eugdpr.org/ (July 7, 2018).

Gefen, D., Pavlou, P. et al. (2002). Egovernment adoption. Paper presented at Americas Conference on Information Systems, Tampa, FL.

Giardullo, P. (2015). Does ‘bigger’ mean ‘better’? Pitfalls and shortcuts associated with big data for social research. University of Urbino Carlo Bo.

Giest, S. (2017). Big data for policymaking: fad or fasttrack. Policy Sci, 50, 367–382.

Groenfeldt, T. (2012). Morgan Stanley Takes On Big Data With Hadoop. Forbes. [Online] http://www.forbes. com/sites/tomgroenfeldt/2012/05/30/morgan-stanley-takes-on-bigdata-with-hadoop/ (July 7, 2018)

Gubbi, J., Buyya, R., Marusic, S., Palaniswami, M. (2013). Internet of Things (IoT): A vision, architectural elements, and future directions. Future generation computer systems, 29(7), 1645–1660.

Hashem, I.A.T., Yaqoob, I., Anuar, N.B., Mokthar, S., Gani, A., Khan, S.U. (2015). The rise of "big data" on cloud computing: Review and open research issues. Information Systems, 47, 98-115.

Henderson, J. C., Venkatraman, N. (1993). Strategic alignment: leveraging information technology for transforming organizations. IBM Systems Journal, 32(1), 472–484.

Höchtl, J., Parycek, P., Schöllhammer, R. (2016) Big data in the policy cycle: Policy decision making in the digital era. Journal of Organizational Computing and Electronic Commerce, 26(1-2), 147–169.

Howard, M. (2001). E-Government Across the Globe: How Will "e" Change Government? Government Finance Review.

Howlett, M. (2015). Policy analytical capacity: The supply and demand for policy analysis in government. Policy and Society, 34, 173–182.

IPO (ND). De zeven kerntaken van de provincies. [Online] http://www.ipo.nl/over-de-provincies/de-zeven-kerntaken-van-de-provincies (September 10, 2018)

Janssen, M., Konopnicki, D., Snowdon, J.L., Ojo, A. (2017). Driving public sector innovation using big and open linked data (BOLD). Information Systems Frontier, 19(2), 189–195.

Jardine, D.D., Teodorescu, D. (2003). Using GIS in institutional research. Chapter 1: An Introduction to GIS: Concepts, Tools, Data Sources, and Types of Analysis. New directions for institutional research, 120.

Jetzek, T., Avital, M., Bjorn-Andersen, N. (2014). Data-driven innovation through open government data. Journal of Theoretical and Applied Electronic Commerce Research, 9(2), 100–120.

Johnston, E. W. 2015. Governance in the information Era: Theory and practice of policy informatics. London, UK: Routledge.

Kamal, M.M. (2006). IT innovation adoption in the government sector: identifying the critical success factors. Journal of Enterprise Information Management, 19(2), 192–222.

Kim, G., Trimi, S., Chung, J.H. (2014). Big-data applications in the government sector. Communications of the ACM, 57(3), 78–85.

Klievink, B., Romijn, B.J., Cunningham, S., Bruijn, H. (2017). Big data in the public sector: Uncertainties and readiness. Information Systems Frontier, 19(2), 267–283.

Knol, W.H.C., Stroeken, J.H.M. (2001). The diffusion and adoption of information technology in small- and medium-sized enterprises through IT scenarios. Technology Analysis & Strategic Management, 13(2), 227–46.

Kwon, O., Lee, N., & Shin, B. (2014). Data quality management, data usage experience and acquisition intention of big data analytics. International Journal of Information Management, 34(3), 387–394.

Labrinidis, A., Jagadish, H. V. (2012). Challenges and opportunities with big data. Proceedings of the VLDB Endowment, 5(12), 2032–2033.

Laney, D. (2001). 3-D data management: Controlling data volume, velocity and variety. Application Delivery Strategies. META Group Inc. [Online] https://blogs.gartner.com/doug-laney/files/2012/01/ad949-3D-Data-Management-Controlling-Data-Volume-Velocity-and-Variety.pdf (February 11, 2018).

Lasswell, H.D. (1951). 'The Policy Orientation' in Lerner, D. & Lasswell, H.D. (eds) (1951). The Policy Sciences, Stanford University Press, Stanford.

Lee, J., Kang, M. (2015). Geospatial Big Data: Challenges and Opportunities. Big Data Research, 2, 74–81.

Li, S., Dragicevic, S., Castro, F.A., Sesterd, M., Winter, S., Coltekin, A., Pettit, C., Jiang, B., Haworth, J., Stein, A., Cheng, T. (2016). Geospatial big data handling theory and methods: A review and research challenges. ISPRS Journal of Photogrammetry and Remote Sensing, 115, 119–133.

Longley, P.A., Goodchild, M., Maguire, D.J., Rhind, D.W. (2010). Geographic information systems and science (ed 3). John Wiley and Sons.

Maltby, D. (2011). Big Data Analytics. Austin School of Information. University of Texas.

Marston, G., Watts, R. (2003). Tampering With the Evidence: A Critical Appraisal of Evidence-Based Policy-Making. The Drawing Board: An Australian Review of Public Affairs. School of Economics and Political Science. University of Sydney.

Mayer-Schönberger, V., Cukier, K. (2013). Big data. A revolution that will transform how We live, work and think. London: John Murray Publishers.

Mullich, J. (2013). Closing the Big Data Gap in Public Sector. Bloomberg Businessweek Research Services. [online] https://www.slideshare.net/SAPAsiaPacific/closing-the-big-data-gapinpublicsector (August 21, 2018).

Nachmias, D., Felbinger. C. (1982). Utilization in the policy cycle: Directions for research. Review of Policy Research, 2(2), 300–308.

Nutley, S., Davies, H., Walter, I. (2002) Evidence based policy and practice: cross sector lessons from the UK. keynote paper for the Social Policy Research and Evaluation Conference, Wellington, New Zealand, 2–3 July.

Nutley, S. M., Walter, I., & Davies, H. T. O. (2007). Using evidence: How research can inform public services. Bristol: Policy Press at the University of Bristol.

Pawson, R. (2006). Evidence-based policy: A realist perspective. London: Sage.

PCL (2017). Strategisch omgaan met data. PCL advies in de aanloop naar de Omgevingsvisie, 8 december 2017. The Province of Utrecht.

Provincie Utrecht (2017). Organogram. [Online] https://www.provincie-utrecht.nl/publish/pages/79711/organogram_topstructuur_oktober_2017.pdf (September 10, 2018).

Provincie Utrecht (ND). Province of Utrecht. [Online] https://www.provincie-utrecht.nl/algemene-onderdelen/serviceblok/english/province-utrecht/ (September 10, 2018).

Rogers, E.M. (2003). Diffusion of Innovations. 5th ed. The Free Press. New York.

Sagiroglu, S., Sinanc, D. (2013). Big Data: A Review. Department of Computer Engineering, Gazi University.

Sanderson, I. (2002). Evaluation, policy learning and evidence-based policy making. Public Administration, 80(1), 1–22.

Sanderson, I. (2006). Complexity, 'practical rationality' and evidence-based policy making. Policy and
Politics, 34(1), 115–132.

SAP (2012). Small and Midsize Companies Look to Make Big Gains With "Big Data," According to Recent Poll Conducted on Behalf of SAP. Walldorf. [online] https://news.sap.com/small-and-midsize-companies-look-to-make-big-gains-with-big-data-according-to-recent-poll-conducted-on-behalf-of-sap/ (February 11, 2018).

Somers, R. (1998). Developing GIS Management Strategies for an Organization. Journal of Housing Research, 9(1). 157–178.

Sweeney, L. (2001) "Information explosion" Confidentiality, disclosure, and data access: Theory and practical applications for statistical agencies. Urban Institute, Washington, DC.

Tene, O., Polonetsky, J. (2013). Big Data for All: Privacy and User Control in the Age of Analytics. Northwestern Journal of Technology and Intellectual Property, 11(5). 240–270.

Vision Zero SF (ND). What is Vision Zero? [online] https://visionzerosf.org/about/what-is-vision-zero/ (November 7, 2018).

Wisdom, J.P., Chor, K.H.B., Hoagwood, K.E., Horwitz, S.M. (2014). Innovation Adoption: A Review of Theories and Constructs. Adm Policy Ment Health, 41, 480–502.

Worral, L. (1991). Spatial Analysis and Spatial Policy using Geographic Information Systems. Belhaven Press London.

Appendix

I List of respondents

ID	Team	Gender	Position
1	GIS	F	Integral Policy Advisor
2	GIS	F	Theme Manager
3	GIS	F	Executive Employee
4	GIS	M	Theme Manager
5	GIS	M	Theme Manager
6	GIS	M	Executive Employee
7	GIS	M	Theme Manager
8	GIS	F	Theme Manager
9	GIS	M	Strategic Advisor
10	GIS	M	Team Leader
11	Soil & Environment	F	Team Leader
12	Traffic & Transport	F	Information Manager
13	Urban Development	F	Program Manager
14	Urban Development	M	Coordinator Area Development
15	Urban Development	M	Senior Policy Advisor
16	External	F	Advisor Dataplatform
17	Province of North-Holland	M	GIS strategist
18	Province of South-Holland	M	Datawarehouse Manager

II List of codes

Attitude towards big data management general
Attitude towards big data management policymakers
Attitude towards big data management Team GIS
Attitude towards big data policymakers
Attitude towards big data Team GIS
Data in policy evaluation
Data structure policy domain
Data usage policymakers
Experience big data Team GIS
Goal big data use
Knowledge of big data policymakers
Open data
Organisational requirements big data
Perceived challenges of adopting big data
Perception data driven policymaking
Possible big data use cases
Possible measures to improve big data adoption

Practical information respondents
Privacy and data security (avg)
Relationship Team GIS and policymakers
Role Province of Utrecht
Role Team GIS
Social network big data related
Tasks and activities policymakers
Tasks and activities Team GIS
Team capabilities required big data
Technical capacities required big data
What is big data perception Team GIS