Measuring, defining success, and the role of datafication within data-driven societies

Master Thesis Applied Ethics
Utrecht University

Christiaan Korterink
3867196
Supervisor: Hafez Ismaili M'hamdi
Second examiner: Rob van Gerwen

August 2018
h.c.korterink@gmail.com

Universiteit Utrecht

Abstract

Datafication, the process of rendering into data, aspects of the world not previously quantified, is a process that municipalities are facing. Moreover, municipalities are encouraged to make the transition to become a data-driven organization. In these processes of datafication, municipalities use different forms of data gathering and analysis. Through profiling, monitoring, data-discovery and predictive analysis several threats for human freedoms occur. In this thesis I analyze the risks towards two forms of human freedom, formulated by Amartya Sen, namely, social opportunities and transparency guarantees.

I argue that there is a danger for municipalities to perceive their inhabitants as algorithmic citizens made up of data. Furthermore, because of this focus on big data and measurable averages, solutions to problems that arise within society are sought within technological possibilities that are available, rather than through human intervention and interpretation. To make sense of data, and to prevent societies from becoming fully automated, this human interpretation and intervention is necessary to value the complexities of human life and to safeguard the human freedoms.

Table of contents

Introduction

There are many problems that arise because of the digitalization within municipalities. Municipalities have to deal with issues of privacy, safety of information, manufacturing, and the quality and ownership of data.[1] Many of these problems have been acknowledged on a broad level, and have been discussed thoroughly in literature. However, there is another worry that I will address in this thesis. This concern occurs because of the digitalization, and is known as *datafication.* Datafication refers to the process of rendering into data aspects of the world not previously quantified. This is not only demographic or profiling data, but behavioral data as well.[2] This datafication is often used for purposes of surveillance. However, there are other purposes, and concerns as well that affect human freedoms. Precisely because there are massive flows of data circulating between devices and institutions, it becomes vital to reflect whether a datafied society enables every individual in society to participate, and does not infringe upon several freedoms that humans should be able to enjoy.

The first step in the process of creating a data-driven society was taking the analog world into the digital environment. It allowed society, large commercial companies and governments, to store more information and to process it more rapidly.[3] The next step was to datafy the information. Because of this datafication, data can be put into a quantified format to be compared and analyzed. It allows analysis across large data sets. To make sense of these large data sets, algorithms play a huge role in data-driven societies. Algorithms are most of the times perceived as neutral technologies that provide the designers with objective results. However, there are some issues concerning the use of algorithms, and relying on their decision-making, that will be addressed later on.

It is striking that in 1986, Langdon Winter introduced the concept of *mythinformation,* a concept that is still present in data-driven societies, and should be critically analyzed. Winter describes mythinformation as follows:

---

[1] Wesseling, H., Postma, R., Stolk, R., 'Datagedreven sturing bij gemeenten, van data tot (gedeelde) informatie voor beter (samen) sturen,' *Vereniging van Nederlandse Gemeenten* (2018) p. 4.
[2] Kennedy, H., Poell, T., Dijck, van, J., 'Data and agency', *Big Data & Society* (2015) p. 1.
[3] Mai, J.E. 'Big Data Privacy: The Datafication of personal information', *The Information Society* (2016) vol. 32, p. 193.

> 'The almost religious conviction that a widespread adoption of computers and communications systems along with easy access to electronic information will automatically produce a better world for human living.'[4]

Cathy O'Neil is a well-known mathematician who criticizes the attitude that most scientists and governors have towards big data, because they trust the outcomes of the algorithms blindly.[5] In this thesis, I will argue in line with O'Neil, that we should be aware of this sanctifying believe in big data. Besides, I will try to formulate courses of action for municipalities to secure human freedoms within a datafied society.

It is important to be aware of the possible meanings that are attached to the term data, and the way it is used. Most of the times, data is seen as evidence for a hypothesis, which will give insights through analysis of big datasets. However, it is often just *data structures* that become visible, that do not have the guarantee to be meaningful in any way.[6] An important aspect of datafying is that this distinction between what data structures are, and what information is, disappears. It is possible to find patterns and correlations between all these elements.[7] Big Data is another term that is commonly used and has many different meanings. In this thesis big data refers to the definition that boyd and Crawford formulated. They define big data as a cultural, technological, and scholarly phenomenon that rests on the interplay of technology, analysis and mythology.[8] Technology is used to maximize computation power and algorithmic accuracy to gather, analyze, link, and compare large data sets. Analysis draws on large data sets, to identify patterns to make claims on an economic or societal level. Mythology refers to the aspect that through the analysis of large datasets a higher form of intelligence is offered, and insights become available that were previously impossible, 'with the aura of truth, objectivity, and accuracy.'[9] It is important to note that analysis of big

---

[4] Winner, L., 'The Whale and the Reactor: A Search for Limits in an Age of High Technology', p. 105.
[5] Weijer, B., van der, 'Wiskundige Cathy O'Neil waarschuwt voor algoritmen: 'Rechten van individu worden niet beschermd', *de Volkskrant* (June, 18, 2018) https://www.volkskrant.nl/wetenschap/wiskundige-cathy-o-neil-waarschuwt-voor-algoritmen-rechten-van-individu-worden-niet-beschermd-~b97a9302/
[6] Smith, B.C., 'Big Data in the Humanities: The Need for Big Questions', *Arts & Humanities Research Council,* visited May, 12, 2018, https://www.sciculture.ac.uk/2013/11/12/big-data-in-the-humanities-the-need-for-big-questions/ (12-11-2013)
[7] Mai, 'Big Data Privacy', p. 193.
[8] boyd, d., Crawford, K., 'Critical Questions for Big Data: Provocations for a cultural, technological, and scholarly Phenonemon', *Information, Communication & Society* (2012) vol. 15, p. 663.
[9] boyd, Crawford, 'Critical Questions for Big Data', p. 663.

datasets can uncover relationships between datasets, which do not tell anything about the causation.[10]

People would purportedly never walk around happily if they had tracking devices on them, without them knowing. However, people walk around with mobile phones that have the ability to track everything the person does. It collects the data about our behavior, when we go to sleep and when we wake up. It also collects data about our whereabouts, where we sleep, and with whom we go to sleep, because different data collections can be compared and analyzed.[11] Personal data have a very high value. Commercial companies try to gain information about individuals to target them effectively with advertisement. Governments and municipalities are trying to find out what they can do with a lot of the new information that is being gathered and created.

Almost every action we undertake, or transaction that we make, results in data. People get accustomed to the fact that they trade their data for services. Digitalization makes this possible, but datafication goes further than that. This datafication provides new possibilities to companies and governments, because it becomes relatively easy to get insights to human behavior, and to monitor this.[12] Distance is needed for personal freedom, but hard to obtain when almost everything is data-driven and measurable. People do not want to be too closely tied, not every action they do should be datafied. In the same time, people do not want to be excluded socially, so they feel obligated to participate in a society that is highly data-driven. The Netherlands Scientific Council for Government Policy wrote a report about the risks of involving Big Data in surveillance. They warn for the risk of negation of distance. Within a society that is highly data-driven, distance between the government and its citizens diminishes to a level that undermines the freedom of individuals.[13] They make an even stronger claim: 'For the government, it is only citizen behavior in relation to the law that should count. In a free society, citizens are not judged according to *who* they are.'[14]

---

[10] Anderson, J., lecture Digital Ethics, Universiteit Utrecht, 01-03-2018.
[11] Schneier, B., 'Data and Goliath: The Hidden Battles to Collect Your Data and Control Your World', *Norwegian Developers Conference,* visited at 11 June 2018: https://vimeo.com/131115865.
[12] Dijck, van, J., 'Datafication, dataism and dataveillance – Big Data Between Scientific Paradigm and Secular Belief', *Surveillance & Society* (2014) p. 198.
[13] Broeders, D., Schrijvers, E., Hirsch Ballin, E., *Big Data and Security Policies: Serving Security, Protecting Freedom,* WRR-Policy Brief 6 (2017) p. 7.
[14] Ibid.

By discussing these matters, and the impact that datafication has on individuals' freedom of choice within data-driven societies, I will try to answer my overall question: How can municipalities secure the human freedoms that are affected by datafication, within a data-driven society?

In the first chapter, I will elaborate on the digital society that we are living in, and the ways that municipalities are being encouraged to become a data-driven institution. There are several problems that occur because of this transition that municipalities have to undergo because of the relationship between the public and the private sector. Other problems that occur happen because of datafication, and as I will argue, are freedom-limiting problems that are underexposed. To fully understand why datafication goes wrong, we need to have a better notion of what freedom contains in this specific context. In the second chapter, I will introduce Amartya Sen's ideas about freedom, and specifically his account of human freedoms, to discuss these matters. Especially relevant for my thesis is the distinction between the narrow view of development, considering technological advance, and the broader view of development, the real freedoms people can enjoy. Furthermore, Sen distinguishes five different forms of human freedoms that will be discussed more thoroughly.

Two of these freedoms will be central to my thesis, because I believe that those are the ones that are most affected by datafication. These two are the human freedoms of social opportunities and transparency guarantees. Social opportunities are contested because the actual choices that individuals can make become limited in a datafied society. Furthermore, a virtue like solidarity is on the line because of the mingling between private and public interests, between governments and private companies. These tensions between the private and the public sector are relevant to understand the possible problems that occur in a highly data-driven society. Therefore, I will analyze the risks of these developments as well. Organizations and institutions that want to become data-driven use practices of monitoring, predictive analysis, profiling and data-discovery. These practices have their own ethical concerns, which I will highlight in the first chapter, before I will analyze them in the third and fourth chapter. In chapter five, I will discuss the role of big data and how this should be supplemented with thick data. Furthermore, I will discuss the role of accountability and trust to come up with possible attitudes towards the datafication within municipalities.

1. Transition towards data-driven organizations

In this chapter I describe the transition that is promoted amongst municipalities, to become a data-driven organization. Furthermore, there are some issues that frustrate this transition, because of the collaboration that municipalities often have with private companies. This is worrisome when public institutions depend on the services of these private firms. I discuss a few examples of this collaboration and the blurring of the divide between private and public, and the role of datafication. Ultimately, these processes and practices of datafication have several freedom-limiting worries that I will highlight.

Municipalities are prone to become overwhelmed. There is a considerable pressure to communicate better and quickly about services, calamities and news with their citizens. This has several reasons. Citizens become more vocal and demanding, and society is quickly transforming into a world that is continuously changing, and where an overload of data is being created.[15] To effectively manage this process, municipalities have to be in charge of relevant techniques to improve the communication with citizens and provide the necessary services. However, there are some concerns for a successful transition. Up till now, municipalities need the expertise of private companies for the development of technologies, and the analysis of big datasets. The possibilities that are created through the development of new technologies, determine which options governments have to solve problems.[16] They rely on private companies, and cannot completely rule out any collaboration with these companies. It is important that municipalities maintain a correct balance between the interference of private companies and the development of municipalities towards data-driven organizations, because people do no want to be ruled by private companies. What we can learn from big datasets, through using techniques like data mining, is available to those with access to the machines, the databases, and the algorithms.[17] It is essential that municipalities are in control of the relevant techniques.[18] In other words, when "the smartest person in the room is the room", [19] much, if not everything, depends on who is in control of the room, and who is the one that oversees it all.

---

[15] Hoedt, den, E., *Deel je rijk - Relevante trends voor overheidscommunicatie*, Report Ministry of General Affairs (2013) p. 9, 10.
[16] Vanolo, 'Smartmentality', p. 891.
[17] Andrejevic, M., 'The Big Data Divide', *International Journal of Communication* (2014) vol. 8, p. 1676.
[18] 'Maak Waar!', *Studiegroep Informatiesamenleving en Overheid* (2017) p. 5.
[19] Andrejevic, 'The Big Data Divide', p. 1676.

KPN is one of the main private companies that contribute in the transformation of municipalities towards data-driven organizations. They work closely together with municipalities, mainly by designing and facilitating the infrastructures for data. A data-driven society can only operate when trusted exchange is possible through digital networks.[20] According to KPN, cities have an extra layer on top of the infrastructure that is invisible. This is what they call a 'data-layer'. This infrastructure connects all the smart devices that people have in their homes, smart lampposts and mobile phones, among other things.[21] This infrastructure is what is at the heart of the 'Internet of Things'. Even if municipalities are in control of the data, and the techniques to analyze big datasets, they will never be able to do so without trusting the services of private companies, like KPN.

1.1 Data gathering and analysis within municipalities

Municipalities are encouraged to use several levels of data gathering and analysis as is shown in figure 1.[22] The report of the Dutch Society for Municipalities stimulates the transition of municipalities to become data-driven organizations. Municipalities should use these different levels of data gathering and analysis. However, each of these practices has some ethical concerns as well, that are underexposed and will be discussed accordingly.

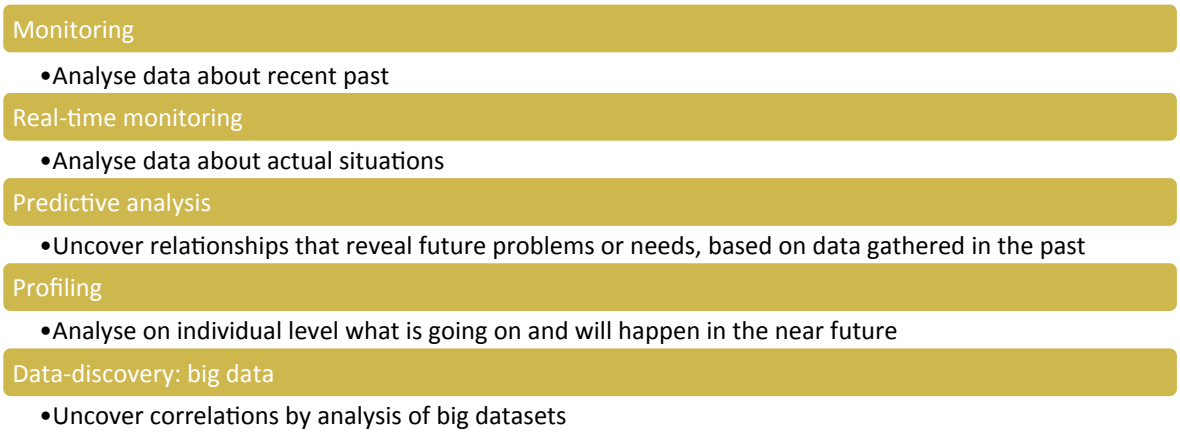

Figure 1

[20] Baloo, J., Brands, E., Steels, F., 'KPN Technology Book: The technology trends KPN has on its radar' (2018) p. 22.
[21] Jean-Pierre Beunen, presentation BigDataGemeenten Scheveningen. Jean-Pierre Beunen is Programme Director Smart Cities KPN. He is also the author of the Whitepaper: 'Smart Cities – A changing world. New technologies and societal paradigms'.
[22] Wesseling, et al. 'Datagedreven sturing bij gemeenten', p. 7.

Monitoring affects the freedom of choice, because people that are being watched, might feel that they have to live up to certain standards, or are required to deliver a certain amount of data, to be able to receive services. Predictive analysis is being practiced on the basis of *patterns of behavior* that are compared to peoples 'data' and the selectors that are chosen. This is closely related to data-discovery by using big data, because they are both concerned with finding, and analyzing correlations. Profiling is a matter closely related as well, but should be discussed from a different angle, because it has its own problems and possible worries towards freedom. These practices are not forbidden, or wrong, in essence, but the ethical considerations are most of the times diminished to discussions about privacy. It is obvious, that privacy, in the sense that several parties know everything about an individual, is a main concern. However, the practices of monitoring, predictive analysis, profiling and data-discovery have more freedom-limiting concerns as well, like freedom of choice.

1.2 Public versus private

In a society that is highly data-driven and datafied, the distinction between private and public gets blurred. People use many different technologies to communicate that cross the border between private and public constantly. 'People sit in their homes, connected to a public network, communicating with private friends, using public wires, exchanging private information, stored on public servers.'[23] In the analog world it was possible to make easily discernible distinctions between private and public. In an information society based on interconnectivity, this seems impossible. People are unaware of this blurring, because in the analog world they would store their data on computers. Nowadays, personal data is stored on mobile devices that people take with them into the public world, outside their houses.[24] As the distinction between public and private becomes blurry, individuals lose their sense of where their data is being stored. Besides their public data might be available for private companies, whom can benefit from gaining these data without the individual knowing, thereby affecting their freedom to shape their lives as they want it to be. Their reality gets framed with choices that are sorted for specific, commercial, purposes.

The transition towards data-driven societies actively promotes a neoliberal political economy and the marketization of public services, wherein city services are used for private profit.

[23] Mai, 'Big Data Privacy', p. 196.
[24] Koops, B.J., 'Digitaal Huisrecht', *Nederlands Juristenblad* (2017) vol. 92, p. 185.

Private companies have a lot to gain with the digitalization of society. They are huge players in the transition.[25] At this moment, the most powerful proponents of these technological advancements and a data-driven society are the stakeholders that have a lot to gain from this transition.[26] Especially for these reasons, it should be the municipality that is in control of the goals that are formulated. Striving towards a technologically advanced and data-driven society is not a bad thing, as long as it puts the citizen at the center of advancements. Moreover, there is a danger that dependency on corporate companies cannot be undone easily.[27] This is risky because it affects the power relations between the private and the public sector. These coalitions between public and private parties are more often than not, arranged without democratic election, draining money from European funds.[28]

1.3 Algorithmic Citizenship

According to Lyon and Stalder, it has been one of the central concerns for modern nation states, to establish stable identities of citizens. They argue that this is a key component for the connection between citizen and state. It is needed to classify individuals in their context flexibly, to determine which administrative procedure to apply,[29] for example to supply loans on rightful grounds. The efficiency of a system that recognizes people on the basis of a personal identifier over different databases, are obvious. The threat of living in a surveillance state has been discussed by many others. However, there are more civic liberties that are on the line, which I will analyze as human freedoms that are limited because of the datafication of society.

There is a tendency to gather as much data as possible, to make risk calculations, and to determine whether someone is the rightful recipient of a loan, or other public good. The worry that rises from risk calculations by algorithms is that people's citizenship becomes 'fluid', or as John Cheney-Lippold explains: citizenship becomes 'jus algoritmi'. The main point that he makes is that one's identity is constantly shifting and as an individual you are not aware of

[25] Vanolo, 'Smartmentality', p. 891.
[26] Andrejevic 'The Big Data Divide', p. 1677.
[27] Kitchin, R., 'The Data Revolution, Big Data, Open Data, Data Infrastructures & Their Consequences', (Sage Publications, 2014) p. 181.
[28] Vanolo, 'Smartmentality', p. 891.
[29] Stalder, F., Lyon, D., 'Electronic identity cards and social classification', in: *Surveillance As Social Sorting: Privacy, Risk and Automated Discrimination'* ed. Lyon, D. (2003) p. 77.

this process.[30] Historically one would be a citizen because one was born in a certain place or city, or because one would register in another municipality. The algorithmic citizen is an identification of citizen-as-status according to data and a transcoding of existing legal limitations, not an identity that one can consciously refer or respond to, because its allocation is both hidden from the individual and changing at every moment.[31] Algorithms present results in data. If algorithms define us, we are no longer defined as human beings made up of atoms, we are who we are in terms of data, another aspect of datafication.[32]

Furthermore, there are social groups that are vulnerable to be disadvantaged. This might be because of their gender, age, socioeconomic status, education level, mode of employment, geographical location or ethnicity.[33] This 'digital social inequality' can be based on one of the determinants above, depending on which of those determinants are chosen.[34] For example, it depends on the determinants that are chosen whether an individual is perceived as trustworthy. This happens because of the flexible assemblage of data that algorithmically becomes one's makeshift index for citizenship.[35] There is no technological system that is hundred percent secure. So even if we would suppose that algorithms are right 99% of the times, the algorithm could be wrong for 1%. On the basis of 10.000 civilians, which is a low estimate of the average level of inhabitants within municipalities,[36] the algorithm will determine the trustworthiness of an individual wrongfully in a 100 different cases. Individuals are being put into categories that they do not belong to. This limits their freedom of choice, because they are not aware of these wrongful categorizations and these categorizations in their turn determine which options individuals have. This is not something a municipality should take for granted, especially when these wrongful characterizations apply to people belonging to vulnerable groups. They become disadvantaged even more, because most of the

[30] Cheney-Lippold, '*Jus Algoritmi*: How the National Security Agency Remade Citizenship', *International Journal of Communication* (2016) vol. 10, p. 1737.
[31] Cheney-Lippold, '*Jus Algoritmi',* p. 1737.
[32] Cheney-Lippold, J., 'We Are Data', (New York University Press, 2017) p. 11.
[33] Lupton, D., 'Digital Risk Society', *The Routledge Handbook of Risk Studies* (2016) p. 7.
[34] Ibid.
[35] Cheney-Lippold, '*Jus Algoritmi',* p. 1737.
[36] In this example, I used 10.000 persons to clarify how many people would be affected if there were only an error margin of 1%. The more accurate estimate of the average municipality within the Netherlands is somewhere between 35.000 – 40.000 civilians. https://www.cbs.nl/nl-nl/nieuws/2017/52/aantal-gemeenten-van-388-naar-380

times, they lack the resources or capabilities to object. This point will be discussed in more detail in chapter 3.

I elaborate on the theory of algorithmic citizens in chapter 3, and argue that this algorithmic citizenship affects freedom of choice and potentially discriminates individuals belonging to certain social groups. Furthermore, I will argue that this fluid citizenship and the role it plays within the practices of predictive analysis, monitoring and profiling, infringes upon the freedom of choice. Besides, I will elaborate on the problems that Pasquale mentions, when he describes highly data-driven societies as 'black-box societies'. This metaphor refers to the data-monitoring systems in society and to a system of which the workings, like algorithms, are mysterious. We are able to observe the input and the output, but we cannot tell why the one results in the other.[37] As I will show, by introducing Amartya Sen's accounts of human freedom, this limits the choices, and the abilities that individuals have to shape their own destiny.

[37] Pasquale, F., 'Black Box Society: The Secret Algorithms That Control Money and Information' (Harvard University Press, 2015) p. 3.

2. Freedom by Sen

Most of the times, data practices are developed with the best of intentions, and the best service for citizens in mind. However, the social impact is hard to foresee, and more often than not, it infringes upon the freedom of individuals. To understand what it is we talk about when we discuss the concept of freedom, I introduce the theory of Amartya Sen as he described it in his work *Development As Freedom.* Development and freedom are connected to each other, because societies are always changing and aiming for progress.

> Development can be seen (…) as a process of expanding the real freedoms that people enjoy. Focusing on human freedoms contrasts with narrower views of development, such as identifying development with the growth of gross national product, or with the rise in personal incomes, or with industrialization, or with technological advance, or with social modernization.[38]

Important in aiming for progress, is to secure the *real* freedoms that people enjoy. It seems apparent that efficiency and economic growth enhance the real freedoms that people enjoy, for example, the rise of personal income might contribute to strengthen the real freedoms people have. However, those claims should be tested because efficiency and economic growth also have the threat of interfering with the real freedoms when there is a gain in efficiency, but a downgrade of the actual choices that an individual has. Sen's ideas about freedom are especially relevant in this context because he distinguishes a broader view of freedom from several narrower views of freedom that mainly focus on things like economic progress, social modernization and technological advancement. These narrower views of freedom can contribute to the development of human freedoms. However, Sen claims, it is important to focus on the ends, and not solely on the means.[39] To make it more specific, human freedoms should be the goal of a municipality, not digitalization in itself, or technological advancement. This is relevant, because what we see in the development of data-driven societies is that the focus is mainly on the technology, instead of on the citizens.[40]

---

[38] Sen, A., 'Development As Freedom' (Oxford University Press, 2002) p. 3.
[39] Ibid.
[40] Dameri, R.P., 'Searching for Smart City definition: a comprehensive proposal', *International Journal of Computers & Technology*, vol. 11, no. 5 (2013), p. 2545.

Real freedoms enhance the actual livings that people manage to achieve. 'The *freedom* to achieve actual livings that one can have reason to value.'[41] According to Sen we should analyze freedom in the form of individual capabilities to do things a person has reason to value.[42] Sen argues that human freedom is at the basis of development, for two main reasons:

1. *The evaluative reason*: assessment of progress has to be done primarily in terms of whether the freedoms that people have are enhanced;
2. *The effectiveness reason*: achievement of development is thoroughly dependent on the free agency of people.[43]

We need these human freedoms not solely as a goal on which we can focus. It is more than a goal in itself. It is the basis for development as well. People are able to participate in social arrangements when their free agency is enhanced. Therefore, we need these human freedoms.[44]

Furthermore, we should not only look at the functionings that people are actually pursuing, but also on the capabilities a person has, but not pursues even though they are available to him.[45] Sen uses the example of someone who chooses to fast. This fasting only gains significance as long as there is an option to eat as well. This is relevant, because choosing to fast is not the same as being forced to starve.[46] In the next chapter, I will argue that because of the processes of datafication, and mainly through practices of predictive analytics, the functioning of choosing is on the line.

According to Sen, people should have reasons to value their lives. We should concentrate on the human freedoms that people need, to achieve an actual living they can value.[47] If we want to develop, the goals should be formulated in line with the human freedoms. It requires the removal of unfreedom, like poverty, poor economic opportunities, systematic social

[41] Sen, 'Development As Freedom', p. 73.
[42] Ibid., p. 56.
[43] Ibid., p. 4.
[44] Ibid., p. 53.
[45] Ibid., p. 75.
[46] Ibid., p. 76.
[47] Ibid., p. 73.

deprivation, and neglect of public facilities as well as intolerance.[48] Sen argues that individuals in possession of adequate social opportunities should be able to effectively shape their own destiny and help each other.[49] Therefore, it is relevant, to discuss whether the datafication of society, and especially within municipalities, contests these social opportunities.

To analyze the freedom-limiting problems that arise because of datafication, I will use two of the five forms of instrumental freedoms that Sen describes. They are instrumental because they concern the way different kinds of rights and opportunities contribute to the expansion of human freedom, not simply as an end, but as means as well.[50] The five freedoms that he discusses are political freedom, economic facilities, social opportunities, transparency guarantees and protective security.[51] Although these freedoms strengthen each other and secure the enhancing of human freedoms, I will use the human freedoms of *social opportunities* and *transparency guarantees* to analyze the freedom-limiting concerns that I have mentioned in the previous chapter. *Political freedoms* and *economic facilities* do enhance the other freedoms, but are mainly focused on the opportunities that individuals should have to determine who should govern and on what principles. Other aspects are freedom of press and the ability to criticize authorities. Economic facilities are the opportunities for individuals to enjoy and to utilize economic resources.[52]

More interesting and applicable to the process of datafication are the *social opportunities*. Social opportunities refer to the arrangements that society create for education, health care and other public services. These services are not only important for the conduct of private life, but also for effective participation within society.[53] It is important to include the set of opportunities that individuals have, also the opportunities that an individual chooses not to pursue, but which are available nonetheless. Datafication has the potential to harm the free agency of the individual through the practices of predictive analysis and monitoring, as I will show in the next chapter.

---

[48] Sen, 'Development as Freedom', p. 3.
[49] Ibid., p. 11.
[50] Ibid., p. 37.
[51] Ibid., p. 38.
[52] Ibid., p. 38, 39.
[53] Ibid, p. 39.

The fourth category of freedom is highly relevant, because, according to Sen, people interact on the basis of some presumption of what they are being offered and what they can expect to get. Therefore, individuals need *transparency guarantees*, to support the underlying basic presumption of trust.[54] Because of the processes of datafication, and the ways governments go about their data management, this human freedom is highly contested.

The last form of freedom is *protective security*. Even when society is doing great, there are people who are vulnerable and need a social safety net.[55] This last form of human freedoms is related to the virtue of solidarity, which is on the line because of datafication as well. However, solidarity is on the line as a possible threat to social opportunities as well, so I will discuss this argument in chapter 3.

How datafication affects these different forms of human freedoms will become apparent in the following sections, where I will refer to these instrumental freedoms, to make clear which one is challenged through the practices of data-driven societies. Moreover, I argue that municipalities are responsible for the protection of these human freedoms, and should try to actively take away possible unfreedoms, like discrimination. The human freedoms, and the rights, opportunities and entitlements belonging to these freedoms, are interconnected.[56] Therefore, it is difficult to separate the possible dangers that arise because of datafication. However, most of the concerns have a main focus. I will analyze the problems in the way they affect social opportunities or transparency guarantees separately, before discussing the interconnections between the two of them.

[54] Sen, 'Development as freedom', p. 39.
[55] Ibid., p. 40.
[56] Ibid., p. 53.

3. Social opportunities contested in an age of datafication

The social opportunities, which every individual should be entitled to, are contested because of the processes of datafication. These opportunities do not only refer to the actual functionings that individuals choose to pursue, they refer to the choices that are available to individuals, whether or not the individual pursues them. Furthermore, social opportunities are important for private life, and for effective participation within society.

3.1 Public goods for sale

One of the things that are problematic is the fact that different prices are attached to services, or to access public goods. Governments have a monopoly in the public services they provide. Therefore, citizens cannot choose where they will go to receive the public services they need.[57] This is one of the concerns that are on the line when private companies have a great involvement in the construction of smart cities. Public goods that were free, and should be available for free, might become for sale.[58] For example, the municipality provides parking spaces within cities. Companies like KPN are interested to design applications that citizens can use, to find available parking spots. However, they will only do so, if there is a financial gain for them.[59] In the end, this infringes upon the social opportunities that individuals have. Social opportunities refer to the arrangements that society creates for public services. When these services, that used to be for free, are only available for money, this affects the actual choices that individuals have.

One of the virtues that is on the line is the virtue of solidarity. Criticasters, like Rob Wijnberg, claim that it is not simply a vast amount of information that creates truth or worth, it depends on the way we select and analyze information, that truth or worth are guaranteed. Wijnberg claims that solidarity is one of the foundations of the democratic society. Because of solidarity, we are able to divide costs and risks equally between different group members.[60] A democratic society achieves this through practicing human freedoms of social participation, in the form of education, which facilitates economic participation. Economic facilities can help

[57] 'Maak Waar!', p. 45.
[58] 'De muren hebben sensoren', *De Groene Amsterdammer* (December 6, 2017) https://www.groene.nl/artikel/de-muren-hebben-sensoren visited June 14, 2018.
[59] Ibid.
[60] Martijn, M., Tokmetzis, D., 'Je hebt wél iets te verbergen, over het levensbelang van privacy' (De Correspondent, 2018) p. 163.

to generate public resources for social facilities,[61] like insurances. Wijnberg describes an example about these insurance companies. Amongst many other companies, insurance companies require more data from their clients than before, for instance data about their health. This way they can measure exactly what one individual needs. To do this more effectively, they use these data to make profiles of their clients and compare those profiles to others. However, it is way more likely that a healthy person would share his data through Fitbit for example, than an individual that is not as healthy, because of diabetes. The norm will be based on the data of the healthy persons who share their data voluntarily. Therefore, the people that are already disadvantaged become disadvantaged even more.[62] In a way, the people that are disadvantaged are criticized as well, because they have more trouble to stay healthy.[63] Not only solidarity is on the line, it influences the ways we make our choices as well. In the long run, it might become a default to use Fitbit, because you should have a good reason not wanting to use Fitbit. As Cathy O'Neill underlines, in the long run it is only the people who do not need any insurance, who can afford one.[64] This signals that participation within society is limited for the ones who are not able to keep up with new trends or the ones who simply lack the capabilities, or the money, to access public goods.

3.2 Defining success

Furthermore, we have to question the way we are 'defined', as was discussed before by introducing 'jus algoritmi'. There is more to this, which affects the social opportunities of the individual as well. Before an individual can be classified as 'trustworthy' or the opposite, there have to be valuations and qualifications to assess whether one is either this or that. These valuations are subjective, but hardly recognizable. Each of these classifications is an algorithmic caricature of the category it represents.[65] This is what Cathy O'Neil refers to as well, when she claims that one important factor of designing an algorithm is the definition of success that is being put into the algorithm.[66] Through defining success, goals are formulated that individuals have to live up to. This influences the social opportunities that citizens have, for what is important in life, is defined by satisfying the algorithm. This infringes upon the

[61] Sen, 'Development as Freedom', p. 11.
[62] Martijn, Tokmetzis, 'Je hebt wél iets te verbergen', p. 163.
[63] Ibid.
[64] Weijer, B., van der, 'Wiskundige Cathy O'Neil waarschuwt voor algoritmen: 'Rechten van individu worden niet beschermd', *de Volkskrant* (June, 18, 2018) https://www.volkskrant.nl/wetenschap/wiskundige-cathy-o-neil-waarschuwt-voor-algoritmen-rechten-van-individu-worden-niet-beschermd-~b97a9302/ visited June, 19, 2018.
[65] Cheney-Lippold, 'We Are Data', p. 10.
[66] O'Neil, C., 'The era of blind faith in big data must end.'

freedom individuals should have to shape their own destiny. When one chooses not to satisfy the algorithm, this should not come at a cost of being sanctioned by being denied a certain loan or access to a public service. Another troublesome feature of these algorithmic decision making processes is when an individual does not satisfy the algorithm, because he is simply not aware which qualifications are determined as successful. This is a matter that influences the transparency guarantees as well, and will therefore be discussed in the next chapter.

3.3 Predictive analysis and self-disciplining

Predictive analytics are used to assess likely future behavior, and to direct actions from these predictions accordingly.[67] Because of this profiling, and the likely percentages that a person is rather trustworthy or not, the person is treated differently, in anticipation of being something this person may or may not be or do.[68] There is a power imbalance between those positioned to make decisions that affect the life chances of individuals, and those subjected to the sorting process. The decisions that are made about these subjects are not really based on an assessment of who or what these people are, but on what they will do in the future.[69] One of the important aspects by relying on predictive analysis is that it seeks knowledge in correlation and prediction, instead of explanation and causation.[70] For example, crime rates do not always tell the complete truth. When municipalities use predictive analysis it becomes an even more difficult matter. In 2016, the municipality of Apeldoorn announced that they were going to predict who would become a criminal. They did not only use crime rates of burglary and theft, they also used sources like schools, family composition, and debt control.[71] According to Kitchin, this results in anticipatory governance, which is doubtful, because it targets attention at particular groups and places, and seeks to prevent behavior that may never occur. In this process, it reshapes how people act through self-disciplining. Moreover, it is often that this predictive analysis goes hand in hand with prejudice and discrimination and creates self-fulfilling prophecies.[72] This reshaping of the self-discipline is worrisome, because this affects the individual in the freedom of the choices they can make, it puts forward goals

[67] Kitchin, 'The Data Revolution', p. 181.
[68] Ibid.
[69] Andrejevic, 'The Big Data Divide', p. 1677.
[70] Ibid.. p. 1679.
[71] 'Apeldoorn voorspelt wie het criminele pad opgaat', *De Stentor* (May, 11, 2016) https://www.destentor.nl/nieuws/apeldoorn-voorspelt-wie-het-criminele-pad-opgaat~a96cdd5d/
[72] Kitchin, 'The Data Revolution', p. 182.

that a 'smart' citizen should strive for. This limits the actual choices that individuals can pursue as well. Being forced to satisfy the algorithm because of possible sanctions otherwise is not the same as deliberately choosing the right thing to do, or to shape your life the way you have reasons to value it.

3.4 Monitoring, scoring individuals and measurable averages

> 'Smartness is becoming a field of social control that makes intrusion in a person's private life quite natural; as a result, we need to pay attention to the goals established in the framework and the label given to smart city projects.'[73]

'Smartness' is a virtue that is highly encouraged in many municipalities in the Netherlands. This smartness refers to individuals who are able to use smart devices that communicate with the municipality. However, one of the side effects is that a lot of personal data is transmitted in the communication as well. Not only do individuals voluntarily share their data, James B. Rule refers to the 'bureaucratic tracking of individual lives' which is seen as social control as well.[74] This is possible because of all the surveillance and monitoring techniques that are present in society at this moment, to gather data. Social control does not necessarily refer to the application of force. As was mentioned before, surveillance is not a main topic in this thesis. However, the ways in which municipalities practice social control is relevant to discuss. James B. Rule refers to social control as the array of processes that encourage the 'right' kinds of actions, and discourage their opposites. These processes include procedures of identifying and rewarding those who have played by the rules in the past, and sanctioning the ones who are prone to cause future trouble.[75] To make these predictions, individuals have to be monitored and their data has to be gathered and stored, which explains the thirst of public and private organizations for all of our personal data.[76]

Because of this monitoring, and the profiles that are created about individuals, governors have more power over their citizens, and they can nudge people to behave in certain ways. There are worries for manipulation by governments that want to influence the lifestyles of their citizens. One extreme example is the current development of the Chinese social credit system.

---

[73] Vanolo, 'Smartmentality', p. 894.
[74] Rule, J.B., 'Privacy, the longue *durée*', in: Roessler, D.B., '*Social Dimensions of Privacy'* (2015) p. 13.
[75] Rule, 'Privacy', p. 13.
[76] Ibid.

We look at the Chinese Social Credit System, and condemn the practices of the Chinese government. They are implementing a scoring system that measures the credibility and determines what people are allowed to do, on the basis of where they rank in society.[77] People who are critical about the Chinese government and post about it online, are being categorized as 'untrustworthy', and face punishments like a ban to fly or denial of buying property.[78] Ultimately, it will only be the obedient people who can participate within society. Why is it questionable that people can no longer behave in criminal ways? According to Morozov, individuals need the freedom to deliberate about the choices that they want to make. He even worries that people will end up with a defective morality when these laws would be made universal.[79] In line with Sen, I would argue that these 'right' choices that are embedded in the system limit the actual functioning of choosing by individuals, because individuals should be able to deliberately choose to live a life that they have reasons to value. Not because they have to satisfy the system of which they sometimes do not even know the workings.

When everything is measured and datafied, every action that is not in line with regulations becomes apparent as well. When the rightness of an action is always embedded within the system, civilians are not tempted to think about their actions anymore. On the one hand, this results in a highly efficient and fully automated society. Individuals do not have to think about the social consequences of their actions, or whether there are better solutions.[80] On the other hand, when systems are not fully automated and individuals have a free choice, this gives them the incentives to think about their actions. They can weigh the values that they believe are important, to them and to society. This results in individuals that have an open and critical view towards their environment and society as a whole. Although efficiency is profitable on the short term, in the long run a democratic society benefits from civilians that are aware of the consequences of the choices they make.[81] This depends on the freedom they have for deliberation about the choices they want to make and it affects the actual set of choices that individuals have available to them.

---

[77] Johnson, S., 'China's 'social credit score' will punish and reward citizens', *Big Think, Smarter Faster* (April, 25, 2018) http://bigthink.com/stephen-johnson/a-look-at-chinas-orwellian-plan-to-give-every-citizen-a-social-credit-score visited 14-06-2018.
[78] Ibid.
[79] Morozov, E., 'Om de wereld te redden, klik hier' (Uitgeverij De Wereld, 2014) p. 338.
[80] Morozov, 'Om de wereld te redden, klik hier', p. 339.
[81] Ibid.

Although the Chinese government has different incentives than our government, they do use the same logic: they give us scores, to predict our future behavior, without us knowing. In China, this system is used to spread the socialist ideology, in the West, these systems are used to prevent non-payment or terroristic attacks.[82] Every action and every lifestyle is diminished to a measurable average, on which people can be judged and sanctioned, and which makes it very visible who the ones are that frustrate progress. This can be used in various directions, depending on the goals that governors formulate, and how they formulate what a successful result is. This affects possible manipulation of lifestyles by the municipality, because people can be nudged to behave in certain ways, in the most extreme case, making way for a police state.[83] It is not impossible that governors in the future can demand of their citizens that they share their data, and use it e.g. for purposes of measuring how sustainable individuals are or how healthy their lifestyle is. The actual choices that individuals have are limited, even when a person is clearly healthy, but does not have the data to back this up. If an algorithm needs the data of a healthy person, and only defines one as healthy when it is communicated that the individual has a daily step count of twelve thousand steps, it affects the lifestyles of individuals because they have to satisfy the algorithm.

It seems obvious that municipalities or governments should aim for monitoring everything in society, to become an efficient organization. However, as I have tried to show in this part, we have to separate the narrow view of progress from the broad view. The narrow view aims at efficiency and control. It seems like progress, but a great deal of freedom of choice is contested and therefore it is questionable, to which extent we should encourage this progress. As Sen argues, individuals need to have the possession of adequate social opportunities to effectively shape their own destiny. These opportunities are lessened when governors only encourage the 'right' choice to satisfy the algorithm.

3.5 Profiling, monitoring and social sorting

Different types of data sets and digital data can be joined together to make risk calculations. Algorithms make these calculations, however, they do not make decisions based on the behavior of individuals. Algorithms contain statements about *average* characteristics and patterns of behavior. At best, they contain information about individuals as part of the random

[82] Martijn, Tokmetzis, 'Je hebt wél iets te verbergen', p. 156.
[83] Vanolo, 'Smartmentality', p. 894.

group, but not about individuals in their own right.[84] Algorithms search for patterns and correlations within large datasets. By doing so, they group together individuals based upon commonalities that might not be easily definable, nor in reality easily recognizable, because of their apparently random nature.[85] They find correlations that might not have any causation,[86] what they uncover are relationships, rather than direct causal connections.[87] Another concern with profiling is that profiling in itself is problematic, because profiles are being constructed through analyzing and comparing the personal data of others.[88] Based on data points of the past, who we are as data becomes an enunciation of another's algorithmic interpretation.[89]

Schauer claims that profiling in itself does not have to be a bad practice. It is even effective to a large extent. We use it in our everyday life to evaluate situations. However, profiling tends to go in the wrong direction when profiling is done according to generalizations based on race, gender or sexual orientation.[90] When people are discriminated because of the profiles that are made about them, we grow concerned. The GDPR states that every 'data subject has the right not to be subject to a decision based solely on automated processing, including profiling.'[91] However, when there is a human in the loop of the decision-making, should this suffice to fulfill this right? Article 22 of the GDPR says something more about this. According to this additional paragraph, this right should not apply when profiling, or the automated decision-making processes, are authorized by governors that are able to safeguard the data subject's rights and freedoms.[92] Therefore, we can argue that profiling in itself is not a bad practice. However, when municipalities occupy themselves with these practices, they should take care of the responsibility to protect human freedoms as well. Human intervention is necessary, at least at the level of safeguarding the rights of the subjects that are affected.

---

[84] Naudts, L., Vedder, A., 'Accountability for the Use of Algorithms in a Big Data Environment', *International Review of Law, Computers & Technology* (2017) p. 5.
[85] Ibid., p. 4.
[86] Wesseling, et al., 'Datagedreven sturing bij gemeenten', p. 18.
[87] Lupton, 'Digital Risk Society', p. 4.
[88] Martijn, Tokmetzis, 'Je hebt wél iets te verbergen', p. 163.
[89] Cheney-Lippold, 'We Are Data', p. 117.
[90] Schauer, F.F., 'Profiles, Probabilities, and Stereotypes' (Harvard University Press, 2003) p. 22.
[91] Art. 22.1, GDPR.
[92] Art. 22.2, GDPR.

As was argued before, the collection of personal data does not have to be problematic in itself. It does not infringe upon the autonomy of individuals, although the processing of this data does have this ability. It infringes upon the freedom and choice individuals have, to profile their selves, as they want. One of the problems with profiling that was not mentioned before, is that people are mostly unaware of the fact that they are being profiled.[93] Another worry is the threat of discriminating treatments because of this profiling,[94] as will be discussed in section 4.4. The *burden of proof* shifts towards the citizen, whenever there is erroneous analysis of data. When people are wrongly associated with a profile, they have to defend themselves and proof that analysis was wrongful.[95] These concerns infringe upon several human freedoms. Vulnerable people who need a safety net rely on the trust that they are characterized rightfully and correct. When profiling goes wrong, this undermines their protective security for they might miss out on services provided by a safety net, like payments, especially because they might not be aware of the wrongful profiles that apply to them.

3.6 Data-discovery: Mistakes and the burden of proof in the digital maze

On the basis of larger datasets, governments can make decisions that are more effective and efficient. That is, at least, the theory. Besides the lack of oversight, transparency, and the possibility for citizens to obtain information about which personal data of them is gathered and analyzed, data can be erroneous.[96] For example, the municipality of Amsterdam found out that at least 7.3% of the basic information that they have about their citizens' addresses, contained mistakes. It is not clear whether these were only 'honest' mistakes, or whether they stem from fraud.[97]

Governmental organizations work together, and share information on a broader level than ever before. There are dozens of databases that have the incentive to solve or reduce societal problems. Data is not stored in these databases, because these databases are connected to each other and information flows from one database to another.[98] The digital government is not an

[93] Hildebrandt, M., 'Profiling and the rule of law', *Identity in the Information Society* (2008) vol. 1, no. 1, p. 63.
[94] Taylor, L.E.M., et al., 'Public Sector Data Ethics: From principles to practice', *Tilburg University* (2017) p. 9.
[95] Broeders, Schrijvers, Hirsch Ballin, 'Big Data', p. 73.
[96] Martijn, Tokmetzis, 'Je hebt wél iets te verbergen', p. 88.
[97] Prins, C., et al., 'iGovernment, Digitizing the Citizen and Government', *WRR-report no. 86.* (2011) p. 25.
[98] Martijn, Tokmetzis, 'Je hebt wél iets te verbergen', p. 80.

institution that has several databases, it has multiple rapidly diversifying information flows.[99] There is a large flow of information, but the central overview lacks most of the times. This is not a government that watches over us like Big Brother. This is a government that has the best incentives, but fails to execute effective governance, because they lost control in guiding the data flows of their citizens.[100] People are quite ignorant about the way this affects the burden of proof, in a negative way. If mistakes are made, who can they turn to, to rectify the matters?[101] Not only individuals are unsure, organizations themselves are unsure as well.

3.7 Digital Maze

In the basic registration files of the municipality, it is possible that complete strangers are attached to each other. The algorithm thinks these people live together, because of mistakes in the basic registrations files that are under control of the municipality. Therefore, individuals lose their right to subsidies, which can result in missing out on a monthly fee of 700 euros.[102] There are many examples of individuals that get lost in this digital maze. Mistakes are corrected in one system, but as long as that system is connected to other public organizations' databases, the same mistake keeps popping back up. One example, that was broadcasted by 'De Monitor', which is a Dutch television program, makes it very clear how individuals are disadvantaged because of this datafication, and the reliance on algorithmic decision-making. Edith takes care of all the financial bookkeeping of her aunt, who suffers from dementia and is 94 years of age. When her aunt moves from one nursery home to another, her files are connected to the files of somebody else, who moves into the same nursery home on that specific day as well. Edith spends months trying to correct this mistake, because this mistake affects the amount of money her aunt receives monthly. The public organizations that are responsible for this mistake, point to each other to solve this issue.[103] This might have several reasons. Individuals get caught up in the diversifying flows of data that they cannot oversee, but the question is whether the organizations themselves are able to handle this. More often than not, there is a disappointing answer. Public servants claim they lost control over the system. At first, this looks like a technological problem, because the technology should be

[99] Prins, et al., 'iGovernment', p. 13.
[100] Martijn, Tokmetzis, 'Je hebt wél iets te verbergen', p. 91.
[101] Prins, et al., 'iGovernment', p. 25.
[102] 'Digitaal Doolhof', *De Monitor, KRO-NCRV* https://demonitor.kro-ncrv.nl/onderzoeken/digitaal-doolhof
[103] 'Digitaal Doolhof', *De Monitor*.

able to rectify a mistake. However, if we look at the deeper concerns, we find that this is a societal problem at the core as well. Goals should not be set to achieve technological or efficiency progress. Goals should be set to develop human freedoms. If there is such a large amount of data flowing from one database to another database, meanwhile affecting and changing the personal data of citizens, we have to question ourselves: who is the one who can stop these processes when things go wrong, and rectify the matters?

The trouble begins with mistakes that are being added in the basic registration files, which happens at the municipality and can happen without any bad intentions. It is not only the office of the tax authorities that digs into these databases, there are over 500 organizations that use the data that is gathered in these files. It can be a seemingly small mistake that in the attachment of all these different databases causes civilians heavy problems. Part of this problem is that individuals are not aware of this, until things go wrong, or when they are denied access to public services. Sometimes it takes years before mistakes are uncovered.[104] Even with the best of intentions in designing algorithms, they can go wrong, and go on for a long time before anyone notices.[105]

It is not satisfactory that the Secretary of State admits that mistakes are being made, due to the technologies that are used, but that 95% of the times, everything is fine. People that are insufficient self-relying, depend on the government, or their municipality. It should not be the other way around.[106] Another problem is that people, who would perfectly manage things on their own, now fall behind, because there is so much information available, and needed for getting access to public services.[107] Therefore, this is not just a problem of digitalization. It is a problem of datafication as well. People do not only fall behind, or have less access to public services, just because they do not have access to digital technologies. They fall behind because there is a lot of data needed. Institutions want to be as complete as possible. Therefore, they require a lot of information of their citizens, which is not always needed. However, this can raise the threshold for people to access public goods.[108]

---

[104] 'Digitaal Doolhof', *De Monitor*.
[105] O'Neil, C., 'The era of blind faith in big data must end', *TED, Ideas worth spreading* (2017) https://www.ted.com/talks/cathy_o_neil_the_era_of_blind_faith_in_big_data_must_end
[106] Zuthpen, van, R., 'Wie niet past, loopt vast: verslag van de Nationale Ombudsman 2016', *De Nationale Ombudsman* (2017) p. 10.
[107] Ibid.
[108] Ibid., p. 12.

4. Transparency guarantees

In the previous chapter I have elaborated on the human freedom individuals should have the right to, described as social opportunities. Individuals need another freedom, transparency guarantees, because people interact on the basis of some presumption of what they are being offered and what they can expect to get. There are several developments because of datafication that threaten this right for the individual that is affected by these processes.

4.1 Black Boxes

Most of the big data-systems that are used by governmental institutions are black boxes. No one really understands how algorithms go from input, to output. Moreover, they remain within these black boxes because of corporate secrecy. Algorithms are constructed by private companies, which hide themselves behind this secrecy. Put to extremes: the government has an algorithmic heart, which cannot be checked by the government itself.[109] Powerful actors deploy strategies of obfuscation and secrecy to consolidate their power and wealth.[110] The implementation of the new General Data Protection Regulation (GDPR) puts some constraints on these strategies of obfuscation and secrecy. Thanks to these new regulations, individuals have to right to ask for clarification about the ways their data is being used.[111] The effectiveness of these regulations is however unclear at this moment, because these regulations are only effective from May 2018 onwards.

4.2 Power Relations private versus public

KPN is one of the main private companies that contribute in the transformation of municipalities towards data-driven organizations. They work closely together with municipalities, mainly by designing and facilitating the infrastructures for data. A data-driven society can only operate when trusted exchange is possible through digital networks.[112] According to KPN, cities have an extra layer on top of the infrastructure that is visible. This is what they call a 'data-layer'. This infrastructure connects all the smart devices that people

[109] This statement is part of an interview that the authors of 'Je hebt wél iets te verbergen', had with Dennis Broeders, one of the authors of many reports of the Netherlands Scientific Council for Government Policy, on the effect of Big Data for governance,. 'Je hebt wél iets te verbergen', p. 113.
[110] Pasqaule, 'The Black Box Society', p. 14.
[111] Art. 13, 14, General Data Protections Regulation.
[112] Baloo, J., Brands, E., Steels, F., 'KPN Technology Book: The technology trends KPN has on its radar' (2018) p. 22.

have in their homes, smart lampposts and mobile phones, among other things.[113] This infrastructure is what is at the heart of the 'Internet of Things'. Even if municipalities are in control of the data, and the techniques to analyze big datasets, they will never be able to do so without trusting the services of private companies, like KPN.

As long as municipalities are dependent on commercial parties, they have to stay in control and take the lead. For example, CityTec, which is owned by Eneco, did not want to share data of the public lights, like the traffic lights, parking installations and light poles. They did agree that the municipalities were the legal owners of the light poles. However, they claimed that CityTec was the commercial owner, and therefore they did not want to share the data of these poles because of economical competitive protection.[114] This is an example of private companies that hide themselves between corporate secrecy. This does not only inflict upon the right that individuals have to gain insight into the data that is gathered about them, it also impedes the relationship between public and private organizations.

Again, the GDPR has implemented some constraints that aid the individual in receiving more insight into the practices of data storing and usage by companies or institutions. Individuals have the right to this access, because they act on presumptions of what they are being offered. Private companies and institutions store data about individuals to combine it with other data sources. These practices have the potential to create new data. Jens Erik Mai describes a concept for how we should understand the ethical concerns that arise from these practices of data-discovery. The problem is not with the gathering of data, for most of the times, people voluntarily share their information. Violations of peoples trust, and when people grow concerned, happen when this data creates new data.[115] There is one very clear and commonly used example of this. In the States, there was a father who went to a Target store, complaining that his daughter had received coupons for maternity clothing and other baby products. After many complaints, it turned out that the company Target knew before anyone else did, that the daughter was pregnant, based on the history of what she had bought.[116] It is possible to

[113] Jean-Pierre Beunen, presentation BigDataGemeenten Scheveningen. Jean-Pierre Beunen is Programme Director Smart Cities KPN. He is also the author of the Whitepaper: 'Smart Cities – A changing world. New technologies and societal paradigms'.
[114] 'De muren hebben sensoren', *De Groene Amsterdammer*
[115] Mai, 'Big Data Privacy', p. 198.
[116] Ibid., p. 192.

question the way the store gathered the data, although the daughter only shared data because of the items she bought. We might grow more concerned about what happens with this data, and the actions that are undertaken on the basis of these data.

4.3 Predictive analytics and selecting the corresponding virtue or determinant

The trustworthiness of a person might be measured on the basis of late payments, or refusal of payments. However, there are other determinants used as well. For example, the neighborhood where one lives is valuated. On the basis of acts committed by other people, profiles are generated that affect the trustworthiness of other individuals. Correlations are used, that might not have any causation, but can affect the individual in a negative way. One of the risks when we make profiles about people is that we stigmatize entire groups. We use generalizations, on the basis of a characteristic of a minority of the group, to make decisions about the entire group.[117] This 'jus algoritmi', or algorithmic citizenship, that Cheney-Lippold talks about refers to a corruption of a social group, which reflects the steering mechanisms of bureaucratic power.[118] It fundamentally affects the citizen in his right to be treated like a human being, with all of its complexities, instead of an identity that is made up solely of data.

Not only is the identification of the algorithmic citizen constantly shifting, the individual is being put in different categories that he is not aware of. As was already noted before, the commonalities of random groups, depending on which determinant is chosen, are often not easily discernible. The groups can often only be identified by those who defined them for a specific purpose, and not by the individuals affected by the statements. The latter might even consider the definition of the group as 'arbitrarily chosen'.[119] Through data mining it is possible to find patterns, but their relevance depends on the questions they address, which depends on who is asking.[120] These opaque processes of algorithmic decision-making and defining citizenship are conflicting with the human freedom of transparency guarantees. People act on the basis of some presumptions that they are being offered, although meanwhile, they might belong to another group than they would expect, involving all the consequences.

---

[117] Schauer, 'Profiles, Probabilities, and Stereotypes', p. 4.
[118] Cheney-Lippold, J., '*Jus Algoritmi',* p. 1737.
[119] Naudts, L., Vedder, A., 'Accountability for the Use of Algorithms in a Big Data Environment', *International Review of Law, Computers & Technology* (2017) p. 5.
[120] Andrejevic, 'The Big Data Divide', p. 1679.

In this way, the whole paradigm of datafication affects the freedom of choice. Because of these incomprehensible processes of data analysis, and the unequal distribution of data knowledge, it affects the capabilities individuals have for using data, and to benefit from giving away one's own data.[121] Even when more parties receive access to data-mining techniques, the power will remain with the ones who have access to the latest technology and the largest databases.[122] This results in individuals who lack the access to these databases and the data that is relevant for them, to shape their own destiny.

Moreover, people need to know how the algorithms that make decisions about their lives go from input to output. There are several challenges for this. As was discussed before, transparency is needed, but companies are not willing to voluntarily share the way the algorithm was build, because of competitive strategies.[123] The ones in control of the algorithms worry that if the rules are known, the algorithms lose their predictive value, because people will be able to game the system.[124] Even when access is gained, it is still hard to deconstruct these algorithms. Getting access to a credit rating system might give insights to formula for assessing and sorting individuals, and how it works in practice. However, it will not necessarily provide full transparency regarding the choices that are being made during the construction of these algorithms.[125] If we have found answers to all of the questions addressed here, we still do not know if the algorithm that we investigate, is the same one as five minutes ago. Algorithms are trained to improve their selves and react to inputs.[126] Additionally, algorithms do not make things fair, if you just apply them. They repeat past practices and automate the status quo.[127] Therefore, there are implications that have to do with the issue of tailored reality as well. It is claimed that algorithms are not based on the current identity of persons, but on a previous version. It infers the interests of the individual on their behavior, and the collection of information, in the past. Because of this profiling, it may affect the choices individuals have. If they receive possible choices on the basis of information gathered

[121] Andrejevic, 'The Big Data Divide', p. 1679.
[122] Ibid., p. 1680.
[123] Kitchin, R., 'Thinking critically about and researching algorithms', *Information, Communication & Society* (2016) vol. 20, p. 20.
[124] Pasquale, 'Black Box Society', p. 35.
[125] Kitchin, 'Thinking critically', p. 21.
[126] Ibid.
[127] O'Neil, C., 'The era of blind faith in big data must end', *TED, Ideas worth spreading* (2017)

in the past, but that information is not relevant anymore, the choices that individuals receive encourage making decisions that maintain old views.[128] These are concerns that have to do with lack of transparency, but are relevant for the social opportunities of individuals as well, because actual choices are limited.

4.4 Discrimination of social groups

Data-driven societies are societies that are no longer formed top-down, but are based on interconnectivity. However, this idea of giving power to the people seems to benefit the ones who are already familiar with digital technologies and exclude the ones who have less access or abilities to use these tools. According to KPN, people become flexible and the data-driven city creates the possibility to effectively use all the data we already possess.[129] Although the focus in the transition towards data-driven cities should be with the citizen, most of the times the focus is with the technology.[130] Companies and governors formulate goals that are more in line with the possibilities that they have because of the available data and the technologies, than with the enhancement of the real freedoms that people should be able to enjoy.

There are social groups that are vulnerable to be disadvantaged. This might be because of their gender, age, socioeconomic status, education level, mode of employment, geographical location or ethnicity.[131] This 'digital social inequality' can be based on one of the determinants above, or because of discrimination related to digital technologies.[132] Due to the datafication of society, the role of algorithms, and the possible discriminations that individuals belonging to certain social groups have to face, this infringes upon their freedom to be who they want to be, and decorate their lives, as they want to. Because algorithms find results on the basis of what the designer thinks is successful, there are biases in the algorithm that may favor men over women, white people over black, depending on the way success is defined.

---

[128] European Economic and Social Committee, 'The ethics of Big Data: Balancing economic benefits and ethical questions of Big Data in the EU policy context' (2016) p. 22.
[129] Whitepaper KPN: Smart Cities: een veranderende wereld. Nieuwe technologie en maatschappelijke paradigma's, p. 7.
[130] Dameri, R.P., 'Searching for Smart City definition: a comprehensive proposal', *International Journal of Computers & Technology*, vol. 11, no. 5 (2013), p. 2545.
[131] Lupton, 'Digital Risk Society', p. 7.
[132] Ibid.

> 'Automated systems claim to rate all individuals the same way, thus averting discrimination. They may ensure some bosses no longer base hiring and firing decisions on hunches, impressions, or prejudices. But software engineers construct the datasets mined by scoring systems; they define the parameters of data-mining analyses; they create the clusters, links, and decision trees applied; they generate the predictive models applied. Human biases and values are embedded into each and every step of development. Computerization may simply drive discrimination upstream.'[133]

Individuals, who have no access to these databases, have no possibilities to understand how information about them influences particular forms of decision making. It depends on what the designer of the algorithm thinks would be a successful result. As long as individuals do not know that they are by accident considered being a security risk, or as unreliable, they are disadvantaged. This form of social sorting, because of data mining, affects the life chances of those people.[134] David Lyon describes this as invisible doors that permit access to or exclude from participation in a multitude of events and processes.[135] The classifications that result from these processes are designed to influence and manage populations and persons, thereby directly and indirectly affecting the choices and chances of data subjects.[136]

One of the major unfreedoms is discrimination, as Sen argues. The challenge is to remove this unfreedom. The question in a data-driven society that is completely datafied is whether this is possible at all. There is a great deal of reliance on automated decision-making and a sanctifying believe in the 'objective' algorithms that make these decisions. As I have tried to show, computerization and algorithms do not make things fair by simply applying them. The only way to deal with discrimination is by critically analyzing the output of the decisions that are made. Therefore it is necessary that humans intervene in the processes of decision-making. I will elaborate on this argument in the following, final chapter.

---

[133] Pasquale, 'The Black Box Society', p. 35.
[134] Andrejevic, 'The Big Data Divide', p. 1681.
[135] Lyon, D., 'Surveillance as social sorting, computer codes and mobile bodies', in: *Surveillance As Social Sorting: Privacy, Risk and Automated Discrimination* (2003) p. 13.
[136] Ibid.

5. Big Data needs Thick Data

It seems comfortable to focus on the narrow forms of freedom, exactly because those are the forms of freedom that are measurable. The good thing about data is that it makes these processes of measuring possible on larger scales than ever before, because of all the data that is created and gathered. However, we have to be critical about what is exactly in the data. Who defines the success that is embedded in the algorithm? To understand this, big data is in need of thick data. Furthermore, we have to question how much of our focus should be on these measurable results.

5.1 Transparency in the results

The commercial market was the first one to embrace data mining. Most of the times, they defined the goals of data mining in terms of competitive advantage.[137] Because of the results and the possibilities for comparison, this can be catalytic for competition. Franc Weerwind is one of the board members of the Dutch Society of Municipalities (VNG). He is in favor of a benchmark tool that is already used by 120 different municipalities. With this tool, municipalities have the possibility to compare the statistics of their sustainability with averages of other municipalities.[138] This opportunity for comparison can motivate municipalities to perform better in certain fields. However, there are concerns as well. Morozov points to this when he talks about the transparency of results. It does look good when you score higher than the average, although this should not be a goal in itself. The real goals should be formulated and articulated a part from measurable averages.[139] It is not just about the 'objective' numbers, it is relevant to examine the reasons for these numbers.[140] What is included in the measurement? It is wonderful if you score higher than averages of other municipalities, but what does this statistic tell you about the situation of the society in its entirety?[141] If one neighborhood has a high crime level, this might be because the police have watched certain neighborhoods more closely than others. If there is a higher level of recorded crimes because of that, it should not come as a surprise.[142] Moreover, performance is monitored and subjected to quality and quantity assurance.[143] Public sector institutions, like

[137] Andrejevic, 'The Big Data Divide', p. 1680.
[138] Linders, B., 'De informatiesamenleving van Franc Weerwind,' *iBestuur Impact* (2018) vol. 8, no. 25, p. 8.
[139] Morozov, E., 'Om de wereld te redden: klik hier', (Uitgeverij De Wereld, 2014) p. 102.
[140] Pasquale, 'The Black Box Society', p. 41, 42.
[141] Morozov, p. 102.
[142] Pasquale, 'The Black Box Society', p. 41, 42.
[143] O'Neill, O., 'A Question of Trust', (Cambridge University Press, 2002) p. 47.

universities, are than all judged, and ranked, in league tables by their performance indicators.[144] The danger with these indicators is that not only the institutions act upon them, but individuals that want to obtain a 'good' rating as well. 'If a certain 'A'-level board offers easier examinations in a subject, schools have reason to choose that syllabus even if it is educationally inferior.'[145] Municipalities that want to obtain good scores on certain levels, can manipulate averages by focusing on the wrong aspects. It is important to question results and uncover underlying assumptions and objectifications, before we perceive data as truth. These 'objective' numbers cannot be seen without the social context, or the biases of the observers.[146] These results might seem transparent, but it is never clear which assumptions were embedded within the designs of the algorithms that were used.

5.2 Big data and the need for thick data

There is a difference between 'big data' and 'thick data'. Big data should not be used without thick data. Big data is presented in simplified ways, which reduces possible interpretations of the data that is analyzed.[147] Another problem is that some aspects cannot be measured by big data. 'Thick data' should supplement the use of big data.[148] Municipalities should not put their trust in big data solely. One of the aspects about big data is that people confide in it for a great deal. Big data is seen as the one and only way to find answers and provide truth. I believe it is not by accident that big data is always spelled out with a capital 'b' and capital 'd'. To act in ways that are more constructive to society, we should not put our trust blindly in the methods of big data. In this, I agree with Cathy O'Neil who refers to big data as 'weapons of math destruction', where she defends herself against claims like: data cannot be wrong, because how can you seriously claim that the objective numbers are wrong? On the other hand, big data provides many possibilities as well that we cannot simply disregard. Datafication is a movement that we cannot deny or stop anymore as well. However, we can choose how we want our society to be datafied. Datafied life undergoes 'the simplifying moves that are needed to convert the messy realities of people's personal attributes and behaviors into the

[144] O'Neill, 'A Question of Trust', p. 48.
[145] Ibid., p. 55.
[146] Pasquale, 'Black Box Society', p. 42.
[147] Dijck, van, 'Datafication', p. 202.
[148] Wang, T., 'Why Big Data Needs Thick Data' *Ethnography Matters* (2016) https://medium.com/ethnography-matters/why-big-data-needs-thick-data-b4b3e75e3d7

objective, tractable language of numbers.'[149] Authority for decision-making is increasingly expressed algorithmically. What is missing from these numbers, are the complexities of human life. This is exactly one of the distinctions between the use of big data and thick data. Big data loses the complexity of the individual. The figure below, which is taken from Tricia Wang's 'Why Big Data Needs Thick Data', shows how thick and big data are different from each other. Besides that, the figure shows how these two should support each other.

Thick Data	Big Data
Relies on human learning	Relies on machine learning
Reveals the social context of connections between data points	Reveals insights with a particular range of quantified data points
Accepts irreducible complexity	Isolates variables to identify patterns
Loses scales	Loses resolution

Figure 2

The municipality of Utrecht is already aware of the importance of these distinctions. They distinguish big data from data-driven, whereas data-driven decisions involve what they define as 'small data', concerning one person or a specific group. The municipality of Utrecht does not like the term big data, and prefers terms like data-driven and information-driven.[150] However, it depends on the definition that is used, because when organizations engage themselves with the practices of profiling and monitoring, they exactly utilize the features of big data: the combining of large data sets.

[149] Jasanoff, S., 'Ordering Knowledge, Ordering Society', in: *States of Knowledge: The Co-production of Science and Social Order*, ed. Jasanoff, S. (London: Routledge, 2004) p. 27.
[150] Wesseling, et al., 'Datagedreven sturing bij gemeenten', p. 18.

5.3 Transparency, accountability and trust

> The just society is not necessarily one in which each individual is treated as an isolated collection of uniquely arrayed attributes demanding individualized attention. Rather, in some even if not in all respects the just society is one in which differences among individuals are often and desirably suppressed in the service of both equality and community.[151]

Over time, there is more awareness for the risks that are posed by the datafication of society. Awareness is one of the most important aspects to create trust and accountability. Trust is not measurable. Trust is a feeling. Therefore, citizens need to have a certain amount of control over their own data, and the ways it is being used. They can only achieve this control when there is a sufficient amount of transparency. However, when there is such a large amount of data flowing from one database to another database, meanwhile affecting and changing the personal data of citizens, we have to question ourselves: who is the one who can stop these processes when things go wrong, and rectify the matters? This calls for accountability as well. The term 'accountability' is used as a synonym for many political concepts like 'good governance, transparency, equity, democracy, efficiency, responsiveness, responsibility, and integrity.'[152] Accountability in the active sense, as a form of responsiveness and responsibility, is easily used, but hard to define.[153] There has been a great tendency towards accountability for some time now. People even worry that because of the datafication that has taken a flight, there has been a shift towards an accountability culture where everything we do, every action we undertake, has to be supported by data.[154] This takes time away from the people who would originally occupy themselves with serving the people, because they spent hours to make their accountability transparent. This shows that there should be a balance between accountability and trust, which transparency can only solve to a certain extent.

To perform effective accountability there are a few steps than can guide municipalities to do this properly, as Taylor et al. describe: 'first, define the rules you want to abide by; second,

---

[151] Schauer, 'Profiles', p. 300.
[152] Bovens, M., 'Two Concepts of Accountability: Accountability as a Virtue and as a Mechanism', *West European Politics* (2010) vol. 33, p. 946.
[153] Ibid.
[154] O'Neill, 'A Question of Trust', p. 46.

monitor what you do; third, correct deviations, and finally be prepared to take responsibility for the whole circle.'[155] By doing so, municipalities automatically have to incorporate transparency and a public-facing element, which are components of accountability. The rules that municipalities should abide by can be formulated in line with the human freedoms that they have to secure.

On the other hand, transparency is seen as the key factor towards solving many of the issues, like profiling. It is true that there has to be accountability when we define accountability as the relationship between an actor and a forum. However, trust is one of the main concepts that municipalities should strive for. This is not the same thing as putting everything down in writing. Although, there are aspects that ask for this writing down specifically to obtain accountability, that is when there are collaborations between the public institutions and the private sector. Before different parties start to work together, for example, when they want to develop predictive analytics tools, they should write down the aims of the project. This does not only force both parties to think about the ethical concerns of the business they are conducting, it will create certain standards that are useful in future experiments and researches, and creates transparency to reflect on the results.[156]

It is important that the municipality stays in control of the development and transition towards data-driven societies that use algorithms for decision-making. 'The construction of the cities of tomorrow runs the risk of becoming a technological issue which will have serious effects on the framing and search for solutions to urban problems'.[157] Again, this refers back to the practices of data mining and the correlations that can be found by combining different data points. This is not problematic in itself, although, this data should not set the agenda. Data becomes relevant in the social context. According to Morozov, governors tend to be busy with rearranging complex social situations, into measurable and datafiable problems with solutions that can be brought forward through analyzing data.[158] Municipalities should not try to solve problems that are not actual problems.

[155] Taylor, et al., 'Public Sector Data Ethics', p. 18.
[156] Ibid., p. 19.
[157] Vanolo, 'Smartmentality', p. 891.
[158] Morozov, 'Om de wereld te redden: klik hier', p. 21.

This focus on the capturing of data, and managing and regulating cities through analyzing this data, promotes a technocratic mode of urban governance. This presumes that all problems are technical, and can be addressed through technical solutions.[159] Big data urbanism suffers from datafication, the believe that all meaningful flows and activity can be sensed and measured, and therefore used to solve problems that arise.[160] Technological solutions are not going to solve the deep-rooted structural problems within cities, because they do not address the root cause. What they can do is enabling the more structural management of the manifestations of these problems.[161]

5.4 Overruling algorithmic decision-making

To analyze the outcomes of the answers that algorithms provide, it is important to include questions about what is included in the measurements. This way, self-fulfilling prophecies can be avoided and it prevents researchers from seeing what they want to see. As was shown, predictive analytics may be practiced with the best of intentions. However, anticipation has other consequences beyond preventing something to happen. Municipalities can gain from predictive analysis to effectively prevent fraud or from preventing people to end up in debt, but they should take care that these practices do not interfere with human freedoms of social opportunities and transparency.

Cheney-Lippold describes an extreme but illuminating example. It is 2013 and a disabled man dials the emergency number, requesting an ambulance three separate times. The operator asks several questions but concludes that the person does not need an emergency ambulance. Three days later, the police find the man in his house, but they are too late, he already passed away. The series of questions that the operator asked the man came straight from an algorithmic triage system. The verbal answers of the person in need were input, the output was to determine whether or not an emergency ambulance would be sent. The operator even told the caller that she did not believe that the system would come up with the wrong answer. The caller's symptoms did not fit the measurable type for deserving emergency treatment. What was taken away from his answers, by simply looking at the data that the caller provided, was

[159] Kitchin, R., 'The Data Revolution, Big Data, Open Data, Data Infrastructures & Their Consequences', (Sage Publications, 2014) p. 181.
[160] Ibid.
[161] Ibid.

the agony in his voice and the repeated pleading for help.[162] What this extreme example shows is that the focus on the data, and the algorithms that make decisions based on the data that is available, lack human intervention. Human complexity and interpretation is lost, and authority for algorithmic decision-making is increasing. Moreover, the humans that are involved in this process should let go of their sanctifying believe in the workings of the system.

As was mentioned before, individuals have the right not to be subjected to automated decision-making, according to article 22 of the GDPR. Therefore, data-driven organizations should insist on human intervention, or review, when algorithms are used. This possibly prevents many of the problems that were addressed in this thesis. The 'human' that should conduct this review has a special duty whereby he must be aware of the biases that are incorporated in the algorithms that are used.

> 'Interpretation is at the center of data analysis. Regardless of the size of a data, it is subject to limitation and bias. Without those biases and limitations being understood and outlined, misinterpretation is the result. Data analysis is most effective when researchers take account of the complex methodological processes that underlie the analysis of that data.'[163]

Interpretation is at the heart of the usefulness and effectiveness of data analysis. Therefore, it is important that the decision-making process it is not solely perceived as a technical procedure. Human review that uncovers mistakes should have the power to override algorithmic decision-making.[164] There are problems with these approaches as well, because there is such a great reliance on algorithmic decision-making, that it might be hard to implement this human review. Although we achieve more efficiency, there has to be a point where we question ourselves whether we can live with the trade-off between these automated decisions and our human freedoms.

---

[162] Cheney-Lippold, 'We Are Data', p. 203, 204.
[163] boyd, d., Crawford, K., 'Critical Questions for Big Data', p. 668.
[164] Danaher, J., 'The Threat of Algocracy: Reality, Resistance and Accommodation', *Philosophy & Technology* (2016) vol. 29, p. 258.

Conclusion

The further datafication of our society is a movement that we cannot deny. However, what we can do is be alert about the goals that we formulate for our society, and the way we want to shape our lives and our society. Algorithms create many possibilities for data mining, profiling, monitoring and predictive analytics. These practices are necessary, and we should use them to the extent that they are profitable. This does not set us free from the obligations to think about the consequences when things might go wrong.

Overall, it has become clear that datafication poses many threats on human freedoms. Especially within the boundaries of the municipality this asks for careful anticipation. There are many concerns because of the collaboration between private and public parties that ask for an active attitude of the municipality to be in control of the technology, and in control of the direction that society goes. The question that this thesis started out with was how they should do that: How can municipalities secure the human freedoms that are affected by datafication, within a data-driven society?

It seems obvious to focus on economic progress and societal development. However, we should be aware that these belong to the narrow view of development. These developments do enhance the real freedoms of individuals, but should not be seen as a goal in itself. By analyzing the broader views of freedom, in particular social opportunities and transparency guarantees, I have showed that datafication has several threats for individuals in a data-driven society. These developments do not enhance the freedom for effective participation within society completely. They even make certain inequalities more visible or they repeat past practices. To make sure that individuals have the availability to access public goods, municipalities should be in control of the relevant techniques and data mining processes. Solidarity is one of the foundations of a democratic society and should be protected from private interference in making public goods available.

Municipalities are encouraged to use the practices of profiling, monitoring, predictive analysis and data-discovery. What I have shown is that these practices all have their benefits for society, but should not be implemented and embraced without any caution. We have to be careful when these profiles are used to make predictions about future likely behavior, based on correlation instead of explanation and causation. Because profiling minimalizes the

complexities of human life, there has to be caution for discrimination, prejudice and self-fulfilling prophecies. The forces of datafication that like to diminish everything and everyone into measurable averages strengthen these practices. Simply applying the forces and practices of big data sets and decision-making through algorithms does not make things fair. The mythinformation thesis posed by Langdon Winter is still accurate, and in line with Cathy O'Neil, I have tried to deconstruct the sanctifying believe in the practices of big data. This is important, because as long as trust remains in the solving powers of big data, not only the problems are identified as technological problems that can be measured. The solutions that are needed are searched for through the practices of big data again. It is possible to overcome these possible discriminations, only when human intervention and interpretation is incorporated in the decision-making processes. Human intervention should have the possibility to overrule and supplement algorithmic decision-making based on big data sets by involving thick data.

We should analyze progress to the extent that actual freedoms of individuals are enhanced, and the free agency of people determines whether true development is effective. We need these human freedoms not solely as a goal on which we can focus. It is more than a goal in itself. It is the basis for development as well. People are able to participate in social arrangements when their free agency is enhanced. Therefore, we need these human freedoms, and municipalities should safeguard them by staying away from formulating goals based on the possibilities that the technology provide. The real goals should be stated a part from measurable averages. To do so, municipalities should have their focus on their citizens, by enhancing social opportunities, transparency guarantees and creating trust.

To define success and to satisfy the algorithm are two major concerns in a datafied society. Not only for the individuals that do not have the ability to live up to certain standards, also for the individuals that are perceived as untrustworthy or unhealthy because they do not have the sufficient data to back it up. Municipalities have to be critical about the invisible doors that limit choices for individuals, because they are limited in their ways to pursue a life they have reason to value.

Not every individual can be treated with all of his own complexities. On the other hand, we should not lose all complexities of human life. The use of big data has great advantages, as long as data is used in service for the people in society, not the other way around. Data flows that affect each other by combining different datasets have to be under control of human intervention. Human intervention is the key to make sure that our society does not become fully automated. Interpretation is at the heart of making sense of data and provides the possibility to prevent discrimination and exclusion. After all, it is the government that should be there for the citizens, human beings for human beings, not a fully automated system that controls their algorithmic citizens.

Literature

Andrejevic, M., 'The Big Data Divide', *International Journal of Communication* (2014) vol. 8, p. 1673 – 1689.

Baloo, J., Brands, E., Steels, F., 'KPN Technology Book: The technology trends KPN has on its radar' (2018).

Bovens, M., 'Two Concepts of Accountability: Accountability as a Virtue and as a Mechanism', *West European Politics* (2010) vol. 33, p. 946 – 967.

boyd, d., Crawford, K., 'Critical Questions for Big Data: Provocations for a cultural, technological, and scholarly Phenomenon', *Information, Communication & Society* (2012) vol. 15, p. 662 – 679.

Broeders, D., Schrijvers, E., Hirsch Ballin, E., *Big Data and Security Policies: Serving Security, Protecting Freedom,* WRR-Policy Brief 6 (2017)

Cheney-Lippold, J., '*Jus Algoritmi*: How the National Security Agency Remade Citizenship', *International Journal of Communication* (2016) vol. 10, p. 1721 – 1747.

Cheney-Lippold, J., 'We Are Data' (New York University Press, 2017)

Dameri, R.P., 'Searching for Smart City definition: a comprehensive proposal', *International Journal of Computers & Technology*, vol. 11, no. 5 (2013), p. 2544 – 2551.

Danaher, J., 'The Threat of Algocracy: Reality, Resistance and Accommodation', *Philosophy & Technology* (2016) vol. 29, p. 245 – 268.

Dijck, van, J., 'Datafication, dataism and dataveillance – Big Data Between Scientific Paradigm and Secular Belief', *Surveillance & Society* (2014) p. 197 – 208.

European Economic and Social Committee, 'The ethics of Big Data: Balancing economic benefits and ethical questions of Big Data in the EU policy context' (2016)

Hildebrandt, M., 'Profiling and the rule of law', *Identity in the Information Society* (2008) vol. 1, no. 1, p. 55 – 70.

Hoedt, den, E., *Deel je rijk - Relevante trends voor overheidscommunicatie*, Report Ministry of General Affairs (2013)

Jasanoff, S., 'Ordering Knowledge, Ordering Society', in: *States of Knowledge: The Co-production of Science and Social Order*, ed. Jasanoff, S. (London: Routledge, 2004)

Kennedy, H., Poell, T., Dijck, van, J., 'Data and agency', *Big Data & Society* (2015) p. 1 – 7.

Kitchin, R., 'The Data Revolution, Big Data, Open Data, Data Infrastructures & Their Consequences', (Sage Publications, 2014)

Kitchin, R., 'Thinking critically about and researching algorithms', *Information, Communication & Society* (2016) vol. 20, p. 14 – 29.

Koops, B.J., 'Digitaal Huisrecht', *Nederlands Juristenblad* (2017) vol. 92, p. 183 – 187.

Linders, B., 'De informatiesamenleving van Franc Weerwind', *iBestuur Impact* (2018)

Lupton, D., 'Digital Risk Society', *The Routledge Handbook of Risk Studies* (2016) p. 1 – 7.

Lyon, D., 'Surveillance as social sorting, computer codes and mobile bodies', in: *Surveillance As Social Sorting: Privacy, Risk and Automated Discrimination* (2003)

'Maak Waar!', *Studiegroep Informatiesamenleving en Overheid* (2017)

Mai, J.E., 'Big Data Privacy: The datafication of personal information', *The Information Society* (2016) vol. 32, p. 192 – 199.

Martijn, M., Tokmetzis, D., 'Je hebt wél iets te verbergen, over het levensbelang van privacy' (De Correspondent, 2018)

Morozov, 'Om de wereld te redden, klik hier' (Uitgeverij De Wereld, 2014)

Naudts, L., Vedder, A., 'Accountability for the Use of Algorithms in a Big Data Environment', *International Review of Law, Computers & Technology* (2017) p. 206 – 224.

Neirotti, P., et al., 'Current trends in Smart City initiatives: Some stylised facts', *Cities* (2014) vol. 38, p. 25 – 36.

O'Neill, 'A Question of Trust' (Cambridge University Press, 2002)

Pasquale, F., 'Black Box Society: The Secret Algorithms That Control Money and Information' (Harvard University Press, 2015)

Prins, C., et al., 'iGovernment, Digitizing the Citizen and Government', *WRR-report no. 86.* (2011)

Rule, J.B., 'Privacy, the longue *durée*', in: Roessler, D.B., '*Social Dimensions of Privacy'* (2015) p. 11 – 31.

Schauer, F.F., 'Profiles, Probabilities, and Stereotypes' (Harvard University Press, 2003)

Sen, A., 'Development As Freedom' (Oxford University Press, 2002)

Stalder, F., Lyon, D., 'Electronic identity cards and social classification', in: *Surveillance As Social Sorting: Privacy, Risk and Automated Discrimination'* ed. Lyon, D. (2003) chapter 4.

Taylor, L.E.M., et al., 'Public Sector Data Ethics: From principles to practice', *Tilburg University* (2017)

Vanolo, A., 'Smartmentality: The Smart City as Disciplinary Strategy', *Urban Studies*, vol. 51, no. 5 (2014) p. 883 – 898.

Wang, T., 'Why Big Data Needs Thick Data' *Ethnography Matters* (2016) https://medium.com/ethnography-matters/why-big-data-needs-thick-data-b4b3e75e3d7

Wesseling, H., Postma, R., Stolk, R., 'Datagedreven sturing bij gemeenten, van data tot (gedeelde) informatie voor beter (samen) sturen,' *Vereniging van Nederlandse Gemeenten* (2018)

Winner, L., *The Whale and the Reactor: A Search for Limits in an Age of High Technology'* (University of Chicago Press, 1986).

Zuthpen, van, R., 'Wie niet past, loopt vast: verslag van de Nationale Ombudsman 2016', *De Nationale Ombudsman* (2017)

Lectures / video's

Anderson, J., lecture Digital Ethics, Universiteit Utrecht, March 01, 2018.

Beunen, J.P., presentation BigDataGemeenten Scheveningen. Jean-Pierre Beunen is Programme Director Smart Cities KPN. He is also the author of the Whitepaper: 'Smart Cities – A changing world. New technologies and societal paradigms'. Attended April, 21, 2018.

'Digitaal Doolhof', *De Monitor, KRO-NCRV*. https://demonitor.kro-ncrv.nl/onderzoeken/digitaal-doolhof visited June, 04, 2018.

O'Neil, C., 'The era of blind faith in big data must end', *TED, Ideas worth spreading* (2017) https://www.ted.com/talks/cathy_o_neil_the_era_of_blind_faith_in_big_data_must_end visited June, 20, 2018.

Schneier, B., 'Data and Goliath: The Hidden Battles to Collect Your Data and Control Your World', *Norwegian Developers Conference*: https://vimeo.com/131115865 visited at June, 11, 2018.

Websites visited:

'Aantal gemeenten van '388 naar 380', *CBS* (December 12, 2017) https://www.cbs.nl/nl-nl/nieuws/2017/52/aantal-gemeenten-van-388-naar-380 visited August 03, 2018.

'Apeldoorn voorspelt wie het criminele pad opgaat', *De Stentor* (May, 11, 2016) https://www.destentor.nl/nieuws/apeldoorn-voorspelt-wie-het-criminele-pad-opgaat~a96cdd5d/

'De muren hebben sensoren', *De Groene Amsterdammer* (December 6, 2017) https://www.groene.nl/artikel/de-muren-hebben-sensoren visited June 14, 2018.

Johnson, S., 'China's 'social credit score' will punish and reward citizens', *Big Think, Smarter Faster* (April, 25, 2018) http://bigthink.com/stephen-johnson/a-look-at-chinas-orwellian-plan-to-give-every-citizen-a-social-credit-score visited June, 14, 2018.

Smith, B.C., 'Big Data in the Humanities: The Need for Big Questions', *Arts & Humanities Research Council,* (12-11-2013) https://www.sciculture.ac.uk/2013/11/12/big-data-in-the-humanities-the-need-for-big-questions/ visited May, 12, 2018

Weijer, B., van der, 'Wiskundige Cathy O'Neil waarschuwt voor algoritmen: 'Rechten van individu worden niet beschermd', *de Volkskrant* (June, 18, 2018) https://www.volkskrant.nl/wetenschap/wiskundige-cathy-o-neil-waarschuwt-voor-algoritmen-rechten-van-individu-worden-niet-beschermd-~b97a9302/ visited June 19, 2018.

Whitepaper KPN: Smart Cities: een veranderende wereld. Nieuwe technologie en maatschappelijke paradigma's.