Big data usage of Dutch municipalities: A presentation of an overall data-readiness framework to denote the differences intra municipalities

Mulder, Marlijn

Big data usage of Dutch municipalities

A presentation of an overall data-readiness framework to denote the differences intra municipalities

M.E. Mulder

11 Jun. 21

Supervisor: Dr. Valérie Pattyn
Second reader: Dr. Alex Ingrams

Abstract

The increased importance of data and its consequences have reached the social debate about the data-governance of the public sector. With recent data scandals in the Netherlands in municipalities, the research sets out to discover why and how differences in data handling of municipalities are to be declared. This research focusses specifically on one government level in the Netherlands, its municipalities. An overall hierarchical data-readiness framework is made based on available data-readiness frameworks consisting of four characteristics; organisational alignment, organisational maturity, organisational capabilities and ethical considerations. The framework is applied by means of a survey and send to different Dutch municipalities. We find that the frameworks hierarchy works and that with the framework the differences between the municipalities become more apparent.

Preface

The research you have before you, is the last thing left to finish my Master's Degree in Public Administration. Coming from a different background, I started this degree with a different mode of thinking. I needed time to adjust to this new way of thinking and that didn't go well from the beginning. Reading the literature for subject was easy, but understanding what to do with them to make a good social research paper was something I had to learn. This Master's thesis subscribes that. It is not perfect, but it is a test that was worth taking, because I learned a lot. Not only about social science research, or my chosen topic, data-readiness in the public sector, but also about myself. The biggest lesson I learned is to trust myself, that I am capable of doing something that I've never done before if I am willing to work hard for it. It is a lesson I will take with me, whatever happens.

I combined this research with the wonderful opportunity of an internship at the Digital Governance Department of the Ministry of Internal Affairs. In the team of the information society I could combine all my interests, societal ethical questions, the impact of technology of the functioning of public organisations and even a bit of International governance. This also gave me the opportunity to receive information, to talk with different people and to join seminars on different topics. With this I gained a broader perspective of the influence digital age on public organisations, the civil servants and the political debate. I am very thankful I got this opportunity. Thank you Mildo and Noor, and my other colleagues for the time, information and talks I got.

My special thanks go to Dr. Valérie Pattyn, you kept me on track when it was needed and pushed me to make the most out of the time I had left. Handing this in, I can't change it anymore, so here we go!

M.E. Mulder

Delft, 11th of June 2021

Table of Contents

1. Introduction

Since in 2015 the Dutch government decided to decentralise its government system, more responsibility of citizens has been put on the shoulders of municipalities. One of the consequences of this decentralisation is that the handling of citizen data is now more a task of municipalities than a task of the central government. In recent years, several municipalities have come in disrepute because of their data practices, breaching data and privacy rules (NOS, 29-04-2020; NOS, 14-05-2021; NOS, 18-05-2021; Bürmann & De Zoeten, 2021; Piekartz, 2021). The question now arises: How can it be that some municipalities are handle data in a wrongfully and others handling data rightfully? What explains the differences between the Dutch municipalities in their data handling? This research sets out to provide a first exploration of an answer to this question.

The assessment of data usage starts with a distinction between 'normal' data and big data (Mayer-Schönberger & Cukier, 2014). Big data requires by its size and collection a different process than normal data (Mayer-Schönberger & Cukier, 2014; Olszak & Mach-Krol, 2018). With its size, the complexity of the data increases, requiring more expertise on data handling (Mayer-Schönberger & Cukier, 2014). The main feature of big data is its ability to work with unstructured databases (Sona, 2014) and therefore being able to find patterns where there were no patterns before. Big data is not about the *why* but about the *what* (Mayer-Schönberger & Cukier, 2014). This quest for correlation instead of causation influences the data process and its implications (Mayer-Schönberger & Cukier, 2014; Giest, 2017). Therefore data-readiness frameworks have been developed to check if an organisation is ready for the change that comes with big data implementation. These frameworks identify three main characteristics of data-ready organisations: Alignment of the organisation with big data goals (Burmeister et al., 2018; Demchenko et al., 2014; Orenga-Roglá & Chalmeta, 2019; Sun & Heller, 2012), the e-maturity (Dwivedi et al., 2012; Gartner, 2017; Klievink & Janssen, 2009; Layne & Lee, 2001; Valdés et al. 2011) of the organisation and the capabilities (Gèzcy, 2015; Tekiner & Keane, 2013) that an organisation has. Most frameworks and models focus on one of the characteristics needed for successful data implementation (Braham, 2019). There are separate frameworks for the ethical questions of big data (Davis & Patterson, 2012; Etlinger & Groopman, 2015). These frameworks discuss the consideration between risks and benefits. The current situation in the Netherlands shows that ethical questions are more relevant than ever when working with big data, especially in the public sector (NLDigitaal 2020; NLDigiBeter). The aim of this research

is to combine the different frameworks into one framework, with special attention to the ethical considerations of organisations in the public sector.

This research also sets out to fully understand one government level: Dutch municipalities. Other data-readiness studies have focussed on central government (Olszak & Mach-Król, 2018), the situation of a certain country (Braham, 2019) or certain sectors, such as the executive sector (Klievink et al., 2017). These frameworks give a general overview, but don't denote *why* a certain overall level is found or why one country is doing better than the other. With the focus on one government level, the differences within the same organisational structure can be identified. The knowledge of the differences could lead to insight in *why* the differences are found. If the 'why' is known, it becomes easier to help municipalities to reach the level of data-readiness desired for their data practices. This will give a new depth to the research on data-readiness levels. This research uses the framework by Klievink et al. (2017) as a starting point, as this framework is focussed on the public sector and is specifically made for the Netherlands. It is therefore applicable to Dutch municipalities without possible issues arising when using a framework made for a different federal system.

In addition to the lack of knowledge on the differences within one government level, there hasn't been direct research done on the possible influencing factors to the level of data-readiness within the government level of municipalities. This research sets out to find possible relevant factors in other research fields and from other sources. These sources are provided by the Ministry of Internal Affairs and Kingdom Relations (In short, Ministry of BZK).[1] The Ministry holds a department specifically focussed on the digital government and hold the Dutch digital data strategy, NLDigitaal (2020) and NLDigiBeter (2020). Information will be gathered from seminars, talks with employees and research of documents.

The research is thus two-sided. It starts with an investigation on the available knowledge about data-readiness to construct an overall data-readiness framework with an ethical element. Then by means of a questionnaire the differences between municipalities could be found. The main research question the becomes: *'How can the differences intra Dutch municipalities be explained by a data-readiness framework?'*

The answer to this question is a framework that provides insight in why differences occur, what the differences are and where they are coming from.

[1] This research was combined with an internship at the Ministry of BZK.

2. Theory

Big data

Big data has been around since the start of the internet (Mayer-Schonberger & Cukier, 2014; Abu-Salih et al., 2021). Big data itself is called the 'fourth industrial revolution' as data becomes more and more valuable for companies, organisations and governments (Mayer-Schönberger & Cukier, 2014; Brown et al., 2011; Abu-Salih et al., 2021; Kaisler et al., 2014). Big data is often conceptualised by the 3V's, Volume, Variety and Velocity (Arockia Panimalar et al., 2017; Giest, 2017; Orengo-Roglá & Chalmeta, 2019). But, also other V's have been added to describe what big data requires (Arockia Panimalar et al., 2017; Orengo-Roglá & Chalmeta, 2019). As these other V's are more about the required systems of big data processing (Arockia Panimalar et al., 2017), the definition used in this paper consists of the 3V's: Volume, Variety and Velocity. These characteristics of big data contain the following:

The *volume* is all about the concept of *n=all* (Mayer-Schönberger & Cukier, 2014). Instead of only focussing on a certain data set, for big data all data is relevant, hence *n=all.* The result is the focus of big data on correlation instead of causation. The *why* is translated into a *what* (Mayer-Schönberger & Cukier, 2014).

The *variety* of data is about the messiness of the data. With the enlarged volume of data comes a need for accepting all sort data entries (Mayer-Schönberger & Cukier, 2014). The data becomes messier. Databases are expected to deal with this messiness, resulting in unstructured databases, such as noSQL (Mayer-Schönberger & Cukier, 2014; Nawin Sona, 2016).

The *velocity* of the data is in what time span the data is collected. For big data, this is mostly real time or near real time data collection (Klievink et al., 2017). Alogrithms processing the data should be able to work with this velocity (Mayer-Schönberger & Cukier, 2014).

These characteristics of data is what makes them big data. 'Normal' data is an information unit in a certain format. For example, if a form asks you to insert your birthday and it only accepts a certain format, month-day-year (Mayer-Schönberger & Cukier, 2014). As the difference between big data and 'normal' data is hard to make in practice (Nawin Sona, 2016) the characteristic: 'If it exceeds an excel sheet' is also added.

Big data promises to find new correlation where there were no found at first, to help understand our own behaviour better and to help organisations on the improvement of their services (Mayer-Schönberger & Cukier, 2014).

Besides positive new findings with the use of big data, big data also asks for a different kind of control. When we gave our data consensually, we knew when and for what our data was used (Mayer-Schönberger & Cukier, 2014). With the increased use of big data, we don't know what our search engine inputs are used for, who they are sold to or what they say about us as a person. Big data thus needs a different approach then 'normal' data when talked about ethical considerations, privacy issues or accountability. According to Mayer-Schönberger & Cukier (2014) the accountability is moved from the individual to the company or organisation using our data. But, we know now, seven years, later that this remains an issue today. Several big data companies, such as Facebook, and Google, have faced lawsuits against their data collection and policies (Criddle, 2020). The main complaint is that their influence is so big that they need to take responsibility as customers are not able anymore to shift to another company.

Big data in the public sector

The question of accountability and responsibility is also apparent for big data usage in the public sector (McNeely & Hahm, 2014). The public sector differs from the private sector that customers have become citizens. The goal of a private organisation is to use big data to its advantage and to keep customers (Mayer-Schönberger & Cukier, 2014). The public sector carries another responsibility as its customers have become citizens: Therefore, they not only need service, but they also need protection. The application of big data by the government raises questions about what conclusions are made with the collected data, e.g. can you shorten someone's subvention on a suspicion of fraud or is that morally wrong?

McNeely & Hahm (2014) describe this as the 'benefits versus risks' consideration that is inherent to big data implementation. Hence, the public sector needs to pay extra attention to the moral questions of big data implementation as the responsibility and accountability of the proper use of personal data is in the hands of the government itself (Mayer-Schönberger & Cukier, 2014). The organisation needs to fulfil certain conditions to be able to handle data in an ethical way, the data-readiness.

Data-readiness

Data-readiness describes an organisations' readiness for big data, specifically aimed at how big data is *used* (Klievink et al., 2017). The level of data-readiness of an organisation depends on its ability of big data implementation, specifically on the four stages of data processing: collection, combination, analysis and use (Klievink et al., 2017) (see figure 1). A higher level of data-readiness is therefore only better if the aim of the organisation is to use big data on a higher level. The aim is to have all four stages of the data process on the same level. To assess the level of data-readiness on the four stages of processing big data, data-readiness frameworks are presented in the literature.

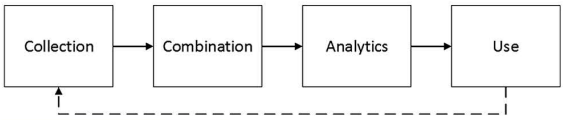

Figure 1. *Four stages of data processing.*

Data-readiness frameworks

For each of the stages of data processing, different frameworks and models exist. There are frameworks aimed at businesses (Burmeister et al., 2018; Ferguson, 2014; Tekiner & Keane, 2013; Gèzcy, 2015), frameworks for analytics of big data (Das & Kumar, 2013; Sun and Heller, 2012), for specific situation or case studie (Barham, 2019) and for relational big data usage as ecosystems (Miller & Mork, 2013; Orenga-Roglá & Chalmeta, 2019; Demchenko et al., 2014). These data-readiness frameworks are applicable to businesses as they assume that data is part of the strategy and aim of the organisation. For the public sector the uncertainty of big data implementation is related to their ability to adjust their organisational aim with big data aims (Klievink et al., 2017). Therefore, a framework aimed at the public sector was deemed to be suitable for the research on municipalities.

The framework by Klievink et al. (2017) provides us with such a framework and was specifically made for the Netherlands, thereby bypassing the possible issues with different federal systems. Therefore, the framework is discussed in more detail.

The framework by Klievink et al. (2017)

The data-readiness framework by Klievink et al. (2017) considers the three characteristics of data-readiness. These were found by literature research. The characteristics are:

Organisational alignment

Organisational alignment is the feature that discusses if an organisation is aligned with the needs for implementing Big Data (Klievink et al., 2017). It assesses if an organisation can implement Big Data without the need for rigorous reforms. For this they use the possible main statutory tasks of a public organisation:

1. Coordination/ project based
2. Research /evaluation
3. Registration /documentation
4. Administration /Management

Organisational maturity

Organisational maturity is defined by Klievink et al. (2017) as 'the maturity of e-governance initiatives within the organisation'. The main component is about the capability of working together on data projects. As this feature follows the e-government growth stages, the stages are hierarchically ordered. These stages are quite common in the literature on e-governance and are therefore almost directly copied from the first e-government growth model by Layne & Lee, (2001).

Organisational capabilities

The last feature of the framework are the organisational capabilities. This feature assesses the capacities an organisation needs to be able to implement Big Data.

The mean of the score of each of the three features determine the overall level of data-readiness of an organisation. Table 1 shows the different capabilities identified and their explanation.

Included in the capabilities is internal and external communication.

Table 1. *Capabilities needed for big data.*

Capability	Explanation
IT governance	Capability to design and develop IT strategy, decision-making and responsibility structures, supporting the organization, including integration of new IT systems
IT resources	Capability to design, develop and maintain suitable IT infrastructure and expertise to facilitate current and new IT systems
Internal attitude	Capability to develop internal commitment and vision for new processes and systems, especially openness towards data-driven decision-making
External attitude	Capability to develop external commitment and support for new processes and systems with important stakeholders
Legal compliance	Capability to design and develop a compliance strategy including process design, monitoring and redesign of processes, especially regarding privacy protection, security and data ownership regulations
Data governance	Capability to design and develop a data strategy including collection, acquisition, quality control and data partnerships
Data science expertise	Capability to bundle/acquire, develop and retain data science knowledge in the organization, especially bundling knowledge on IT, business, statistics and mathematics

Considerations on the framework

The framework by Klievink et al. (2017) is *techno-optimistic* (Vydra & Klievink, 2019). Meaning that it

holds a positive position towards big data usage in the public sector. This is also one of the assumptions made (Klievink et al., 2017).

Following up on this positive view the framework doesn't consider the ethical questions that can be raised about big data in the public sector, as was also discussed earlier. The current social debate within the Netherlands requires a response to the questions of accountability and responsibility (NOS, 29-04-2020; NOS, 14-05-2021; NOS, 18-05-2021; Bürmann & De Zoeten, 2021; Piekartz, 2021). The ethical questions are therefore added into the proposed framework. The characteristic 'ethical considerations' will be made with the available knowledge on big data ethics.

Lastly, the framework by Klievink et al. (2017) presents an overall data-readiness by calculating the mean of the different characteristics. The research was conducted with different executive organisations. The result is an overall idea on the level of data-readiness, of the Dutch executive public sector, but doesn't indicate where and why certain differences between the responses occur. Instead of following this calculation of the level of data-readiness, the proposal is to change the framework to a hierarchical framework as can also be found in e-government and data maturity models (Dwivedi et al., 2012; Gartner, 2017; Klievink & Janssen, 2009; Layne & Lee, 2001; Valdés et al. 2011). First the ethical considerations are discussed and then the complete framework is presented.

Ethical considerations

Specialised ethical big data frameworks (Davis & Patterson, 2012; Etlinger & Groopman, 2015) focus on how to implement big data practices in a responsible way. They describe the relationship between consumer and company based on trust. Wrong data usage can damage the trust and cause the customer to switch to another company (Davis & Patterson, 2012; Etlinger & Groopman, 2015). To contain the trust, ethical questions about data usage should be part of the brand strategy (Davis & Patterson, 2012; Etlinger & Groopman, 2015). But, this doesn't solve the issue for public organisations and especially municipalities. As discussed before, the relationship between the government and citizen is different from the company to the customer. Different that the government has the obligation to protect its citizens. Trust in government is not only necessary for reputation but also for democratic accountability. The ministry of BZK has set out to research what values should be ensured when working with data and AI applications within a public organisational context. Therefore, they are also applicable to

municipalities. The research on public values was conducted in collaboration with the Utrecht University and resulted in 30 relevant values on 6 different aspects (see figure 2).

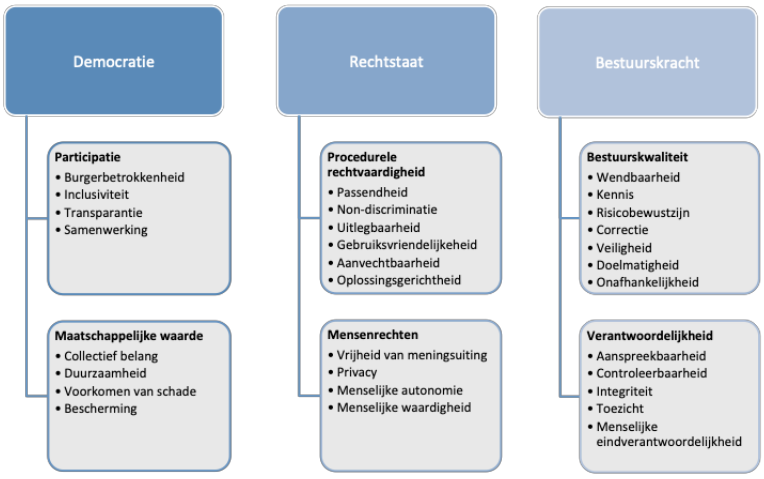

Figure 2. *Code for Good Digital Governance (CODIO).*

Another tool for ethical big data usage is the Ethical Data Assistant. This tool was developed to help civil servants asking moral questions when they are doing a big data project (DEDA, 2019). The code for good digital governance was used to find the relevant values on when something is called 'good governance'. From this research, the important question found is: *Is the data we are collecting really necessary for the goal of this project?* This is therefore the highest levels as it requires municipalities to ask met-questions about their own policies. The lowest level of the ethical considerations characteristic is defined by the legal obligations of all organisations to follow the GDPR. Level 2 consists of what is legally mandatory in the Netherlands and then the voluntary tools are presented. The reasoning of these levels can be found in appendix A.

Proposed framework

The proposed framework is hierarchical, as this will help to denote where the differences between the municipalities can be found. It also means that only if all conditions of all characteristics are fulfilled, then the municipality reaches the next level. This also means that if a municipality has reached a certain level, all lower levels are automatically fulfilled; they are sufficient for the conditions of the lower levels. The complete framework is presented in table 2. The framework of Klievink et

Table 2. *Hierarchy model of the proposed data-readiness framework*

	Organisational alignment	Organisational maturity	Organisational capabilities	Ethical considerations
Level 5	Using data for policy making	Demand-driven, joint up government	External support	Do we need to use this data at all?
Level 4	Using data for analysis of situations	Inter-organisational integration	Internal support	Conscious of public values and their implications
Level 3	Using data for non–operational goals	Nationwide portal	Data expertise	Ethical data tools for projects
Level 2	Improving data processes	Integrated organisations/teams	Data governance	Privacy tools used
Level 1	Collection of data	Stove-pipe organisation	IT governance	Confirm GDPR

al. (2017) used five levels from very low to very high, for consistency reasons were these five levels changed to level 1 to level 5.

Determinants of data-readiness

With the new proposed framework, the differences between the municipalities can be made visible. As there is no research done focussed on one government level, the theoretical background for the possible determinants comes from different fields and sources. The possible determinants on the reached level are discussed.

1. Municipality size

In data science, the general idea about complexity is that if complexity increases, more data structure is needed (Demchenko et al., 2014). The complexity from a municipality is mainly caused by every citizen, as this citizen requests services from the municipality. One of the main possible influencing variables is therefore the *municipality size.*

2. Government funding

The other possible influencing factor is based on resources gathered at the Ministry of BZK. Information gathered from seminars, talks with employees and research in documents led to the idea that government funding for big data projects could be of influence on the level of data-readiness. When a big data project is funded, the suspected outcome is that this project will have a spill-over effect to other big data projects within the same municipality. The result could be a higher level of data-readiness. This is therefore the other possible influencing variable.

3. Other possible determinants

Other possible variables related to *municipality size* and *government funding* could be the amount of people working on big data projects (*team size*). If a municipality can constitute a larger team, civil servants have the possibility to become specialised, to make a career within the department or to work on their own interest (Martocchio, 2006) possibly resulting in a higher level of data-readiness. The other variable is based on the technology improvement idea of (Martocchio, 2006) the longer one works (*years working with big data*) with a certain technology, the more experienced one gets and therefore simply knows more. This could also result in a higher level of data-readiness. These variables are suspected to be correlating with level of data-readiness, but to also correlate with the other variables. If a municipality is larger, it has a need for more civil servants, possibly resulting in a larger data team. If a municipality

received funding for a big data project a few years ago, the spill-over effect, if any, should be visible.

These variables are thus expected to influence the level of data-readiness and therefore to denote the differences within the municipalities.

Hypotheses

To answer the main research question: 'How can the differences in the level of data-readiness intra Dutch municipalities be explained with a data-readiness framework?' three hypotheses are presented.

Null-hypothesis

As no research is done on the hierarchy of the proposed framework and there is a lack of knowledge on hierarchical models for the public sector specifically, the null-hypothesis suspects that the hierarchy of the framework doesn't work. Meaning that it is possible to fulfil all conditions on a certain level, without having to fulfil the conditions of the precedent levels.

The null-hypothesis (h0) is: *The proposed hierarchy of the data-readiness framework is not hierarchical in reality.*

Hypothesis 1 (h1) follows from this: *The proposed hierarchy of the data-readiness framework is hierarchical in reality.*

The following hypotheses are focussed on how the differences between the municipalities can be explained.

Hypothesis 2 (h2): *If the municipality has a larger number of citizens, they will achieve a higher level of data-readiness.*

Hypothesis 3 (h3): *If the municipality received government funding for one of its big data projects, they will achieve a higher level of data-readiness.*

The other variables, *team size* and *years working with big data,* are controlled for in the survey but don't have a separate hypothesis as they are expected to be correlating with the main variables.

3. Research Design

Survey

The design used in this research is a survey design. This design was chosen as it gives the researcher the possibility of a controlled environment without the need for a control group (Toshkov, 2012). The survey is send to different municipalities, the responses are collected and analysed. This closely resembles the research design of Klievink et al. (2017). Their research consisted of 41 interview questions. The questions also had subsections so it took an hour to complete the interview. For practical reasons as the current Covid-19 situation and the shorter time span of the research, the questions were transformed into a questionnaire that respondents could independently answer on their own speed (Bryman, 2016).

Survey Design

The survey consisted of 25 questions, five for each characteristic. Each question corresponds with a certain readiness level of a characteristic. Therefore, the questions are also asked in hierarchical order. A combination of closed, multiple-choice, multiple-box and open questions was chosen to prevent morally desirable answers. The combination of open and closed questions helps to understand attitudes of respondents better (Bryman, 2016). The concept of big data is introduced in the beginning as in figure 3. Before question 16, about the availability of data knowledge, the concept of 'data knowledge' is explained.

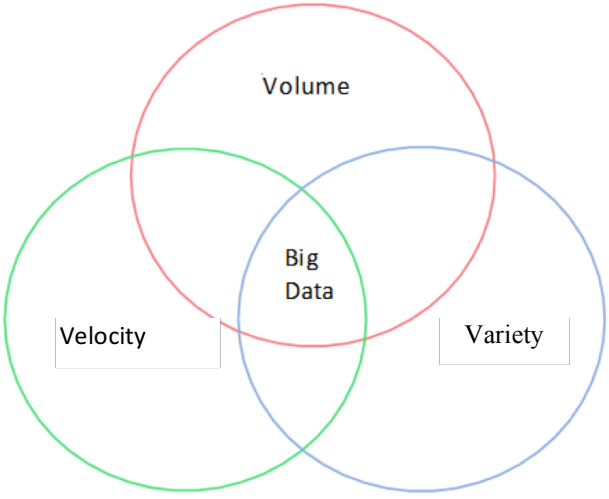

Figure 3. Representation of the 3V's: Volume, Velocity and Variety

The survey is aimed at big data practices. But, the lower levels of the data-readiness hierarchy are legal obligations, therefore the concept of big data is not mentioned in these questions (see Appendix E for the complete questionnaire). Question 16 asks about the relevant knowledge about the data usage process, this question is also used to check the fulfilment of the condition for level 1 of the characteristic *organisational alignment*, shown in table 2, pg. 12.

The questions about the characteristic 'ethical considerations' were constructed based on knowledge on the measuring methods of public value (Kelly et al., 2002) and human rights

(Landman, 2004). These insights were combined with the presented values of the Code for good digital governance (Meijer & Ruijer, 2021). Instead of asking open questions, we decided upon multiple-box questions. This was done to provide respondents the possibility to answer what they thought was right. As moral questions have no direct yes or no answer, this provided the right in between of meeting the respondent without asking for morally desirable answers.

Due to the considerations of different questions and possible issues with the measuring of public values and human rights, a pre-test was done. Two civil servants of one municipality and one civil servant of the Ministry of BZK agreed to do a pre-test of the survey. This pre-test was done to make sure that the questions were understandable, the questions followed logically on each other and that the amount was doable within 10 minutes. The feedback received lead to the conclusion that the questions were understandable, but that the measuring of public values and human rights remains a topic of discussion. The feedback given was implemented before the survey was published.

Operationalisation

This section elaborates on the operationalisation of the variables used in this research. All variables are operationalised by the presence or absence of the requested concepts (Toshkov, 2012). We start with the four characteristics of the proposed framework. Then the municipality size and government funding are discussed. Lastly, we operationalise the other possible variables; team size and years working with big data.

Framework operationalisation

Organisational alignment

Klievink et al. (2017) operationalise organisational alignment by combining the statutory tasks with IT alignment (Klievink et al. 2017). This result in four statutory tasks with big data

Table 2. *Hierarchy model of the proposed data-readiness framework*

	Organisational alignment	Organisational maturity	Organisational capabilities	Ethical considerations
Level 5	Using data for policy making	Demand-driven, joint up government	External support	Do we need to use this data at all?
Level 4	Using data for analysis of situations	Inter-organisational integration	Internal support	Conscious of public values and their implications
Level 3	Using data for non–operational goals	Nationwide portal	Data expertise	Ethical data tools for projects
Level 2	Improving data processes	Integrated organisations/teams	Data governance	Privacy tools used
Level 1	Collection of data	Stove-pipe organisation	IT governance	Confirm GDPR

alignment. We combine this with the big data use process: collection-combination-analysis-use (Klievink et al., 2017). This is operationalised by asking about statutory tasks and combining it with questions on what data is used for. This give the opportunity to find a possible alignment with big data implementation, but also shows if our assumption that municipalities work with data is confirmed.

Organisational maturity

Klievink et al. (2017) follow the e-government maturity models when operationalising the organisational maturity. As this characteristic is already hierarchical, the survey follows the same steps, from stove-pipe organisation to demand-driven joint up government, presented in figure 4. The e-government maturity model is based on the ability of horizontal integration (Layne & Lee, 2001). As horizontal integration is related to communication between organisations and departments, the maturity model was transformed into questions about communication of big data projects. Communication issues with transparency are still identified as the biggest challenge of big data implementation for governments (Matheus et al., 2020).

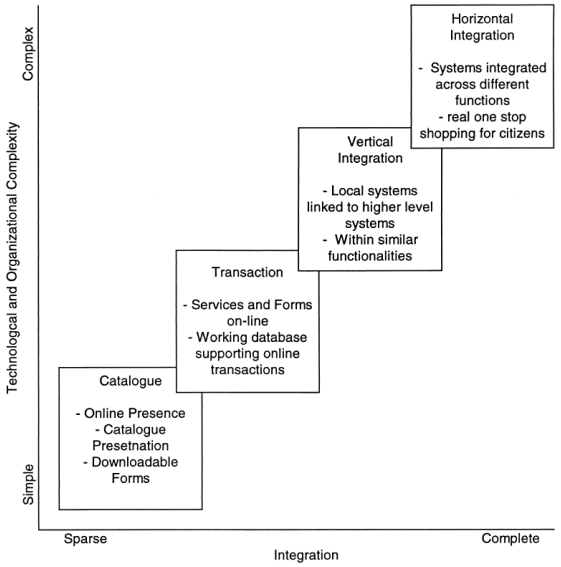

Figure 4. *Dimension and stages of e-government development (Layne & Lee, 2001).*

Organisational capabilities

Based on the literature research done by Klievink et al. (2017) and the extended thesis report of Romijn (2014), the organisational capabilities identified by Klievink et al. (2017) were transformed in a hierarchical order. Starting with IT governance as a decent IT infrastructure is the first condition for the ability to implement data practices (Mayer-Schönberger & Cukier, 2014). Other frameworks also press the importance of an IT strategy (Klievink et al., 2017;

Taylor, 2021). The second capability needed for successful big data implementation is a strategy to deal with the data (Taylor, 2021). The strategy is operationalised by asking about the different parts of the data process (collection-combination-analysis-use). If the process of data is supported throughout the whole organisation then implementation is more likely to work out (Taylor 2021). The next step in the capabilities is the expertise on big data. If expertise is available on at least all four stages of data processing, that is enough for the reaching of the next level. Other expertise can also be relevant, such as juridical (Taylor, 2021; Davis & Patterson, 2012).

Ethical considerations

The ethics of big data for the public sector are different from the private sector as the government has responsibility and accountability when implementing big data practices. Related to responsibility is the question of human rights, what one can and cannot do towards another human being (Amnesty, n.d.; United Nation, 2021). The question of accountability is related to the question of 'generating public value' and of ensuring public values (Kelly & et al. 2002; ROB, 2021). The operationalisation, and thereby the measuring of human rights and public values are always under dispute. But, some have tried to come up with measurements of human rights (Landman, 2004) and public values (Kelly et al., 2002). For human rights this start with the legal obligations, such as privacy. This has been installed recently with the implementation of the GDPR in the EU (EU, Guide, 2020). The operationalisation thus starts with the legal aspect, conform GDPR. One of the requirements is to have a privacy policy to be accountable for the handling of personal data (EU, 2020). To ensure personal privacy, the government of the Netherlands has developed self-assessment tools to help government organisation to ensure this right (AP, 2021). Next, tools are developed to help ensuring values within data projects. The DEDA (Ethical Data Assistant) is the most common example. If the DEDA or other ethical frameworks are used on project basis then the municipality reaches the next level. Other options are the 'Code for good Digital Governance' (Meijer & Ruijer, 2021), International UN and EU guidelines, such as the implications of the GDPR. The fourth level is determined by the awareness of public values. Kelly et al. (2002), describe that awareness of public values is essential for generating public values as an organisation. The values chosen for awareness are based on knowledge provided by the Ministry of BZK on what values are from their experience more frequently ensured (inclusivity, transparency, safety, and protection) and ones that are less apparent in data projects (Collaboration, Freedom of speech,

end responsibility, integrity, non-discrimination and disputability). At least one from all six features of the CODIO is presented in the survey. A minimum of four values is considered 'aware of public values'. The operationalisation of the last level is based on the question: is the data necessary? For this we have looked at data maturity models (Gartner, 2017) and current initiatives known by the Ministry of BZK. The model of Gartner prescribes to reach the highest level (6), the 'Chief Data Officer (CDO) sits on board'. In that way, the data practices are ensured when the policy making happens. The initiative the Ministry found was similar. The Dutch association of municipalities is trying to implement an ethical commission for data projects. Therefore, ethical initiatives on the highest level are the operationalisation that ensure the 'is this data needed?' question is asked.

Variables

Municipality size

The municipality size is determined by splitting the Dutch municipalities in five respective groups (Bantema et al. 2021), presented in table 3.

Table 3. *Calculation table of municipality distribution in the Netherlands*

Municipality groups	**# citizens**	**Real count**	**Real percentage**	**Response count**	**Response percentage**
Group A	<15.000	43	12%	-	-
Group B	15.000-30.000	128	36%	-	-
Group C	30.000-60.000	116	33%	-	-
Group D	60.000-100.000	33	9%	-	-
Group E	>100.000	32	9%	-	-
		352	100%	-	100%

All municipalities were sorted in the five groups (CBS, 01-01-2021) and received a percentage of the total municipalities (table 3). The municipality is then translated to a size group using the CBS data (CBS, 01-01-2021). The percentages of the distribution of the responses is then compared to the one in reality. In this way, the sample size of the municipalities, the one that responded are checked on their representability for municipalities in the Netherlands.

Government funding

Government funding is when a municipality received funding for a big data project. Big data as defined in the theory; by its volume, variety and velocity. The application for the funding should at least use big data in its goals, this to prevent possible conflict with advanced data-analytics.

Years working with big data

The respondent is asked if they know when they started using big data within their municipality. The current year minus the year answered is the years working with big data. If no answer is given, the variable cannot be determined for the municipality.

Team size

Team size is operationalised by using an ordinal measuring method. Team sizes range from less than 10 to more than 50 total working on big data, this is based on knowledge provided by the Ministry of BZK.

Method of analysis

The analysis of the required responses consists of different parts. First the representation of the responses is checked by filling in table 2. Secondly the hierarchy models are filled using a score card (see appendix B). This provides an answer to the absence or presence of a certain condition. With this knowledge, the hierarchy models are summed and then tested using percentages for conditions fulfilled, conditions unfulfilled and conditions answered with 'not applicable' or 'I don't know'. If the percentages follow the hierarchy, the hierarchy can be confirmed. The overall data-readiness level is than determined by the condition: What is the highest level where *all*

Table 2. *Hierarchy model of the proposed data-readiness framework*

	Organisational alignment	Organisational maturity	Organisational capabilities	Ethical considerations
Level 5	Using data for policy making	Demand-driven, joint up government	External support	Do we need to use this data at all?
Level 4	Using data for analysis of situations	Inter-organisational integration	Internal support	Conscious of public values and their implications
Level 3	Using data for non–operational goals	Nationwide portal	Data expertise	Ethical data tools for projects
Level 2	Improving data processes	Integrated organisations/teams	Data governance	Privacy tools used
Level 1	Collection of data	Stove-pipe organisation	IT governance	Confirm GDPR

conditions are fulfilled? A table of the level of data-readiness and the different municipalities is then made. The variables are analysed by descriptive statistics, such as crosstabs and by calculating the mean.

Execution of the research

The survey was send by email to contacts known by the Ministry of BZK. These contacts were civil servants known to be working with data and possibly big data practices. The email explained the aim of the research and requested to send the survey to civil servants from other municipalities. In this the responses could be increased.

The survey and the research were also promoted during several seminars about data practices of municipalities. Lastly, the survey was also promoted with a linked-in post. The survey was open for two weeks before the analysis began.

4. Findings

Hypothesis 1: Does the hierarchy of the proposed framework work?

For every response, a hierarchy model was filled (see table 2). There were two municipalities that were two times represented (see appendix D). One of these was the test municipality (Appendix D, municipality 1). As those respondents filled in a different questionnaire, they are not considered when determining the hierarchy of the framework. There are thus 13 responses considered for the answering of the first hypothesis. The other municipality (Appendix D, municipality 2) with two responses is considered when determining the hierarchy as these responses also count for the validity of the framework. Table 4 (pg. 24) presents the fulfilment percentages of the conditions. This shows that none of the respondents fulfilled the condition of level 5 on the characteristics of organisational maturity and organisational capabilities. But it also shows that 92% of the respondents fulfilled the legal required condition of confirming with the GDPR (Ethical considerations, level 1). The hierarchy of the framework expects that the higher the level, the lower the percentages in comparison to the earlier levels. So, for organisational alignment this is the case as the condition for level 5 is only reached by 31% of the respondents and the condition for level 1 is reached by 77% of the respondents. This is also the case for the other characteristics.

Table 2. *Hierarchy model of the overall data-readiness framework*

	Organisational alignment	Organisational maturity	Organisational capabilities	Ethical considerations
Level 5	Using data for policy making	Demand-driven, joint up government	External support	Do we need to use this data at all?
Level 4	Using data for analysis of situations	Inter-organisational integration	Internal support	Conscious of public values and their implications
Level 3	Using data for non–operational goals	Nationwide portal	Data expertise	Ethical data tools for projects
Level 2	Improving data processes	Integrated organisations/teams	Data governance	Privacy tools used
Level 1	Collection of data	Stove-pipe organisation	IT governance	Confirm GDPR

When we consider the percentages of the conditions that are unfulfilled, we expect the opposite: Larger percentages of respondents on the conditions for the higher levels, and lower percentages for the conditions on the lower levels. This is the case for the organisational alignment characteristic, on the highest level 62% indicates that the condition for level 5 is unfulfilled and only 15% for the condition on level 1. Interestingly, on the other characteristic of organisational maturity, organisational capabilities and ethical considerations we see that this is not the case. There is for example a decline between level 3 and level 4 on the characteristic of ethical considerations. Just as there is a decline between level 2 and level 3 on the organisational maturity and between level 4 and 5 on the organisational capabilities. But,

we have seen on the fulfilment of the conditions for each level that there is a hierarchy, so what explains the unfulfilment of the levels?

Table 4. *Overall fulfilment percentages of conditions by respondents*

Characteristic/ Level	**Organisational alignment**	**Organisational maturity**	**Organisational capabilities**	**Ethical considerations**
level 5	31%	0%	0%	8%
level 4	54%	31%	8%	15%
level 3	54%	38%	31%	15%
level 2	69%	38%	31%	77%
level 1	77%	38%	31%	92%

Percentages of unfulfilled conditions

Characteristic/ level	**Organisational alignment**	**Organisational maturity**	**Organisational capabilities**	**Ethical considerations**
level 5	62%	38%	46%	38%
level 4	31%	15%	62%	15%
level 3	23%	8%	8%	31%
level 2	23%	15%	8%	0%
level 1	15%	8%	15%	0%

Percentages of condition 'not applicable' "I don't know'

Characteristic/ Level	**Organisational alignment**	**Organisational maturity**	**Organisational capabilities**	**Ethical considerations**
level 5	8%	62%	54%	54%
level 4	15%	54%	31%	69%
level 3	23%	54%	62%	54%
level 2	8%	46%	62%	23%
level 1	8%	54%	54%	8%

Total check

Characteristic/ Level	**Organisational alignment**	**Organisational maturity**	**Organisational capabilities**	**Ethical considerations**
level 5	100%	100%	100%	100%
level 4	100%	100%	100%	100%
level 3	100%	100%	100%	100%
level 2	100%	100%	100%	100%
level 1	100%	100%	100%	100%

The sudden declines of the percentages between certain levels can be explained by the percentages of respondents indicating that they thought the question was 'not applicable' to their situation, they didn't know what to answer or they didn't answer the question. If sudden declines in the unfulfilled conditions correspond with inclines in the not applicable and unknown condition answers. For example, on the characteristic of ethical considerations, the decline between level 3 and level 4 is met with an increase in the conditions not applicable or not known. Notable in the percentages of the conditions answered to be not applicable or not known is that more than 50% of our respondents indicate that they don't know or don't find the question applicable to their situation. As the questionnaire stated the definition of big data and then asked if the respondent was known with big data projects, and if they then indicated that they weren't, they got a message that even if big data was not applicable it was still useful to answer: 'not applicable' in the questions were specifically was asked about big data. 8 Out of 13 respondents (62%) indicated that they were not familiar with big data projects within their municipality, or that they didn't use big data in their municipality. This high percentage can explain why the percentages of 'not applicable' or 'I don't know' are that high.

Also notable is that the organisational capabilities characteristics, scores the lowest of all characteristics. The literature puts the emphasize of data-readiness on the capabilities (Gèzcy, 2015; Tekiner & Keane, 2013). The lack of capabilities in the respondents, even on level 1

(what is needed for data in general), is concerning as the organisational alignment (what the collected data is used for) scores higher. Municipality 2 is the most apparent example (table 5).

Table 5. *Filled hierarchy model of municipality 2*

Characteristic /level	**Organisational alignment**	**Organisational maturity**	**Organisational capabilities**	**Ethical considerations**
Level 5	Using data for policy making	Demand-driven, joint up government	External support	Do we need to use this data at all?
Level 4	Using data for analysis of situations	Inter-organisational integration	Internal support	Conscious of public values and their implications
Level 3	Using data for non–operational goals	Nationwide portal	Data expertise	Ethical data tools for projects
Level 2	Improving data processes	Integrated organisations/ teams	Data governance	Privacy tools used
Level 1	Collection of data	Stove-pipe organisation	IT governance	Confirm GDPR

But, the percentages of the conditions that are fulfilled, still indicate that even though respondents were not familiar with big data projects, their municipality followed the hierarchy. Therefore, the hierarchy of the framework is confirmed.

Hypothesis 2: Does the size of the municipality influence the level of data-readiness?

To answer the hypothesis the Dutch municipalities were divided in five groups. Table 6 shows the municipalities size group and their representation in the survey responses[2]. No collected responses were in group A and only one response was a municipality of group B. Group C is well represented in our responses and group D and E are overrepresented in the collected responses. As the question is if size influences the level of data-readiness, the

[2] Note: The test-municipality (1) is now part of the responses. As two responses for the test municipality were collected they are merged into one level of data-readiness. The municipality that was two times represented is now one response, represented in one level of data-readiness. Therefore we speak of 13 municipalities in the responses.

overrepresentation of larger municipalities can give us qualitative insight in the relationship between the level of data-readiness and the size of the municipality.[3]

Table 6. *Calculation of representation distribution of collected responses*

Municipality groups	# of citizens	Count in reality	Percentage in reality	Count	Percentage questionnaire
Group A	<15.000	43	12%	0	0%
Group B	>15.000 - 30.000	128	36%	1	8%
Group C	>30.000-60.000	116	33%	4	31%
Group D	>60.000-100.000	33	9%	2	15%
Group E	>100.000	32	9%	6	46%
Total		352	100%	13	100%

Table 7 shows the municipality set out to the municipality size group. They are met with the level of data-readiness that was found in the questionnaire. Notable is that level 1 is the most represented (light blue) and that no level 3 nor level 5 was found. Table 7 also shows that in group C no higher level than 2 is found. The only municipality size group where a higher level is found, is in group D. As the largest municipalities (group D and E) don't reach a higher level per se, it seems this hypothesis needs to be rejected. But, the level of data-readiness was calculated by the lowest complete fulfilled level.

[3] Two of the responses collected represented collaborations of smaller municipalities. These municipalities work together to make it easier to help each other with resources. The sum of these smaller individual municipalities exceeds the 100.000 citizens. As they answered the questionnaire on behalf of the collaboration, they are considered as one group. We do this instead of dividing the smaller municipalities and counting them in the smaller group where their individual municipality should be.

Table 7. *Cross tab of municipality size group with municipality.*

Nested is the level of data-readiness.

Municipality	Data-readiness level	Municipality size group A	B	C	D	E
	1		1			
	2					2
	3			2		
	4				4	
	5					1
	6			1		
	7			1		
	8					1
	9					2
	10					1
	11					1
	12				1	
	13			1		

When table 5 is compared to table 7, the following is notable. Municipality 2 is in the largest municipality size group, E and scores level 2. Table 5 shows *why* this large municipality scores level 2. Their response indicated the lack of ethical tools when implementing big data projects. On the other characteristics, the municipality scores higher. Municipality 2 is compared with municipality 3, group C with level 2 (presented in table 8). Municipality 3 also indicated that ethical tools were not used in big data projects. The municipality only differs on one aspect with municipality 2, namely the inter-organisational integration. The lack of difference indicates that municipality size doesn't influence the level of data-readiness. But, the results of the hierarchy show that the larger municipalities are working with big data, while the collaborative municipalities and others that indicated their lack of knowledge about big data practices in their municipality only work with 'regular' data, shown by the reaching of level 1 (see for an overview appendix D).

Table 5. *Filled hierarchy model of municipality 2*

Characteristic /level	**Organisational alignment**	**Organisational maturity**	**Organisational capabilities**	**Ethical considerations**
Level 5	Using data for policy making	Demand-driven, joint up government	External support	Do we need to use this data at all?
Level 4	Using data for analysis of situations	Inter-organisational integration	Internal support	Conscious of public values and their implications
Level 3	Using data for non–operational goals	Nationwide portal	Data expertise	Ethical data tools for projects
Level 2	Improving data processes	Integrated organisations/ teams	Data governance	Privacy tools used
Level 1	Collection of data	Stove-pipe organisation	IT governance	Confirm GDPR

Table 8. *Filled hierarchy model of municipality 3.*

Characteristic /level	**Organisational alignment**	**Organisational maturity**	**Organisational capabilities**	**Ethical considerations**
Level 5	Using data for policy making	Demand-driven, joint up government	External support	Do we need to use this data at all?
Level 4	Using data for analysis of situations	Inter-organisational integration	Internal support	Conscious of public values and their implications
Level 3	Using data for non–operational goals	Nationwide portal	Data expertise	Ethical data tools for projects
Level 2	Improving data processes	Integrated organisations/ teams	Data governance	Privacy tools used
Level 1	Collection of data	Stove-pipe organisation	IT governance	Confirm GDPR

Hypothesis 3: Does receiving government funding for a big data project influence the level of data-readiness?

To answer this hypothesis, the records of the Ministry of BZK were checked to find municipalities that received government funding.[4] There are no funds specifically aimed at municipalities,[5] but there are innovation funds where municipalities can apply for. From these records only three municipalities that answered the questionnaire were found (more municipalities received funding, but they didn't answer the questionnaire). Table 9 shows the government funding set out to the municipality and is then met with the data-readiness level.

Table 9. *Municipalities that received government funding. Nested with the level of data-readiness*

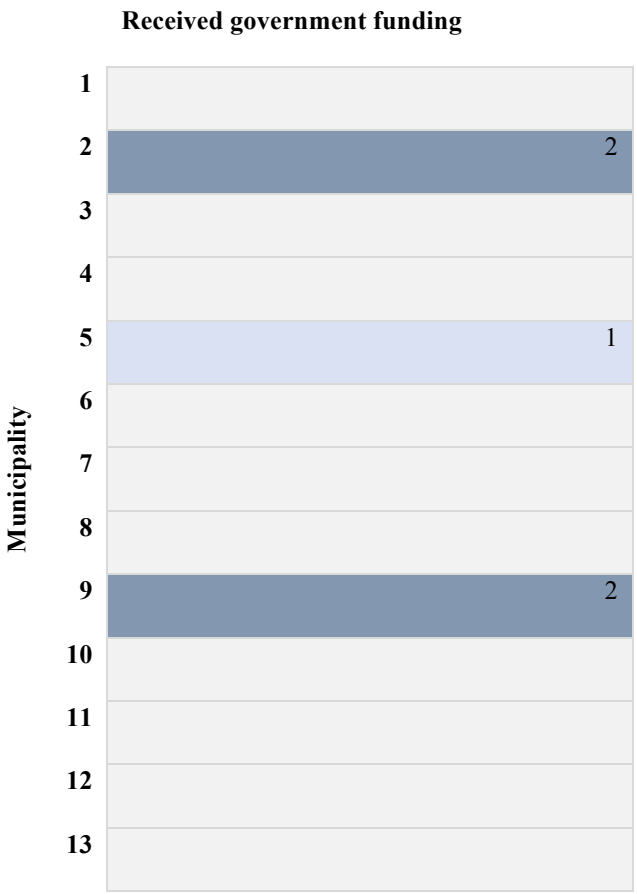

Table 9 shows that the municipalities receiving government funding didn't score the highest level of data-readiness. But, their mean of level is higher (1,7) than from the ones that didn't receive any government funding (1,4) and higher from the mean of all data-readiness levels of the municipalities (1,5) (see table 10). As the government funding is project based, it could be that the project itself would reach a higher level, but this doesn't work out for the overall municipality. We assumed that government funding would help municipalities to increase their budget for the implementation of big data projects. With the knowledge of this project, municipalities would be able to continue current, and to implement future big data projects. This 'spill-over effect' of the government funding is not directly shown. Notable is the municipality group the

[4] This data was confidentially gathered and can therefore not be presented in this report. If one would like to know, please contact the researcher.

[5] At the time of writing the ELSALabs are created, these 'labs' or platforms are specifically aimed at the knowledge distribution and collaboration of municipalities on digital governance questions.

government funding is given to (see table 10). The three respondents that also received the government funding are all in group E, the largest municipality group. This could explain why the expected spill-over effect is minimal. Larger municipalities work with more teams (all three indicated that they work with several teams). This increases the need for inter-organisational communication, also increasing the possibility of non-communication. The hierarchical models of the municipalities (Appendix D, municipality 2,5 and 9) show that either the communication channels are unknown (municipality 9) or that they are lacking (municipality 5). Only municipality 2 indicated that communication between teams happened on a regular basis. The hierarchy models also show that both municipality 2 and 9 (Appendix D) use data for policy making.

Table 10. *Cross table of municipality size group with municipality, nested data-readiness and hatched in red responses with government funding.*

	Municipality size group				
Data-readiness level	A	B	C	D	E
Municipality					
1		1			
2					2
3			2		
4				4	
5					1
6			1		
7			1		
8					1
9					2
10					1
11					1
12				1	
13			1		

Other variables

Two other variables were considered to influence the level of data-readiness: *years working with big data* and *size of the team working with big data.*

Four municipalities (Appendix D, municipality 2,3,4, and 8) indicated how long they work with big data. They are shown in table 11 (3 years = purple; 6 years = orange). The other respondents answered the question with 'not applicable' or 'I don't know'. The municipalites that answered 'not applicable' are also in table 11 in the colour yellow. Those that answered the question answered in two ways, either municipalities worked with big data since 2015, or since 2018. The year 2015 can relate to the decentralisation of the Dutch government system that put more responsibility at the municipality. Why municipalities started implementing big data in 2018 specifically, remains unknown. Notable in the distribution of how long municipalities work with big data is that only one of them (municipality 8) reached a level 1. The others reached at least a level 2, while the ones that indicated that big data was not applicable to them only reach level 1. This could indicate that municipalities working longer with big data helps with the reaching of a higher level.

The variable team size was researched by using a crosstab (table 12) of team size with data-readiness level. Interestingly the municipality with level 4 doesn't have the largest team (<20), the largest team is found in level 2 (>50). But, all municipalities that reach level 1 are working with a team of less than 10 people. An explanation could be that a small team start with implementing data practices and as the work broadens, they find new members.

Table 11. *Table of municipality size group with municiplity with indicated years working with big data.*

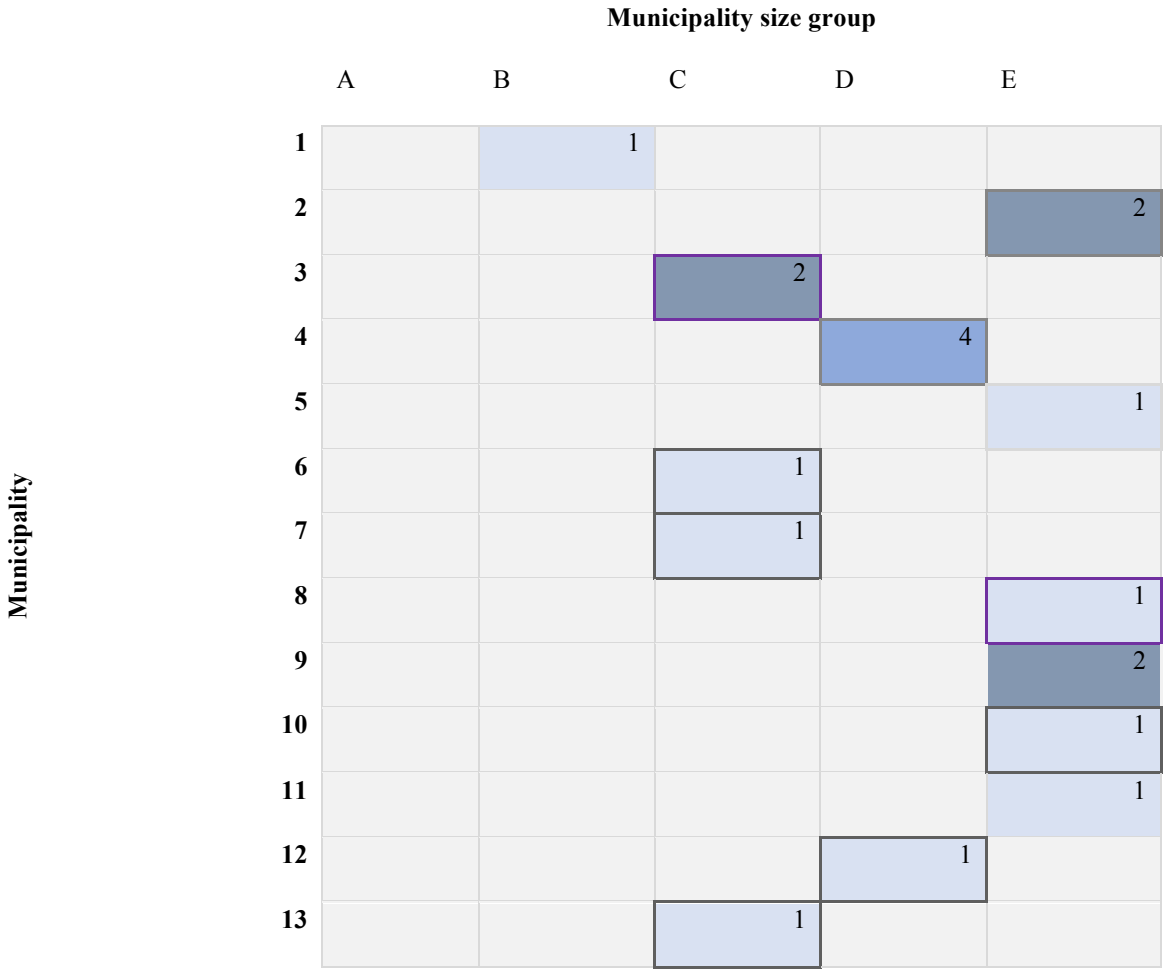

Table 12. *Cross table of team size with data-readiness level*

		Team size					
		<10	<20	<50	>50	Not applicable	Total
Data readiness level	Level 1	4	0	0	0	5	9
	Level 2	1	0	1	1	0	3
	Level 3	0	0	0	0	0	0
	Level 4	0	1	0	0	0	1
	Level 5	0	0	0	0	0	0
Total		5	1	1	1	5	13

The analysis of the results indicates four main conclusions. First, the hierarchy of the proposed framework and the appended questionnaire are a confirmed hierarchy. Secondly, the municipality size doesn't seem to have a direct influence on the level of data-readiness, but larger municipalities are overall further developed on the characteristics of the framework. Thirdly, a municipality receiving government funding, doesn't seem to influence the level of data-readiness. The municipalities receiving government funding are also the larger municipalities, resulting in a higher level reached within different characteristics of the framework. Then, the results of the *years working with big* data show that 5 of the 13 municipalities don't use big data at all. They reach level 1 as this level was made to present the legal aspects that are obligatory for municipalities in the Netherlands. The other possible variable needs to be researched with more responses so possible correlations can be found. For now, the other, possibly mediating, variables seem to have some sort of relationship with the level of data-readiness and possibly the main variables, municipality size and government funding.

5. Conclusion

This research started with the question: '*How can the differences in the level of data-readiness intra Dutch municipalities be explained with a data-readiness framework?'* The conducted research consisted of three parts. First, a data-readiness framework was made. Literature research found that there was no overall data-readiness framework. Frameworks only focussed on one aspect of data-readiness, such as e-government maturity models (Dwivedi et al., 2012; Gartner, 2017; Klievink & Janssen, 2009; Layne & Lee, 2001; Valdés et al. 2011), big data ethical models (Davis & Patterson, 2012; Etlinger & Groopman, 2015), if the organisations goals are aligned with big data (Burmeister et al., 2018; Demchenko et al., 2014; Orenga-Roglá & Chalmeta, 2019; Sun & Heller, 2012) and if the right capabilities are in place (Gèzcy, 2015; Tekiner & Keane, 2013). These models and frameworks were therefore combined into one overall data-readiness framework. This framework was based on the data-readiness framework by Klievink et al. (2017). This framework was specifically made for the Dutch public sector, thereby solving possible application problems due to different federal systems, in advance. The data-readiness framework by Klievink et al. (2017) assumed a *positive* view on the big data usage for public organisations. In the current social climate with several scandals about privacy (NOS, 29-04-2020;), personal data (NOS, 14-05-2021; NOS, 26-05-2021) and fake social media accounts (Piekartz, 2021; NOS, 18-05-2021), this assumption cannot be made. The ethical considerations *before* the implementation of big data practices cannot be left out when asking about data-readiness. Therefore, a fourth characteristic was added to the three existing characteristics presented by Klievink et al. (2017). This fourth characteristic is called 'ethical considerations' and combines legal obligations, such as working confirm the GDPR, with values considered important when implementing big data practices, such as transparency and ownership. The framework by Klievink et al. (2017) also changed in the calculation of the overall data-readiness. Other frameworks, such as the e-governance maturity model (Layne & Lee, 2001), are hierarchically set up, that the next level can only be reached by fulfilling the precedent level. This hierarchy was implemented in the new framework by finding a necessary condition for each of the four characteristics to make the next level of the characteristic sufficient for the precedent level. As a result, the framework tells the overall level of data-readiness, but also pinpoints where possible improvements can be made to reach a higher level. This is a unique approach to the question of overall data-readiness.

The framework was tested by a survey. 13 different municipalities answered the questionnaire, and resulted in 15 different filled 'hierarchy models' (appendix C for an example and D for the filled hierarchy models). In the analysis, the hierarchy of the framework was confirmed by comparing the different percentages of answers. The hierarchy models helped with the answering on how the differences between the municipalities can be explained. The main possible variables influencing the level of data-readiness were the *size of the municipality* and if the municipality received *government funding for a big data project.* These variables assume that the complexity of the municipality increases with its size, resulting in a need for more efficient systems to process all data, and that project based government funding would result in a spill-over effect to the rest of the municipality. With the current respondents, the size of the municipality doesn't seem to influence the level of data-readiness directly, but larger municipalities have more developed characteristics if they use big data. If a municipality received *government funding* also didn't seem to have a direct effect on the level of data-readiness.

Overall, the results of the questionnaire show that almost half of the responses don't use big data practices at all. They reach level 1 as this level is made on what is legally mandatory in the Netherlands. This indicates that municipalities use data, but not big data (yet). This result is no different from the result of the research done by Klievink et al. (2017). The overall data-readiness was then 'medium' on a scale from very low to very high. That is interesting as four years since the publication of the framework by Klievink et al. (2017) have passed. With the current increase of data projects and initiatives and the social call for ethical data usage, we expect that it will not be long before big data will be part of the daily practice of municipalities. If so, this framework pinpoints where the municipality needs to improve to reach a higher level of data-readiness.

6. Discussion

For this research, a questionnaire was used to fill in the hierarchy models. The questionnaire was shortened compared to the interview questions by Klievink et al. (2017), that consisted of 41 questions. The shortening of the questionnaire to 25 content questions caused independence for the respondents to answer them, but also resulted in questions and comments about when e.g. big data was applicable to the municipality. Similarly, the questionnaire combined explicit questions about big data practices with general questions about data processing. For further research, a mixed-method approach can help understand the answers of the questionnaire better and to possible pinpoint where the bottle neck of the municipality is.

Another comment on the questionnaire was about the rigor of the definition of big data. As was found in the analysis, municipalities use data practices, but not big data per se. As data-readiness models are mainly used for big data practices and not for general data or advanced data practices, these general data practices were not considered as possible responses. Half of the respondents indicated that they didn't use big data, but also indicated that they used general data practices. Therefore, we suggest exploring if the data-readiness models and frameworks can also be applied to general data practices. If so, then this could provide useful insights in the data usage of Dutch municipalities. With more insight comes more view on the issues and could therefore prevent data breaches and leaks. This would also improve the relationship between citizens and the government.

What was also found, but not discussed in this research are the differences between responses from the same municipality. Further research could also be focussed on the differences intra municipalities: Do civil servants in a certain function know what they need to know? Is a privacy officer up to date about the GDPR implications for data collection? This research would require a different approach than what was conducted for this report.

Further research can also be focussed on the relationship between AI and big data. Assumed is that big data and AI are related as with AI implementation more data is gathered, eventually resulting in big data (Demchenko et al., 2014). An example could be to not only look at government funding for big data projects, but also for AI projects that involve large amounts of data.

Lastly, this research is the first to focus on one government level. Instead of comparing Dutch executive government organisations (Klievink et al., 2017) to find an *overall* data-readiness of

the Dutch public sector, one government level is chosen and compared to find why they differ. The first possible factors that influence the data-readiness and can therefore declare the differences, are thus a start for more research to be conducted on government levels.

References

Abu-Salih, B., Wongthongtham, P., Zhu, D., Chan, K.Y., & Rudra, A. (2021). *Introduction to Big data Technology*.

Amnesty. (n.d.). *Mensenrechten in de pocket.* Amnesty Nederland.

Arockia Panimalar, S., Varnekha Shree, S., Veneshia Kathrine, A. (2017). The 17V's of Big Data. *International Research Journal of Engineering and Technology (IRJET), Volume 4,* issue 9-2017. pg. 329-333. Retrieved from, https://www.irjet.net/archives/V4/i9/IRJET-V4I957.pdf

Autoriteit Persoonsgegevens [AP] (n.d.) *Data protection impact assessment (DPIA).* Retrieved from, on the 6th of June 2021, https://autoriteitpersoonsgegevens.nl/nl/zelf-doen/data-protection-impact-assessment-dpia

Bantema, W., Westers, S., Hoekstra, M., Herregodts, R., Munneke, S., (2021). *Een wankel fundament onder een stevige praktijk.* Politiekunde *109*. Politie & Wetenschap.

Barham, H. A. (2019). Development of a readiness assessment model for evaluating big data projects: Case study of smart city in oregon, USA. (Order No. 13880955, Portland State University). *ProQuest Dissertations and Theses,* 323. Retrieved from https://www-proquest-com.ezproxy.leidenuniv.nl/dissertations-theses/development-readiness-assessment-model-evaluating/docview/2298099532/se-2?accountid=12045

Brown, B., Chui, M., and Manyika, J. Are you ready for the era of 'big data'? McKinsey Quarterly 4 (Oct. 2011), 24–35. Retrieved from, https://www.mckinsey.com/business-functions/mckinsey-digital/our-insights/big-data-the-next-frontier-for-innovation#

Bryman, A. (2016). *Social research methods*. Oxford: University Press.

Bürmann, S., & Zoeten, N. de, (2021). *Raad voor Openbaar Bestuur wil minister voor Digitale Zaken.* Nieuwsuur. Retrieved from, https://nos.nl/nieuwsuur/artikel/2382262-raad-voor-het-openbaar-bestuur-wil-minister-voor-digitale-zaken

Burmeister, F., Drews, P., and Schirmer, I. (2018). Towards an Extended Enterprise Architecture Meta-Model for Big Data–A Literature-based Approach. In Proceedings of 24th *Americas Conference on Information Systems, New Orleans*, 2018, 1-10.

CBS (2021). *Regionale kerncijfers Nederland.* Retrieved from, https://opendata.cbs.nl/statline/#/CBS/nl/dataset/70072ned/table?dl=3B993

Criddle, C., (28-10-2020). *Facebook sued over Cambridge Analytica data scandal.* BBC News. Retrieved from, https://www.bbc.com/news/technology-54722362

Das, T.K. and Kumar, P.M. Big data analytics: A framework for unstructured data analysis. *International Journal of Engineering and Technology 5*, 1 (Feb.-Mar. 2013), 153–156.

Davis K., & Patterson D., (2012). *Ethics of Big Data.* O'Reilly Media Inc. Sebastopol, CA. ISBN: 978-1-449-31179-7.

Demchenko, Y., de Laat, C., and Membrey, P. (2014). Defining architecture components of the big data ecosystem. *In Proceedings of the International Conference on Collaboration Technologies and Systems* (Minneapolis, MN, May 19–23). IEEE Press, pg. 104–112.

Digitale Overheid (2021). *Home.* Ministery of Internal Affairs. Retreived from https://www.digitaleoverheid.nl/

Dwivedi, Yogesh, Weerakkody, Vishanth, & Janssen, Marijn. (2012). Moving towards maturity. *The Data Base for Advances in Information Systems, 42*(4), 11-22.

Elinger, S., & Groopman, J., (2015). *The Trust Imperative: A framework for ethical Big Data Use.* Altimeter.

EU, Guide. (2020). *Complete guide to GDPR compliance.* Retrieved from, https://gdpr.eu/

EU. (2020). *What is GDPR, the EU's new data protection law?.* Retrieved from, https://gdpr.eu/what-is-gdpr/

European Commission. (2021). *What does the general data protection regulation (GDPR) govern?* Retrieved from, https://ec.europa.eu/info/law/law-topic/data-protection/reform/what-does-general-data-protection-regulation-gdpr-govern_en

Ferguson, M. Architecting a Big Data Platform for Analytics. *IBM White Paper*, Oct. 2012; http://www-01.ibm.com/common/ssi/cgi-bin/ ssialias?htmlfid=IML14333USEN

Gartner. (2017). Published on O'Reilly. *Appendix A. Gartner Maturity Model for Data and Analytics.* Retrieved from, https://www.oreilly.com/library/view/streaming-change-data/9781492032526/app01.html

Gèczy, P. (2015). Big data management: Relational framework. *Review of Business & Finance Studies 6,* (3). pg. 21–30.

Giest, S. (2017). Big data for policymaking: Fad or fasttrack? *Policy Sciences, 50*(3), 367-382.

Kelly, G., G. Mulgan and S. Muers. 2002. *Creating Public Value: An Analytical Framework for Public Service Reform*. London: UK Cabinet Office.

Klievink, B., Romijn, B., Cunningham, S., & de Bruijn, H. (2016). Big data in the public sector: Uncertainties and readiness. *Information Systems Frontiers*, 19 (2), 267–283.

Klievink, B., & Janssen, M. (2009). Realizing joined-up government - dynamic capabilities and stage models for transformation. *Government Information Quarterly, 26*(2), 275–284. doi:10.1016/j.giq.2008.12.007

Landman, T. (2004). Measuring Human Rights: Principle, Practice, and Policy. *Human Rights Quarterly, 26*(4), 906-931. Retrieved June 11, 2021, from http://www.jstor.org/stable/20069767

Layne, K., & Lee, J. (2001). Developing fully functional E-government: a four stage model. *Government Information Quarterly, 18*(2), 122–136.

Matheus, R., Janssen, M., & Janowski, T. (2021). Design principles for creating digital transparency in government. *Government Information Quarterly*, *38*(1), [101550]. https://doi.org/10.1016/j.giq.2020.101550

Martocchio, J.J. (2006). *Research in Personnel and Human Resources Management*: Vol. 1st ed. JAI Press Inc.

Mayer-Schönberger, V., & Cukier, K. (2014). *Big data: A revolution that will transform how we live, work, and think*. Boston, NY: Houghton Mifflin Harcourt.

McNeely, Connie L, & Hahm, Jong-on. (2014). The Big (Data) Bang: Policy, Prospects, and Challenges. *The Review of Policy Research, 31*(4), 304-310.

Meijer A., & Ruijer, E., (april 2021*). Code Goed Digitaal Openbaar Bestuur (CODIO): Borgen van waarden bij de digitalisering van het openbaar bestuur.* USBO Advies, Universiteit Utrecht. Retrieved from, https://www.uu.nl/sites/default/files/CODIO%20-%20Definitief%20-%20Opgemaakt%20-%20210416.pdf

Miller, H.G. and Mork, P. From data to decisions: A value chain for big data. *IT Professional 15,* 1 (Jan.- Feb. 2013), 57–59

NL DIGIbeter. (June 2020). *Agenda Digitale Overheid.* Ministry of Internal Affairs. Retrieved from https://www.digitaleoverheid.nl/wp-content/uploads/sites/8/2020/07/nl-digibeter-2020.pdf

NL DIGITAAL (March 2020). *Data Agenda overheid.* Ministry of Internal affairs. Retrieved from https://www.digitaleoverheid.nl/wp-content/uploads/sites/8/2020/04/7100-BZK_Publicatie_NL_DIGItaal_DAO-UPDATE-2020-T.pdf

NOS. (29-04-2020). *Privacy waakhond legt Enschede 600.000 euro boete op vanwege wifitracking.* Retrieved from, https://nos.nl/artikel/2378665-privacywaakhond-legt-enschede-boete-op-van-600-000-euro-vanwege-wifitracking

NOS. (14-05-2021). '*Nederland dreigt grip op digitale wereld te verliezen'.* Retrieved from, https://nos.nl/artikel/2380689-nederland-dreigt-grip-op-digitale-wereld-te-verliezen

NOS. (26-05-2021). *Excuses minister Bijleveld voor informatieverzameling burgers door defensie.* Retrieved from, https://nos.nl/artikel/2382430-excuses-minister-bijleveld-voor-informatieverzameling-burgers-door-defensie

NOS. (18-05-2021). *Gemeenten speuren anoniem op sociale media: 'Weten niet dat ze grenzen overgaan.'* Retrieved from, https://nos.nl/artikel/2381211-gemeenten-speuren-anoniem-op-sociale-media-weten-niet-dat-ze-grenzen-overgaan

Olszak, Celina, & Mach-Król, Maria. (2018). A Conceptual Framework for Assessing an Organization's Readiness to Adopt Big Data. *Sustainability (Basel, Switzerland), 10*(10), 3734.

Orenga-Roglá S., & Chalmeta R. (2019) Framework for Implementing a Big Data Ecosystem in Organisations. *Communications of the ACM, 62* (1) , pp. 58-65.

Piekartz, H., (2021). *Gemeenten kijken op grote schaal en in het geheim mee met burgers op sociale media.* Volkskrant. Retrieved from, https://www.volkskrant.nl/nieuws-achtergrond/gemeenten-kijken-op-grote-schaal-en-in-het-geheim-mee-met-burgers-op-sociale-media~bb40d4c7/

Raad openbaar Bestuur (2021). *Sturen of gestuurd worden; over de legitimiteit van sturen met data.* Adviesrapport ROB.

Sona, N. (2014). *Chapter 14: Big data models and the public sector.* In A. Aggarwal 'Managing Big Data Integration in the Public Sector. Pg. 245- 266. Advances In Public Policy and Administration (APPA) Book Series.

Sun, H. and Heller, P. (Aug. 2012). Oracle Information Architecture: An Architect's Guide to Big Data. *Oracle White Paper*, Retrieved from, https://d2jt48ltdp5cjc.cloudfront.net/ uploads/test1_3021.pdf

Taylor, K., (n.d.) *Data Governance Maturity model: Gartner & IBM.* Retrieved from, on the 6th of June 2021, https://www.hitechnectar.com/blogs/data-governance-maturity-models-explained/

Tekiner, F., and Keane, J.A. (2013). Big Data Framework. In Proceedings of IEEE *International Conference on Systems, Man, and Cybernetics, Manchester, U.K*. 13-16 Oct. IEEE Press, 1494-1499.

Toshkov, D., (2012). *Research Design in political science.* Palgrave Macmillan.

United Nations. (n.d.). *Global Issues; Human rights.* Retrieved from on the 6th of June 2021, https://www.un.org/en/global-issues/human-rights

Utrecht Dataschool (nov. 2019). *De Ethische Data Assistant (DEDA).* Universiteit Utrecht. Retrieved from, https://dataschool.nl/deda/

Valdés, G., Solar, M., Astudillo, H., Iribarren, M., Concha, G., & Visconti, M. (2011). Conception, development and implementation of an e-government maturity model in public agencies. *Government Information Quarterly, 28*(2), 176–187. doi:10.1016/j. giq.2010.04.007.

Vydra, S., & Klievink, B. (2019). Techno-optimism and policy-pessimism in the public sector big data debate. *Government Information Quarterly, 36*(4), 101383.

Appendices

Appendix A: Extensive considerations on the hierarchy framework

Question 21:

Do you know where you can find the privacy policy of your municipality?

This question relates to knowing where you can find something is also knowing that it is important. We intended a non-confirmative question that implies a certain knowledge.

Question 22:

Did you ever use, or are you familiar with one of the following tools? You can choose more than one.

Options:

- BIO-SA tool (Baseline Information security Government)
- PriSA (Privacy Self Assessment)
- Privacy maturity model
- Rijkspia (Privacy Impact Assessment of Rijksoverheid)

All tools provided here are easy accessible and easy to use self assessment tools. They are specifically meant for government organisations wanting to know if they meet the privacy regulations (CIP, 2019). The CIP (Centrum Informatiebeveiliging en Privacybescherming) offers several products, from self-assessment tools to workshops on privacy rights and ensurance of those rights. The CIP is specifically aimed at government organisations and therefore it suits the questionnaire.

Question 23:

Do you use additional tools for Big Data projects?

Options:

- Yes
- No
- I don’t know

- Not applicable

To understand if a municipality is working with ethical tools on project basis, we developed this question, with two follow up options. When people crossed ‘I don’t know’ or ‘Not applicable’ they are directed to the next question. Otherwise, if yes they get the following question:

Question 23.Yes:

Can you indicate if you use one of the following tools?

Options:

- DEDA (Ethical Data Assistant)
- Code for good public governance (CODIO)
- International guidelines (UN or EU)
- Principles of good public governance (NL)

These options were chosen for they are well-known for those noting the importance of values in digital government. The Ethical Data Assistant is a project tool specifically made by the University of Utrecht (DEDA, 2018). The CODIO is also made in collaboration with the university of Utrecht and based on research specifically aimed at finding the values that are relevant for good public governance (Meijer & Ruijer, 2021). International guidelines from the UN and EU have influence on the policies made in the Netherlands, as these policies are executed by lower government levels, these guidelines are relevant for municipalities. In accordance with those guidelines, the Netherlands also has its own guidelines that correspond with its own government goals. These can also be used as guidelines for the ethical considerations.

If respondents answer ‘no’ they get the question:

Question 23 No:

Could you indicate why you don’t use additional tools?

The following reasons are mentioned:

- I am not familiar with the available tools
- I am familiar but I cannot work with these tools
- There is no knowledge available
- We use another system, namely [enter own text]

Because of the diversity of tools it can be hard to navigate which tool to use for what, making it hard to work with tools in general.

Besides the diversity of the tools themselves, they also come with guidelines or workshops to use them in the right way. This can also make the barrier higher for people to use those tools. A certain kind of knowledge about tools is then to be expected. If that knowledge is not available, it becomes harder to use the tools.

Lastly, we want to have the option of respondents using different systems than those we proposed.

Question 24.a: Could you indicate what values are ensured when a big data project is executed?

The answers of this question are based on the CODIO values. But, this research comes up with 30 relevant values, and that is not workable in a questionnaire. We decided to make a selection of the available values. We indicate why we choose these and not the others.

- Cooperation
- Inclusivity
- Transparency
- Security
- Freedom of speech
- End responsibility
- Protection
- Integrity
- Non-discrimination
- Contestability

- Other, namely [enter text]
- Not applicable

In short, we choose ones that were obvious, such as transparency, inclusivity, non-discrimination and security and others that are less obvious, or most times forgotten, such as freedom of speech, protection and contestability. The ‘obvious’ are those that are apparent in the communication of the Dutch central government towards other public organisations. The ‘less obvious’ are those that are relevant (according to the CODIO) but there is almost no attention to those individually.

We then ask respondents to indicate what values they think are important for the future (question 24.b). The same options are used for this question.

Question 25: Could you indicate if your municipality uses one of the following ethical initiatives for its handling of big data?

Options are:

- Ethical commission for data projects
- Ethical officer
- Ethical frameworks
- Other [text entry box]
- We don’t use ethical initiatives
- Not applicable

These options are based on known initiatives, e.g. the VNG (collaboration of Dutch municipalities) has installed an ethical commission for all its data projects.

Ethical officers are known in different municipalities and with the variety of framework available we expect that some municipality transferred a framework to its organisation. Other options are welcome to be known and not using them is an answer too.

Appendix B: Assessment scorecard and decision rules

Alignment assessment

Q6. If big data is used for one of the options, the organisation scores on level 1. Also true if Q16 is fulfilled.

Q7. Automated processes in executive tasks scores level 2.

Q8.Only ‘regularly’, ‘often’ and ‘always’ count as sufficient for the condition to be fulfilled. Scores level 3.

Q9.Only ‘regularly’, ‘often’ and ‘always’ count as sufficient for the condition to be fulfilled. Scores level 4.

Q10. Only ‘regularly’, ‘often’ and ‘always’ count as sufficient for the condition to be fulfilled. Scores level 5.

Maturity assessment

Q11. Only ‘regularly’, ‘often’ and ‘always’ count as sufficient for the condition to be fulfilled.

If ‘one team’ is chosen, they also score on this level.

Q12. If yes, scores on level 2.

Q13. If yes, scores on level 3.

Q14. If ‘regularly’, ‘often’, ‘always’ selected, scores on level 4.

Q15. If yes, scores on level 5.

Capability assessment

Q16. If expertise available on: ‘collection’, ‘combination’, ‘analysis’ and ‘usage’, then condition fulfilled.

Q17. If all four score >50%, then level 2. If one <50%, then no score. If ‘not applicable’ dependent on answer to Q19. If condition fulfilled, then yes, if not fulfilled, then no.

Q18. Both need to score >50%, then level 3. If one <50%, then no score. If ‘not applicable’ dependent on Q19. If condition fulfilled, then yes, if not fulfilled then no.

Q19. Only ‘regularly’, ‘often’ and ‘always’ make the condition fulfilled. If selected, then level 4.

Q20. If yes, then level 5.

Ethical assessment

Q21. If yes, then level 1.

Q22. If at least 1 selected, then level 2.

Q23. If yes AND tool selected in 23a., then level 3.

Q24. If >2 are selected AND if >2 are selected in 24b., then condition fulfilled. Scores on level 4.

Q25. If at least 1 selected, then level 5.

Overall assessment

The level where not all conditions are fulfilled determines the level of data-readiness.

Appendix C: Example of Assesment

Assessment score card

Table I. *Example score card of assesment*

Framework feature	Question and corresponding level	Answer	Score (yes/no on level)
Alignment	6. Level 1	Management & administration; indirect coordination	Yes
	7. Level 2	Automated post system	Yes
	8. Level 3	'Sometimes'	No
	9. Level 4	'Never'	No
	10. Level 5	'Not applicable'	No
Maturity	11. Level 1	'One team'	Yes
	12. Level 2	'Yes'	Yes
	13. Level 3	'No'	No
	14. Level 4	'Sometimes'	No
	15. Level 5	'I don't know'	No
Capabilities	16. Level 1	'Juridical knowledge'	No
	17. Level 2	All four: 'Not applicable'	No
	18. Level 3	'IT strategy 40%'; 'IT infrastructure 80%'	No
	19. Level 4	'Sometimes'	No
	20. Level 5	'No'	No
Ethical considerations	21. Level 1	'Yes'	Yes
	22. Level 2	'No'	No
	23. Level 3	'No'; lack of expertise	No
	24. Level 4	'Not applicable'	No
	25. Level 5	'Not applicable'	No

Overall data-readiness score

As the organisation scores the lowest in the feature capabilities on level 1, the organisation has a data-readiness of level 1 Not all conditions for level 1 are fulfilled. The assessment shows where the organisation now should work on to reach the same level on all features.

The example filled hierarchy thus reaches a level of 1, as only the conditions on the level of 1 are fulfilled.

Table II. *Example filled hierarchy model*

Characteristic /level	**Organisational alignment**	**Organisational maturity**	**Organisational capabilities**	**Ethical considerations**
Level 5	Using data for policy making	Demand-driven, joint up government	External support	Do we need to use this data at all?
Level 4	Using data for analysis of situations	Inter-organisational integration	Internal support	Conscious of public values and their implications
Level 3	Using data for non–operational goals	Nationwide portal	Data expertise	Ethical data tools for projects
Level 2	Improving data processes	Integrated organisations/ teams	Data governance	Privacy tools used
Level 1	Collection of data	Stove-pipe organisation	IT governance	Confirm GDPR

Appendix D: Hierarchy models of municipalities

Legend

Fulfilled condition	
Unfulfilled condition	
Not applicable	
I don't know/no answer	

Municipality 1

Two answers recorded: 1a and 1b

	Organisational alignment	Organisational maturity	Organisational capabilities	Ethical considerations
Level 5	Using data for policy making	Demand-driven, joined up government	External support	Do we need to use this data at all?
Level 4	Using data for analysis of situations	Inter-organisational integration	Internal support	Conscious of public values and their implications
Level 3	Using data for non–operational goals	Nationwide portal	Data expertise	Ethical data tools for projects
Level 2	Improving data processes	Integrated organisations/teams	Data governance	Privacy tools used
Level 1	Collection of data	Stove-pipe organisation	IT governance	Confirm GDPR

Table 1. *Result representation municipality 1a (test)*

	Organisational alignment	Organisational maturity	Organisational capabilities	Ethical considerations
Level 5	Using data for policy making	Demand-driven, joined up government	External support	Do we need to use this data at all?
Level 4	Using data for analysis of situations	Inter-organisational integration	Internal support	Conscious of public values and their implications
Level 3	Using data for non–operational goals	Nationwide portal	Data expertise	Ethical data tools for projects
Level 2	Improving data processes	Integrated organisations/teams	Data governance	Privacy tools used
Level 1	Collection of data	Stove-pipe organisation	IT governance	Confirm GDPR

Table 2. *Result representation municipality 1b (test)*

Municipality 2
Two answers were recorded: 2a and 2b
This municipality received government funding for big data projects.

	Organisational alignment	Organisational maturity	Organisational capabilities	Ethical considerations
Level 5	Using data for policy making	Demand-driven, joined up government	External support	Do we need to use this data at all?
Level 4	Using data for analysis of situations	Inter-organisational integration	Internal support	Conscious of public values and their implications
Level 3	Using data for non–operational goals	Nationwide portal	Data expertise	Ethical data tools for projects
Level 2	Improving data processes	Integrated organisations/teams	Data governance	Privacy tools used
Level 1	Collection of data	Stove-pipe organisation	IT governance	Confirm GDPR

Table 3. *Result representation municipality 2a*

Characteristic/ level	Organisational alignment	Organisational maturity	Organisational capabilities	Ethical considerations
level 5	Using data for policy making	Demand-driven joined- up government	External support	Do we need to use this data at all?
level 4	Use big data for analysis of situations	Inter-organisational integration	Internal support for big data projects	Conscious of public values and their implications
level 3	Using data for non-operational goals	Nationwide portal	Data expertise	Ethical data tools for projects
level 2	Improving data processes	Integrated organisation/teams	Data governance	Privacy tools used
level 1	collect data	Stove-pipe organisation	IT-Governance	Confirm GDPR

Table 4. *Result representation municipality 2b.*

Municipality 3

Characteristic/ level	Organisational alignment	Organisational maturity	Organisational capabilities	Ethical considerations
Level 5	Using data for policy making	Demand-driven, joined up government	External support	Do we need to use this data at all?
Level 4	Using data for analysis of situations	Inter-organisational integration	Internal support	Conscious of public values and their implications
Level 3	Using data for non–operational goals	Nationwide portal	Data expertise	Ethical data tools for projects
Level 2	Improving data processes	Integrated organisations/teams	Data governance	Privacy tools used
Level 1	Collection of data	Stove-pipe organisation	IT governance	Confirm GDPR

Table 5. Result representation municipality 3.

Municipality 4

	Organisational alignment	Organisational maturity	Organisational capabilities	Ethical considerations
Level 5	Using data for policy making	Demand-driven, joined up government	External support	Do we need to use this data at all?
Level 4	Using data for analysis of situations	Inter-organisational integration	Internal support	Conscious of public values and their implications
Level 3	Using data for non–operational goals	Nationwide portal	Data expertise	Ethical data tools for projects
Level 2	Improving data processes	Integrated organisations/teams	Data governance	Privacy tools used
Level 1	Collection of data	Stove-pipe organisation	IT governance	Confirm GDPR

Table 6. *Result representation municipality 4.*

Municipality 5

Received government funding

Characteristic/ level	Organisational alignment	Organisational maturity	Organisational capabilities	Ethical considerations
level 5	Using data for policy making	Demand-driven joined- up government	External support	Do we need to use this data at all?
level 4	Use big data for analysis of situations	Inter-organisational integration	Internal support for big data projects	Conscious of public values and their implications
level 3	Using data for non-operational goals	Nationwide portal	Data expertise	Ethical data tools for projects
level 2	Improving data processes	Integrated organisation/teams	Data governance	Privacy tools used
level 1	collect data	Stove-pipe organisation	IT-Governance	Confirm GDPR

Table 7. *Result representation municipality 5.*

Municipality 6

	Organisational alignment	Organisational maturity	Organisational capabilities	Ethical considerations
Level 5	Using data for policy making	Demand-driven, joined up government	External support	Do we need to use this data at all?
Level 4	Using data for analysis of situations	Inter-organisational integration	Internal support	Conscious of public value1s and their implications
Level 3	Using data for non–operational goals	Nationwide portal	Data expertise	Ethical data tools for projects
Level 2	Improving data processes	Integrated organisations/teams	Data governance	Privacy tools used
Level 1	Collection of data	Stove-pipe organisation	IT governance	Confirm GDPR

Table 8. *Result representation municipality 6.*

Municipality 7

Characteristic/ level	Organisational alignment	Organisational maturity	Organisational capabilities	Ethical considerations
level 5	Using data for policy making	Demand-driven joined- up government	External support	Do we need to use this data at all?
level 4	Use big data for analysis of situations	Inter-organisational integration	Internal support for big data projects	Conscious of public values and their implications
level 3	Using data for non-operational goals	Nationwide portal	Data expertise	Ethical data tools for projects
level 2	Improving data processes	Integrated organisation/teams	Data governance	Privacy tools used
level 1	collect data	Stove-pipe organisation	IT-Governance	Confirm GDPR

Table 9. *Result representation municipality 7.*

Municipality 8

Characteristic/ level	Organisational alignment	Organisational maturity	Organisational capabilities	Ethical considerations
level 5	Using data for policy making	Demand-driven joined-up government	External support	Do we need to use this data at all?
level 4	Use big data for analysis of situations	Inter-organisational integration	Internal support for big data projects	Conscious of public values and their implications
level 3	Using data for non-operational goals	Nationwide portal	Data expertise	Ethical data tools for projects
level 2	Improving data processes	Integrated organisation/teams	Data governance	Privacy tools used
level 1	collect data	Stove-pipe organisation	IT-Governance	Confirm GDPR

Table 10. *Result representation municipality 8.*

Municipality 9

Characteristic/ level	Organisational alignment	Organisational maturity	Organisational capabilities	Ethical considerations
level 5	Using data for policy making	Demand-driven joined- up government	External support	Do we need to use this data at all?
level 4	Use big data for analysis of situations	Inter-organisational integration	Internal support for big data projects	Conscious of public values and their implications
level 3	Using data for non-operational goals	Nationwide portal	Data expertise	Ethical data tools for projects
level 2	Improving data processes	Integrated organisation/teams	Data governance	Privacy tools used
level 1	collect data	Stove-pipe organisation	IT-Governance	Confirm GDPR

Table 11. Result representation municipality 9.

Municipality 10

Characteristic/ level	Organisational alignment	Organisational maturity	Organisational capabilities	Ethical considerations
level 5	Using data for policy making	Demand-driven joined- up government	External support	Do we need to use this data at all?
level 4	Use big data for analysis of situations	Inter-organisational integration	Internal support for big data projects	Conscious of public values and their implications
level 3	Using data for non-operational goals	Nationwide portal	Data expertise	Ethical data tools for projects
level 2	Improving data processes	Integrated organisation/teams	Data governance	Privacy tools used
level 1	collect data	Stove-pipe organisation	IT-Governance	Confirm GDPR

Table 12. *Result representation municipality 10.*

Municipality 11

Characteristic/ level	Organisational alignment	Organisational maturity	Organisational capabilities	Ethical considerations
level 5	Using data for policy making	Demand-driven joined- up government	External support	Do we need to use this data at all?
level 4	Use big data for analysis of situations	Inter-organisational integration	Internal support for big data projects	Conscious of public values and their implications
level 3	Using data for non-operational goals	Nationwide portal	Data expertise	Ethical data tools for projects
level 2	Improving data processes	Integrated organisation/teams	Data governance	Privacy tools used
level 1	collect data	Stove-pipe organisation	IT-Governance	Confirm GDPR

Table 13. *Result representation municipality 11.*

Municipality 12

Characteristic/ level	Organisational alignment	Organisational maturity	Organisational capabilities	Ethical considerations
level 5	Using data for policy making	Demand-driven joined- up government	External support	Do we need to use this data at all?
level 4	Use big data for analysis of situations	Inter-organisational integration	Internal support for big data projects	Conscious of public values and their implications
level 3	Using data for non-operational goals	Nationwide portal	Data expertise	Ethical data tools for projects
level 2	Improving data processes	Integrated organisation/teams	Data governance	Privacy tools used
level 1	collect data	Stove-pipe organisation	IT-Governance	Confirm GDPR

Table 14. *Result representation municipality 12.*

Municipality 13

Characteristic/ level	Organisational alignment	Organisational maturity	Organisational capabilities	Ethical considerations
level 5	Using data for policy making	Demand-driven joined- up government	External support	Do we need to use this data at all?
level 4	Use big data for analysis of situations	Inter-organisational integration	Internal support for big data projects	Conscious of public values and their implications
level 3	Using data for non-operational goals	Nationwide portal	Data expertise	Ethical data tools for projects
level 2	Improving data processes	Integrated organisation/teams	Data governance	Privacy tools used
level 1	collect data	Stove-pipe organisation	IT-Governance	Confirm GDPR

Table 15. *Result representation municipality 13.*

Appendix E: Complete questionnaire

Questionnaire

Vragenlijst Big Data gebruik in Nederlandse gemeenten

Start of Block: Introductie

Introductie

Beste lezer,

Welkom bij deze vragenlijst omtrent Big Data-gebruik onder Nederlandse gemeenten. Alvast bedankt voor uw tijd voor het invullen.

Het doel van deze vragenlijst is om inzicht te verkrijgen in hoe en waarom Nederlandse gemeenten wel of geen gebruik maken van Big Data.

In het uiteindelijke verslag zal niet worden verwezen naar individuele gemeenten, maar alleen gesproken worden over algemene verschillen en trends. Voor dit onderzoek is het wel van belang te weten van welke gemeente u bent en welke functie u vervult.

Toestemming gegevens

Nu u dit weet, kunt u akkoord gaan met dat uw gegevens opgeslagen en verwerkt worden voor dit onderzoek? De gegevens worden binnen Rijksoverheid opgeslagen en verwijderd twee jaar na publicatie van dit onderzoek.

1. o Ja (1)
2. o Nee (2)

Introductie

Dank u wel. De vragenlijst begint vanaf hier.

Er zijn 25 vragen met daartussen informatie over begrippen. Bovenin ziet u uw voortgangsbalk.

U kunt de survey ook opslaan en later afmaken.

Druk op de pijl onderin om met de vragen te starten.

Page Break	

1 Welke gemeente vertegenwoordigt u?

_____

Page Break	

2 Wat is uw functie binnen de gemeente?

_____

Page Break	

Uitleg Big Data

Big Data is een breed begrip dat veel definities kent. Voor deze vragenlijst gaan we uit van de 3V's: Volume, Variëteit en Verandering. Als deze alle drie voorkomen in de gebruikte data, dan is er sprake van Big Data (zie figuur).

Deze figuur wordt gedurende de vragenlijst ook rechts bovenin weergeven.

Page Break	

3 Bent u bekend met Big Dataprojecten en -strategieën van ${1/ChoiceTextEntryValue}?

3. o Ja, met beide (1)
4. o Ja, met Big Data-projecten (2)
5. o Ja, met Big Data-strategieën (4)
6. o Nee, met beide niet (3)
7. o Nee, wij maken (nog) geen gebruik van Big Data (5)

Page Break	

NB

U heeft aangegeven dat uw gemeente nog geen gebruik maakt van Big Data. Voor dit onderzoek is het nog steeds nuttig als u alle vragen doorloopt. Als de vraag specifiek verwijst naar Big Data en uw gemeente voldoet niet aan de definitie en de vraag, kiest u dan voor 'niet van toepassing'.

Page Break	

4 Sinds wanneer maakt ${1/ChoiceTextEntryValue} gebruik van Big Data voor projecten?

8. o Sinds: (4) _____

9. o Niet van toepassing (5)

Page Break	

5 Kunt u een schatting geven van het aantal mensen dat meewerkt aan Big Data-projecten binnen uw gemeente?

10. o minder dan 10 (1)
11. o Minder dan 20 (2)
12. o Minder dan 50 (3)
13. o Meer dan 50 (4)
14. o Meer dan 100 (5)
15. o Niet van toepassing (6)

End of Block: Introductie

Start of Block: Organisatie

6 Voor welk van de volgende doeleinden worden Big Data-processen ingezet binnen ${1/ChoiceTextEntryValue}? U kunt meerdere toepassingen kiezen.

16. ☐ Management & Administratie (1)
17. ☐ Evaluatie & Onderzoek (2)
18. ☐ Registratie & Documentatie (3)
19. ☐ Indirecte Coördinatie (4)
20. ☐ Project basis (5)
21. ☐ Geen van bovenstaand (6)
22. ☐ Niet van toepassing (7)

Page Break	

7 Kunt u een voorbeeld geven van een (deels) geautomatiseerde dienstverlening binnen uw gemeente?

_____

Page Break	

8 Hoe vaak wordt bij de gemeente reeds bestaande data ingezet voor doeleinden die buiten de dienstverlening vallen? Denk hierbij aan het handhaven van wetten of het doen van onderzoeken.

23. o Nooit (1)

24. o Soms (2)

25. o Regelmatig (3)

26. o Vaak (4)

27. o Altijd (5)

28. o Weet ik niet (6)

Page Break	

9 Hoe vaak wordt bij de gemeente reeds bestaande data ingezet voor het analyseren van een bepaalde situatie?

1. o Nooit (1)
2. o Soms (2)
3. o Regelmatig (3)
4. o Vaak (4)
5. o Altijd (5)
6. o Weet ik niet (6)

Page Break	

10 Hoe vaak wordt bij de gemeente reeds bestaande data ingezet voor het maken van nieuw beleid?

1. o Nooit (1)
2. o Soms (2)
3. o Regelmatig (3)
4. o Vaak (4)
5. o Altijd (5)
6. o Weet ik niet (6)

End of Block: Organisatie

Start of Block: Communicatie

11 Hoe vaak vindt uitwisseling van Big Data-projectdoelen en -resultaten plaats tussen de verschillende teams binnen uw gemeente?

1. o Nooit (1)
2. o Soms (2)
3. o Regelmatig (3)
4. o Vaak (4)
5. o Altijd (5)
6. o Er is één team (6)
7. o Weet ik niet (7)
8. o Niet van toepassing (8)

Page Break	

12 Bestaat er een platform waarop de inzichten en resultaten van Big Data-projecten van uw gemeente *intern* worden gedeeld?

1. o Ja, namelijk: (1) _____
2. o Nee (2)
3. o Weet ik niet (3)
4. o Niet van toepassing (4)

Page Break	

13 Bestaat er een platform waarop de inzichten en resultaten van Big Data-projecten van uw gemeente *extern* worden gedeeld?

1. o Ja, namelijk: (1) _____
2. o Nee (2)
3. o Weet ik niet (3)
4. o Niet van toepassing (4)

Page Break	

14 In hoeverre wordt de informatie over Big Data-projecten gedeeld met inwoners en relevante stakeholders?

1. o Nooit (1)
2. o Soms (2)
3. o Regelmatig (3)
4. o Vaak (4)
5. o Altijd (5)
6. o Weet ik niet (6)
7. o Niet van toepassing (7)

Page Break	

15 Zijn er Big Data-projecten binnen uw gemeente waarvan de broncode openbaar beschikbaar is?

1. o Ja (1) _____
2. o Nee (2)
3. o Weet ik niet (3)
4. o Niet van toepassing (4)

End of Block: Communicatie

Start of Block: Mogelijkheden

Uitleg expertise

De volgende vraag gaat over expertise binnen uw organisatie. Onder expertise vallen voor deze vragen:

1. Werknemers aangenomen voor een specifieke functie.
2. Werknemers die geschoold zijn om een bepaalde functie te kunnen vervullen.

16 Heeft uw gemeente eigen data-expertise in huis? Zo ja, op welke van de volgende gebieden?

1. ☐ Verzamelen van data (1)
2. ☐ Combineren van data (2)
3. ☐ Analyseren van data (3)
4. ☐ Gebruiken van data (4)
5. ☐ Juridische kennis over datagebruik of privacywetgeving (5)
6. ☐ Beoordelen van dataprojecten (6)
7. ☐ Anders (7) _____
8. ☐ Niet van toepassing (8)

Page Break	

17 Heeft uw gemeente een Big Data-strategie waarin beschreven wordt hoe de gemeente omgaat met de vier procesonderdelen van Big Data (hieronder weergegeven)? Hier wordt een vergelijkbare strategie bedoeld als bijvoorbeeld de Nederlandse datastrategie.

	Geen	Volledig uitgewerkt	Niet van toepassing
Verzamelen van data ()			
Combineren van data ()			
Analyseren van data ()			
Gebruiken van data ()			

Page Break	

18 Heeft uw gemeente een IT-strategie of -infrastructuur voor het gebruik van Big Data?

	Geen	volledig uitgewerkt	Weet ik niet/ Niet van toepassing
IT-strategie ()			
IT-infrastructuur ()			

Page Break	

19 Hoe vaak heeft u overleggen met managers of managementteams binnen uw eigen gemeente over Big Data-projecten?

1. o Nooit (1)
2. o Soms (2)
3. o Regelmatig (3)
4. o Vaak (4)
5. o Altijd (5)
6. o Weet ik niet (6)
7. o Niet van toepassing (7)

Page Break	

20 Heeft ${1/ChoiceTextEntryValue} een strategie om inwoners op de hoogte te houden van Big Data-projecten en hun ontwikkelingen?

1. o Ja, namelijk: (1) _____
2. o Nee (2)
3. o Weet ik niet (3)
4. o Niet van toepassing (5)

End of Block: Mogelijkheden

Start of Block: Ethiek

21 Weet u waar u het privacybeleid van uw gemeente kunt vinden?

1. o Ja (1)
2. o Nee (2)

Page Break	

22 Heeft u gebruik gemaakt of bent u bekend met een van de onderstaande tools? U kunt meerdere tools kiezen.

1. ☐ BIO-SA tool (Baseline Informatiebeveiliging Overheid) (4)
2. ☐ PriSA (Privacy Self Assessment) (5)
3. ☐ Privacy volwassenheidsmodel (6)
4. ☐ Rijkspia (Privacy Impact Assesment van het Rijk) (7)

Page Break	

23 Maakt u gebruik van aanvullende tools voor Big Data-projecten?

1. o Ja (1)
2. o Nee (2)
3. o Weet ik niet (3)
4. o Niet van toepassing (4)

23 Kunt u aangeven van welke tools u gebruik maakt?

1. ☐ De Ethische Data Assistent (DEDA) (1)
2. ☐ Code Goed Digitaal Openbaar Bestuur (CODIO) (2)
3. ☐ Internationale richtlijnen (UN of EU-richtlijnen) (4)
4. ☐ Beginselen goed openbaar bestuur (NL) (5)
5. ☐ Anders, namelijk: (3)

_____

23 Kunt u aangeven waarom u geen gebruik maakt van aanvullende tools?

1. ☐ Niet bekend met beschikbare tools (1)
2. ☐ Tools zijn bekend, maar niet werkzaam (2)
3. ☐ Gebrek aan expertise (3)
4. ☐ We hebben een ander systeem, namelijk: (4)

_____

Page Break	

24 Welke van de onderstaande waarden worden gewaarborgd wanneer een Big Data-project wordt uitgevoerd binnen ${1/ChoiceTextEntryValue}?

1. ☐ Samenwerking (1)
2. ☐ Inclusiviteit (2)
3. ☐ Transparantie (3)
4. ☐ Veiligheid (4)
5. ☐ Vrijheid van meningsuiting (6)
6. ☐ Privacy (7)
7. ☐ Bescherming (8)
8. ☐ Integriteit (9)
9. ☐ Non-discriminatie (10)
10. ☐ Anders, namelijk: (11)

_____

Page Break	

25 Kunt u aangeven of ${1/ChoiceTextEntryValue} gebruikt maakt van een van de volgende ethische initiatieven voor de omgang met Big Data?

1. ☐ Ethische commissie voor dataprojecten (4)
2. ☐ Ethical officer (7)
3. ☐ Ethische kaders (8)
4. ☐ Anders (5) _____
5. ☐ Wij maken geen gebruik van ethische initiatieven (9)
6. ☐ Niet van toepassing (10)

End of Block: Ethiek

Start of Block: Afsluiting

Afsluiting

Dit was de laatste vraag van de enquete.

Afsluiting

Zou u op de hoogte willen worden gehouden van de resultaten van dit onderzoek?

1. o Ja (1)
2. o Nee (2)

Page Break	

Afsluiting

Vul hieronder dan uw e-mailadres in:

_____

Afsluiting

Dank u wel voor uw deelname.

Als u nu doorklikt wordt de vragenlijst afgesloten en uw antwoorden verwerkt.

Page Break	

Afsluiting

Dank u wel voor uw deelname. Als u nu doorklikt dan wordt de vragenlijst afgesloten en uw antwoorden verwerkt.

End of Block: Afsluiting