DATA-DRIVEN WORK AND PRIVACY AT DUTCH MUNICIPALITIES —
A DOUBLE-EDGED SWORD

Exploring the Perceived Data-driven Development Impact on Privacy in Dutch Municipalities

Matthijs H.G. von Piekartz

Department of Behavioural, Management and Social Sciences, University of Twente

Public Administration Master Thesis

Dr. Ringo Ossewaarde
Dr. Claudio Matera

January, 2022

Abstract

Purpose: Through the increasing prevalence of data-driven work within organisations, including Dutch local governments, its effects have become a popular topic in literature. Theory shows that while data-driven work can create value for both public and private organisations, key issues such as maintaining individuals' privacy have to be overcome to implement data-driven work successfully. These conditions result in a situation where citizens can benefit from their local government implementing data-driven approaches while negative consequences also (in)directly affect them. Because sets of information fully power data-driven work, personal data usage may harm citizen privacy and the control individuals have over their data. Therefore, this study aims to investigate what phenomena can be found through recent main data-driven developments at local governments and their possible effect on citizen privacy as perceived and experienced by the civil servants from the Dutch municipalities.

Method: Interview questions were created based on the research questions and theoretical framework for which respondents were sought from Dutch municipalities through purposive sampling. This inductive qualitative research had participants answer 10 open-ended questions, which were coupled with probes and additional questions dependent on the answers given for the main questions. Respondents were found through the professional networking website called 'LinkedIn', for which a criterion was that they were (heavily) involved with data-driven work activities within their organisation. This search led to a sample size of 15. The interview data were analysed by hand and reflected on through the theoretical framework. The coded data was examined by doing a content analysis that helped answer the research questions.

Results: The interview data showed that the civil servants perceived and experienced benefits and chances, including financial incentives and better (data-driven) analytics, predictive, and decision-making methods. The main observations from the risk and challenges are data quality issues, (the lack of) necessary competencies and capacity, unconnected data streams, people assuming data is always objective, and privacy concerns. Organisational phenomena reported include data-driven work implementation being ongoing but not yet finished, (inter) organisational cooperation, and the usage of informative campaigns to ensure data-driven and

privacy awareness. Respondents further report the use of data ethics aids, privacy officers, and information security officers. Some of the factors driving these data-driven work and privacy developments are technological, external, and internal factors. The technological factors include (improved) access to new technologies for data analysis. External factors include a degree of hype surrounding data-driven work and legislation such as the GDPR and its corresponding fines. The main internal influences observations include data-driven work implementation as part of a long-term strategy, decision-making, information access, work optimisation, value creation, and working ethics improvements. Further, the privacy impact changes include positive, negative, and unclear phenomena. The intrinsic motivation of respondents and awareness regarding privacy concerns were evident positive changes. The negative changes include increased personal data collection and the risk of privacy leakage or misuse of personal information. As part of the unclear impact observations, deductive disclosure can occur when data is combined or is in a small data set. The repurposing of personal information can further negatively impact individuals' privacy.

Conclusion and policy implications: This study finds that based on the content analysis results, data-driven work developments at Dutch municipalities are increasing, leading to more personal data being collected. This research adds new knowledge by analysing current changes within these organisations using recent literature. Using this approach, positive, negative, and unclear impacts on the privacy of individuals have been found. The perceived negative and unclear effects are in line with the literature from the theoretical framework used. Moreover, the misuse of personal data, privacy leakage, deductive disclosure, or data purpose changes represent negative or unclear privacy changes. In contrast to the literature, positive changes experienced by civil servants include a heightened awareness for privacy concerns, awareness campaigns, intrinsic drivers for integrity, accountability to maintain individuals' privacy and adherence to recent privacy legislation such as the GDPR. While further research can investigate privacy impact changes, increased internal and external cooperation (between various municipalities) would help address and identify current and future privacy concerns.

Keywords

Data-driven work, privacy, Dutch municipalities, local government, digital transformation, GDPR, personal data, content analysis, deductive disclosure

Contents

1. Introduction

1.1 Problem overview and motivation

As the governance of states grows in complexity, governmental organisations have been looking for ways to manage this myriad of processes. This new development also means civil servants need to govern their accompanying information usage and growth to achieve organisational goals. Some of these ways include working with technologies such as data analysis, machine learning, and artificial intelligence (AI). These techniques aim to get meaningful insights into massive amounts of structured or unstructured data. This data is partially generated by citizens who have access to an increasing number of appliances such as computers, smartphones, tablets, and other smart devices. These digital devices are often connected to the internet and generate bits and bytes of information. However, the governments themselves, who have expanded governance to a broader range of areas, also created information growth—recording information in greater depth and detail than ever before.

Nowadays, more information is generated every year than people have generated during hundreds of years before. The explosion of growth helped create new information management and technological advancements during the last few decades. Similarly, the last few years have simultaneously seen increased productivity and efficiency for both individuals and organisations through better use of data (Malomo & Sena, 2017). However, according to McAfee et al. (2012), this increase also presents organisations with new challenges to gain meaningful insights into this data. These data-driven work developments are not only present in national governments and commercial parties but also within local governments such as municipalities.

For organisations in the public sector, data-driven work offers great opportunities because, in theory, there is no competition or 'price' to be determined for data. An important question is whether, and if so, under what conditions, this data can be used, especially when trying to protect the privacy of individual citizens. Post (2001) and Solove (2008) both argued that ever since the early 20th century, the term privacy has been challenging to define. However, whenever the term

privacy is used in combination with an increase in data collection, most tend to agree that the privacy of individuals has come 'under attack'. This threat occurs due to increased personal data usage in data-driven work (Warren & Brandeis, 1890; Hildebrandt, 2006). Further, other works on this theme of personal data usage in data-driven work by Solove (2008) and Zhang (2018) underline that the dawn of the information age has made privacy a frontline issue where an alarmed public is aware of current privacy issues but seems to readily give out data regardless.

Within this context of data-driven and privacy developments, local governments collecting large amounts of data can now gain valuable insights into these data sets in ways that were not possible before. For example, data from the wireless network points throughout cities can give regional governments insights into where the largest concentrations of people are on any given day (Data Protection Authority [AP], 2021). This indication can be helpful to decide in which areas to deploy resources or close off during events like the COVID-19 pandemic. According to Kim et al. (2014), analysing this data means that local governmental organisations such as the municipalities can also collect and analyse massive volumes of data to improve their decision-making processes. These data-driven decisions can, for instance, be used to create value for both citizens and the organisation. Meanwhile, the data-driven work developments within municipalities also raise the question of whether the privacy of individual citizens is impacted (Gruschka et al., 2018). This privacy concern was, for example, shown by the wireless location data collection project in the Dutch city of Enschede that detected concentrations of people, of which parts of the project were deemed as breaching citizen privacy by the Dutch data protection authority (AP, 2021). For this reason, Dutch civil servants were interviewed in this thesis to determine the privacy impact of data-driven work-related developments at Dutch municipalities.

1.2 Research questions

This research focuses on whether data-driven work within Dutch municipalities impacts citizens' privacy. More specifically, whether current main data-driven work developments affect the privacy of individual citizens as perceived by civil servants working at Dutch municipalities. The

bulk of the literature used to research the problem overview, motivation, and theoretical framework focuses on articles from the past decade. However, this thesis researches the current status of data-driven work and privacy at Dutch municipalities and seeks to discover the presence of the data-driven work and privacy themes that can be found in these organisations at the time of writing. Thus, the perceptions and experiences documented through interviews. These subjective metrics are analysed through inductive content analysis and chosen due to scope limitation to gain a detailed current overview of the broader subject instead of diving further into the in-depth technical details associated with analysing data-driven work using a metric system. This design is usually used when the phenomena are fragmented or difficult to define, such as broad concepts, like data-driven work and privacy (Elo & Kyngäs, 2008; Erlingsson & Brysiewicz, 2017).

This focus results in the following central research question: "Do data-driven work developments by the Dutch municipalities change the privacy of Dutch citizens as perceived and experienced by civil servants?" The topic is researched from a civil servant perspective as privacy changes resulting from data-driven work at local governments do not necessarily mean that individuals can still exert a degree of control over their personal data or their inputs are acknowledged. The three sub-questions are: (i) "What are the data-driven work developments?", (ii) "What are data-driven work privacy developments?" and (iii) "How do municipalities perceive citizen privacy changes in regards to data-driven work developments?". In this context, all research questions focus on the main recent developments instead of all developments. In this study, the terms (Dutch) citizen and individual in its plurality are used interchangeably. Further, the EU's General Data Protection Regulation (GDPR) applies to all residents and thus individuals in the Netherlands (Voigt & Von dem Bussche, 2017). The research explores observations on how local governments perceive these possible changes when dealing with data. While the term 'data-driven' work is a relatively new but established concept, the concept of privacy is ambiguous and complex. Thus there are various definitions for privacy mentioned in the literature, and it is also a central theme within the GDPR. This legislation is part of the

theoretical framework of this thesis as it applies to Dutch municipalities when processing personal data and is a common driver for data and privacy protection.

1.3 Approach

Due to its complexity, qualitative observational research was done through a content analysis. The decision to go for qualitative research was made to help describe complex subjects such as perceived data-driven work and privacy changes and put their corresponding observations into more manageable parts. This approach aims to summarise collected interview data systematically so that conclusions can be drawn (Beiske, 2007; Saunders et al., 2007; Erlingsson & Brysiewicz, 2017).

Regarding scientific and social relevance, this analysis adds to previous research as, according to Hallinan et al. (2012), there is a lack of clarity among the general public about the shape, size, and especially consequences of data flows concerning their privacy. They state that citizens' superficial understanding of this environment feeds uncertainty, leading to a general acceptance of the release of data. The authors moreover mention that individuals in the past often viewed giving up information as an unavoidable necessity. Further, other literature such as that by Gruschka et al. (2018) state that the increasing scale of technology usage can adversely impact privacy, indicating a privacy issue. The research questions add to previous research that it researches current privacy changes relating to current main data-driven work developments within Dutch municipalities. Due to the ever-increasing impact, scale and new applications of data-driven work, the literature suggests that more insights into the ever-evolving state of citizen's privacy are required (Hallinan et al., 2012; Gruschka et al., 2018).

This research happens in the context of data-driven work and privacy development in Dutch municipalities. First, the definitions and main recent developments of data-driven work in general and within Dutch municipalities were explored in this environment. Second, data-driven work-related privacy changes as perceived by civil servants were analysed. Privacy, as

mentioned earlier, is a very ambiguous concept and can be challenging to define. Thus, the focus is data-driven work changes that impact people's privacy, such as an increase in personal data collection.

This thesis uses primary and secondary data to help understand social phenomena occurring because of data-driven work and privacy developments in Dutch municipalities. For this purpose, the primary data in the form of interviews helps to understand the observations regarding the perceptions of civil servants at Dutch municipalities on the impact of data-driven work-related privacy changes. At the same time, secondary data is used to understand the contexts these respondents are in. The secondary data will be used in the form of publicly available existing literature where relevant literature and studies, mainly from the last decade, have been used. The primary data will be collected through the aforementioned semi-structured interviews with relevant government actors who interact with data-driven work. This reasoning is expanded upon in the Method chapter (Ch. 3).

This interview data will be analysed using a content analysis that focuses on participants' experiences. Content analysis is used in this research because human experiences are complex, multifaceted and carry meanings on multiple levels; the content analysis is used in this thesis to evaluate patterns within the interview data to find deeper underlying interpretations (Erlingsson & Brysiewicz, 2017). These interpretations in this interview data are the perceptions and experiences of Dutch civil servants regarding data-driven development-related privacy developments. Because of the large amount of textual data created by the interviews, the analysis is made more specific by the research questions to draw conclusions that are not too broad and puts the focus on the privacy aspect of main data-driven work developments at Dutch municipalities and the things of direct relevance to these. The content analysis also finds noteworthy due to data contradicting or affirming literature and the frequency of certain words or phrases related to data-driven work and privacy; thus, it combines a small element of quantitative research within the qualitative research. Moreover, this content analysis focuses on a very specific timeline, which in this case is 2021, when the interviews have taken place. Having the

timeline focus on the present-day helps to create a snapshot of events as they are currently instead of focusing on the past or future. The limitations of using content analysis for this research are that due to the large quantity of text and amount of literature that must be read, it can be time-consuming and difficult to find specific answers to research questions (Elo & Kyngäs, 2008; Krippendorff, 2018).

For this part, the theory from the theoretical framework was translated into concrete features, which served as the first labels for a coding scheme to base the interview questions on and codify the interview transcripts. The criteria for interview participants were that these interviews were with people who work at Dutch municipalities and are involved with data-driven activities through their work (Further detailed in the methodology in Ch. 3). For example, this included civil servants involved with digitalisation, data collection, data storage, and data analytics projects.

1.4 Structure

The approach of this thesis is different from classic public administration theses because it includes a technological aspect. This technological aspect was included to understand the impact of data-driven work developments within the public domain. Because of that, the thesis starts by presenting the main definitions and current developments in data-driven work (and its usage in Dutch municipalities). This part (Ch. 2.1) introduces several identified benefits and risk developments of data-driven work and some common organisational considerations. Aside from exploring the main recent data-driven work developments aspect of the study, the research also focuses on understanding the concept of privacy (Ch. 2.2). In this chapter, the theoretical framework is selective and also includes a few parts of the GDPR. These parts are included because the GDPR also mentions privacy, and the regulation is applicable in the Netherlands. The GDPR is a compliance framework that standardises the legislation for personal data processing done by public and private organisations across the European Union. An explanation follows the theoretical framework as to the study's method used for data collection (Ch. 3). Then,

in Ch. 4, the analysis of the study is done. The thesis concludes with a conclusion and discussion chapter (Ch. 5), which includes the limitations of this study, some recommendations for future research, and details of possible implications for the public sector. The final part of the thesis closes with the references used for this study and an appendix containing the interview invitation (App. I), interview protocol (App. II), and an addendum on the transcripts of the in-depth interviews with civil servants from 15 different Dutch municipalities (App. III).

2. Theoretical framework

The research is about collecting stories about the perceptions and experiences of civil servants working at Dutch municipalities through interviews and analysing the resulting textual interview data using content analysis. As the content analysis aims to systematically and objectively describe phenomena such as the data-driven work-related privacy phenomena in municipalities, the theoretical framework helps retest existing data in a new context and provides a framework that helps connect and organise interview data and shed light on observations that might otherwise be overlooked or misinterpreted. Thus, the purpose of this chapter is to use the literature to help understand previous knowledge about the context of this study's interview participants. The chapter is divided into four parts: the thesis topic's background and relevance (i), a closer look at data-driven work (ii), the situation of data-driven work's privacy impact (iii), and a conclusion (iv). The background and relevance section (i) explores the mentions of the topic in literature and establishes expectations as to what will be found in the interview data. Further, in the defining data-driven work and its developments within Dutch municipalities (ii) section, recent literature explores the concept of data-driven work and its developments. The third section focuses on the privacy impact of data-driven work-related changes (iii). Finally, a conclusion (iv) summarises the key findings of this chapter.

2.1 Background and relevance

There is a parallel to be found in the increasing proliferation of data-driven work in public and private organisations and the COVID-19 pandemic, which also tells a story of a global development akin to the speed at which various data-driven work changes within the public and private sectors are occurring. Another resemblance with the (at the time of writing) current pandemic is that its implications have been hinted at for decades but still managed to catch both private and public organisations unaware (Desouza & Jakob, 2017; Voigt & Von dem Bussche, 2017). Thus, the title of this thesis is a play on words by analogy; 'a double-edge sword' provides a characteristic of data-driven developments within Dutch municipalities, exploring its various positive and negative changes and consequences, including a possible privacy impact.

The civil servants working with data-driven work for Dutch municipalities find themselves in an environment where while developing quickly, data-driven work implementation is still often in its infancy (Boyd & Crawford; 2012, Desouza & Jakob, 2017; Van der Weerdt & De Vries, 2014). According to Gruschka et al. (2018), despite being at this stage, data-driven work already provides significant advantages to organisations, including (local) governments. These advantages include better data capitalisation leading to improved decision-making abilities and the streamlining of internal processes. Therefore, governments of countries leading in IT innovation have scrambled to start data-driven projects to improve organisational efficiency and create value for citizens while promoting key societal aspects such as security, (digital) economy growth and public affairs engagement (Kim et al., 2014).

This uptick in data-driven work can also be found in the (local) Dutch public sector. According to an article created for the AP (2020), "the Netherlands is at the forefront of digitalisation within the EU" (p. 2). This growth is also present in Dutch municipalities that are (planning on) using data-driven work to quantify and rationalise new or existing processes. Any resulting data-driven decisions primarily aim to increase the effectiveness of management decisions. The usage of data-driven work in decision-making is becoming one of the highest priorities of the Dutch government. An interest that has been further increased by the rapidly expanding amounts of data collected by these organisations through the internet, social media, sensor data, and mobile devices. At the same time, the cost of storing and processing these large data sets has fallen drastically (Manzoor, 2016). Despite this increase in interest, data-driven work in Dutch municipalities still appears to be often a small-scale operation. For example, a statement that underlines this phenomenon can be found in an article by Van der Weerdt and De Vries (2014), who mention that it is "often it is even a 'hobby' or at least a small-scale experiment by an enthusiast who, for example, sits in the IT or marketing department" (p. 20).

An increase in the importance of privacy protection can be noted as data flows and technology play a more prominent role in shaping social structures. A key concern for civil servants arising

from new data-driven developments is privacy; much of the digital data used relates directly or indirectly to individuals. A data-driven approach results in increased access to personal information that, when used, can break data protection laws and threaten the privacy of citizens (Gruschka et al., 2018). Two articles also underline this privacy issue and highlight the importance of keeping data-driven work-related privacy developments in mind. For example, Hildebrandt (2006) argues that data can be used for extensive profiling of citizens based on existing data, impacting individuals' autonomy. In addition, Zhang (2018) states that if large-scale data usage does not adequately protect user data during this process, it will impact data protection and the data subject's privacy.

Despite the importance of citizens' privacy not being underestimated in the reviewed literature, civil servants are still trying to cope with the issue that people's understanding of these issues still lacks (Hallinan et al., 2012). Surveys by Eurobarometer in 2019 note that only a slight majority of respondents (57%) know of the existence of a public authority whose job it is to protect their data. Luckily, this awareness is growing, showing an increase of 20 percent compared to an earlier study in 2015. However, only a fifth of the respondents in 2019 knew precisely which public authority is responsible for protecting citizen data. This lack of awareness does not seem to be because citizens do not care; surveys from Ireland, for example, show that 84% of respondents regarded their privacy as 'very important'. In order to illustrate citizen perceptions on this issue, this concern follows closely behind crime prevention which was mentioned as being very important by 87% of respondents (European Commission, 2015, 2019). Thus this increase in data-driven work interest in the local public sector has been paired with (amongst other developments) an increase in privacy issues, which, while suffering from awareness issues, is still considered to be of importance to citizens.
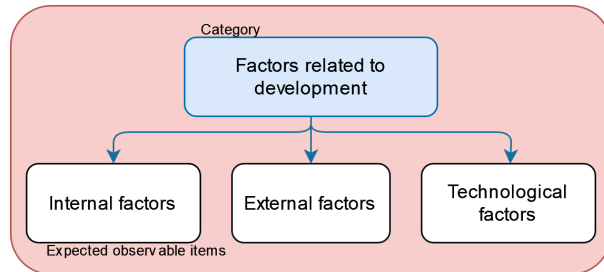
Figure 1. The influencing factors associated with data-driven work-related privacy developments.

2.2 Defining data-driven work and its developments

The scope of this thesis focuses on data-driven work at Dutch municipalities. Knowing what data-driven work is, its features, and how it can be identified supports the thesis's qualitative research coding scheme. The term data-driven work has recently been popularised within the public sector. While there is a broad diversity of ways to define data-driven work, it does not appear to be as ambiguous as the term privacy (further discussed in Ch. 2.3). Regardless, in literature, no single clear option can be found to define data-driven work. However, a common definition of being 'data-driven' found in works by Voigt and Von dem Bussche (2017), and Gruschka et al. (2018) are when organisations base decisions upon information inferred from (large) data streams through analysis instead of relying on intuition. Data-driven work gathers information from society, which are converted into data sets, then analysed and interpreted using domain knowledge to convert its outcomes into actionable insights, leading to a (better) informed decision. Through systematically collecting, analysing, examining, and interpreting available data, these informed decisions help create value for society. Further, combining people and the analysis of data can help find arguments to support the decisions made by people or find solutions that were otherwise overlooked. As previously indicated in this thesis, this better data-supported decision-making is not always guaranteed by simply implementing data-driven work. Thus, the term data-driven work in the public sector is not synonymous with a better decision-making process. As mentioned in Ch. 2.1, it also includes many risk (unintended) factors and possible challenges (Mandinach, 2012; Provost & Fawcett, 2013).

Other sources try to identify what is data-driven work by describing and comparing its processes or observable items. One of these approaches is described by Lai and Schildkamp (2013), who state that data-driven work can be identified by looking at the process. They state five steps that describe data-driven work processes; these include (i) purpose, (ii) data collection, (iii) analysis, (iv) interpretation, and (v) action. The purpose step is to find out what the question will answer. Further, the data collection step helps determine whether the data can help answer the problem. The analysis step is to find out what the results meant and what their implications are. Finally, the action step aims to determine how the interpretation and the analysis can help support a decision. These steps can be compared to existing or planned processes to decide whether or not they are actually 'data-driven' or not (Lai & Schildkamp, 2013). In summary, the conceptualisation of data-driven work has been approached in many different ways. However, most reviewed literature tends to agree that 'data-driven' is when organisations use analysing (large) data sets to form decisions instead of relying on knowledge or intuition.

In the Netherlands, the main benefit and chance for implementing data-driven work in the (local) government is gaining more knowledge about the data they are dealing with on a day to day basis. Kim et al. (2014) also give us some insights into why Dutch municipalities should nowadays implement data-driven work in their organisation. The authors state that "elected officials, administrators, and citizens all seem to recognise that being able to manage and create value from large streams of data from different sources" (p. 85). Both the public and private sectors can derive value from these large amounts of data. This value is gained through better and faster insights into social challenges as data-driven work makes it possible to follow specific social trends and developments in real-time. According to Klievink et al. (2017), governmental organisations worldwide have a long history of collecting this data on citizens, and data-driven approaches could be seen as an avenue to use that data further. Civil servants can make more substantiated decisions using the knowledge gained from data-driven work (i.e. with descriptive or predictive data analysis). This data-driven decision making leads to more focused and complete decisions where the intended value is easier to achieve. This advantage is corroborated

by McAfee et al. (2012). In this article the authors even go as far as stating that "the evidence is clear: data-driven decisions tend to be better decisions" (p. 9), referencing identified better overall organisational performance. According to an article by Anderson (2008), this outcome is possible because data-driven algorithms can sometimes find patterns where traditional analysis cannot. Anderson further argues that "we can throw the numbers into the biggest computing clusters the world has ever seen and let statistical algorithms find patterns where science cannot" (para. 13), showing the difference with traditional methods of analysis. For example, buying in for the social domain or cybersecurity as a municipality can be further substantiated because municipalities can better understand the consequences and reasons for supporting a decision after in-depth data analyses. This data-driven better decision making is in contrast to the current situation, where most insights are gained through superficial knowledge and experience. Another key advantage is that data-driven approaches can merge various data streams and mine them for data-driven insights (Manzoor, 2016; Malomo & Sena, 2017).

In the public sector, data is mainly sourced from their organisation, partners, or publicly available sources. However, Van der Weerdt and De Vries (2014) make a distinction between municipalities and other public organisations, stating that the former "do have large amounts of structured data" of their own (p. 32) while the latter uses unstructured data from a variety of sources. This data, for example, can be used for better resource allocation. They further argue that, above all, it can also indicate where no support is required in the specific field of competence the data project is placed in, and residents can be motivated to resolve their own issues. These valuable insights into real-life situations form the inspiration for effective data-driven work implementation and management (Kim et al., 2014; Malomo & Sena, 2017).

Aside from data-driven decision making, there also is a financial aspect at play. According to McAfee et al. (2012), the private sector benefits identified include more data-driven organisations being on average 6% more profitable and 5% more productive than their competition. These data-driven work implementations in the private sector have helped alert public servants to its potential benefits and use for saving money through, for example, more

efficient resource allocation (Malomo & Sena, 2017). In line with the topic of efficient resource deployment, in the public sector, financial benefits have been identified by Brown et al. (2011). The article states that implementing data-driven approaches in the European public sector would result in a cost reduction of 20 percent, leading to the value creation of billions. In total, in the developed economies of Europe, government administrations could save more than 100 billion euros in operational efficiency improvements alone by using data-driven work. Even without including potential benefits from detecting fraud, increasing tax revenues or reducing errors through data-driven work (Brown et al., 2011; Malomo & Sena, 2017). However, these data-driven work developments within (local) public sectors do not only contain benefits, and the following paragraph introduces several main risks identified in the literature.

To better understand the current situation around data-driven work and its role within Dutch municipalities, the theoretical framework also includes several risks and challenges which can be present during the usage of data-driven work. These factors can sometimes have harmful and unintended consequences on the entire organisation (Spillane & Diamond, 2007). Various authors have written about the risks of data-driven work, such as the technological, organisational, public, and personal perspectives. In one of these sources, Bopp et al. (2017) identified a risk that most organisations have in common: they are under immense pressure to become data-driven even if the value or goal is not always clear. This pressure is increased by the complexity of embedding new technology in an organisation, which may require radical changes in its existing processes. Aside from the costs for the initial setup mentioned earlier and receiving support from senior management, a culture shift is also needed. While being true in both the private and the public sectors, Malomo & Sena (2017) mention that this point is unfortunately not yet sufficiently appreciated by key stakeholders within local governments. This lack of prioritisation by some stakeholders can create a capacity problem as data-driven work implementation requires employees willing to know how to (efficiently) work with these systems.

The 'people' factor is reinforced by public organisations requiring flexibility due to the rapid development of data-related technology; it can prove difficult to anticipate what technological changes might occur soon. This unpredictable future means that organisations may need to deal with unforeseen risks. Thus, to successfully implement data-driven technologies, local governments need to develop internal skills that facilitate both modelling complex phenomena and analysing them. This capacity requirement makes skills to work with data-driven techniques necessary to implement a data-driven approach within local governments (Manzoor, 2016; Malomo & Sena, 2017). Kim et al. (2014) also mention this lack of competent professionals as one of the key risks when preparing to implement a data-driven approach, reporting that "searching for and finding competent data scientists (...) is difficult and expensive for most organisations" (p. 80).

According to a 2011 report published by McKinsey, in the US alone, some 140,000 to 190,000 specialist data scientists and a further 1.5 million managers will be needed by 2018. In addition, most local governments are grappling with the issue of how to exploit the different data they hold and hire people with the necessary analytical capabilities so that they can move toward data-driven decision making (Brown et al., 2011; Malomo & Sena, 2017). In 2017, the Centre for Crime Prevention and Safety released a study about the information position of Dutch municipalities (Goudriaan, 2017). This author similarly mentions that data-driven work projects at municipalities are often small in scale and that this is also related to a capacity problem being present in many of the municipalities investigated in their study. Within a small municipality, the data protection related department often consists of just one person, who in some cases is not even working full-time at the municipality in question.

Aside from implementation issues, the reviewed literature also mentions data protection and data management when discussing data-driven work risk developments. Just as citizens can benefit from public organisations' increased efficiency when successfully implementing data-driven work, the opposite can occur when it is not implemented correctly. According to Bopp et al. (2017), when facing problems with implementing data-driven work, the implemented data-driven

systems should focus on helping to empower organisations by keeping in mind their unique data requirements. The authors argue that "the design of information systems should work towards empowering organisations in ways that make sense for their unique data needs and those of their constituents" (p. 3608). The article further helps underline that if this is not the case, the approach does not lead to better decision making through a process of data-driven decision making, monitoring and evaluation. Instead, it can lead to data drift, data fragmentation, and the possible erosion of organisational autonomy (p. 3611).

The data drift mentioned can cause a situation where the collected data does not serve the organisation's goals. While data fragmentation means the data collected is not of a high enough quality to base management decisions on. The erosion of autonomy happens when external actors exert influence and change decisions and prioritizations about data work within an organisation. The article further states that an incorrect implementation can lead to information being gradually distorted when disseminating information (Bopp et al., 2017). This risk is also identified by Zhang (2018), stating that distortion can harm the reliability and authenticity of the data and concludes that "therefore, ensuring the authenticity and reliability of data is extremely important" (p. 276). Together, these risks of improper implementation can create a cycle of increasing disempowerment within the organisation, making it more challenging to achieve its goals (Kim et al., 2014; Bopp et al., 2017).

The risk of disempowerment is not only present within organisations but also plays a role for individual citizens. During a 2019 Eurobarometer study, 62% of respondents indicated they felt that they have no or only partial control over personal information provided to organisations mentioned they are concerned about. Despite showing a decrease of five percent compared to an earlier study in 2015, this result still represents a clear majority of respondents. If organisations are too focused on quantification and rationalisation and data is not screened effectively, this can deceive organisations. Thus, negative effects can occur for individual citizens, such as decontextualisation concerns (Zhang, 2018). This decontextualisation happens when processes create unexpected unethical side-effects such as racial profiling. Gruschka et al. (2018) also

mentioned this privacy aspect, underlining the usefulness of processing large quantities of data and reported the "high privacy risks when operating on personal data" (p. 5027).

New technologies have raised privacy concerns numerous times for over a century now. However, the rapid advancement and, according to Solove (2008), "profound proliferation of new information technologies during the twentieth century—especially the rise of the computer" (p. 4) have made people's privacy a frontline issue. A common example of data-driven work (unintentionally) using sensitive personal data is when an algorithm used by a large US-based retailer successfully predicted that a person was pregnant based on their search queries alone (Gruschka et al., 2018). This example shows that data-driven methods have to consider more than just the information collected and stored alone but also the data eventually processed or predicted. According to Zhang (2018), the combining of data through (predictive) data-driven methods indicates the existence of a so-called 'slippery slope', stating that the issue is reinforced by citizens freely trading off their data on the internet for (more convenient) access to certain services or because they themselves make it publicly available on social media. He further argued that criminals could use this data and further harm people's privacy, aside from causing economic losses. This situation leads to a chain of events culminating in a negative outcome. The author continues by stating that being unaware of how this personal data can harm one's own or others' privacy can cause even further forms of data leakage, which can subsequently impact individuals' privacy.

In general, citizens' seem aware that data leaks occur but still consider the 'tradeoff' to be attractive (Solove, 2008; Gruschka et al., 2018). However, this lack of clarity about privacy risks creates uncertainty, feeding citizens' fears about data processing and its possible consequences for both society and the citizen. Hallinan et al. (2012) state that one of the key factors to protecting privacy is data protection. The authors argue "that the public seem to have resigned themselves to the increasing release of data as a necessity of life in the modern world" (p. 264) and that "whilst not unaware of dangers and the existence of structures through which data processing and protection operate, there is a lack of understanding as to how and why they

operate" (p. 270). A lack of understanding means there is a limited basis upon which they can make decisions when choosing whether to share personal data. This combination of factors results in a case of bounded rationality, where cognitive limitations, imperfect information, and time constraints lead to a sub-optimal decision in agreeing to a tradeoff where access is traded for personal information. To change this situation where citizens make decisions on limited data, they need to adapt and increase their data literacy and awareness. However, efforts to improve data literacy are presently primarily limited to civil servants and scientific researchers (Hallinan et al., 2012; Zhang, 2018).

Protecting data also means focusing more on the proliferation of data leaks in the local public sector. According to the Centre for Big Data Statistics (2019), municipalities submitted 1,525 reports of data leaks in 2019 alone. Because of that, these account for almost 33% of all incidents in the public domain. In 2019, the AP (2020) revealed that municipalities reported the most data leaks within the public sector. They were reporting nearly 30% more data leaks than in the year before. In contrast, municipalities reported around only 10% more data leaks in 2019 than the year before. Accidentally "sending or handing over personal data to the wrong recipient" (p. 13) is the most common occurrence of data leaks, taking place when people send messages to the wrong recipient. Municipalities are also more often victims of phishing, where cybercriminals disguise their emails as those of different organisations. This data protection risk leading to further privacy risk must be especially monitored if the municipality is the controller or processor of citizens' data.

Another issue when preventing personal data from leaking is that anonymisation is a popular prevention method but comes with its own limitations. While the goal is to fully anonymise this data to prevent individuals from being identified or tracked within large data sets, this is difficult to achieve fully. While these techniques are also mentioned later in this chapter, in practice, it appears privacy cannot be protected through only the anonymisation of data. Zhang (2018) notes that "the protection of privacy cannot be effectively achieved only through anonymous protection" (p. 276). In many cases, individuals can still be tracked using these data points

because these large data sets often still result in a unique combination, from which a set of identifiers can correspond to a single individual (Gruschka et al., 2018). This argument is supported by Malomo and Sena (2017), who mention that "as the volume of data grows, 'deductive disclosure' becomes a major risk as (...) de-identified data can easily be re-identified" (p. 9, as cited in Ohm, 2010). A well-known case of re-identification happened in 2006, where a Netflix movie recommendation method competition data set allowed for the re-identification of anonymised data. In this competition, Netflix made a data set with over 500,000 customers' rating data available. While this data set only content subscriber IDs and movie ratings, researchers could still re-identify individual customers with a high probability by combining this data with publicly available information. This outcome results in the risks where efforts to protect data can still be de-anonymised using what is known as a 'background knowledge attack' (Gruschka et al., 2018, as cited in Narayanan & Shmatikov, 2008, pp. 111–125). This issue, coupled with the lack of consequence awareness among citizens, can lead to information leakage when using anonymised data sets (Zhang, 2018). These factors within data-driven work can be managed by taking steps within, for example, Dutch municipalities to mitigate risks and capitalise on its benefits.

The other organisational considerations include how municipalities themselves now often collect, store, and process large quantities of (personal) data. As a result, local governments will have to find a way to manage and supervise the usage of their large data sets as well. Further, while this text focuses on municipalities in the Netherlands, this is not just a local or national issue but also within the EU; there are concerns regarding the usage of large amounts of data through data-driven work. This widespread use of data means that various organisations and states can cooperate in implementing regulations that protect the privacy of their citizens, which is reflected in the GDPR. Broad support within the organisation is necessary to implement data-driven work within the processes of the municipality successfully, as cooperation between various departments is necessary (Van der Weerdt & De Vries, 2014).

This cooperation must also take place outside of the municipality as it often also tends to the needs of communities that may stretch beyond its administrative boundary. According to Manzoor (2016), this information position creates a situation where the organisation is unlikely to have all the necessary data to entirely understand complex phenomena within the community. This conflicts with most data-driven approaches, which assume that the organisation has access to all the required information. As a result, several (local) governmental organisations will need to work together and share the data (Malomo & Sena, 2017). Kim et al. (2014) contribute to this notion of cooperation by stating that governments should "collaborate with ICT big brothers" (p. 57), large private companies such as IBM, Facebook, Google or Amazon, who can give access to large datasets (including multivariate data). This data can include census, housing, and medical data.

Public servants, administrators, and citizens all appear to identify the benefits of managing and creating value from large data sets in various (unstructured/structured) forms. However, Kim et al. (2014) and Zhang (2018) mention a need for caution and that governments need to take a step-by-step approach to find what is realistically possible and set the correct goals for their organisation to promote data-driven work implementation. Zhang further states several options to manage data-driven work within governmental organisations. The most common action local governments can take, as briefly mentioned before, would be to anonymise the data. This anonymisation recommendation can also be found in an article by Desouza and Jakob (2017), who argues that "anonymisation and pseudonymisation techniques are necessary (...) in order to prevent unintended information leakage" (p. 2002). This anonymisation is done by transforming the data so that it can still be used to gain meaningful insights into data sets but also in a way that it can no longer be used to identify or trace the person initially tied to the data. Identifying data can include values like a citizen service number (BSN), bank account number, or IP address. However, whether its usage has a conflict with any of the privacy regulations or is at risk of re-identification is something that should be assessed on a case-by-case basis (Desouza & Jakob, 2017; Zhang, 2018).
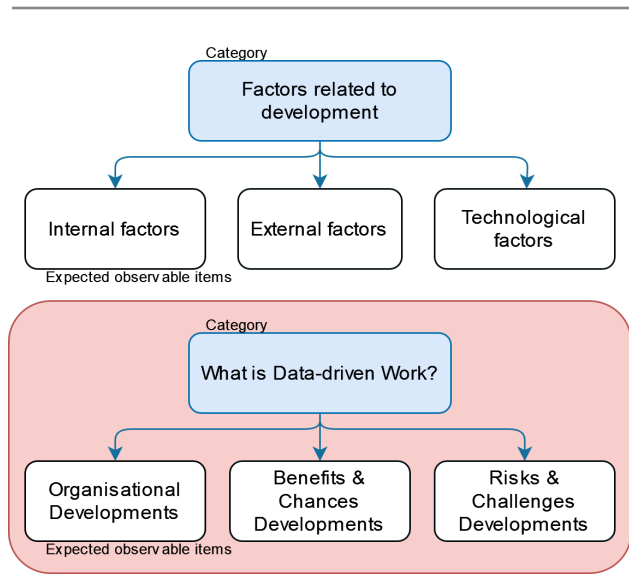
Figure 2. The data-driven developments.

2.3 Investigating data-driven work's privacy impact

Public servants also face the challenge of conceptualising privacy and its impact when creating policy for data-driven work or engaging with it in day-to-day activities. This challenge includes civil servants having to determine what types of data can lead to privacy risks and when to classify data as sensitive/personal. According to Solove (2008), privacy is a concept in disarray, and nobody can precisely articulate what it means. The author further states that it is a sweeping concept that includes many aspects such as control, freedom, and personal data protection. Privacy is often seen as a person's right to keep their personal matters a secret and also means someone can perform an action or think about something without others knowing about it. Warren and Brandeis (1890) first described this fundamental freedom and human right as the 'right to be left alone'. This description means that privacy is maintained when individuals are not observed on influences from the outside. However, this conceptualisation only focuses on the 'right to life', which protects from physical interference, and 'liberty' means freedom from restraint. Later there was also attention to people's intellect and feelings.

Over the last century, current definitions of privacy and the scope of legal protection have broadened from just covering the right to life to now encompassing more rights and freedoms and one's own identity. Examples include the right to (in)tangible personal possessions and civil privileges. More recently, data-driven work-related literature often raises privacy concerns in both the private and public sectors, as described in the previous chapter (Warren & Brandeis, 1890; Hildebrandt, 2006). Post (2000) further underlines the complexity of conceptualising privacy, calling it a value "so complex, so entangled in competing and contradictory dimensions (...) engorged with various and distinct meanings" (p. 2087). In his article, he even states his despair at whether the concept can be usefully addressed at all.

Having been given some insights into the complexity of privacy, one common theme identified through analysing the relevant literature during the thesis research is control over personal data. Van der Weerdt and De Vries (2014) state that "there are more and more voices from citizens who want to have control over their own data" (p. 16), and this degree of control individuals have over their privacy will serve as the key identifier for this thesis. Solove (2008) also included this degree of control, who used definitions from other authors to create six general types that are commonly used in privacy analysis. These six types include: (i) the right to be let alone, (ii) limited access to the self, (ii) secrecy, (iv) personhood, (v) intimacy, and (vi) control over personal data (p. 204). The 'right to be let alone' was briefly mentioned earlier and focused on the right to be free of physical interference. Second, limited access to the self is the right to prevent unwanted access by other people. The fourth type, personhood, is about protecting people's individuality, dignity, and personality. The intimacy aspect is about being able to limit access to intimate relationships or life. Finally, the control over personal information type is the ability to control one's personal data. Fuchs (2011) states that the privacy discussion is often centred around the degree of control individuals have on any data stored and whether public or private organisations use the data. Its importance is made evident by how often the 'degree of control' was mentioned in one way or another when attempts were made at conceptualising privacy. In an earlier article, Solovo (2004) also states that control over information is used to

define privacy. Schoeman (1984) distinguishes "the measure of control of an individual over personal information, intimacy, and visibility" (p. 2) as one of three definitions of privacy.

Similarly, Gormley (1992) also mentioned control as one of four types of privacy, defining privacy as "citizens' ability to regulate information about themselves" (p. 1337). Thus, this thesis limits the scope of privacy to the 'degree of control' citizens have over personal data as the conceptualisation for privacy as the same aspect can be found in the EU's GDPR. This reasoning is further discussed in the next section, the method chapter. Control is used in favour of consent, which, according to Cullen & Mayer-Schönberger (2014), has "problems associated with making consent meaningful and practical are widespread and well documented" (p. 16), and the current focus on consent has shifted the responsibility for data protection to the data subject instead of the controller or processor. Instead of a technical, individual, collective, or market perspective, this thesis also uses the regulatory view of privacy. Focusing on the earlier outlined degree of control individuals have through investigating whether and to which degree this information is being used by Dutch municipalities.

Regarding the GDPR's relevance to this research, in 2016, the EU took the next step in data protection and replaced the Data Protection Directive from 1995 with the 2016/679 GDPR (Voigt & Von dem Bussche, 2017). After a two year adjustment period, the GDPR first came into force in May 2018, where all organisations subject to this regulation had to be compliant (European Council, 2020). This legislation is relevant to all European Economic Area (EEA) organisations, the EU, and any international organisations that process data from EU citizens. Because of that, the GDPR affects nearly every country in the world (Gruschka et al., 2018). It is important to note that this new framework is not a directive but a regulation. Because of that, the legislation is directly applicable in all EU member states and does not need to be implemented by national legislators first. This aspect helps to create data protection uniformity throughout the EU (Voigt & Von dem Bussche, 2017). The approach of using this regulation to help support qualitative research about the impact of (increased) personal data usage on privacy was chosen because this

legislation is also active in the Netherlands and plays a crucial role in regulating the privacy of citizens living in Dutch municipalities.

Another reason why the scope of this thesis focused on the GDPR is that data-driven work can include the personal data of individuals, and the GDPR only applies when personal data is being processed. According to Art. 4 of the GDPR, personal data is "any information concerning an identified or identifiable natural person ('data subject')" (par. 1). The GDPR applies to both controllers, and to a lesser degree, to the processor of personal data. The regulation becomes applicable for organisations the moment any processing takes place. According to Art. 4 No. 1 of the GDPR, 'personal' data is deemed personal when any information can be related to an individual or help identify them GDPR (Voigt & Von dem Bussche, 2017). This information includes indirect identifiers like IP addresses, pictures showing people, and phone numbers, while direct identifiers include information like full names or ID numbers. If such identifiers are not present, the data is considered 'anonymous' as mentioned before, and processors of this anonymous data are outside the scope of the GDPR (Recital 26) (Gruschka et al., 2018).
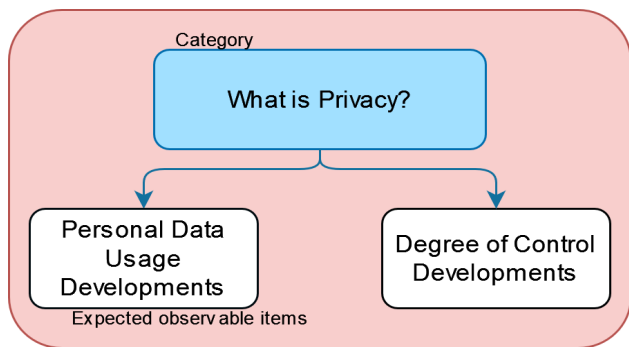
Figure 3. The data-driven work related privacy developments.

In the case of data-driven work that uses personal data within Dutch municipalities, the local government is the 'controller' and sometimes 'processor' of information of the 'data subject', the individual. Suppose municipalities do not put data protection principles and safeguards to protect personal data in place, such as using pseudonymisation or anonymisation. In this situation, they could breach the privacy rights of the individuals involved (Gruschka et al., 2018).

Therefore, in many situations where municipalities use data-driven work that processes personal data, these organisations will act as the controller. According to GDPR Art. 4 No. 7, a controller can be a legal or natural person, agency or public organisation that either alone or together with others decides the purposes and means of processing personal data (Voigt & Von dem Bussche, 2017). Further, the regulation applies to any controllers (and processors) that process the personal data of EU citizens, regardless of whether the controller or processor is in the EU and the processing takes place in the European Union. In this case, the term 'processing', according to Art. 4 No. 2, includes when (a set of) operations is performed on personal data, whether or not that is done through automated means. According to Voigt and Von dem Bussche (2017), this indicates that any usage of data is considered processing.

The GDPR uses a risk-based approach to data security; if controllers and processors do not ensure compliance with the GDPR or cannot prove their compliance (burden of proof), they could be subject to fines. If these conditions are not met, this is directly enforceable with a fine of up to 20 million EUR or up to 4 percent of an organisations' global turnover (Voigt & Von dem Bussche, 2017). These fines can be given for not complying with the regulation, and a resulting fine is always based on the articles which have been breached. Furthermore, if the regulation is breached and individuals have incurred (in)material damages due to infringement, they also have the right to compensation under the GDPR (Gruschka et al., 2018). This sanction means that preventing incidents and avoiding fines can be an organisational driver for privacy protection.
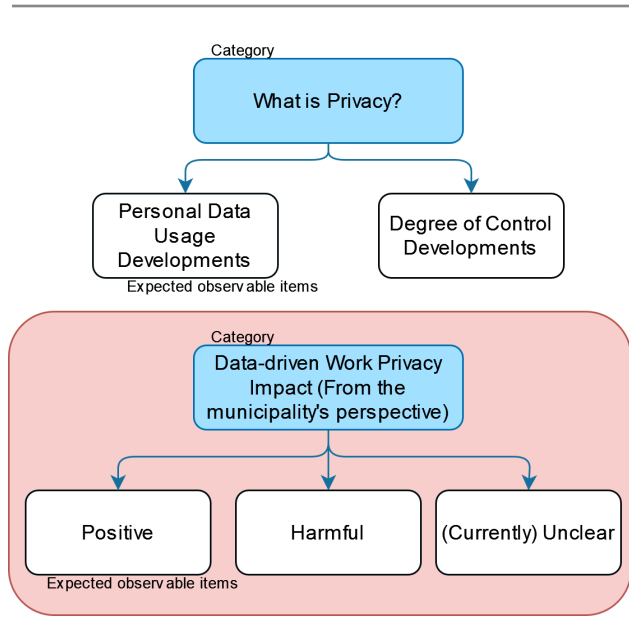
Figure 4. The perceived impact of data-driven work related privacy developments.

To more clearly describe how the theoretical framework guided the creation of the methodology, a visualisation of the key elements from the theoretical framework was created. The model below (Figure 5) visualises concepts and relationships discussed above and attempts to present these aspects, such as data-driven work and privacy aspects from the theory, in a single model.
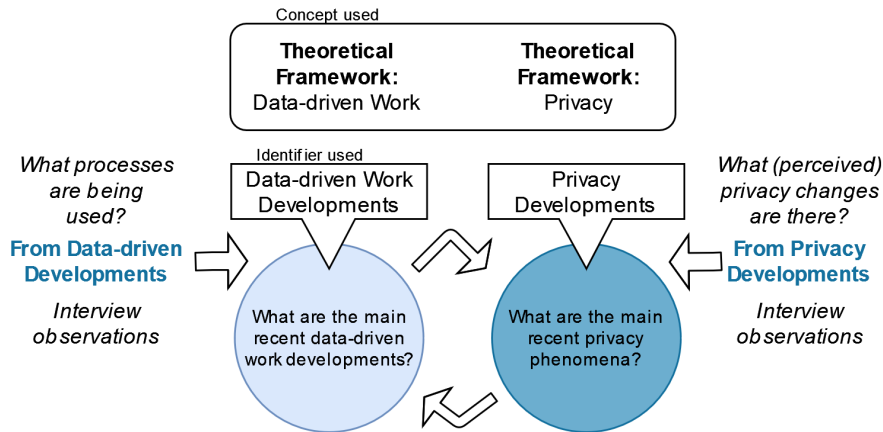
Figure 5. Visualised model of the theoretical framework.

This figure shows the data-driven work and privacy context the interviewed civil servants at Dutch municipalities are in. The expected observations include perspectives from the literature

on data-driven work, privacy and its focus on personal data and recent legislation such as the GDPR (Schoeman, 1984; Gormley, 1992; Solove, 2008).

2.4 Conclusion

This section aimed to identify the literature's key theoretical expectations, insights, and claims that helped inform and guide the content analysis. These insights were used to review existing knowledge about civil servants working in Dutch municipalities in the context of data-driven work implementation and related privacy factors. The background and relevance section identified three leading data-driven work development factors (Fig. 1), technological (i), internal (ii), and external (ii). These factors are driven by citizens having higher expectations and municipalities needing to solve increasingly complex problems. Furthermore, an increasing amount of data is being stored and collected by these local governments, for which new data-driven methods are implemented to gain new insights into this information. First, these data-driven methods are made possible by the newly widespread availability of technologies (i) that can be used to help tackle these complex issues and support data-driven decision-making processes. This development encompasses methods for new descriptive analyses that can analyse and describe data and predictive analytics that can identify the likelihood of future outcomes. Second, the internal factors (ii) identified are that municipalities want to develop data-driven work out of political reasons, as part of a long-term strategy, or to fit other organisational ambitions. Third, external factors (iii) include new legislation such as the GDPR that mandates better privacy frameworks (as mentioned in Ch. 2.1 and 2.2). Similarly, (inter)national developments can cause municipalities to want to follow and not be 'left behind' and keep up with the latest developments. Thus a certain degree of hype can also be a concrete, identifiable feature of external factors relevant to data-driven work developments.

The main recent data-driven developments (Fig. 2) include benefits and chances (i), risks and challenges (ii), and organisational developments (iii). The identified benefits and chances (i) features consist of saving costs, improving organisational efficiency, creating value for the

organisation and citizens, and helping to provide answers for complex solutions. Data-driven implementations aim to improve decision-making by relying on a more profound understanding through data analysis rather than superficial information and experience. Thus, proper implementations of data-driven work developments benefit both municipalities and citizens, for example, by offering better services or more efficient resource usage. However, various risks and challenges (ii) also have been identified. For example, municipalities can experience capacity and competency issues when implementing data-driven work. These issues relate to culture mismatches in the local governments where digital literacy and related competencies are low. Further, there are not always funds available for new data-driven methods, creating another implementation barrier. Another challenge is that the quality of the data collected by the municipality may now always be of the right quality. The municipalities then need to commit to considerable (time) investments to make data-driven work viable. Moreover, these data-driven work results do not always have the required nuance that municipalities sometimes need when dealing with citizens. A mismatch between people and data can cause disempowerment issues, where the data does not match the unique data needs or data literacy is low. This problem can cause the data-driven work implementation to be a negative factor rather than a positive one. Main organisational developments (iii) include that the successful implementation of data-driven work through dealing with these risks can help reduce any adverse effects. For example, municipalities could refer to the legislation and transform personal data into data without personal identifiers to address privacy risk developments. The main recent developments include changing the organisational culture to fit data-driven work, (de)centralisation efforts, and prioritising data-driven work to balance solving social issues and maintaining citizen privacy.

Regarding privacy developments, the concept of privacy is defined as a person's right to keep their personal matters a secret and is in this thesis focused on the degree of control (ii) someone can exert over their personal information (i). In the context of data-driven work, two main features of privacy developments were identified in the literature (Fig. 3). These developments stem from municipalities collecting, processing, and storing an increasing amount of personal data. Personal data (i) means any information which can be related to an identified or identifiable

natural person. The earlier mentioned increased data collection, which includes the increased collection of sensitive information, has led to increased privacy concerns. Despite the consent mechanism mentioned in the GDPR (Ch. 2.3), where data owners are asked permission to process and store their information, this is not a central aspect due to data owners not knowing what they consent to. Instead, while data protection (such as access restriction, (pseudo) anonymisation and aggregation) methods are often applied to personal information, personal data usage is changed along with the degree of control (ii) a citizen can exert over their data. Thus these privacy developments lead to a possible impact on citizens' privacy. This privacy impact is the difference between the situation before and after a data-driven work related to privacy developments.

The impact of data-driven work developments within Dutch municipalities also includes three features identified from the literature (Fig. 4). These main aspects are positive (i), harmful (ii), and (currently) unknown (iii) impacts. Positive impacts (i) can include increased attention to privacy and data protection. Similarly, this increased attention has led to tighter legislation such as the GDPR, resulting in organisations including privacy by design. Further, data protection has also become more important because of the legislative requirements, incidents at other municipalities and intrinsic motivations such as ethical concerns. This data protection includes new anonymisation efforts such as (pseudo) anonymisation and aggregation. Other concrete features include the harmful impacts (ii), such as increased data leaks and increased personal information collection and processing. The identified negative features also include further decreased control over personal information because of complex data streams. Due to this complexity, it can be challenging to know whom to approach about personal data. Studies have shown that very few people are aware of the watchdogs that should investigate and possibly reprimand or fine offenders. Moreover, some (currently) unclear developments (iii) are identified; for example, municipalities can have shifting intentions in the future. Therefore, the personal data collected can be used in the future by persons with different purposes than during the initial data collection. Similarly, combining data and the proliferation of new predictive

analyses can lead to re-identifying (pseudo) anonymised data, which also has a (currently) unclear privacy impact.

In summary, various Dutch municipalities are (planning on) using data-driven work. Data-driven work in this context means that an organisation bases decisions upon information inferred from (large) data streams through analysis instead of relying on intuition for quantification and rationalisation in new or existing processes. The theoretical framework is used in the content analysis in two parts. In the coding process, the theoretical framework provides general categorisations regarding what to expect from the interview data before these coding labels are made more specific through the content analysis. Further, during the analysis and organising section of the content analysis, the theoretical framework helps compare, affirm, or contrast the observations from the interview data, which help create the findings and conclude by answering the research questions and stating policy implications for Dutch civil servants.

3. Method

3.1 Research method and design

In order to explore the impact of data-driven work developments on the privacy of Dutch municipalities, this thesis concentrates on the philosophy of pragmatism. The main component of this research philosophy is the research problem (Saunders et al., 2007). Pragmatic philosophy, in this case, is used due to its focus on addressing real-world issues through research. Further, an inductive approach was used to understand privacy issues due to the recent data-driven work developments and identify emergent phenomena such as evaluative and technical approaches.

The study follows a cross-sectional time horizon because it was only conducted at one point in time. Moreover, because one-time data through a single method can address the research question, a mono-method choice and purposive sampling were chosen due to the study's design involving inductive-qualitative research (Bell et al., 2018). To ensure validity and the accurate value of the sample used and prevent research bias, as many steps as possible of the research process have been detailed (Savin-Baden & Major, 2010). Respondents include Dutch civil servants found through the social network website called 'LinkedIn' working at a Dutch municipality involved with data-driven work or municipalities planning to do so in the future. The method of finding interview participants meant there was an assumption that the respondent does (or will do) something related to data-driven work and that this data-driven work is likely to be somehow related to privacy. This assumption also highlights a significant limitation upfront, where their perceived data-driven work is not actually data-driven or does not relate to privacy. However, the criteria for respondents aim to limit this risk of reduced relevancy in interview responses. Further, another limitation is that theoretical saturation might be challenging to achieve with 15 respondents as it is difficult to gauge how much new information is collected in advance.

3.2 Data collection method

Based on the qualitative-inductive approach, the data collection for the content analysis was carried out through semi-structured interviews. The used approach does not just include objective facts but focuses on the subjective interpretation of an individual's experiences and proven observations to theorise about possible phenomena developments (Beiske, 2007). The aim of these interviews is to help analyse the experiences and perceptions of civil servants about data-driven work and related privacy developments at Dutch municipalities from their perspective. In total, 15 civil servants working for Dutch municipalities were selected to participate in these interviews ($n = 15$) because they are involved with data-driven work within their organisation. This homogeneous group, only people working for Dutch municipalities, includes respondents with three job descriptions: data advisors, analysts, and managers. These three functions, data manager, data analyst, and data advisor, were selected to collect interview data from multiple perspectives related to the same data-driven work process within municipalities. The expectations regarding the interpretation of the interview data are that respondents are unhindered and unbiased as far as that is possible when sharing their perceptions and experiences with due regard for the literature from the theoretical framework (Ch. 2), setting up expectations for the interviews and questions.

Respondents were found by searching for civil servants who (are planning to) work with data-driven work within Dutch municipalities. These respondents have been approached through the professional network website called 'LinkedIn' and were selected because they worked at a Dutch municipality with data-driven work in some capacity. This last criterion was often filled as they used the terminology 'data-driven work' in their profile or job description. Further relevant respondents were found by approaching people with a job description such as data analyst, data manager, data advisor, or data officer. These respondents were then asked whether they would like to participate through a connection request. Then, upon receiving a positive response and setting up the interview, the respondents were further informed about the research and asked for their consent to record the interview following the interview protocol.

The interviews had a length of between 30 minutes and 70 minutes. Because of the COVID-19 pandemic, all interviews were conducted through online conferencing programs (Microsoft Teams, Google Hangouts, and Skype). The main variance between interviews was that some respondents gave very concise short answers, whereas others were lengthier in their responses. Further, other differences were based on the function as data analysts focused more on the technical side of the data-driven work processes, which took longer to explain than the non-technical policy-centred answers predominantly given by data managers and data advisors. Overall, all respondents were enthusiastic about the subject, and it was possible to gain relevant observations from all respondents.

With permission from the respondent, an audio recording was made of every interview; this was used to create a transcription later. Aside from a recording, a notebook was used to write down any interesting statements or keywords that helped form follow-up questions to learn more about a specific topic or statement during the interview. The 10 main questions used in the interview are included in App. II. Aside from the 10 questions, frequent follow-up questions were also included, dependent on the respondent's answers. Aside from open-ended questions, probes were also used (i.e. "Can you elaborate on that idea?", "Could you clarify that?" or "Could you give an example?"). Not every additional probe was not recorded in the interview protocol, but they are included in the transcripts (App. III). Regarding ethical considerations, participation was entirely voluntary, and any (personal) information given was anonymised, stored securely, and discarded at the end of the research.

Further, two respondents expressed the desire not to have their names and their organisation mentioned in the research. Therefore, the participants have been given identifiers based on an abbreviation of their function description and a number. The municipality's names have also been anonymised by assigning each municipality its letter and the corresponding number, showing how many respondents work at the same municipality. This aggregation means that the function and province information from the respondents can still be used in the analysis while no

personal information has been included that could lead to the identification of the individual participants (Beiske, 2007).

Respondents	Function type	Municipality
(1) DM1	Management	A1
(2) DAN1	Analyst	D1
(3) DA1	Advisor	B1
(4) DAN2	Advisor	E1
(5) DM2	Management	B2
(6) DAN3	Analyst	A2
(7) DAN4	Analyst	A3
(8) DM3	Management	B3
(9) DAN5	Analyst	E2
(10) DM4	Management	B4
(11) DM5	Management	A4
(12) DM6	Management	C1
(13) DM7	Management	C2
(14) DA2	Advisor	A5
(15) DA3	Advisor	F1

Table 1. Respondents involved in the interviews

The steps involved in creating the coding scheme also help guide the process for the content analysis. The coding scheme for the interviews is centred around the expectations for the interviews from the literature in the theoretical framework and was created through translating observable features of the theoretical concepts (data-driven work, privacy, and its various facets) into concrete questions. The categories were ordered similarly to the theoretical framework to help maintain a coherent storyline throughout the thesis. These features were put in a coding scheme that helped form the interview questions with four main aspects, data-driven work developments (i), privacy developments (ii), factors related to the context of how these

data-driven work and related privacy developments develop (iii), and the data-driven work-related privacy changes impact (iv). The coding scheme focuses on how civil servants experience and perceive data-driven work and its main current developments, the factors related to these phenomena, and their privacy impact. The questions regarding data-driven work (i) and privacy developments (ii) are formed using the risks & challenges, benefits & chances, and (personal) data usage facets. The concrete features for factors related to these phenomena (iii) are divided by internal, external, and technological factors. Finally, the privacy impact developments (iv) are faceted by three positive, harmful, and (currently) unclear privacy developments. The latter indicates a possible perceived privacy change that can not yet be clearly identified using the current setup. All the questions were formulated in English and later translated to Dutch, which was the primary language of all the respondents (Beiske, 2007; Savin-Baden & Major, 2010).

Category	Code	Definition	Keywords
Data-driven work Developments (i)	Risks & Challenges	The respondent experiences risks or challenges to implementing data-driven work	Capaciteits problemen, organisatie cultuur problemen, tekorten vaardigheden, aantonen waarde, financiele tekorten, data kwaliteit, gebrek aan nuance
	Benefits & Chances	The respondent experiences benefits or chances to implementing data-driven work	Efficiency, effectiviteit, kosten besparingen, waarde creatie, oplossingen
	Organisational Developments	The respondent experiences organisational changes when implementing data-driven work	Decentralisatie, centralisatie, samenwerking, digitalisering, formele afspraken, prioritisatie
Privacy Developments (ii)	Personal Data Usage	The respondent indicates personal data usage in data-driven work	Gebruik persoonlijke informatie, gebruik ge(pseudo)anoni miseerde data,

			toegangsbeperkingen
	Degree of Control	The respondent perceives a decrease in the degree of control over personal information	Toestemming data eigenaar, controle over eigen informatie,
Data-driven Work & Privacy Development Factors (iii)	Internal Influences	The respondent experiences data-driven work developments being driven from within the organisation	Intrinsieke motivatie, strategie, masterplan, politiek, organisationale ambities
	External Influences	The respondent experiences data-driven work developments being driven from outside of the organisation	Externe incidenten, nationale ontwikkelingen, buitenlandse ontwikkelingen, wetgeving, hype vorming
	Technological	The respondent indicates that data-driven work developments are driven by (new) technological factors	Nieuwe technieken, implementatie stadium, toegevoegde waarde besef, meer informatie verzameling
Data-driven Work Privacy Impact (iv)	Positive	The respondent perceives that data-driven work improves personal data protection	Data (pseudo) anonimisatie of aggregatie,aandacht voor privacy, privacy by design, strengere wetgeving
	Harmful	The respondent perceives that data-driven work decreases personal data protection	Privacy lekken, toegenomen persoonlijke data verzameling, verminderde controle over persoonlijke data
	(Currently) Unclear	The respondent perceives that data-driven work effects are unclear	Veranderende intenties, her identificatie door combineren data

Table 2. The qualitative interview coding scheme is based on identified features from the theoretical framework and the author's expectations.

3.3 Data analysis method

This content analysis was done by identifying patterns regarding the data-driven work and privacy developments from the perspective of the civil servants at Dutch municipalities. The introduction (Ch. 1) and theoretical framework (Ch 2.) parts further substantiated the research questions and helped create expectations regarding the existing knowledge in literature about the context of data-driven work (i) and privacy (ii). This research describes its driving factors, recent main developments and related privacy impacts. First, the concept used in this research for data-driven work (i) is that work is 'data-driven' when an organisation bases decisions upon information inferred from (large) data streams through analysis instead of relying on intuition through quantification and rationalisation within new or existing processes. The corresponding data-driven work developments can be identified through organisational considerations, benefits and chances, and risks and challenges development indicators. Further, the factors related to these developments include their internal, external, and technological aspects.

The literature regarding the concept of privacy (ii) helped create an understanding of the importance of privacy developments and its associated concerns when using data-driven work within Dutch municipalities. Privacy in this context means a person's right to keep their personal matters secret which is narrowed down in this thesis as the degree of control they have over their own data. The concept of privacy in this thesis is any information that can be related to an identified or identifiable natural person. The (improper) implementation of data-driven work that collects and uses sensitive information from citizens could cause data protection issues in(directly) impacting citizen privacy, changing the degree of control they have over this personal information. This finding means the amount of personal data used and control over this information are seen as related. However, instead of combining the two variables, both are used as indicators because separate personal data and control phenomena could be identified from perceptions during the qualitative research. These privacy impact features include three different indicators, positive, negative, and (currently) unclear impacts. In this context, the privacy impact is the difference between the situation before and after a data-driven work (privacy-related)

development. In literature, some of these changes seem to have a clear positive or negative impact but can sometimes also include unclear impacts where it is unknown whether the change is (now or in the future) positive or negative. To use this approach to find phenomena from the civil servant's perspective at Dutch municipalities, the interview data, as mentioned earlier, was used (Savin-Baden & Major, 2010). While using the coding scheme from Table 2, the following questions were kept in mind:

- *"How do civil servants at the Dutch municipalities talk about 'data-driven work', how do they characterise it?"*
- *"What are respondents (and their organisation) trying to accomplish with data-driven work?"*
- *"What is the view of respondents on data-driven work's privacy developments?"*
- *"What are the experiences of respondents from balancing data-driven work and privacy?"*
- *"Do respondents understand the potential privacy impact of data-driven work?"*
- *"Do respondents seem reluctant when answering specific questions (like when discussing data privacy issues)?"*

Five steps were used to perform the content analysis coding process based on the interview data: (i) interview data was prepared and organised, (ii) interview data were reviewed and explored, (iii) the data coding scheme was expanded using, the (iv) codes that were assigned to the data, and (v) recurring phenomena were identified (Elo & Kyngäs, 2008).

All the collected material was entirely in Dutch and thus also transcribed and coded in that language. Before starting, any existing personal preconceptions about the research were reflected on and suspended as much as possible to be fully open to the experiences and perspectives of the respondents. As part of the first step (i), the audio recording transcription was coded manually and organised to ensure accuracy and increase familiarity with the information. During the next step (ii), the entire transcriptions were scanned to mark and scratch out parts irrelevant to the study. This step also included potential coding labels and reviewing the first potential

phenomena. Third (iii), the coding system was expanded using the reviewed interview transcripts. This step makes adjustments and adds more labels on top of the existing coding scheme used to translate theoretical concepts into interview questions. Next (iv), open coding took place where coding labels were added to the text. The transcripts were thoroughly scanned at that stage, and anything that was considered potentially relevant or interesting was highlighted. Further, parts of the text that matched the codes were also highlighted, and code labels were made more specific as new commonalities emerged. The last step (v) included axial coding, where the number of labels was further reduced, and a hierarchical order was created to convert the various coding labels into (frequently) recurring phenomena. Moreover, another significant part of this step included narrowing down the naming and defining of these generated phenomena. This process involved formulating precisely what each phenomenon means and how it gives meaning to the data collected. For example, 'capacity' and 'cooperation' were common phenomena identified. Therefore, it was also defined what kind of capacity (skills or willingness to use data-driven work) or cooperation (between data-driven work processors or citizens) was meant during this stage (Savin-Baden & Major, 2010; Krippendorff, 2018).

For the analysis part of the content analysis, the interview data analysis focuses on participants' perceptions and experiences. Content analysis is used in this research because human experiences are complex, multifaceted and carry meanings on multiple levels; the content analysis is used in this thesis to evaluate patterns within the interview data to find more profound underlying interpretations (Erlingsson & Brysiewicz, 2017). In this research, these interpretations in this interview data are those of Dutch civil servants regarding data-driven work-related privacy questions. The research questions are more targeted towards data-driven work-related privacy questions to help narrow down the content analysis of a large amount of textual interview data. Further, a small amount of quantitative research includes infrequent remarks about the frequency of certain words used during the interviews that are relevant to the coding categories mentioned above. Moreover, to make the content analysis more specific and compare the current situation to previous knowledge, this content analysis focuses on the present-day rather than the past or future (Elo & Kyngäs, 2008; Krippendorff, 2018). This

analysis aims to provide a detailed summary of civil servants' perspectives at the Dutch municipalities. As the interpretative focus analyses people's experiences rather than solely relying on pre-existing theoretical expectations, detailed experiences about each pattern are examined before moving on to more general developments identified. The content analysis balances the descriptions from identified phenomena with interpretations from the researcher, focusing on respondents' experiences and their meanings. Further, sometimes quotes from the respondents are used to engage with possible credibility issues (Bell et al., 2018; Savin-Baden & Major, 2010).

3.4 Conclusion

This chapter describes the method used to help answer the research question. To improve transparency and reliability, each step of this research method was described as extensively as possible. A qualitative approach through a content analysis was used in this thesis by conducting semi-structured interviews with civil servants from Dutch municipalities through purposive sampling consisting of 15 respondents ($n = 15$). The respondents were selected through the professional networking website 'LinkedIn' because they are (planning to be) involved with data-driven work within their municipality. The interviews and their open-ended questions were structured using a coding scheme. Further, the literature from the theoretical framework helped translate theory to substantiate the research questions and helped create an understanding of previous knowledge regarding the context of data-driven work and privacy both in general and in Dutch municipalities. These initial theoretical expectations were translated into interview labels and questions, which are further refined during the content analysis. The interviews themselves are centred around what open-ended questions would be of use to investigate the experiences of civil servants and what ultimately would help answer the research question.

This approach supports finding out how civil servants perceive and experience data-driven developments within Dutch municipalities, their related factors, the accompanying privacy developments, and possible future changes. The interview data was investigated through content

analysis to find out based on the perspectives and experiences of civil servants what current main data-driven developments take place and how these possibly affect citizen privacy. The content analysis itself found in the following analysis chapter (Ch. 4) involved reading through the text numerous times and categorising them. This analysis aimed to help the researcher to gain a deeper understanding of the interview data and compare individual respondents' answers with each other and previous knowledge and theories found in the literature, which help answer the research questions and state policy implications.

4. Analysis

In this chapter, the results of the interview data analysis are presented. This chapter aims to help answer both the main research questions and its three sub-questions by describing and quantifying phenomena through content analysis. The structure of this chapter is created around the key topics of the research, which are later used to answer the research questions as not all the results can be tied to a specific research question. This approach is because the problem overview presents a broad issue; this was made more specific and now gives more comprehensive results. For each topic, general observations about what the interview data showed are included. These observations are recurring patterns, trends, statements from the interviews affirming or contrasting each other, or individual quotes that help answer the research question. The results are supported with relevant interesting direct quotations and limited information about the respondents and their environments. Further, the categories found in the content analysis are also compared to the previous knowledge from different time periods stated in the theoretical framework. The quotations are cited together with the pseudonymised identifier given to each respondent (e.g. "DA1" or "DM2"), which was based on their role and an incremental number as stated in the methods chapter (Ch. 3.2).

The first things presented (Ch. 4.1) are the key observations regarding data-driven work development phenomena as perceived by civil servants. These are themed around data-driven trends within Dutch municipalities, as seen through civil servants' work in these developments. Second (Ch. 4.2), findings regarding the perceived importance of privacy and its development phenomena are presented. The third part (Ch. 4.3) includes an analysis of the interview data regarding the perceptions and experiences about influences of data-driven work and privacy development phenomena. The fourth part (Ch. 4.4) is about the perceived and experienced internal and external organisational developments and the identified phenomena regarding the data-driven work privacy impact. These findings include observations about the collected data, the context the respondents are in, and a comparison of the (dis)similarities compared to the

results from previous knowledge stated in the theoretical framework. Finally, the conclusion (Ch. 4.5) sums up the chapter and summarises the main points of the content analysis.

4.1 Dutch municipalities going digital — Data-driven work advantages and disadvantages

During the interviews, respondents often reported that the discussion about data-driven work implementation within Dutch municipalities was started years, if not decades before. However, an aspect of these discussions vital to this research is how various people perceive data-driven work and define it. Consequently, section 4.1 of the content analysis aims to investigate the perceived and experienced changes brought on by Dutch municipalities "going digital" and focuses on data-driven work and its implementation and how civil servants define data-driven work and perceive its advantages and disadvantages.

Defining data-driven work is also not easy for Dutch municipal civil servants

In line with recent literature, data-driven work also proved challenging to define for the interviewed civil servants; the interview data shows respondents reporting various notions of the concept of data-driven work. When asked about data-driven work within their respective municipalities, the respondents tended to answer in line with recent literature (Floridi, 2012; Voigt & Von dem Bussche, 2017; Gruschka et al., 2018) and view its concept in the broadest sense of the word and agreed that this concept is difficult to define. Specific criteria noted to engage with data-driven work at a municipality include having the data, capacity and skills available within the organisation. The general term was used in the context of digital information being used to perform activities. However, data-driven work can further be defined by its types, as pointed out by one of the respondents who, when discussing the uses of data-driven work within his municipality, commented that, "Datagedreven werken is natuurlijk een relatief hol begrip. Dus wat is dan het werkwoord van datagedreven werken? Dus wat mij betreft gaat het over datagedreven handhaven bijvoorbeeld. Of datagedreven voorspellend onderhoud of data gedreven besluitvorming of datagedreven beleidsvorming of je moet een werkwoord daarachter

zetten zeg maar. Om het wat inhoud te geven." ("Data-driven work is, of course, a relatively hollow concept. So what is the verb of data-driven work? So as far as I'm concerned, it's about data-driven enforcement, for example. Or data-driven predictive maintenance or data-driven decision-making or data-driven policy-making or you have to put a verb behind that, so to speak. To give it some substance."; DM6). This comment sheds light on the complex situation civil servants are in when engaging with data-driven work. In this example, the respondent even refers to data-driven work as a "hollow" concept, having no real value, worth, or effectiveness. It gains in substance and thus in its usefulness for the people implementing it only when data-driven work is actually implemented. Noteworthy here is that there are similar situations of respondents interviewed for this study where civil servants working for Dutch municipalities are struggling with defining data-driven work. These situations are discussed under section 4.3, which discusses the perceived influences of data-driven work and its related privacy factors. In short, civil servants tend to perceive the concept of data-driven work differently. However, some prefer adding a verb such as analysis, monitoring, or prediction to be more specific about what data-driven work involves.

Data-driven work implementation is reported to bring extensive benefits and chances

The respondents extensively mentioned the perceived advantages of data-driven work implementations during the interviews. This interview data shows four main phenomena about benefits and chances developments within Dutch municipalities as perceived by the respondents. These are the financial (i), analytic (ii), predictive (iii), and decision-making (iv) benefits and chances. First, what stands out are the financial benefits noted by respondents include better resource allocation and tax collection which are often mentioned as a side-note in literature but are predominantly the primary benefit mentioned by nine respondents when discussing data-driven work related benefits. The financial aspect ties into the other three times as better resource allocation and tax collection was reported to be due to better data-driven (predictive) analytics and decision-making (DM1; DA1; DAN3; DAN4; DM5; DM6; DM7; DA2; DA3). For example, one respondent, a data analyst, mentioned tax enforcement benefits, "De prioriteiten zijn natuurlijk wel eerst financieel voornamelijk [...] begin je met de partij die de belasting oplegt

zeg maar 15.000 adressen belasting op en in ons systeem komen wij op 16.000 ofzo. [...] Ja dat is wel de moeite waard om eens te onderzoeken waar het verschil van duizenden adressen zit. [...] En het bijkomende dan. Als je dan gaat onderzoeken komen die bijvangsten allemaal zeg maar." ("The priorities, of course, are primarily financial [...] you start with the party who imposes the tax, say 15,000 addresses are taxed and in our system we arrive at 16,000 or so. [...] Yes, it is worthwhile to investigate where the difference of thousands of addresses lies. [...] And the bonus is then. If you then start to investigate, all of these additional financial catches, so to speak, come into play."; DAN1). The data analyst in this quote gives insights into both the financial benefits from increased tax collection thanks to improved data-driven analytics but also gives insights into the massive scale of the datasets they are involved with. Further, this situation creates a financial incentive for the acquisition of more personal data. For this example, the context was that the analysis and monitoring of a system showed a discrepancy that allowed for extra tax collection. In short, the example showcased that the financial benefit, in this case, outweighed the effort of having to check thousands of entries to find 'financial catches'. What stands out is the willingness to invest extra time in data-driven work in exchange for monetary benefits as less emphasis is put on these financial incentives in recent literature.

Data-driven analytics are also perceived to provide better insights into existing and newly gathered information. The respondents reported that data-driven analysis makes it easier to monitor policy. One respondent, a data manager (DM4), stated that monitoring key performance indicators (KPIs) within the local society is used to elevate the local environment to a higher level. According to the respondent, this increased understanding of what is going on assists them in making the municipality cleaner, safer, and more liveable. However, in contrast with the other phenomena, predictive data-driven work has, according to the respondents, not been extensively implemented within Dutch municipalities yet. While implementation projects exist using artificial intelligence and machine learning, only fraud detection, which is mentioned earlier in this subchapter, and electronic invoicing are mentioned in the context of predictive algorithms.

One respondent talks about this electronic invoicing when discussing potential chances. This algorithm aims to predict which creditors are going to be late with sending their invoices, "Wat nu wel bij ons op tafel ligt om misschien iets mee te gaan doen is machine learning. Hoe heet dat, een random forest voorspelmodel op e-facturatie. Want we zien dat er aanbieders zijn die consequent te laat declareren, dus die het een jaar geleden ondersteuning geven, maar die factureren pas een half jaar later of een jaar later. Wij willen ervoor zorgen dat dit sneller gebeurt, zodat we niet aan het eind van het jaar bepaalde aanbieders achter de broek moeten zitten van he. Kom op. We moeten gefactureerd worden. En als we dan een beeld hebben welke aanbieders consequent te laat factureren, zouden met zo'n voorspel model of kan die kunstmatige intelligentie vanuit dat machine learning die aanbieder eruitpikken." ("What is now on the table for us to perhaps do something with is machine learning. What is that called, a random forest prediction model on e-invoicing. Because we see that there are providers who consistently send invoices too late, so they did the work a year ago, but they only invoice six months later or a year later. We want to make sure that this happens more quickly so that we don't have to go after certain providers at the end of the year. Come on. We need to be invoiced. And if we then have a picture of which providers are consistently billing too late, we could use such a prediction model or artificial intelligence based on machine learning to pick out those providers."; DAN2) These predictive algorithms provide context about the quickly evolving situation the civil servants that implement data-driven work are in. These data-driven predictive methods such as machine learning and artificial intelligence carry possibilities that have not been dealt with before and thus require the development of new procedures and processes to be implemented. Noteworthy here is the willingness to talk about predictive algorithms, which exemplify the context of civil servants needing and being willing to adapt to the ever-evolving nature of data-driven work.

Data-driven decision-making is another benefit and chance which every respondent discussed. Every civil servant mentioned it gives some insights into its importance. The interview data further shows that the respondents agree that making decisions using data is seen as being preferable to making decisions primarily based on superficial information and experience alone. Further, the insights from monitoring, analytics, and predictive data-driven work were also noted

to help support decision-making. One data manager remarked that, "Laat mij zeggen twijfels kun je wegnemen door het onderbouwen met cijfers en meningen. Emoties kun je wegnemen door daar feiten, cijfers euh door daar gegevens tegenover te zetten. En ik denk dat dat goed is." ("Let me say that doubts can be dispelled by substantiating them with figures and opinions. Emotions can be dispelled by countering them with facts, figures, and data. And I think that is good."; DM5). This statement underlines how decision-making not only helps form new decisions but can also support existing ones. The remark from the civil servant about the perceived benefits of data-driven decision-making countering decisions made based on emotion and experience highlight the importance of data-driven work implementation. In summary, according to civil servants, the perceived benefits and chances are mainly centred around financial incentives, improved analytics, monitoring, and decision-making.

Respondents report multiple data-driven work related risks and challenges including privacy concerns

The privacy-related data-driven work concerns are part of a broader collection of risks and challenges. The questions about the risks and challenges of data-driven work developments yielded a more comprehensive range of answers from respondents than the ones about benefits and chances. These observations regarding perceived data-driven work risks and challenges include data quality (i), competencies and capacity (ii), unconnected data streams (iii), objectivity assumptions (iv), and privacy risks and challenges (v). One of the phenomena identified during the interviews was the concern frequently mentioned by respondents about data quality used for data-driven work within Dutch municipalities and their partners. These partners are further discussed later in this section. The data quality concern phenomenon is about having data sets with erroneous entries, creating a challenge for data-driven work that uses structured or unstructured data and risks for future use. Respondents saw data quality as one of the primary challenges for data-driven work (developments) within the municipality:

DM7: "Data kwaliteit is superbelangrijk voordat er analyses gedaan worden wordt daar ook eerst naar gekeken [...] wat mij betreft de voorwaarden nummer één om überhaupt een stap

verder te kunnen." ("Data quality is extremely important before analyses are carried out, and this is also looked at first [...] as far as I am concerned, the number one condition for being able to take a step forward at all")

DA2: "Dat die data kwaliteit vaak niet op orde is. En als je data niet goed is, kan je niet zo goed onderzoek doen. Dus eigenlijk vind ik dat je zou moeten beginnen met de kwaliteit op orde brengen." ("That data quality is often not up to scratch. And if your data is not good, you can't do very good research. So actually, I think you should start by getting the quality right.")

DAN5: "We zijn al een tijdje bezig met het datagedreven en het belangrijkste is de datakwaliteit natuurlijk. Om datagedreven te kunnen werken, heb je je goede data nodig en op het moment dat je geen goede data hebt dan kun je daar ook geen conclusies aan hangen. Daar kun je ook geen mensen op aanspreken. Of prognoses doen en dat soort dingen." ("We have been working on data-driven for a while now, and the most important thing is of course data quality. To be able to work in a data-driven way, you need good data and if you don't have good data, you can't draw conclusions from it. You can't call people to account for that either. Or make forecasts and things like that.")

These quotes emphasise the importance of data quality which is at the foundation of any data-driven work process. However, minding data quality internally to enable data-driven work within the municipality is not sufficient as one respondent noted, "Ik merk ook collega's van andere gemeenten dat in veel gemeenten de datakwaliteit nog niet hoog genoeg is om daar echt eigenlijk onderzoek op te kunnen doen op academisch niveau." ("I also notice from colleagues from other municipalities that in many municipalities the data quality is not yet high enough to be able to do real research on it at an academic level."; DA2). According to this respondent, a data analyst, as multiple data streams are often required for data-driven work, the necessity for maintaining good data quality also expands itself to the municipality's partners. One of the interviewed data analysts remarked another factor concerning data quality, pointing out erroneous values in a data set may lead to resistance from the organisation's lower layers. The respondent states, "Het is natuurlijk ook wel vervelend als je als er achter komt dat in jouw bestand 500 fouten zitten zeg maar ofzo. En dat zorgt soms wel eens voor wat weerstand [...] als

je de mensen hebt die mutaties invoeren enzo. Die voelen zich dan wel eens snel zelf aangesproken van wat krijg ik nou de schuld dat daar al die fouten in in zitten?" ("Of course, it is also annoying if you find out that your file contains 500 errors or so. And that sometimes causes some resistance [...] if you have the people who enter mutations and so on. They can quickly feel that they are being addressed and say: "What do I get the blame for all these errors?"; DAN1). This perceived reluctance of admitting human error problems with data-driven work highlights the importance of (internal) cooperation. This quote is unexpected as recent literature focuses on highlighting the advantages of internal and external cooperation as stated by multiple sources of knowledge such as by Van der Weerdt & De Vries (2014), Kim et al. (2014), and Malomo & Sena (2017). This reluctance further contrasts with the upper management layer of the municipalities where they want to cooperate and move forward towards higher quality data sets.

Aside from data quality, the second important phenomenon has two closely related aspects: the challenge of having the competencies and capacity within Dutch municipalities to enable data-driven work (implementation). These competencies include the lack of data literacy among internal stakeholders to read, understand, create, and communicate data as information. The lack of capacity is the difficulty of attracting and maintaining enough data analysts, scientists, advisors, and managers (DAN3). When asked about the causes of these capacity and competency challenges, most respondents pointed to the speed at which the digital transformation progresses, making it difficult for some people to keep up with new systems. One respondent saying, "Maar 50 procent van de Nederlanders heeft voldoende digitale vaardigheden om mee te gaan met innovatie en ontwikkeling [...] alleen 50 procent betekent dat maar één op de twee die vaardigheden heeft, en dat merk je ook in je eigen organisaties. [...] Dan heb je altijd mensen tussen zitten die gewoon iets niet snapt of niet goed gebruikt." ("Only 50 percent of the Dutch have sufficient digital skills to keep up with innovation and development [...] only 50 percent means that only one in two have those skills, and you also notice that in your own organisations. [...] There will always be people who just don't get it or don't use it properly."; DA1). A different respondent (DM3) commented on this topic by stating that most municipalities are on a staffing freeze, making it challenging to clear funds to increase capacity by hiring more data

analysts. This challenge creates situations where data-driven projects are done by people who do it as side-projects, during their breaks or by a trainee, which causes the project to be abandoned when their traineeship period is over. Another unexpected observation not stated in the reviewed literature is that the high average age of civil servants at respondents' municipalities was also frequently mentioned in the context of skills and experience necessary for data-driven activities. However, a data manager (DM5) interviewed noted that this phenomenon is not age-related and stated that younger people in the organisation similarly lack the required skills.

Respondents tended to believe the governance of data was also lagging behind the implementation of data-driven work. According to an interviewed data manager (DM4), this is mainly due to the lack of formal agreements about data management. Leading to another recurring point of agreement significant to the research question about data-driven developments within municipalities is the challenge that so-called 'data islands' create. A source of information within the municipality or one of its close data partners that have non-existent or minimal external connections. These data islands form a challenge because the information needed for data-driven work is unavailable or otherwise incompatible. One respondent, a data analyst, stated the following, "Wat er nu niet is, is de feedback terug naar van het ene naar de andere. Van de ene leverancier van de data naar de andere. Dus ik stuur nu data naar een black box in en ik moet maar hopen en vertrouwen dat het goed gaat." ("What is not there now is the feedback from one to the other. From one supplier of data to another. So now I send in data to a black box, and I just have to hope and trust that it goes well."; DAN1). Further, this respondent, who has an advisory role for data-driven work processes at multiple Dutch municipalities, also mentioned that this lack of feedback between the parties using a data set also contributes to data quality issues as the people creating the inputs do not get feedback about the quality of the data. This feedback also includes comparing and exchanging files based on a standard format; respondents stated the creation of common data formats as a challenge and current topic of interest. For example, databases could have a formatting system only focused on mutations, not comparisons.

The cooperation challenges civil servants face are also due to the varying interests of their partners. A pattern of challenges when connecting data streams within a municipality also include the challenge that municipalities implement the national and municipal policy that directly affects residents, ranging from waste management to issuing passports. The interviewed civil servants noted having to connect different data sets per issue as not every case is relevant for their partner. A data manager stated, "Het is steeds weer bepalen aan de hand van een vraagstuk wie meedoen. Bijvoorbeeld als je zegt langer veilig thuis wonen. Dan zegt de scholengemeenschap. Nou super geinig, ik doe even niet mee. [...] Maar op het moment dat het gaat over jongens en meisjes in het onderwijs waar jongens het minder goed doen dan meisjes in het onderwijs. Ja, dan doen in één keer scholen wel mee en de GGD's en de BSO's en het centrum jong en kinderopvang. Maar op het moment dat je zegt brandveiligheid, zeggen zij. Ja, dit is voor ons minder relevant. Dus je bepaalt steeds aan de hand het vraagstuk wie er instapt en uitstapt." ("It is always a question of deciding who will participate. For example, if you say 'live safely at home for longer'. Then the school community says. Well super nice I'm not participating for a while. [...] But when it comes to boys and girls in education, where boys do less well than girls in education. Yes, then all at once schools do join in, and so do the Municipal Health Centres, and the BSOs, and the Young Childcare Centre. But the moment you say fire safety, they say: Yes, this is less relevant for us. So based on the issue who gets in and who gets out."; DM2) This mismatch between data partners' needs also vary between municipalities which have different data needs depending on their location and size. According to another interview, the civil servant remarked this also tends to create friction, "Dat ons tempo en onze behoefte heel anders is dan de kleine gemeentes en daar ontstaat nog wel eens frictie door dus. En daar zijn we nu ook over aan het nadenken." ("That our pace and our needs are very different from those of the small municipalities, and that sometimes creates friction. And we are thinking about that too."; DM3). This quote stands out as it gives insights into the growing complexity for cooperation between Dutch municipalities of various sizes as municipalities may also be at different stages of data-driven work implementation, widening the divide two municipalities have to overcome to work together.

Civil servants strive for implementing data-driven work that gives insights with 100% reliability but is not quite there yet. Respondents noted a difference between the strategy for data-driven work and the reality of it. One respondent stated on this topic, "Het is nog niet 100 procent betrouwbaar. Het is voor 99 procent betrouwbaar en dat geeft inzicht, maar doet dat net niet helemaal." ("It is not yet 100 percent reliable. It is 99 percent reliable and that gives insight, but it just doesn't quite do it."; DM5). In contrast, another respondent (DM4) does call data reliable but points out the lack of nuance. While the strategy assumes data is objective, the data is not, and the civil servants engaging with the data may lack the required competencies to interpret the data correctly. The general challenge identified by respondents as a data-driven development within Dutch municipalities tended to be centred around how data is seen as always objective. At the same time, the inputs for the data are not always reliable, or the outputs can be misinterpreted. These challenges and risks for the future are often mentioned in the context of recent incidents in the Netherlands, such as predictive algorithms that turn out to be unintentionally discriminatory towards minorities (DM4; DM3; DM2; DAN2). Furthermore, situations abroad are also cited as a risk for relying too much on data, such as the mass surveillance projects mentioned as limiting its citizens' privacy (DM1; DA1; DAN4).

In analysing the interview data, privacy also emerged as a perceived challenge for current data-driven work and risk for future data-driven work. The factors leading to some data-driven work-related privacy phenomena, discussed in Ch. 4.4, lead to increased (personal) data usage that drives privacy concerns among civil servants working for Dutch municipalities. For each respondent, privacy was one of the main themes when discussing emergent challenges for the data-driven work they were involved in. Two respondents noted specific examples about how privacy formed a challenge or even the main obstacle to their work. They discussed these changes:

DM3:   "Een ontwikkeling die ik zie is dat wij echt last hebben van privacywetgeving. Dus je wil mensen graag helpen door gegevens uitwisselen? Dus bijvoorbeeld iemand die bij ons wijkteam komt en een probleem heeft en bijvoorbeeld ook nog naar de sociale dienst moeten. En

ook nog naar een psychiater. En die dus elke keer dat hele verhaal weer opnieuw moet vertellen. En allemaal dat soort zaken. En DPA's [Data Processing Agreement]. We hebben om de haverklap discussie met de functionaris gegevensbescherming die voor alles een DPA wil." ("One development I see is that we are really troubled by privacy laws. So you want to help people by exchanging data? So, for example, someone who comes to our neighbourhood team and has a problem and also has to go to social services, for example. And to a psychiatrist as well. And who has to tell the whole story over and over again. And all that sort of thing. And DPAs [Data Processing Agreement]. We are having discussions at every turn with the data protection officer who wants a DPA for everything.")

DAN1: "Dat was mijn grootste hobbel in het hele proces eigenlijk. Om die privacy, om de systemen zo te beveiligen enzo. En zo ook te doen dat het helemaal privacygevoelig is en ook de kans op datalekken nihil is of heel klein zegmaar." ("That was actually my biggest obstacle in the whole process. The privacy, securing the systems and suchlike. And to do it in such a way that it is completely privacy-sensitive and the chance of data leaks is nil or very small, so to speak.")

These comments seem to affirm the reviewed literature (Schoeman, 1984; Gormley, 1993; Solove, 2008) and provide further evidence for the existence of a growing phenomenon about privacy concerns. While the first respondent is concerned about privacy slowing down the internal processing due to the need for data processing agreements, the other expresses concern about data leakage. These quotes highlight how privacy concerns are both a current issue through slowing down current processes and a possible risk for the future where personal data can leak. In a nutshell, data quality, competencies and capacity, cooperation, and objectivity assumptions enable data-driven work, which is the cause of many privacy concerns that are the subject of this thesis.

Cooperation is seen as the key in managing data-driven work and its related privacy developments in a rapidly changing environment

As a part of this research's aim of understanding the broader context of perceived data-driven work related privacy changes, it is important to track organisation-wide changes due to

data-driven work implementation. To highlight these developments, the interview data analysis showed three main phenomena that happened due to data-driven work implementation within the respondents' organisations. These include the following themes, data-driven work implementation progress (i), internal and external cooperation (ii), and cultural changes (iii).

All Dutch municipalities have implemented data-driven work in some form; however, the exact implementation phase varies from municipality to municipality. When asked about the current state of data-driven work implementation in their municipality, respondents either stated that, according to them, implementation is in an early form or that implementation is in a more advanced state but is still an ongoing process. However, remarkably none of the respondents reported that they have not started implementing any data-driven work whereas recent literature primarily focuses on the emergence of data-driven work in the (local) public sector such as by Klievink et al., (2017), and Malomo & Sena (2017). The first group of respondents that is still in an early implementation state, reported the following:

DM7:   "We zitten nog steeds eigenlijk in die pioniersfase. En hoe we uit de pioniersfase komen dat moeten we nog steeds afwachten." ("We are actually still in that pioneering phase. And how we get out of the pioneering phase is something we still have to wait and see.")

DAN3:   "We zijn eigenlijk twee jaar geleden zijn ze echt gestart met de zeg maar een data traject zoals dat zo mooi heet van het sociaal domein. Daar zijn ze gestart om echte data analyse op te bouwen." ("Two years ago, we really started the so-called data trajectory of the social domain. They started to build up real data analysis.")

DM1:   "En sinds half september is er dus wat verbreed naar de afdeling enthousiast krijgen voordat datagedreven werken en daarin te ondersteunen. En sinds januari is er een organisatiebreed programma datagedreven werken gestart." ("And since mid-September, there has been some broadening to getting the department excited about data-driven work and supporting them in it. And since January, an organisation-wide data-driven work programme has started.")

The results for the factors perceived as driving this data-driven work implementation trend are reported in Ch. 4.3. Each of the respondents above stated that they have recently started data-driven work, or it has become a priority. The other part of the respondents are further along with the implementation of data-driven work and often note earlier data-driven activities such as business analytics. For instance, a data manager (DM6) stated that their municipality was already doing business analytics using computer data for 15 to 20 years. Another respondent, a data-driven work advisor, reported that their organisation neared 'maturity' concerning data-driven work implementation, "Iets verder dan de kinderschoenen. Bij de kinderschoenen ga ik ervan uit dat je dat je nog probeert duidelijk te maken dat het nodig is. En we zijn verder dan de kinderschoenen in dit geval bij ons we hebben programma's. We zijn er over aan het nadenken met teams die al dingen actief doen. Privacy teams hebben al een paar jaar bijvoorbeeld, ook landelijk. En we zijn ons ook echt bewust van wat er moet gebeuren en wat ook landelijk gebeurt wat we kunnen gebruiken. Ik denk misschien zelfs al dat we langzamerhand richting onze volwassen schoenen gaan, dat je echt iets hebt staan waar je op kan bouwen." ("A little further than the children's shoes. With the children's shoes, I assume that you are still trying to make it clear that it is necessary. And we are beyond infancy in this case with us we have programmes. We are thinking about it with teams that are already doing things actively. Privacy teams have been around for a few years, for example, nationwide. And we are also really aware of what needs to be done and what is also happening nationally that we can use. I think perhaps we are gradually moving towards adulthood, so that you really have something you can build on."; DA1). The respondent gives insights regarding the implementation progress of data-driven work within that municipality and shows that a privacy team, while just being beyond the infancy stage of the implementation, says something about the perceived importance of this privacy impact even in the relatively early stages of the implementation process. Further implementation steps found in the interview data include storing data in the cloud, implementing machine learning, developing a new dashboard to access data, and using data warehouses to store information. The respondents also mention creating policy documents, documentation, and creating guidelines for data governance as principal future steps.

Another interesting result was that every respondent reported the importance of internal cooperation concerning data-driven work in their municipality. For example, according to one respondent (DAN1), a data analyst, the various departments within their municipality cooperated on data-driven work activities. Another analyst (DAN3) reported the same thing and added that data-driven work exceeded a single domain and concerned every aspect of the organisation. However, departments only often worked together on specific projects, with the scope of the data-driven work having to match the scope of the departments' activities. A municipality's external cooperation consists of different organisations. These organisations include other municipalities, political parties, commercial parties, and the Statistics Netherlands agency (CBS). The CBS is often mentioned as a primary strategic partner. Another common cooperation includes the Association of Netherlands Municipalities (Vereniging van Nederlandse Gemeenten, VNG). Further cooperation focuses on municipalities working on the same project in the same area. These cooperations between municipalities focus on developing shared data-driven work processes but also on identifying and addressing common (privacy) risks and challenges (DAN1; DAN3; DM4). Sometimes this cooperation is set up as a "data alliance" where multiple data-driven work partners exchange information. The degree of external cooperation also depends on whether the municipality has been focusing on centralisation or decentralisation. However, in one interview the civil servant (DM4) reported this fracture in various parties accessing data storage and continued by saying that this leads to the compartmentalisation of information.

The perceived cultural change phenomena are centred around implementation changes and prioritisation differences. One of the key elements of cultural changes important here include creating support for data-driven work by showcasing value creation and hosting programs that highlight the importance of data. Aside from showcasing value creation through data-driven work, another prioritisation is political. Political decisions drive the public sector's policy and the civil servants interviewed have to carry out the plans voted in by elected officials. Thus, prioritisation largely is decided by the national government or municipal council in the case of the Dutch municipalities. According to one respondent, some municipalities use an 'incident

ladder' to decide which part of data-driven work gets implemented first, "Daar (in een andere gemeente) hebben ze een soort incidenten ladder. Hoe zwaarder het incident, hoe meer data adviseurs je erop zet, hoe meer je informatie beveiligingsexperts er dan opzet." ("There (in another municipality), they have a kind of incident ladder. The more serious the incident, the more data security experts you put on it."; DA1). This incident ladder is a system that helps prioritise data protection activities. Further, when asked about the timeline of these data-driven work implementation projects, respondents either stated these were predominantly short-term or that their municipality focused more on the long-term applications of data-driven work. In summary, municipalities are at various stages of the data-driven work implementation process, which heavily rely on both internal and external cooperation to work. This cooperation includes creating cultural changes to help foster acceptance of the new data-driven work methods.

4.2 The importance of privacy — Privacy as a public good

The importance of privacy plays an important role in investigating the data-driven work-related privacy changes in Dutch municipalities. This section aims to find out how civil servants perceive and experience the discussion about privacy issues and concerns within their municipality. Identifying these privacy changes is part of the method which aims to help understand the context of the discourse about privacy importance, which as a public good, is seen as something that should be available to all members of society. With personal data usage inherent to data-driven work-related privacy changes, the most interesting views from respondents about this aspect are included in this section.

The importance of privacy in the context of the increased proliferation of data-driven work in Dutch municipalities

When asked about recent privacy phenomena regarding data-driven work implementation, respondents tended to believe that the changes in data-driven work made privacy more important. The aspects used to discuss privacy included ethics and privacy protection measures. What stands out is that respondents tended to agree that privacy considerations and their

protection are growing increasingly important. For example, while talking about the role of privacy in data-driven work developments, one respondent (a data manager) said the following, "Ja, nee, dit is heel belangrijk. [...] dan zorgen we ervoor dat daar geen persoon herleidbaar en gegevens in staan. En daar wordt heel secuur naar gekeken. En door de vragende partij niet altijd met evenveel begrip op gerekend, want ze konden net niet in detail vraagstukken beantwoorden die ze wilden. Maar ja, als wij niet meer info mogen geven, dan doen we dat niet. Dat is echt een belangrijk onderwerp. Ja, en ook wat je naar buiten brengt en wat niet. En in welke vorm. Daar wordt veel over gepraat." ("Yes, no, this is very important. [...] then we make sure that there is no personally identifiable information in there. And that's looked at very carefully. And the requesting party doesn't always respond with the same level of understanding, because we just couldn't answer the questions they wanted in detail. But then again, if we're not allowed to give more info, we won't. This is a really important issue. Yes, and also what you bring out and what not. And in what form. That is often talked about."; DM7). In this quote, the privacy of citizens is prioritised over getting the best answers from data-driven results. This statement is important as it again shows the difficult position these civil servants are in where the benefits and advantages must be weighed against the right of all citizens to have their privacy protected. In contrast, despite their possible privacy impact, these data-driven work implementations can also help improve the municipalities' services for its citizens. More privacy impact results as reported by the respondents are mentioned in Ch. 4.4. In summary, as data-driven work increases, the related privacy concerns become more critical due to factors inherent to data-driven work, such as increased (personal) data collection.

How civil servants perceive and identify privacy concerns

Two respondents reported the usage of an ethical data assistant to anticipate privacy changes when asked about identifying privacy concerns (DA2; DM6). This data ethics decision aid (DEDA) is a tool developed for data managers, advisors and analysts to recognize these issues. One of these respondents (DA2) tended to believe this could help identify ethical problems and help create a paper trail in case problems arose with personal data usage later on. However, while the other respondent agreed that it was a useful tool, they also added that, "Maar daarmee ben je

nog niet meteen ook honderd procent zeker dat je nog zo ethisch werkt. Als je het allemaal netjes invult en zeg maar daar opvolging aan geeft. Uiteindelijk is mensenwerk." ("But that does not mean that you are one hundred percent sure that you are working ethically. If you fill it all in neatly and follow it up, so to speak. At the end of the day, humans are involved"; DM6). Another factor that helps identify privacy changes is having a privacy officer and a chief information security officer (CISO) working in a municipality. What stands out here is that, in total, 12 out of 15 respondents reported having at least a privacy officer or CISO. In one case, the privacy officer is shared with another municipality (DAN3). These examples emphasise the human element involved in how civil servants perceive data-driven work-related privacy concerns. In short, privacy factors identification happens with a decision aid such as DEDA but the identification of privacy concerns still primarily lies with the civil servants themselves or a privacy or compliance officer.

The usage of personal data in data-driven work is seen as a important factor in data-driven work related privacy changes

A significant factor in privacy concerns is that when data-driven work implementation efforts increase, this often also means that more personal data is collected to support these new processes. While other data streams used by Dutch municipalities include open data or data that can not otherwise be traced back to any specific individual, it is clear from the responses that every municipality collects personal data for their data-driven work processes. When asked about these developments regarding data-driven work, respondents tended to believe that more data is becoming available and more personal information is being used:

DA3: "Ik kan me voorstellen dat van stel dat je heel veel meer persoonlijke data gebruikt. Dan je kunt zeggen van ja, dat is wel nieuw als. En we houden nu veel meer dingen bij dan voorheen." (""I can imagine that if you use a lot more personal data. Then you can say, yes, that is new if. And we keep track of a lot more things now than we did before.")

DM3: "Voor mij zijn er een aantal dingen die relevant zijn. Dus het ene is dat steeds meer data beschikbaar komt en dat data het dus ook makkelijker maakt om te sturen." (""For me, there are a

number of things that are relevant. So one is that more and more data is becoming available and so data also makes it easier to manage.")

This drive to acquire more data is also highlighted in Ch. 4.4 which gives some examples of personal data collection through methods such as Wi-Fi point and garbage container tracking. However, respondents agreed that this information, with few exceptions, should be stored and used in its anonymized, pseudonymized or aggregated form. Some systems designed theoretically to not include personal information according to respondents in practice can still create situations where civil servants can see or have to anonymize or aggregate this data. For example, in one interview the civil servant reported how their garbage collection process collected citizen address information. He further mentioned that, "Er is altijd nog wel eens een medewerker die heeft nog een toegang tot een BP een persoonsregistratie ofzo om nog stiekem even te kunnen kijken." ("There is always an employee who still has access to a BP, a person registration or something like that, to be able to take a look on the sly."; DAN1). However, two respondents noted that all the data acquired by them is strictly anonymized. Of these, one respondent reported the following, "Wij hebben alles geanonimiseerd. Zowel de persoon in die een uitkering heeft, dat soort dingen, het is allemaal geanonimiseerd. Alleen de consulent kan eigenlijk pas zien wie het uiteindelijk is" ("We have anonymized everything. Both the person on benefits, things like that, it's all anonymized. Only the consultant can actually see who it is in the end"; DA2). This quote stands out because while it underlines the increased acquisition of personal information, the respondent notes strict data control measures through the anonymization of personal information. Nonetheless, the consultant involved with the source data can still see all the personal information. Thus, despite the anonymization efforts, there is still a privacy leakage risk if anything goes wrong at the consultant's level or due to the possibility of deductive disclosure through re-identification.

This anonymized data creates situations where the personal data is re-identified; two examples given by a respondent was the abnormally high usage of a garbage collection point or the frequent use of a garbage container at a residential address that formally has no residents. In

these cases, civil servants can change a use case where no personal data is typically used in the data-driven work process to a situation where the information can be traced back to a specific individual. Something similar happens when data is anonymized or aggregated. There are two phenomena where this anonymised data can be traced back to a particular individual, "Op een moment dat jij op het adres niveau op kleiner gebied, in een postcodegebied met met vier cijfers twee letters deze postcode 6 ook wel genaamd. Als je op een adres niveau een kenmerk toevoegt graag en je gaat stoppen, dan heb je persoonsgegevens te pakken. Of op het moment dat jij zegt van in die straat woont iemand met een bijstandsuitkering. Dan loop je naar die straat en dan kun je m aanwijzen dan heb je persoonsgegevens te pakken. Voordat je het weet ben je bij de persoonsgegevens, want je wilt iets over de adressen zeggen, over de gebieden zeggen. Dus dus de 9 van de 10 keer zijn persoonsgegevens." ("At a time when you are at the address level in a smaller area, in a postcode area with four digits two letters, this postcode 6 is also called. If you add an attribute at address level like this and you stop, then you have personal data. Or the moment you say someone with social security benefits lives in that street. Then you walk to that street, and you can point to him, then you have personal data. Before you know it, you have reached the personal data because you want to say something about the addresses, something about the areas. So 9 times out of 10, it is personal data."; DAN4). In this example, the respondent believed that combining or using a small data set could lead to the re-identification of an individual. In this case, the individual was the person receiving benefits in a specific zip code area. This respondent reported another unexpected situation where a small data set and multiple parts of aggregated data led to the re-identification of an individual, "Oud studiegenoot had dat. Die zijn pa die was paardenfokker. Die woont in het buitengebied ergens in Brabant en op CBS buurtniveau dat het kleinste ding dat ze doen. Daar stond voor dat vlak je stond nooit het gemiddelde inkomen. Daar stonden vijf boerderijen. Drie mensen uit het dorp die in buitengebied waar komen wonen en hij, maar daar kun je gewoon weten hoeveel die verdiende." ("An old classmate had that. His dad was a horse breeder. He lives in the countryside somewhere in Brabant, and at CBS neighbourhood level, that is the smallest thing they do. In front of that plane, you never stood the average income. There were five farms there. Three people from the village who lived in the outlying area and he, but you could just tell how much he earned.";

DAN4). In this instance, the individual is identified because he was an outlier in a small data set. The respondent perceived this re-identification using anonymized data as negatively impacting the identified individual's privacy. What is unexpected is that the anonymous data used for data-driven work remains personal and can be traced back to an individual with these criteria. To summarise, the respondents perceive that more personal data is being collected. However, the increased usage of this personal information can impact citizens' privacy because of privacy concerns such as privacy leakage, human error, or anonymization issues leading to the re-identification.

4.3 The myriad of influences — Factors related to data-driven work and privacy changes

In order to understand the context in which these perceived data-driven work and privacy changes are happening, this section focuses on factors that, according to the interviewed civil servants, are influence factors that can serve as drivers for data-driven work implementation in Dutch municipalities and its related privacy factors. These factors include the technological changes, where new technology makes data-driven work possible but also is key to corresponding privacy concerns. Aside from these reported technological influences, internal factors from within the organisation and external factors from the outside are discussed in this section. The respondents further suggest that they play an important role in influencing data-driven work implementation and related privacy changes within municipalities. First, the reported main technological factors related to data-driven work developments are discussed, which drive the data-driven work trends through new or improved access to data, algorithms, and hardware. The second section analyses perceived external changes, including developments within society, other organisations, changes in legislation, or political influences. The third part is about the internal effects that are perceived to drive changes from within the organisation.

Technological developments clear the way for data-driven work implementation

When respondents shared their perspectives about technological influences for data-driven work and related privacy changes, respondents agreed that many new technologies enable data-driven work. These phenomena include brand-new (improved access to) data collection, storage, protection, and processing methods or hardware such as surveillance technology. The first observation about new technology influences is about increased access to data from the interview data. Three respondents reported increased access to a greater volume or variety of data. When asked about data-driven work developments within their municipality, the respondents said the following, "Voor mij zijn er een aantal dingen die relevant zijn (DAN1; DAN3; DM3). Dus het ene is dat er steeds meer data beschikbaar komt en dat data het dus ook makkelijker maakt om te sturen." ("For me, there are a number of things that are relevant. One is that more and more data is becoming available and that data therefore also makes it easier to manage."; DM3). Respondents tend to believe this increase in data is not just by collecting and storing more data at municipalities but also by connecting more existing databases from multiple sources. Further, another key technological phenomenon is the existence of more complex and sophisticated threats that causes a greater need for data protection to prevent data leakage. Despite this development, one of the civil servants (DAN1) interviewed further reported that while these new attacks formed a challenge for the future, most data leaks are caused accidentally and are not due to attacks from outside the organisation.

While access to new methods such as algorithms for data analysis has been around for a long time, improved access to more advanced techniques, including predictive algorithms, open up new possibilities. One respondent, a data manager, reported about how these methods have become more accessible due to faster computer systems, "Vroeger kon je ook allerlei dingen met algoritmes. Alleen dan als je het dan wilde gaan uitrekenen, was de computer drie jaar bezig om die analyse te maken. Dus je ziet gewoon dat de technieken om het analyses te doen veel sneller en krachtiger zijn geworden." ("In the past, you could also do all sorts of things with algorithms. Only then, if you wanted to calculate it, it took the computer three years to make the analysis. So you just see that the techniques for doing analyses have become much faster and more

powerful."; DM2). These responses stand out as they show how these new technological developments cause data-driven work changes within the organisation. Whereas interview data shows that advanced data analytics sees widespread implementation now, respondents report that these methods have already started being used decades ago. Two respondents, data managers, reported the municipalities started this data-oriented approach in the 1970s and over a decade ago, respectively:

DM5:   "Het gaat om werken en dus het is eigenlijk weer een soort nieuwe evolutie, laat ik het maar even zo zeggen eigenlijk doen we altijd al werken en vooral als gemeenten. Maar we zijn ook al heel vroeg begonnen, jaren 70 al met grote dataverwerking toe te passen." ("It's about working, and so it's actually another kind of new evolution, let me put it this way: we've always done work, and especially as municipalities. But we also started very early, in the 1970s, to apply big data processing.").

DM6:   "Data gedreven werken deden we natuurlijk al heel lang. Weet je wat het is soms kan je zeggen een oude wijn in nieuwe zakken zeg maar. Bijvoorbeeld BI is natuurlijk ook een vorm van data gedreven werken, want dat doen we al 15 à 20 jaar dus en dat doen we nog steeds." ("Of course, we have been doing data-driven work for a long time. You know what it is sometimes you can say an old wine in new bags. For example, BI is also a form of data-driven work, because we have been doing that for 15 to 20 years and we are still doing that.").

As these examples show implementation has been underway for many years, more widespread implementation is also happening with machine learning and artificial intelligence technologies, including analytic and predictive algorithms. Further, while these innovations are frequently discussed in the interviews, the respondents state that these methods do not yet see widespread implementation in Dutch municipalities. Nevertheless, the literature shows these technologies are already becoming more readily available, creating a phenomenon where municipalities are (preparing for) early stages of implementation (Manzoor, 2016; Klievink et al., 2017; Malomo & Sena, 2017). When asked about whether their organisation had started with implementing machine learning, one respondent answered the following, "We hebben gegevens en daar gaan

we analyses mee doen. En dan heb je daarbinnen dan nog de twee grote, de een die zegt van er moet machine learning bij. Daar ben ik dan weer niet van omdat ik het gewoon niet of nog niet kan." ("We have data and we are going to analyse it. And then you have the two big ones, the one that says machine learning should be added. I'm not one of them, because I just can't do it, or not yet."; DAN4). This response affirms earlier indications that predictive algorithms are a point of discussion within municipalities while other respondents already report the limited implementation of machine learning, as discussed in Ch. 4.1. These statements are significant as they play a role in the unclear privacy impact changes mentioned in Ch. 4.4 due to the combination of various data sets and related deductive disclosure privacy concerns.

The interview data also shows a change in tracking and surveillance technology that further influence data-driven work developments and related privacy changes through the acquisition of personal data. Three respondents report the usage of surveillance and tracking methods in Dutch municipalities, such as tracking through cameras, tracking chips, and Wi-Fi points. The key point of interest is that these technological developments are only mentioned in interviews when discussing privacy incidents. When asked about whether incidents are drivers for data-driven work-related privacy considerations at their municipality, the respondents answered the following:

DA1: "Dat gebeurt als een incident plaatsvindt. en tot slot hebben we een vraagstuk werd vrij recent is is een wifi tracking. De Autoriteit Persoonsgegevens heeft recentelijk een boete uitgedeeld aan de gemeente Enschede. En dat heeft bij veel gemeenten ervoor gezorgd dat alle slimme camera's en wifi trackers in ons land om maar iets mee te maken had uit werd gezet." ("That happens when an incident occurs. and finally, we have an issue that is quite recent is wifi tracking. The Authority for the Protection of Personal Data recently issued a fine to the municipality of Enschede. And that has caused many municipalities to switch off all smart cameras and wifi trackers in our country.")

DA2: "Ook een ander voorbeeld is bijvoorbeeld de Autoriteit Persoonsgegevens die de gemeente Enschede volgens mij een paar maanden terug op hun donder heeft gegeven voor de

wifi telsysteem." ("Another example is, for example, the Netherlands Authority for the Protection of Personal Data, which, I believe, gave the municipality of Enschede a scolding a few months back for the wifi counting system.")

DM1: "En ik geloof dat in Tilburg de hele binnenstad camera's heeft opgehangen en dus dat zijn wel onderwerpen waar je het even over hebt." ("And I believe that in Tilburg, the entire city centre has hung up cameras, so those are some of the issues you are talking about.")

These respondents tended to believe these new tracking technologies enable data-driven work developments but are also a cause for discussion and privacy changes. The quotes show a debate within the Dutch municipalities about the usage of surveillance systems but, most importantly, the reactive response to incidents happening at other municipalities. For example, the quote of DA1 shows a situation where civil servants are faced with a situation where other municipalities have been reprimanded or fined for the usage of surveillance technologies. This incident led to the reassessment of their internal usage of these technologies. One of the main observations from this section is that this reactive response contrasts the analysed views from civil servants from Ch. 4.1, where internal and external cooperation was viewed as a factor to prevent challenges and risks through which privacy risks are identified beforehand. The results for privacy impact as perceived by the respondents are stated in Ch. 4.4. Aside from the technological influences, it is important to note that there are also external influences reported in the interview data.

External factors that are viewed as driving changes
The first phenomenon concerning the external factors of data-driven work within Dutch municipalities is how data-driven work is seen as a trend or hype. Respondents tended to agree that data-driven work was recently popularised and could be viewed as a trend that impacts the municipality through prompting implementation (plans) or discussion. Two respondents, both data managers, reported on how data-driven work is regarded as a trend or hype and how this impacts an organisation:

DM2: "Ja, er zitten twee dingen in de eerste kant is gewoon dat het een soort hype op dit moment is. Dus kijk op het moment dat artikelen als Harvard Business Review schrijven data is het nieuwe goud. Dat soort teksten, ja, dat doet iets." ("Yes, there are two things in it. The first side is just that it's kind of a hype right now. So look at the moment when articles like Harvard Business Review write data is the new gold. That kind of writing, yes, does something.").

DM7: "Kijk, wat ik merk in gemeente land is dat de term data gedreven werken is. Is ineens heel hot zullen we maar zeggen. We moeten allemaal data gedreven werken en je moet oppassen dat je dat niet gaat doen omdat het een hot item is zeg maar." ("Look, what I notice in the municipality is that the term data driven working is. Suddenly it's very hot, shall we say. We all have to work in a data-driven way, and you have to be careful not to do that just because it's a hot item, so to speak.").

These quotes give an example of civil servants being aware of the trend driving data-driven work popularity. The decision of whether to start with data-driven work is sometimes perceived as "not a choice" (DA3). The interview data showed that implementation must be done to prevent the organisation from falling behind other government organisations. Further, the interview data shows respondents tended to believe the public sector was already lagging behind commercial parties for data-driven work. One respondent (DM3) noted Dutch public organisations, including municipalities, should cooperate on a national or European level and compete with large technology firms.

Another main factor is the political influence through policy and legislation developments. When asked about how data-driven work implementation was prioritised, respondents who mentioned political drivers tended to agree that policymakers needed more data to improve decision-making processes. For example, a data manager (DM5) reported that they noticed politicians have an increasing need to support decisions using data and facts. However, a single respondent provided an alternative view about this information need and instead reported that through political influence, data-driven decisions were not automatically preferred, "Ik denk dat je er rekening mee moet houden dat in gemeenten de politiek echt wel een heel groot aspect is. En soms kan het

er ook gewoon helemaal niet uit om iets goed te onderbouwen voor de politiek. [...] soms laten cijfers zonder verhaal iets zien wat niet gezien wil worden. En dus die onderbouwing is heel belangrijk. En soms is het makkelijker om een besluit te nemen op basis van een goed verhaal dan op data." ("I think you have to take into account that in municipalities politics is really a very big aspect. And sometimes it's just not possible to substantiate something properly for the politicians. [...] Sometimes figures without a story show something that does not want to be seen. And so that substantiation is very important. And sometimes it's easier to make a decision based on a good story than on data."; DM1). In this instance, decisions are based on superficial knowledge and experience instead of data and contrast with the general aim of municipalities to support decisions using data-driven work. In short, this example gives some key insights into perceived internal resistance or reluctance to implement data-driven work because the result of the implementation may not always align with policymakers' views.

The other phenomenon observed from analysing the interviews is the effect of legislative changes on data-driven work. What stands out is that the respondents nearly always mention the GDPR or other privacy legislation when discussing data-driven work related privacy changes. Out of the 15 respondents, 13 mentioned legislation when asked about data-driven work and related privacy change factors. Further, these responses were exclusively found in the context of privacy changes for data-driven work within municipalities. One respondent, a data manager, reported legislation as the most important driver for privacy considerations, "Ten eerste heb je gewoon wetgeving, je hebt te voldoen aan de AVG [GDPR]. En dat vergt al heel wat inspanning hoor. Omdat gewoon goed te doen. Dus daar zouden we nu vooral mee bezig." ("First of all, you have legislation, you have to comply with the AVG. And that requires a lot of effort. Because just do it properly. So that's what we should be doing right now."; DM6). Another respondent, a data analyst, also mentioned the corresponding fines of the GDPR as an important influence, "Het is gewoon wetgeving. Voor de AVG mocht het ook al niet, maar toen hing er geen boete boven hoofd en met de AVG wel" ("It is just legislation. Before the AVG it was also not allowed, but then there was no fine, and with the AVG there is"; DAN4). This fine shows another case of a perceived external influence factor that changes the importance of data-driven work-related

privacy factors, which cause civil servants to be more careful to avoid penalties. In summary, the prevalence of data-driven work implementations due to external influence factors are largely attributed to an element of hype that popularised data-driven work and new privacy legislation such as the GDPR.

Internal factors are also seen as playing a key role in driving data-driven work implementation

While the perceived external influence factors set the narrative for why data-driven work is implemented, the internal factors play a large role in its feasibility. When asked about the reason for starting data-driven within their municipality, respondents reported internal phenomena causing data-driven work and privacy changes. The main observations for influences mentioned include following a strategy (i), getting more information (ii), work optimizations (iii), value creation (iv), creating awareness for data-driven work (v), and working ethically (vi). These six themes are analysed below using quotes from the interviews that underline the role each factor plays in data-driven work implementation.

Aside from policymakers needing information to improve their decision-making process, the interview data shows data managers, analysts, and advisors also have this information need. Respondents tend to agree that data can help optimise processes and help support decisions that are not only based on superficial information and experience. One respondent (DM3) reported the desire to use data-driven work to be more objective, respond to data signals, and work more effectively. Another internal influence reported is that these data-driven work and privacy changes are a part of a data-driven work implementation strategy. When asked about whether these plans were short-term or long-term, respondents gave varying results. One respondent (DM4) reported that their municipality has a long-term master plan to ultimately raise the public space's liveable environment to a higher level. In contrast, another civil servant interviewed experienced that currently, their municipality was focusing on short-term goals. However, they mentioned that they thought the emphasis should instead be on medium-term plans, "Wat je nu vaak ziet is dat de nadruk op korte termijn is op het op het maken van een toepassing. Door een

voorspelmodel of een dashboard. Maar de nadruk zou moeten liggen op de naar de middellange termijn op het op het neerzetten van een werk opzet wat mij betreft. En dat is nog wel echt een stap die wij als gemeente nog kunnen maken" ("What you often see now is that the short-term focus is on making an application. By creating a prediction model or a dashboard. But in my opinion, the emphasis should be on setting up a working structure in the medium term. And that is a step that we as a municipality can still make."; DM6). This quote gives insights into that some municipalities still have a short-term focus for data-driven work implementations. At the same time, there is a desire within the organisation to have a longer-term strategy.

Further, the interview data also shows value creation for both the citizens and the municipal council and creates awareness for data-driven work as the main internal influences. Respondents that mentioned value creation tended to agree that there is a need to quickly show off the value of data-driven work for financial support. For example, one data advisor also reported this need to prioritise showing the value creation to the municipal council, "En ja en vervolgens is natuurlijk wel lastig om bijvoorbeeld een raad, onze gemeente of ook een ook een bestuur om daar heel veel middelen en mensen vrij te maken. Het is wel echt iets wat ze wat. Wat je merkt dat dat dat in ieder geval. Ja, de raad. Die stopt liever hun geld in een in een in een grote complex in de stad dat voor iedereen zichtbaar is en iedereen wat aan heeft dan dat zij vinden dat de administratieve huishouding van de gemeente op orde is." ("And then, of course, it is difficult for a council, our municipality or an executive board, for example, to free up a lot of resources and people for this. It is really something. What you notice is that in any case. Yes, the council. They would rather invest their money in a large complex in the city that is visible to everyone and that is of benefit to everyone than that they think that the administrative management of the municipality is in order."; DA3). In another interview, a data manager (DM5) added that data-driven work "must be worth it" before proceeding with implementation. These perspectives contrast with the interview data results mentioned in 4.3, where respondents reported that data quality issues must be prioritised when implementing data-driven work.

Value creation is also mentioned as a key part of creating awareness about data-driven work. The interview data shows that value creation, awareness creation programs, surveys, and newsletters are used to create awareness for data-driven work. One respondent reported support and awareness for data-driven work is still quite low and must be created to generate support for data-driven projects, "Draagvlak en bewustzijn is vrij laag nog. Het is echt een van de kritieke dingen in data gedreven werk op dit moment nog mensen overtuigen van het nut." ("Support and awareness is still quite low. It really is one of the critical things in data-driven work at the moment - convincing people of the usefulness of data"; DA2). In response to this phenomenon, other respondents report projects to raise awareness about data-driven work are reported to have been created to solve this problem. For example, a civil servant working as a data analyst experienced success from awareness campaigns focused on both data-driven work and privacy. They stated the goal was that this awareness further spreads through the organisation, "Daarbij is het wel mooi als je ze weet te vinden dus eigenlijk is het een soort olievlek dat ze elkaar allemaal besmetten. Zoals ze nu doen dat ze bij ons data bewustwording sessies organiseren. Dus echt wel in gesprek van wat is data? Wat kan je er mee? Waarom heeft het meerwaarde?" ("In addition, it is nice if you know how to find them, so it is actually a kind of oil slick that they all contaminate each other. The way they do it now is to organise data awareness sessions with us. So they really do talk about what is data? What can you do with it? Why does it have added value?"; DAN3). Respondents further report that a privacy officer and CISO also support this awareness creation. This quote emphasises the struggle civil servants working at Dutch municipalities face to show the added value of data-driven work implementations within their organisation. This challenge to add value in the earlier stages of data-driven work implementations is made more apparent if Dutch municipalities lack a long-term strategy for data-driven work implementations as mentioned in Ch. 4.1, during the early stages, civil servants working with data-driven work are still predominantly working on its processes such as data collection and maintaining data quality.

When asked about data-driven work and privacy drivers, respondents tended to mention both ethics and accountability as playing an important role. One respondent even went as far as reporting ethics being the single most important factor, "Ik denk dat het toch wel ethiek het

belangrijkste is, bijna alleen het je moet je dan wel altijd er bewust van zijn. Ik denk dat heel veel mensen ethiek ook onbewust doen. Ik denk dat het helpt als je er ook bewust even over nadenkt" ("I think ethics is the most important thing, almost, but you have to be aware of it all the time. I think a lot of people do ethics unconsciously. I think it helps if you also think about it consciously"; DM1). This quote can be seen as noteworthy as, on the one hand, this shows how important some civil servants consider ethics to be, while on the other hand, this means the 14 other respondents found other aspects to be even more critical. Another respondent (DA3) added to this sentiment by stating that good data handling and accountability are just as important to them as the technical aspects of data-driven work. When discussing accountability, three respondents mentioned recording data-driven work actions to create a paper trail in case something goes wrong. For example, a data advisor reported creating a paper trail to prevent privacy-related conflicts in the future, "Ik doe veel data bewerkingen en ik zie veel data aan. En je moet wel opletten dat je achteraf niets iets op je flikker krijgt. Of het iets is en daarom is bijvoorbeeld die paper trail waar ik het over had is ook gewoon prettig daardoor. Dan voel je je gesteund en weet je dat je dat je daar geen gedoe mee kan krijgen." ("I do a lot of data processing, and I see a lot of data. And you have to be careful that you don't get something on your back afterwards. If it is something, and that's why, for example, that paper trail I mentioned is also nice because of that. Then you feel supported, and you know that you can't get into trouble"; DA2). What is important here is that this practical consideration of a paper trail is exclusively mentioned when discussing privacy in the context of data-driven work drivers. Further, there were also observations from the interview data about the privacy impact caused by the data-driven work developments within Dutch municipalities. In short, the drive to use data to support decisions, optimise work, create value, create awareness for data-driven work, and work ethically are five factors that are perceived as the main internal factors that make data-driven work implementation possible.

4.4 Conflict and progress in a connected world — The privacy impact

As discussed in the earlier sections of the analysis chapter, the privacy issues related to data-driven work implementation plays a prominent role in Dutch municipalities. This section aims to investigate the privacy impact of the data-driven work implementation changes at municipalities as perceived and experienced by the respondents. According to the civil servants, these privacy changes include positive changes, negative changes, and several unclear changes were found. These unclear changes are situations that could change citizens' privacy but are, in contrast to the other observed phenomena, cannot clearly be defined as positive or negative.

The positive data-driven work-related privacy developments differ from recent literature

What is unexpected is that in contrast to recent literature, this content analysis found perceived data-driven work privacy developments from interviews with civil servants at Dutch municipalities that have a positive impact on protecting people's privacy. When asked about the drivers for privacy considerations in the context of data-driven work implementation, respondents tended to agree that the importance of positive intrinsic motivations when working with privacy was the main priority. For example, in one of the interviews, the respondent, a data manager, mentioned the necessity for good ethics when working with privacy, "Ethiek. Ik denk dat in het sociaal domein in ieder geval wel. Ja, ik denk overal wel in de gemeente hoor het wel vaak ethiek is en je gaat niet bij een gemeente werken omdat je iemand iets vervelends aan wil doen. [...] Ik denk dat het toch wel ethiek het belangrijkste is." ("Ethics. I think that in the social domain at least. Yes, I think everywhere in the municipality it is often ethics, and you don't join a municipality because you want to do something bad to someone. [...] I think ethics is the most important thing."; DM1). This quote shows how some civil servants almost take the intrinsic motivations when dealing with personal data for granted and view it as nothing out of the ordinary. Another data manager reported a different intrinsic motivation, integrity, being important when considering the privacy of individuals, "Het is een vorm van integriteit dat je gewoon rekening houdt met de privacy. Privacy van gegevens van mensen die je gebruikt." ("It is

a form of integrity that you just take privacy into account. Privacy of data of people you use.";
DM3). This example makes it apparent that at least some civil servants consider integrity when
dealing with personal data as something that is nothing out of the ordinary. Further, as found in
Ch. 4.2, creating a paper trail is also part of these intrinsic motivations and an organisational
change that has a clear positive impact on individuals' privacy.

These intrinsic motivation phenomena such as ethics, integrity, and accountability, can also be
related to when respondents reported on the prioritisation of data-driven work or privacy, "Ja,
nee, dit is heel belangrijk. Nee, want ik wat ik net al zei, vanuit [...] worden databases gemaakt.
Specifiek voor vragen die er leven en dan zorgen ervoor dat daar geen persoon herleidbaar en
gegevens in staan. En daar wordt heel secuur naar gekeken. En door de vragende partij niet altijd
met evenveel begrip op gerekend, want ze konden net niet in detail vraagstukken beantwoorden
die ze wilden. Maar ja, als wij niet meer info mogen geven, dan doen we dat niet. Dat is echt een
belangrijk onderwerp. Ja, en ook wat je naar buiten brengt en wat niet. En in welke vorm. Daar
wordt veel over gepraat." ("Yes, no, this is very important. No, because as I just said, from [...]
databases are created. Specifically for questions that arise and then make sure that there is no
person traceable and data in there. And that is looked at very carefully. And the requesting party
doesn't always respond with the same level of understanding, because they just couldn't answer
the questions they wanted in detail. But then again, if we're not allowed to give more info, we
won't. This is a really important issue. Yes, and also what you bring out and what not. And in
what form. There's a lot of talk about that.; DM7). This statement stands out because it highlights
a situation where the civil servant is faced with a dilemma of providing optimal services to their
customers through accurate information or looking after their citizens' best interests and
safeguarding their data. Further, the accountability aspect tended to be discussed in the context of
recent legislation such as the GDPR. Seven out of 15 respondents tended to agree that
privacy-oriented legislation had positively impacted individuals' privacy and helped increase
awareness among municipalities (DM1; DM7; DAN3; DAN4; DAN5; DM6; DA3).

According to the interview data, advanced analytics and predictive algorithms are reported to have led to incidents in Dutch municipalities where individuals were discriminated against based on their background or social status. When asked about privacy drivers, respondents tended to agree that these incidents led to more internal discussions about preventing discrimination and protecting the privacy of individuals. One civil servant interviewed reported that privacy protection is also picked up by the municipal council, "Je ziet ook dat de raad door de gemeenteraad er erg op let. Want ik heb allerlei moties van partijen om bijvoorbeeld na te gaan of algoritmes bij ons niet discrimineren." ("You also see that the council pays a lot of attention to it. Because I have all kinds of motions from parties to check, for example, whether algorithms do not discriminate in our area."; DA3). This limit on data-driven work in favour of privacy also includes how there always needs to be a human being involved in automated processes. One respondent stated that in their municipality, there always needed to be a human being involved in any automatic decision-making process (DM3). This measure is also mentioned as a requirement in the GDPR. It helps keep the person maintaining the system and prevent privacy challenges and threats such as the leakage of personal data. To further avoid leakage privacy through personal data leaks, respondents tended to believe data protection measures help protect privacy. One respondent (DM5) reports the usage of an ethical hacker to test their data protection. These ethical hackers simulate real-life situations that civil servants have to deal with. The ethical hacker is just the latest tool in the toolbox created to defend municipalities from external digital attacks. Aside from the phenomena for positive privacy changes perceived by the respondents, negative changes were also identified from the interview data. In summary, 10 out of 15 respondents mentioned intrinsic motivations such as ethics, integrity, and accountability or awareness efforts to counter discrimination and ensure a human element in (semi) automated data-driven decision-making when discussing maintaining people's privacy for data-driven work-related privacy developments and concerns. These perceived internal developments contrast recent literature, focusing on the main data-driven work-related privacy developments on data protection such as cybercrime and privacy concerns such as privacy leakage (Ch. 2.3).

Most civil servants perceive data-driven work-related privacy changes that have a negative impact on people's privacy

Whereas positive changes help protect citizens' privacy, the perceived negative data-driven work-related privacy changes can have far-reaching consequences, leading to situations detrimental to citizens' privacy where their personal information is misused or exposed to the public view. The perceived negative impact changes phenomena includes respondents (DM3; DM7; DAN5; DM6) reporting the degree of control changes for private information when data-driven work in Dutch municipalities. Respondents tended to believe one of the main negative changes comes from the increase in personal data collection. For example, another respondent stated the tracking devices used in garbage collection containers impact individuals' privacy, "Dus ja, dan kom ik wat meer in de privésfeer van de bewoner terecht. Ja, eindelijk denk ik dat het wel impact heeft, denk ik." ("So yes, then I more end up in the private sphere of the resident. Yes, finally I think it does have an impact."; DAN1). They continued by stating that it was important that these tracking devices do not go too far, "En het gevaar is dat we niet moeten doorslaan. Dat we China worden ofzo." ("And the danger is that we should not go overboard. That we become China or something."). In another interview, the civil servant interviewed (DM5) added to this perspective that municipalities should prefer low-tech solutions such as handheld tally counters to high tech solutions that can negatively impact individuals' privacy, such as Wi-Fi tracking points logging personal data. These examples indicate that despite that sometimes far-reaching data-driven work data acquisition methods within Dutch municipalities, municipal civil servants are still cautious when implementing data-driven work and struggling to find the line between when data-driven work is preferred or an unexpected low-tech solution due to the added privacy challenges and risks related to data-driven work which can also be traced back to past data-driven work or privacy-related incidents at municipalities.

Further, opposed to the statements under the positive changes section where respondents viewed legislation as a driver for privacy considerations, privacy regulations are also seen as a barrier and potential risk. One respondent stated they found it challenging to balance privacy legislation and efficient data-driven work in their municipality because careful personal data handling can

slow down or stop data-driven processes. In these cases, they stated that, "Wetgeving kan soms ook gewoon een beetje in de weg zitten." ("Legislation can also sometimes just get in the way."; DM1). These accounts stand out because they further emphasise the dilemma that civil servants face when balancing people's right to privacy and optimally implementing data-driven work in a way that can result in better services to citizens. Another important observation from the interview data is the perceived negative impact of privacy leakage. In this situation, personal data leaks from the municipality's databases. As also stated in the previous section, respondents tended to agree that data protection plays an important role in preventing privacy leakage. However, one respondent reports that data protection is challenging because it becomes increasingly complex, "Vergeet niet trouwens de security kant ook een hele belangrijke ontwikkeling de laatste jaren. Die nemen steeds meer toe en wordt steeds complexer. Dus we zijn altijd in de positie en ervoor moeten zorgen dat we dat wel goed doen." ("Don't forget that the security side is also a very important development in recent years. They are increasing and becoming more and more complex. So we are always in a position and have to make sure that we do it right."; DA3). Further, adding to this complexity is the threat from outside attacks where hackers attempt to steal (personal) data or make it public (DA1). Another phenomenon that can lead to privacy leakage, harming privacy, is the human aspect of data handling. According to one respondent, this challenge and risk is difficult to manage, "Het gebeurt best veel [data lekken] en het is vaak het menselijke aspect wat de boosdoener is. Je kan nog moeilijk echte concrete. Acties op ondernemen, mensen verder proberen te trainen en in te lichten maar dat houdt op een gegeven moment op wat je kan doen." ("It happens quite a lot [data leaks], and it is often the human aspect that is the culprit. It's still difficult to take real concrete action. You can't take any real concrete action, try to train and inform people further, but at a certain point, what you can do stops."; DA1). This quote shows how civil servants perceive privacy leakage; even with the best data protection measures, the human element still plays a key role in maintaining people's privacy. In a nutshell, following recent literature, an increase in personal data usage results in a range of changes indicated by respondents that can negatively impact citizens' privacy through data protection issues such as human error or hacking efforts by external parties causing privacy leakage or because of misuse from civil servants.

The future for data-driven work related privacy factors is often seen as unclear

The unclear privacy impact changes give insights into the perceived (near) future of main data-driven work-related privacy developments. Several perceived changes impacted privacy that were neither clearly positive nor negative. When asked about the drivers for privacy considerations when implementing data-driven work in their municipality, a part of the observations could not be clearly classified as positive or negative. However, one respondent reported that the larger volume of data generates more discussions about privacy considerations within their organisation (DM7). This increase in data-driven work was also reported as resulting in fears for the future privacy impact. For example, a data manager stated that while the intentions for data collection could be presently good, they can be different in the future. They stated, "In Nederland zijn uiteindelijk statistisch gezien de meeste Joden afgevoerd omdat de gemeentelijke administraties gewoon goed op orde waren en daarom netjes overal bijstond of je Joods was of gemengd was en dat vind ik wel grappig. Zit daar eigenlijk het meer ethische aspect? Dat wordt vaak heel praktisch. En er zitten wel echt een paar hele zware dilemma's in." ("In the Netherlands, statistically speaking, the majority of Jews were eventually deported because the municipal administrations were in good order and therefore it was written neatly everywhere whether you were Jewish or mixed, and I find that amusing. Is that where the more ethical aspect lies? That often becomes very practical. And it does contain some very serious dilemmas"; DM3). In this example, the data was collected for a different purpose than the reason for which the (personal) data was later used. The remarkable aspect here is the reference to the Second World War, which saw repurpose data causing some citizens to pay the ultimate price. Another observation from the interview data is the possible re-identification of anonymised data. One respondent reported (DM5) that while anonymised data should negate privacy issues, they stated that combining this anonymised information could lead to the re-identification of an individual. However, the interview data gave no insights into whether this re-identification causes a clear positive or negative privacy impact. In short, several developments have a clear impact on people's privacy. While the privacy impact tentatively seems to be perceived as negative, the situation depends on future circumstances that can not yet be ascertained and thus

can not be described as having a negative or positive privacy impact. While the repurposing and combination of personal data can have negative consequences, the interview data does not indicate that repurposing or combining is done now or planned soon.

4.5 Conclusion

This chapter aimed to provide an overview of the main observations about the respondents' perspectives found in the interview data. What stands out is that in line with the reviewed recent literature, the concept of data-driven work is perceived by civil servants as challenging to define and is being applied in a broad sense of the word and used in various contexts. These contexts observed include both data-driven processes such as analytical purposes and data-driven decision-making such as policy-making based on data instead of experience and superficial information. The main reported data-driven work developments include benefits and chances, risks and challenges, and organisational developments. The benefits and opportunities include financial incentives, better (data-driven) analytics, predictive, and decision-making methods. Further, the main observations from the risk and challenges are data quality issues, (the lack of) competencies and capacity, unconnected data streams, people assuming data is always objective, and privacy concerns. Moreover, the perceived organisational phenomena reported are varying degrees of implementation progress, ranging from just starting to data-driven work being partially implemented and used. Other reported observations include that for data-driven work, municipalities' cooperate internally throughout the various departments and with different (public and private) organisations such as the Association of Dutch Municipalities (VNG) and Statistics Netherlands (CBS). Cultural developments can also be found in these organisational changes, focused on acquiring support and the required competencies for data-driven work implementation.

Another section stated the results for respondents' perspectives regarding the privacy developments at Dutch municipalities. An important reason for the increase of privacy considerations is that municipalities increase data acquisition for data-driven work; this data is

sourced from new data streams (e.g., the garbage containers and Wi-Fi points mentioned in the interviews) and existing databases. From these databases, (personal) information is repurposed for data-driven work (e.g. the municipal addresses and buildings database mentioned in the interviews). With the increase of data-driven work in these organisations, privacy is perceived to become more important and is experienced to be a subject of discussion among civil servants. Respondents report using a data ethics decision aid to investigate the correct method of dealing with their data. Further, every respondent mentions including either a privacy officer or a Chief Information Security Officer (CISO) within their organisation. The data used is often aggregated or (pseudo)anonymised. However, people still use personal data at the municipalities due to a lack of access control or anonymisation. Even after anonymisation, re-identification can happen due to a small data set. For example, irregularities in a data set can sometimes lead to re-identification when some people are clear outliers in the data set. Also important to note is that a part of the acquisition of the data for data-driven work already has happened when data-driven implementation starts off by using existing databases. Thus there are indications that there is a shift in purpose as the data is being used for a different purpose (data-driven methods) than what it was originally collected for.

The aspects that civil servants all mentioned as driving the data-driven work and privacy developments were grouped into three categories, technological, external, and internal factors. The technological factors include access to new methods or improved access to these, more computer processing power and an increase in tracking systems such as cameras and tracking chips in garbage containers. The observed external factors include data-driven work being a trend or hype, increasing its popularity. Further, policy and legislation developments increase the focus on data-driven work. Both due strategies include plans for data-driven work implementation in the municipalities, but also due to new legislation such as the GDPR has created new rules and fines that have to be kept into account. What stands out is that 13 out of 15 respondents mentioned legislation in the context of data-driven work-related privacy developments. Other observations were centred around the internal factors of data-driven work and privacy. Finally, the main aspects perceived by civil servants driving data-driven

work-related privacy factors were following a strategy, getting more information, work optimisations, value creation, creating awareness for data-driven work, and working ethically.

The last section is about the perceived privacy impact related to data-driven work developments. What was unexpected is that respondents reported data-driven work changes that were perceived to have a positive impact. Their intrinsic motivations are observed to play a central role in this positive impact, including the drive for integrity and accountability. Further, awareness programs and incidents at other municipalities help raise awareness for data protection and people's privacy when dealing with personal data. Respondents also stated that these data-driven processes should never be fully automated, and a human should always be involved in the process or the eventual decision-making. The negatively perceived privacy impact observations from the interview data are the increase in the tracking of people, more data protection complexity, hackers and human errors leading to data leakage. Moreover, respondents stated they find it challenging to balance legislation (GDPR) and efficient data-driven work, and the legislation can be seen as a barrier. Finally, there were observed perceived phenomena that are neither clearly positive nor negative. First, the increase in data-driven work leads to more privacy considerations. Another observation includes how changing the purpose of data can lead to privacy issues later. Moreover, even in larger data sets, the computer could combine various inputs of anonymised data to re-identify an individual later as these methods become more readily available and more data streams have been connected.

5. Conclusion and discussion

5.1 Conclusion

The main research question for this study was, "Do data-driven work developments by the Dutch municipalities change the privacy of Dutch citizens as perceived by civil servants?" The current research establishes data-driven work-related privacy phenomena perceived by civil servants working at Dutch municipalities as a double-edged sword: both privacy awareness and privacy concerns increase with data-driven work proliferation. From the perspective of civil servants, the research has shown that data-driven work developments by the Dutch municipalities do change the privacy of their citizens. These privacy changes are seen as being driven by a surge in the popularity of data-driven work implementations within the public sector. The other perceived factors that influence data-driven work implementation are financial incentives, new technologies, and (improved) data-driven approaches such as decision-making, analysis and prediction methods. According to the literature and interview data, the main data-driven work developments see municipalities trying to gain more insights from collected information and making decisions based on data instead of relying on superficial knowledge and experience.

Further, the 'lessons learned' from this research are first: in contrast with literature from recent years, such as from Klievink et al. (2017), and Malomo & Sena (2017), which instead focus on the emerging role of data-driven work in the (local) public sector, data-driven work implementation is now prevalent in all Dutch municipalities involved in this study. However, while data-driven methods such as predictive or analytics algorithms are already used, their implementation is still ongoing. The literature further indicated that data-driven work increases the collection and processing of personal information, which through data loss of control, privacy leakage, misuse, and a shift in intended purpose can have an unclear or negative impact on the privacy of individuals. A key finding from this study is that this research affirms those claims. The driving force, (increased) personal data usage for data-driven work, stems from municipalities having incentives to collect more personal information from a world that already sees an ever-increasing quantity of data and repurposing existing databases for data-driven work.

The content analysis contributes new knowledge in contrast to the literature used for this study; aside from negative privacy impacts, perceived phenomena that clearly positively impact individuals' privacy were also reported. These observations include an increased awareness of privacy concerns through, for example, past incidents (i.e. privacy leakage or the misuse of personal data incidents) at their own or other organisations and programs that focus on raising privacy awareness. Another interesting finding was that both the interview data and literature affirmed that internal and external cooperation on privacy issues supports the early detection and management of data-driven work-related privacy changes, both positive and negative. While respondents report widespread cooperation on all layers of the organisations, they also independently mention developing similar processes from scratch and the desire to increase cooperation with other Dutch municipalities and other stakeholders such as other private and citizens themselves. A final lesson learned from this research regarding positive impacts on privacy is that civil servants often experience having an intrinsic motivation to maintain privacy, which is perceived to be driven by integrity and accountability when dealing with personal information. The leading factors perceived by civil servants to push these privacy considerations mentioned are the earlier incidents and new privacy legislation, such as the GDPR and its corresponding fines.

5.2 Discussion, future research and limitations

This discussion section aims to put the thesis work in a broader context and create a dialogue between the literature and findings. In this thesis, the literature corroborates much of the observations regarding the perceptions and experiences of the interviewed civil servants but also sometimes provides alternating views. The scientific gap is filled by analysing the current main developments regarding the evolving state of citizen's privacy, as also reported in recent literature by Hallinan et al. (2012) and Gruschka et al. (2018), in that citizen privacy does indeed play an increasingly important role in the data-driven work developments within Dutch municipalities and that the impact of these developments has both positive and negative effects. While data-driven work has increased in importance due to new technologies and methods,

making it possible for municipalities to offer more services to their citizens, this development also leads to citizens having increasingly higher expectations. These two-sided findings are in line with the literature, where Boyd & Crawford (2012), Desouza & Jakob (2017), and Van der Weerdt & De Vries (2014) mentioned data-driven work increasing in the public sector despite still being in its infancy—making use of the decrease in the cost of storing and processing large data sets as indicated by Manzoor (2016). This change appears to have been made possible through new technology and methods such as improved processing power, storage and data analytics.

Further, while previous research by Van der Weerdt and de Vries (2014) mentions that data-driven work implementation is often on a small scale (p. 20), this study shows that contemporary data-driven work at Dutch municipalities is considered to be an important part of the organisation, clearly exceeding the size of a "side-project" or "hobby" mentioned in the literature. Further, the interview data shows data-driven work and associated privacy developments are mainly caused by, as mentioned earlier, (improved) access to new technologies and methods. Other development factors include internal factors such as financial incentives, organisational strategy to innovate and the desire to deliver services to citizens through improving processes and decision-making. The main external factors are legislation such as the GDPR and its corresponding fines and data-driven work, having seen a significant increase in popularity. These results further build on existing evidence that the data-driven work-related privacy considerations are also increased as data-driven work implementation increases. However, the increase in personal data usage and its anonymised variants create (re)identification risks and challenges for the individuals who belong to the information. In line with these results, Malomo and Sena (2017) and Gruschka et al. (2018) also reported on this risk of deductive disclosure as the volume of data used by organisations grows.

This interview data also gives a clearer understanding of how civil servants perceive the relationship between data-driven work implementation and its privacy impact in Dutch municipalities. These results show that not all privacy developments in the context of data-driven

work implementation are adverse despite the positive factors found in this study not yet being reported in existing literature used for the theoretical framework. The negative impact phenomena reported in the interview data include the increased amount of personal data collected and processed by Dutch municipalities and the corresponding challenges and risks due to lack of control individuals have over data, possible privacy leakage and misuse. This personal information acquisition for data-driven work does not only happen through the incentivization of new data collection like in the example of Wi-Fi point tracking but also by the repurposing and connecting of existing databases such as the municipal addresses and buildings database with the social services database. These observations from the interview data are in line with mentions in literature by Hildebrandt (2006), Hallinan et al. (2012), and Gruschka et al. (2018), which similarly report privacy concerns due to increased personal data collection and processing or data protection challenges and risks.

Future research can focus on the privacy impact phenomena that had an unclear impact on individuals' privacy. This research can clarify whether these privacy changes are positive or negative and give insights into the probability of re-identification through (combining) anonymised data sets used by Dutch municipalities and their privacy impact. Moreover, this study spent much time defining data-driven work, privacy and its developments. Instead of a generalist approach, a follow-up study could further research how this data-driven work changes impact privacy on an individual level. For example, this can be done by conducting qualitative or quantitative research using the inhabitants of a municipality as a sample instead of its civil servants. Further, future research could also analyse a specific project or case instead of analysing data-driven work within municipalities in general.

This study also includes limitations; there was a limited sample size ($n = 15$) compared to quantitative analysis due to longer, more in-depth interviews. While the sample size limits the generalizability of the results, this approach provides new insights into the privacy phenomena at Dutch municipalities in the context of data-driven work implementations. Further, this study clearly shows the increasing importance and usage of personal information, but it gives no

insights into the differences among the various geographic areas of the Netherlands. To prevent the possible re-identification of the respondents, this geographic and other personal information was not included. Another limitation was that data-driven work and privacy are complex concepts that are difficult to define. Thus, the approach to analyse privacy changes using personal data usage and degree of control changes may result in limited insights into the privacy impact of data-driven work implementations in Dutch municipalities on individuals. Third, while the study gives insights into personal data usage at Dutch municipalities, the degree of control over personal information mentioned by Schoeman (1984), Solovo (2004), and Fuchs (2011) played a smaller factor than the personal data usage part. The degree of control over personal data was only discussed for specific examples of data sets, and the research instead had a broad focus on all information being used at a municipality for data-driven work developments.

5.3 Policy implications for the public sector

Three groups of people must take action from three layers within Dutch municipalities, policymakers, key actors involved in the implementation of data-driven work, and finally, all other civil servants that are not intimately involved in the implementation process but are still intended to in one way or another interact with data-driven processes within the organisation. The key policy implications for the (local) public sector are that data-driven work implementations change privacy factors, such as increased personal data usage. Therefore, municipalities should be aware that this can have an impact on people's privacy. Paradigm shifts such as the transition from a knowledge-based to a data-driven paradigm affect all layers of a Dutch municipality. Thus, this transition is a long-term process. Despite being discussed for decades, long-term (spanning several years or more) orientated data-driven work and privacy-related implementation plans are few and far between, especially within cooperative partnerships between municipalities. Without cooperation, situations occur where two municipalities that are close to each other are both 'reinventing the wheel', so to speak, and working on developing the same data-driven work implementations (from scratch). This cooperation would not only benefit collaboration for data-driven developments but could also

improve the exchange of information regarding privacy factors such as data anonymisation, data protection, and creating privacy awareness programs. According to the content analysis, respondents often perceived the privacy incidents at other organisations and municipalities as an important driver for thinking about and discussing privacy concerns. An increase in cooperation between municipalities could lead to the early identification of privacy risks instead of responding retroactively to privacy incidents such as privacy leakage and data protection issues, thus minimising their privacy impact. Further, future technological developments regarding predictive methods and data combining must also be kept in mind when creating policy regarding data-driven work and privacy in the (local) public sector, as this can have an unplanned unclear privacy impact. For example, deductive disclosure through combining data or repurposing data usage leading to identifying individuals may soon become a topic of importance. This importance can occur because respondents in this study perceived privacy leakage risks, which can seriously affect people ranging from discrimination issues to prosecution. Future policy should also help emphasise positive privacy impacts and further support raising awareness about privacy concerns, promote privacy protection competencies and intrinsic motivations such as integrity and accountability.

At the top layers of the municipality, policymakers can push for more privacy awareness programs and an organisational structure that better represents the balance between the importance and capitalisation of the benefits of data-driven work and the protection against its negative effects like its effect on citizens' privacy. This balancing could be done by increasing the capacity of the privacy protection or data protection officer roles, i.e. by ensuring this is a full-time role instead of a part-time position or a shared role between multiple municipalities. Further, at the middle level, civil servants who frequently engage in data-driven work or its implementation could push for maximising cooperation between the various Dutch municipalities or other public and private organisations and try to prevent needlessly developing data-driven work processes independently or exclusively collaborating depending on the topic. Finally, at all layers of the municipality, any civil servant involved in data-driven work processes should participate in data-driven work and its related privacy aspect awareness programs. These

programs help spread the necessary skills and support that enable successful data-driven work implementation throughout the organisation and improve awareness of privacy concerns. This awareness helps civil servants identify data-driven work-related privacy issues and thus prevent accidental privacy leakage or the misuse of personal data due to human error. In summary, these policy implications for the public sector indicate that municipalities should remain vigilant about the data-driven work-related privacy factors when implementing data-driven work. Moreover, municipalities can take action to minimise factors that cause negative privacy impact (i.e., preventing privacy leakage or misuse) while at the same time creating policy to promote positive privacy changes such as through cooperation or privacy awareness programs.

References

Autoriteit Persoonsgegevens. Cijfers datalekken 2019. (2019, September). https://autoriteitpersoonsgegevens.nl/nl/onderwerpen/beveiliging/meldplicht-d atalekken/overzichten-datalekken/cijfers-datalekken-2019.

Anderson, C. (2008). The end of theory: the data deluge makes the scientific method obsolete. Wired Magazine 23.06. 2008. *URL: https://www. wired. com/2008/06/pb-theory/(09.10. 2017)*.

Beiske, B. (2007). Research methods: Uses and limitations of questionnaires, interviews, and case studies. GRIN verlag.

Bell, E., Bryman, A., & Harley, B. (2018). Business research methods, Oxford university press.

Boete gemeente Enschede om wifitracking. Autoriteit Persoonsgegevens. (2021, April 29). https://autoriteitpersoonsgegevens.nl/nl/nieuws/boete-gemeente-enschede-om wifitracking.

Bopp, C., Harmon, E., & Voida, A. (2017, May). Disempowered by data: Nonprofits, social enterprises, and the consequences of data-driven work. In Proceedings of the 2017 CHI conference on human factors in computing systems (pp. 3608-3619).

Boyd, D., & Crawford, K. (2012). Critical questions for big data: Provocations for a cultural, technological, and scholarly phenomenon. Information, communication & society, 15(5), 662-679.

Brown, B., Chui, M., & Manyika, J. (2011). Are you ready for the era of 'big data'. McKinsey Quarterly, 4(1), 24-35.

Cullen, P., & Mayer-Schönberger, V. (2014). Data protection principles for the 21st century: revising the 1980 OECD guidelines. Redmond, WA: Microsoft Corporation.

Desouza, K. C., & Jacob, B. (2017). Big data in the public sector: Lessons for practitioners and scholars. *Administration & Society*, *49*(7), 1043-1064.

Elo, S., & Kyngäs, H. (2008). The qualitative content analysis process. *Journal of advanced nursing*, *62*(1), 107-115.

Erlingsson, C., & Brysiewicz, P. (2017). A hands-on guide to doing content analysis. African Journal of Emergency Medicine, 7(3), 93-99.

European Commission. (2015). Special Eurobarometer 431: Data protection. Accessed 20 April 2021.

European Commission (2019). Special Eurobarometer 487a (June 2019). http://ec.europa.eu/public_opinion/archives/ebs/ebs_431_sum_en.pdf Accessed 20 April 2021.

Fuchs, C. (2011). Towards an alternative concept of privacy. Journal of Information, Communication and Ethics in Society.

Gormley, K. (1992). One hundred years of privacy. Wis. L. Rev., 1335.

Goudriaan, F. (2017). Op weg naar een optimale informatiepositie. https://hetccv.nl/onderwerpen/informatiepositie-gemeenten/kleine-gemeenten/

Gruschka, N., Mavroeidis, V., Vishi, K., & Jensen, M. (2018, December). Privacy issues and data protection in big data: a case study analysis under GDPR. In 2018 IEEE International Conference on Big Data (Big Data) (pp. 5027-5033). IEEE.

Hallinan, D., Friedewald, M., & McCarthy, P. (2012). Citizens' perceptions of data protection and privacy in Europe. Computer law & security review, 28(3), 263-272.

Hildebrandt M (2006) Privacy and Identity. In: Claes E, Duff A and Gutwirth S (eds) Privacy and the Criminal Law. Oxford, UK: Hart, pp. 43–60.

Kim, G. H., 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., & de Bruijn, H. (2017). Big data in the public sector: Uncertainties and readiness. Information systems frontiers, 19(2), 267-283.

Krippendorff, K. (2018). *Content analysis: An introduction to its methodology*. Sage publications.

Malomo, F., & Sena, V. (2017). Data intelligence for local government? Assessing the benefits and barriers to use of big data in the public sector. Policy & Internet, 9(1), 7-27.

Mandinach, E. B. (2012). A perfect time for data use: Using data-driven decision making to inform practice. Educational Psychologist, 47(2), 71-85.

Manzoor, A. (2016). Emerging role of big data in the public sector. In Big Data: Concepts, Methodologies, Tools, and Applications (pp. 2458-2478). IGI Global.

McAfee, A., Brynjolfsson, E., Davenport, T. H., Patil, D. J., & Barton, D. (2012). Big data: the management revolution. Harvard business review, 90(10), 60-68.

Post, R. C. (2000). Three concepts of privacy. Geo. LJ, 89, 2087.

Provost, F., & Fawcett, T. (2013). Data science and its relationship to big data and data-driven decision making. Big data, 1(1), 51-59.

Saunders, M., Lewis, P., & Thornhill, A. (2007). Research methods. Business Students 4th edition Pearson Education Limited, England.

Savin-Baden, M., & Major, C. H. (Eds.). (2010). New approaches to qualitative research: Wisdom and uncertainty. Routledge.

Schildkamp, K., Lai, M. K., & Earl, L. (Eds.). (2013). Data-based decision making in education: Challenges and opportunities (Vol. 17). Springer Science & Business Media.

Schoeman, F. D. (Ed.). (1984). Philosophical dimensions of privacy: An anthology. Cambridge University Press.

Solove, D. J. (2008). Understanding privacy.

Spillane, J. P., & Diamond, J. B. (Eds.). (2007). Distributed leadership in practice. New York, NY: Teachers College, Columbia University.

Van der Weerdt, C. A., & Vries, A. D. (2014). Dienstverlening verbeteren met Big Data: Een verkenning voor gemeenten.

Voigt, P., & Von dem Bussche, A. (2017). The eu general data protection regulation (gdpr). A Practical Guide, 1st Ed., Cham: Springer International Publishing.

Warren, S. D., & Brandeis, L. D. (1890). The right to privacy. *Harvard law review*, 193-220.

Zhang, D. (2018, October). Big data security and privacy protection. In *8th International Conference on Management and Computer Science (ICMCS 2018)*. Atlantis Press.

Appendix I - Interview invitation

Dear Sir/Madam,

On LinkedIn I have noticed that you work at the municipality of <name> as <position>. Would you perhaps be interested in helping me by giving an interview for my Master's thesis research? My name is Matthijs, and I am a Master's student in Public Administration at the University of Twente. The reason why I am contacting you is that I am currently working on the data collection for my thesis research on data-driven work and privacy developments within municipalities in the Netherlands.

In my case, this data collection is a part of qualitative research in the form of interviews. For these interviews, participants are needed who are (now or in the future) in contact with data-driven work within a (Dutch) municipality. Are you or perhaps someone else within your organisation in a position to be interviewed by me for my research (approximately 30-45 minutes)?

An interview can occur physically, via a communication platform (Microsoft Teams, Skype or Google Hangouts), or via a telephone call. With your permission, the interview will be recorded in order to make a transcript later. The research will not contain any participants' personal information (name, address, etc.). If there are any questions, I will be happy to answer them. I am looking forward to hearing from you.

Kind regards,

Matthijs von Piekartz

Master student Public Administration, University of Twente

Appendix II - Interview protocol

As part of my master's thesis research at the University of Twente, I am investigating how data-driven work developments within Dutch municipalities may have changed the privacy of individuals. This interview is part of content analysis and aims to gather information about this from the perspective of municipalities. For this purpose, respondents were sought who within Dutch municipalities who are now (or may in the future) involved with data-driven work

With your permission, I would like to record this interview for transcription. All information you give me will be treated confidentially, and no personal information will appear in the report (e.g. name, address, name of organisation). This interview will consist of semi-structured open-ended questions. If you have no more questions, I would like to begin.

Part I: Data-driven work developments interview questions (Question area of interest followed by the question asked during the interview):

Number	Category	Questions
Q1.	Recent developments	What is your view on recent data-driven developments in the municipality?
Q2.	Activities types	Can you tell me something about the various data-driven activities that are taking place or will take place?
Q3.	Reason to start	What is your vision on why the municipality is starting with data-driven work?
Q4.	Reflection	What is your view of the prioritisation of different data-driven activities?
Q5.	Cooperation	Can you tell me something about how there is cooperation with other parties for data-driven work?
Q6.	Types of data	Can you tell me something about the types of information used for data-driven work?

Table 3. Interview questions about data-driven work

Part II: Privacy developments interview questions (Question area of interest follow by the question asked during the interview):

Number	Category	Questions
Q7.	Privacy developments	What is your view on recent privacy developments at the municipality?
Q8.	Privacy in the data-driven process	Can you tell us something about the role of privacy within data-driven work at the municipality?
Q9.	Drivers	Can you tell me something about the role of privacy in data-driven work at the municipality?
Q10	Reflection	What is your opinion on the current strategies, tools, interventions that may be used to ensure citizens' privacy?
Q11	Recommendations	What are your recommendations for preserving the privacy of citizens now and in the future?

Table 4. Interview questions about privacy

Part III: Extra questions to ask for more details (Question area of interest follow by the question asked during the interview):

Number	Category	Condition	Questions
Q3A.	Long -term / short term	This question was asked if respondents mention any future intentions for data-driven work	Do these activities tend to be long-term or short-term focused?

Q9A.	Incident drivers	This question was asked if the GDPR or recent incidents were mentioned	Do drivers such as privacy incidents in other municipalities or new legislation such as the GDPR also play a role in this?
Q10A.	Balance between privacy and data-driven work	If balance or struggles with choosing between privacy and data-driven work are asked the following probe question was asked	How do you balance optimal data-driven work with privacy?

Table 5. Extra conditional interview questions for more details

Part I: Data-driven work developments interview questions (Question area of interest follow by the question asked during the interview):

Q1. Recent developments (What is your view on recent data-driven developments in the municipality?)

Q2. Activities types (Can you tell me something about the various data-driven activities that are taking place or will take place?)

Q3. Reason to start data-driven work implementations (What is your vision on why the municipality is starting with data-driven work?)

Q3A. This question was asked if respondents mention any future intentions for data-driven work, Long -term / short term (Do these activities tend to be long-term or short-term focused?)

Q4. Reflection activity importance (What is your view of the prioritisation of different data-driven activities?)

Q5. Cooperation with civilians, departments, external organisations (Kunt u iets vertellen over hoe er wordt samengewerkt met andere partijen voor data-gedreven werk?)

Q6. Types of data (Can you tell me something about the types of information used for data-driven work?)

Part II: Privacy developments interview questions (Question area of interest follow by the question asked during the interview):

Q7. Privacy developments (What is your view on recent privacy developments at the municipality?)

Q8. Privacy in the data-driven process (Kunt u iets vertellen over de rol van privacy binnen data-gedreven werk bij de gemeente?)

Q9. Drivers to maintain privacy (Can you tell me something about the role of privacy in data-driven work at the municipality?)

Q9A. This question was asked if the GDPR or recent incidents were mentioned. Incident drivers / legislation (Do drivers such as privacy incidents in other municipalities or new legislation such as the GDPR also play a role in this?)

Q10. Reflection (What is your opinion on the current strategies, tools, interventions that may be used to enforce citizens' privacy?)

Q10A. If balance or struggles with choosing between privacy and data-driven work are asked the following probe question was asked: Balance between privacy and data-driven work (How do you balance optimal data-driven work with privacy?)

Q11. Recommendations (What are your recommendations for preserving the privacy of citizens now and in the future?)

These were all the questions I wanted to ask you, thank you very much for your time. Do you have any additions of your own that might be useful for the research?

Appendix III - Interview transcripts

Due to possible re-identification, anonymised interview transcripts are not included in this file and can be requested from the following email address: m.h.g.vonpiekartz@student.utwente.nl.