Rationality and politics of algorithms: Will the promise of big data survive the dynamics of public decision making?

Van der Voort, H.G.; Klievink, A.J.; Arnaboldi, M.; Meijer, A.J.

Government Information Quarterly 36 (2019) 27–38

Contents lists available at ScienceDirect

Government Information Quarterly

journal homepage: www.elsevier.com/locate/govinf

Rationality and politics of algorithms. Will the promise of big data survive the dynamics of public decision making?

H.G. van der Voort[a,*], A.J. Klievink[a], M. Arnaboldi[b], A.J. Meijer[c]

[a] *Faculty of Technology, Policy and Management, TU Delft, Jaffalaan 5, 2628 BX Delft, the Netherlands*
[b] *Department of Management, Economics and Industrial Engineering, Politecnico di Milano, Italy*
[c] *Utrecht School of Governance, Utrecht University, the Netherlands*

A B S T R A C T

Big data promises to transform public decision-making for the better by making it more responsive to actual needs and policy effects. However, much recent work on big data in public decision-making assumes a rational view of decision-making, which has been much criticized in the public administration debate. In this paper, we apply this view, and a more political one, to the context of big data and offer a qualitative study. We question the impact of big data on decision-making, realizing that big data – including its new methods and functions – must inevitably encounter existing political and managerial institutions. By studying two illustrative cases of big data use processes, we explore how these two worlds meet. Specifically, we look at the interaction between data analysts and decision makers. In this we distinguish between a rational view and a political view, and between an information logic and a decision logic. We find that big data provides ample opportunities for both analysts and decision makers to do a better job, but this doesn't necessarily imply better decision-making, because big data also provides opportunities for actors to pursue their own interests. Big data enables both data analysts and decision makers to act as autonomous agents rather than as links in a functional chain. Therefore, big data's impact cannot be interpreted only in terms of its functional promise; it must also be acknowledged as a phenomenon set to impact our policymaking institutions, including their legitimacy.

1. Big data analytics improves decision-making, but how?

The promise of big data analytics to "predict the present" (Choi, 2012) and even the future (Goel, 2010) make big data potentially invaluable for decision-making. The literature on big data shows an abundance of these potentials. For example, big data analytics regarding customer preferences and behaviours and market trends can improve business intelligence and business decisions (H. Chen, Chiang, & Storey, 2012; P. Simon, 2013). Big data can be used to extract trends where none could previously be found, and it can support evidence-based policymaking. In the public sector, big data can help officials make better decisions (OMalley, 2014) and improve overall government efficiency and effectiveness (Milakovich, 2012). Big data can provide support information for better-informed policymaking (Janssen & Kuk, 2016), based on near real-time insights into societal patterns and citizen needs (Y.-C. Chen & Hsieh, 2014b) and through improved policy evaluation (Schintler, 2014).

There is much literature on the potential of big data or barriers to its use (e.g. H. Chen et al., 2012; Choi, 2012; Höchtl, Parycek, & Schollhammer, 2016; OMalley, 2014). Yet, when implemented in real public decision-making processes, big data both serves and challenges institutions. Indeed, the meeting of big data with public institutions may well have unforeseen consequences. Big data could potentially affect existing roles within organizations regarding the use of knowledge, and it may alter decision-making, agenda setting, policy formulation, incentive structures, capabilities and many more of the processes through which public policymaking is nowadays shaped (Klievink, Romijn, Cunningham, & Bruijn, 2017).

The current emphasis on the potential merits of big data relies on a functional theory of decision-making; simply, that big data leads to better information and therefore to better decisions. So, while it is commonly suggested that big data can improve decision-making, it remains implicit how decision-making will be improved in practice (see, e.g. Arnaboldi, Busco, & Cuganesan, 2017). This implicit functional theory likely lacks explanatory value, as it neglects the institutions that shape the process from data generation to the decisions taken. This may hamper our ability to explain or prevent disappointments in big data implementation projects.

This paper therefore focuses not on the potential of big data, but rather on how big data is used for public decisions, in decision-making processes. We challenge the abovementioned implicit theory by examining different, even competing, analytical views. Moreover, we advance the field by applying these views to empirically study the process of big-data use in decision-making. Through these conceptual and empirical steps, we seek to answer the question of how big data impacts public decision-making.

* Corresponding author.
E-mail address: h.g.vandervoort@tudelft.nl (H.G. van der Voort).

https://doi.org/10.1016/j.giq.2018.10.011
Received 29 December 2017; Received in revised form 15 October 2018; Accepted 15 October 2018
Available online 21 October 2018

After introducing the main concepts ‘big data’ and ‘decision-making’ in Section 2, we build an analytical framework by juxtaposing two theoretical dimensions: one being the nature of public decision-making processes and the other being the salience of either the data provision side or the data use side. This yields four theses and four corresponding research questions. The questions, which Section 3 discusses in detail, are the following:

1. What opportunities might big data provide data analysts to provide better information to decision makers?
2. What opportunities might big data provide decision makers to better absorb information from data analysts?
3. What opportunities might big data provide data analysts to pursue their own interests while providing information to decision makers?
4. What opportunities might big data provide decision makers to pursue their own interests while absorbing information from data analysts?

Section 4 introduces the research approach, which is based on two exploratory case studies: a dashboard for tracking criminal incidents in Tilburg (Netherlands), and the Digital Traces project in Milan (Italy). Sections 5 and 6, respectively, present these cases. The focus of this qualitative part is on big data use processes, encompassing the sequence of activities from data collection to decision-making (i.e., how big data is used) (Janssen, Van der Voort, & Wahyudi, 2017). In such processes, we view the roles of both data analysts, as information providers, and public decision makers as critical to the impact of big data. Section 7 analyses the two cases through the lens of our four theses. We then wrap up with conclusions in section eight and a discussion in section nine.

2. Big data and decision-making as a process

‘Big data’ is a big concept, covering a wide array of recent developments in the production and processing of data. Big data is often defined in terms of a number of ‘V's’, the early definitions focusing on three V's by describing big data as high-Volume, high-Velocity and high-Variety information assets that demand cost-effective, innovative forms of information processing for enhanced insight and decision making (Gartner, n.d.; Kitchin, 2014: 68). Over the years this definition, has been extended, with different definitions zooming in on different features. Each captures the idea that big data is relevant for decision-making, and extracts several aspects of the big data concept. In this paper, we follow Klievink et al. (2017) who defined big data not in terms of characteristics of the data, but in terms of the characteristics of their use (Klievink et al., 2017). They noted five characteristics that differentiate big data processing from conventional data processing:

1. use and combining of multiple, large datasets from various sources, both external and internal to the organization;
2. use and combining of structured (traditional) and less structured or unstructured (nontraditional) data in analysis activities;
3. use of incoming data streams in real time or near real time;
4. development and application of advanced analytics and algorithms, distributed computing and/or advanced technology to handle very large and complex computing tasks; and
5. innovative use of existing datasets and data sources for new and radically different applications than the data were gathered for or spring from.

These five characteristics help to define big data, building on earlier conceptualizations (Adrian, 2011; M. Chen, Mao, & Liu, 2014; Davenport, Barth, & Bean, 2012; Gantz & Reinsel, 2011; Hota, Upadhyaya, & Al-Karaki, 2015; Janssen & Kuk, 2016; Mayer-Schönberger & Cukier, 2013; P. Simon, 2013). Nonetheless, apart from a powerful promise, the full effects of big data on decision-making have rarely been systematically covered.

Big data appears to have huge potential to create value for businesses and governments. However, it remains unclear what institutional mechanisms and factors may explain how big data creates business value (Janssen et al., 2017) and public value (Chatfield & Reddick, 2017). Janssen et al. and Chatfield and Reddick focused, respectively, on factors and conditions concerning data quality and public value creation. Our focus here is more process oriented. We view decision-making as a process leading to a certain choice (Simon, 1960). Besides the choice itself, this process involves problem formulation (intelligence) and generation of alternatives (design) (Simon, 1960). Simon argued that if decisions are complex enough, the phase of the generation of alternatives is already significant for the eventual choice, because this phase involves a selection, and the way alternatives are framed impacts the further decision.

Many authors, including Simon and Janssen, present the decision-making process as a sequence of activities (e.g. Y.-C. Chen & Hsieh, 2014a; Marx, 2013). Some acknowledge, however, that during these activities many actors are of crucial importance for the output (the decision or decision quality) and outcome (value). Actors' involvement thus affects the dynamic of the decision-making process. This casts a shadow on the explanatory value of models that depict decision-making as a neat sequence of activities. Some scholars therefore prefer to focus on actor dynamics.

Somewhere in decision-making processes, data are prepared for use in making decisions. This suggests that information is extracted from data by attaching meaning to the data and sometimes by data personalization, moving data up in the familiar knowledge pyramid (Ackoff, 1989). Could the actors in these activities be guided by their own interests? There are multiple views on how activities and actors influence decision-making. The next section elaborates on these, introducing theoretical views on the role of big data in decision-making processes.

3. Better decision-making thanks to the data revolution: a framework and four theses

There are multiple ideas about the way big data affects public decision-making. These originate from various disciplines, from engineering to management, legal studies and public administration. Their outlooks range from optimistic to outright sceptical regarding the impact that big data can have on public decision-making. The conflicting views can be traced to fundamentally different starting points for assessment. These different starting points spring from two dimensions: different perspectives on decision-making processes and different logics underlying the relationship between decision-making processes and big data. This section develops an analytical framework based on these two dimensions, which then covers a breadth of potential views on the impact big data could have.

3.1. Processing data for decision-making: an information logic and a decision logic

Big data entails a radical change in the way knowledge is constructed. It involves different competencies, different work processes and, as a consequence, different complexities (Kitchin, 2014). It also involves new functions and players, and new relations between players. According to Crawford, Miltner, and Gray (2014), “Data sets, including predictive data, may lead to new concentrations of power and they are never methodologically removed from human design and bias.” In this regard, Uprichard (2015) raised relevant ‘who’ questions: Who is generating big data? Who owns all or parts of it? Who can access and process it? Who does so in practice? Who receives analytics and findings, and for what purposes? Who profits the most? Who profits the least? Answers to these questions reveal the interests that underlie the process from data generation to decision-making (Uprichard, 2015).

These observations and questions suggest that the choice of focal actor might prove crucial in understanding the impact of big data on decision-making; and also that the impact can hardly be comprehended from a single point of view. Data collectors, for example, will perceive it differently than data users. The first analytical dimension of our model is thus the point from which the big data process is viewed. We identify two logics. First is the information logic, which can be considered mainly the viewpoint of those at the very start of the data use process. Second is the decision logic, representing the point of view of decision makers.

The people at the start of the process – henceforth termed ‘data analysts’ for convenience – oversee data collection. They have a supporting role in decision-making. Their main concern, at least in a formal sense, is to serve decision makers by providing high-quality information via data. According to the information logic, big data provides opportunities because the sheer quantity of data, alongside advanced processing capacities and new methodologies, enhances information quantity and quality. Big data is thus thought to serve decision makers at least by providing more and better information. However, from a decision maker's viewpoint, the decision logic, the advantages of this are not self-evident. Their role is to process information into choices that can be demonstrated as effective and legitimate. For this, the amount and quality of information may not be the main limitation. According to the literature on psychology and public administration, many vital qualities needed for good decisions may be bounded by human shortcomings, such as individual capacity to process information and decide rationally (see, e.g., Tversky & Kahneman, 1981), but also by coordination capacity within and between organizations and capacities in terms of time and money (Simon, 1997). Due to these ‘bounded rationalities’, decision-making processes tend to be characterized by small, incremental steps wherein means sometimes prevail over ends (Lindblom, 1959; Wildavsky, 1978). Summarized, decision makers must somehow control their information sources, in order to make sense of them and to account for the information they apply.

As implied, the information logic considers source variety to be a vital mechanism for information quality. Nardi and O'Day (1999) famously used an ecological metaphor to describe the environment in which information is ideally produced and decisions are made (Nardi & O'Day, 1999). They framed information ecologies as “systems of people, practices, values, and technologies in a particular local environment”. As with other evolutionary theories (e.g., Aldrich, 1999; Stonier, 2012), scholars using the information ecology metaphor claim that within those ecosystems diversity of people and technologies is key, to enable the co-evolution of people and technology for decision purposes, both everybody's own, but also those in broadly established common institutional settings, such as a library or a school (e.g., Fidel, 2004; Lueg, 2001). However, from a decision logic, variety is problematic. Once decision makers become responsible for decisions for a common purpose, their task is to weigh ideas against each other. In the ecological metaphor, decision-making can be viewed as a selection environment for these ideas. Though variety of information may serve decision quality, for decision makers the selection process represents the main challenge, as this where they must align interests, values and ideas (de Bruijn & ten Heuvelhof, 2008).

From the information logic, key choices are primarily methodological. Multiple organizational options are available to better serve those involved and bring in new technology. In the decision logic, choices between options are acknowledged as engaging competing values, which may be rather ideological in character, for instance, efficiency, safety and sustainability. The management literature presents managing such competing values as a core leadership skill (Quinn, Bright, Faerman, Thompson, & McGrath, 2014). In the institutional literature, too, value conflict is presented as central to decision-making (Steenhuisen & van Eeten, 2008; Stewart, 2006; Thacher & Rein, 2004).

Table 1 summarizes the two logics as described.

Table 1
Two logics of information use for decision-making.

	Information logic	Decision logic
Role	Supporting decisions	Control information sources
Quality mechanism	Variety	Selection
Main choices	Methodological	Ideological

3.2. Public decision-making as a rational and as a political process

The second dimension in our framework regards assumptions on the way public decision-making develops. The presumed promise of big data for decision-making usually hinges on the idea that decision-making is served by better information. Höchtl et al. (2016) used the policy cycle of Nachmias and Felbinger (1982) to explain how big data can help to improve policy decision-making. Their argument is mainly a positive one. They encourage the use of big data to improve policy-making and look for opportunities and for solutions to the identified challenges. Höchtl and co-authors (2016) see big data as a basis for improved public decision-making, as it could make more high-quality information available and thereby elevate decision quality. This argument rests on two assumptions: (i) that the data revolution yields better information and (ii) that better information leads to better decision-making. The implication here is that decision-making is a constructive sequence of clearly differentiated activities; for example, data generation, information provision and decision-making. Such a sequence is in fact a common element in theories on public decision-making, for instance, in the policy cycle (e.g., Helbig, Dawes, Dzhusupova, Klievink, & Mkude, 2015; Jann & Wegrich, 2007; Lasswell, 1956). Moreover, everyone involved in the process is assumed to work, by and large, towards some common purpose. This implies that data are collected to serve decision makers, and decision makers are willing and able to adopt the new flow of information.

However, public administration scholars have challenged both assumptions. The assumed sequence of activities is at odds with more political theories of decision-making. The political view presents decision-making as an interaction between multiple actors, with no actor having the means to impose its desires on the others (Scharpf, 1993; Scharpf, 1997). The actors are mutually dependent, and it is hard to secure their commitment to a common problem or solution (Hans de Bruijn & Heuvelhof, 2008). Decisions, then, are made in a complex policy battle wherein some actors participate and others do not, wherein some win and others lose, and wherein winners have no guarantee of winning the next round too (Cohen, March, & Olsen, 1972; Koppenjan & Klijn, 2004; Teisman, 2000). The question of whose problems and solutions are identified becomes relevant in such a network, as the answer suggests who has succeeded in imposing their ideas on others.

The presumption of a common goal, too, is at odds with politically oriented theories. Decision makers do their job in an environment in which actors have plenty of incentive to behave strategically and use any power source they feel expedient, such as formal authority, information disclosure, naming and shaming, and shirking (Axelrod & Hamilton, 1981; Jansen & Meckling, 1976; Kuit, 2002). Strategic behaviour and informal dynamics add to the forces underlying public decision-making, again challenging the idea of a neat sequence of process steps. Uncertainties then arise from the potential behaviour of actors, in addition to scientific knowledge gaps (Klijn & Koppenjan, 2015: 72; Scharpf, 1997; Van Asselt, 2000). Using a decision logic, Kogan argued that governments may ‘misuse’ evidence-based decision-making, by selecting the evidence that supports prior constructed goals and politically driven priorities (Kogan, 1999). According to Feldman and March, policymakers' information requests have a symbolic dimension (Feldman & March, 1981). Though it is impossible for policymakers to digest all relevant information, they feel that it would be

Table 2
Two views on decision-making.

Rational view	Political view
Focus on what: activities	Focus on who: actors
Focus on predefined process steps	Focus on real-life transactions
Focus on common goals	Focus on individual goals

illegitimate not to request it. As a result, it is hard to still the information hunger. However, it is questionable whether the information – evidence-based or not – will be absorbed. This implies that decision makers are responsive to their own needs. They may shop around for knowledge to legitimize their decisions. This idea contradicts the sequence of predefined activities towards some common purpose. As a result, instead of a neat sequence of activities, the policy cycle may well contain multiple iterations between a diversity of process steps, making it more like a plate of spaghetti than a cycle (Klijn & Koppenjan, 2015: 72).

When assessing how big data affects public decision-making, it is thus useful to distinguish between two views on public decision-making: the rational view and the political view. The rational view represents a clear process in which big data can enhance the various steps in which information is required. The political view represents an erratic, dynamic process, in which political or other goals partially determine when, where and how there is use for big data. Table 2 summarizes these views.

3.3. Towards a framework

To answer our question about the impact of big data on decision-making, both of the logics and views presented may be required. From an information logic, the big data era promises new insights because of the abundance of data and extensive processing capacity. Algorithms defined by experts play a large part herein. The key issue for data analysts to determine is how big data can provide better information for decision-making. Yet, the question of what is ‘better’ depends on the other dimension. From a rational viewpoint, big data provides opportunities for enriching decision-making. From a political viewpoint, big data provides more autonomy to serve decision makers as data analists prefer.

With big data, decision makers are confronted with a new method for establishing policy-relevant information. Information could potentially come at a different rhythm than before, particularly faster, even real time. From a decision logic, how will decision makers deal with this new reality? How might they better absorb information from analysts? Again, ‘better’ can be interpreted differently. From a rational viewpoint ‘better’ means more absorption of information from analysts. From a political viewpoint ‘better’ means more freedom to absorb information the way the decision makers prefer.

The dependencies between these logics and views suggest that it makes sense to juxtapose the two dimensions. Table 3 presents the result and serves as our analytical framework for the empirical explorations.

We now briefly explore the four quadrants.

3.3.1. The ‘information optimization’ thesis (Quadrant 1)

The ‘information optimization’ thesis suggests that big data impacts decision-making by enabling data analysts to provide better information to decision makers. This quadrant emphasizes the methodologies of big data and the process steps taken from data analytics to decision-making. Big data empowers analysts to improve their services to decision makers. The core impact of big data on decision-making lies in providing meaningful data for a specific purpose. This is where data analysts and decision makers meet.

From the rational viewpoint, data analysts provide data according to decision makers' requests. This is where the data analyst's job stops and the decision maker's job starts. This interface might prove important. In data collection and information generation, choices are made in two main steps. The first is the selection or aggregation of data and calculation of indicators; the second is its visualization (Kitchin, 2014: 106–109). Analysts elaborate data starting from the decision maker's need, proposing the most appropriate level of aggregation and the best indicators. The choice of how to visualize the data then appears to be rather straightforward once the indicators are defined. From a rational viewpoint the visualization is chosen to suit the user of information, to facilitate understanding and interaction.

3.3.2. The ‘decision optimization’ thesis (Quadrant 2)

The ‘decision optimization’ thesis suggests that big data impacts decision-making by enabling decision makers to absorb information by data analysts. Quadrant 2 emphasizes the power of big data to facilitate ‘evidence-based’ policymaking. The idea that big data leads to better decision-making because of its volume, variety and velocity relies on the assumption that the more ‘evidence based’ the information is, the more decision makers will accept it. Underlying this assumption are two more assumptions. First, policymaking as a sequence of activities implies that a process is finished before the next process starts. This would mean that the problem should be identified before a decision is taken. Otherwise the evidence-based information could miss the boat. Second, decision makers would be willing to follow up on the information that data analysts supply without questions or iterations. Strictly, from a rational viewpoint there is no iteration between hard evidence and political wishes.

These assumptions are hardly made explicit in the literature. However, they are inherent in the thesis that big data leads to better decision-making.

3.3.3. The ‘politics of algorithms’ thesis (Quadrant 3)

The ‘politics of algorithms’ thesis suggests that big data impacts decision-making by enabling data analysts to pursue their own interests while providing information to decision makers. Quadrant 3 presents the information logic and the political view. Those adopting a political view will emphasize the construction of information by actors in a political and unstructured interaction process wherein data are not perceived as neutral. The earlier-mentioned observations of Crawford et al. (2014) and Uprichard (2015) suggest that political choices are inevitable to get the big data process going. As an example, in Ireland each of the 88 planning authorities has its own land use and zoning classification system (Kitchin, 2014: 157). This implies that integration of data for decision-making requires a prevalence of one system over

Table 3
Four perspectives on the impact of big data on decision-making.

	Information logic	Decision logic
Rational/analytical view	**Quadrant 1: information optimization** Data analysts can rationalize processes by providing better data in a more accessible format.	**Quadrant 2: decision optimization** Decision makers can rationalize processes by retrieving exactly the information they need for their decisions.
Political view	**Quadrant 3: politics of algorithms** Smart algorithms provide data analysts the opportunity to influence outcomes while providing information to decision makers	**Quadrant 4: information market** Big data provides decision makers the opportunity to limit their dependence on information from specific sources

another. The use of big data here requires choices that are inherently political.

This has serious consequences for the idea of big data-enabled 'real-time' decision-making (see, e.g. Höchtl et al., 2016). Although big data analysts serve decision makers in a formal sense, the analysts may have substantial informal power if the decision maker is unaware of the intricacies of big data processing. Indeed, vital choices are increasingly made by data analysts, outside the decision maker's reach, because the latter seldom fully understand the algorithms that give meaning to the data. The interaction between data analysts and decision makers then can be described as the 'politics of algorithms'. That is, politically sensitive choices are (wittingly or not) made by analysts and remain implicit to others through inherently complex algorithms. These others include the decision makers who make choices based on the outcomes of those algorithms applied to big data sets.

3.3.4. The 'information market' thesis (Quadrant 4)

The 'information market' thesis suggests that big data impacts decision-making by enabling decision makers to pursue their own interests while absorbing information from data analysts. In quadrant 2 decision makers are assumed to absorb evidence-based information. However, according to the decision logic, decision makers are hardly expected to absorb information from data analysts, even if it is evidence-based. Instead, they are thought to be interested in trading their political wishes for evidence from analysts, regardless of the origin of the evidence. As stated, decision makers are assumed to be responsive to their own needs. As such, big data can be seen as a new source of information, among other sources, and to provide added flexibility in legitimizing decisions. In other words, decision makers are clients in an information market and big data only makes this market larger.

4. Research approach

4.1. Four lenses, four questions

Table 3 presents four theoretical lenses, each with its own assumptions. These assumptions differ and sometimes contradict. None is authoritative in and of itself, for various reasons. Quadrants 1 and 2 mainly contain theoretical contributions focusing on the potential of big data. Empirical evidence of this is still rare, however, especially in public decision-making. Quadrants 3 and 4 have hardly been applied to big data. From each of the four perspectives, the main question, "how does big data impact public decision-making practices", must be answered with a slightly different slant. This leads to the more targeted questions, here reiterated from section one, as follows:

1. What opportunities might big data provide data analysts to provide better information to decision makers? (Quadrant 1)
2. What opportunities might big data provide decision makers to better absorb information from data analysts? (Quadrant 2)
3. What opportunities might big data provide data analysts to pursue their own interests while providing information to decision makers? (Quadrant 3)
4. What opportunities might big data provide decision makers to pursue their own interests while absorbing information from data analysts? (Quadrant 4)

These are relatively open question that allows for an explorative study. To answer these questions, we use a qualitative case study approach, because this approach allows an in-depth view on the relations between the different relevant variables around big data and decision making (Yin, 2009). We present two case studies of big data use processes for decision-making in the public domain. Both case studies are empirically guided by the four questions. Both cases are fuelled by big data; however, they differ with regard to the motives behind big data's use. The first case study concerns a data analysis 'dashboard' for combatting crime in the Dutch city of Tilburg. This goal is not contested in any way, so the case is relatively functional. It elaborates on how such a dashboard functions as a device to facilitate analysts and decision makers in doing their jobs. The second case study concerns a digital monitoring system for the city of Milan. This system's aim was rather political. It was developed on the eve of a mayoral election to scrutinize Milan's position as an international city. This case study focuses on the interaction between data analysis and decision-making.

By asking these questions, we hope to determine which theses yield the most fruitful answers. It is not our intention to compare the cases. Instead, we look at whether the cases provide support for the theses, and what the findings mean for future empirical studies.

4.2. The case studies

4.2.1. A burglary, robbery and theft dashboard in Tilburg

Tilburg is a city in the southern Netherlands. Our research concerned a data analysis dashboard to combat burglaries and thefts. The basic idea was that better information would enable the city to take more specific and appropriate measures to fight crime. A special working group towards this goal began in 2013 and was disbanded in 2015, as the need for a concerted effort to reduce burglary, robbery and theft had disappeared. But the dashboard has remained in use. Primary data collection for this case consisted of interviews and document analysis. Interviews were conducted with the knowledge broker and the policy consultant in 2014, with an additional interview with the knowledge broker in 2017, and additional information obtained from the data analyst via e-mail. All available documents were collected and analysed: a policy plan for the activities of the group implementing the data science, an evaluation of a set of policy measures and minutes of meetings. Secondary data collection consisted of an analysis of the Dutch Kwaliteitsinstituut Nederlandse Gemeenten (KING, 2014), the only other available study on this data dashboard. Data collection and analysis focused on implementation of the technology, institutional changes and use of the data in police and government processes. The data were analysed with coding software and axially coded. The final case description was sent to the respondents to confirm the accuracy of the description.

4.2.2. Digital traces in Milan

The Milan case regards construction of a digital monitoring system for the city of Milan. One of the authors studied this case as a participant observer. The researcher was involved as a mediator between analysts and decision makers. This provided a valuable in-depth understanding of the process, but also introduced significant methodological challenges, making the case illustrative only. The research presented here is part of the larger Urbanscope project (www.urbanscope.polimi.it), a research laboratory for experimentation in collection, organization, analysis and visualization of cross-domain geo-referenced data. Observation of the process started in 2014, when the idea of the monitoring system was initially raised, and continued until July 2016. The primary data source was direct participation in meetings and minutes. There were 43 meetings in the observation period, with an average duration of three hours. Half of the meetings involved both analysts and decision makers; half only analysts. These data were complemented by document analysis and eleven face-to-face interviews, five with analysts and six with decision makers.

Two important observations should be made regarding our qualitative studies. First, we applied our propositions to three simple process steps: the generation of data, the generation of information from the data and decision making. Second, we limited ourselves to two ideal type actors, those being the analyst and the decision maker. As noted in the discussion on the 'politics of algorithms', the relation between the data revolution and the quality of decision-making very much relies on the relation between these two 'actors'.

5. Burglary, robbery and theft dashboard in Tilburg

5.1. Context of the case

Reducing the prevalence of high impact crimes, such as burglary, robbery and theft, is a key priority for urban governments. Tilburg is a large producer of soft drugs and has a relatively high crime rate. At the same time, fiscal pressure has forced the city to cut spending. It has therefore sought to use data as an innovative approach to crime control. With the use of data, the government hopes to better target interventions, thus reducing the need to increase the resources devoted to public safety. The approach is also well aligned with the city's policy to operate on the basis of information and evidence. According to Tilburg respondent 1, "Information-led actions is our normal approach and is not up for discussion."

5.2. Description of the big data use process

5.2.1. Data collection

The first step in the analysis of crime patterns was collection of relevant datasets. Using data is complicated, since the data are collected and managed by different organizations, such as city hall and the national police. The city uses datasets from the department of urban safety, such as crime reports, but also datasets from other agencies about issues as varied as housing prices, school truancy and zoning decisions (Tilburg respondents 1 and 2). In addition, the city has been granted access to national police datasets. In particular, it uses police datasets to analyse crime patterns: a dataset on burglary, robbery and theft incidence; a dataset on the value of stolen goods; and a dataset on the modus operandi of criminals.

The idea underlying the approach is that new information can be produced by combining different types of data (Tilburg respondent 1). Access to the data from the police could be obtained by signing an agreement to guarantee the privacy of persons to whom these datasets relate. As such, the city managed to combine datasets that had not previously been combined for urban safety.

5.2.2. Presenting the data on a dashboard

The data were used to build a 'dashboard' with key information about, among other things, burglaries, robberies, thefts, violence and soft drug production in the city (Tilburg respondents 1 and 2). Fig. 1 shows this dashboard. On the left side of the dashboard is a map of the city indicating the number of burglaries, robberies and thefts per 1000 inhabitants in the various neighbourhoods. The colour red is used to indicate a relatively high prevalence, with yellow neighbourhoods having a lower rate. Graphs at the top right of the dashboard show changes over time, thus indicating whether the crime rate is rising or falling. A spreadsheet at the lower right presents absolute numbers of burglaries, robberies and thefts.

The map provides a starting point for obtaining a more detailed picture of burglary, robbery, theft, violence and soft drug production incidents. Clicking on a neighbourhood produces a detailed map with information on incidents, including modus operandi and the value of stolen goods. In addition, the postal codes of known career criminals can be viewed (the so-called 'top 100' criminals). Users can combine this information with information about the type of housing and neighbourhood. A good understanding of both the city and the nature of the criminal acts is required to make sense of the diverse geographical data.

5.2.3. Using the dashboard for strategic and operational safety work

Broad analyses of social and criminal trends inform strategic public safety plans for the city. These plans provide a framework for tactical measures to reduce criminal incidents in the various neighbourhoods. The dashboard constitutes a key input for these strategic and tactical plans (Tilburg respondents 1 and 2). Tilburg respondent 2 highlighted the value of the information:

> "Before we had the dashboard we did not have a good view of the hotspots, and we basically used a scattergun approach.… Now we look specifically at the hotspots through the area-based approach, the analyses at the neighbourhood level and the area-focused approach from BRT [BRT is the city working group on burglary, robbery and theft]. Now we aim at specific targets."

At a later date, policy analysts can also use the dashboard to evaluate whether the measures taken indeed lowered the crime rate.

The respondents indicated that the dashboard was used in different ways for strategic and operational safety work (e-mail, 24 July 2017). It

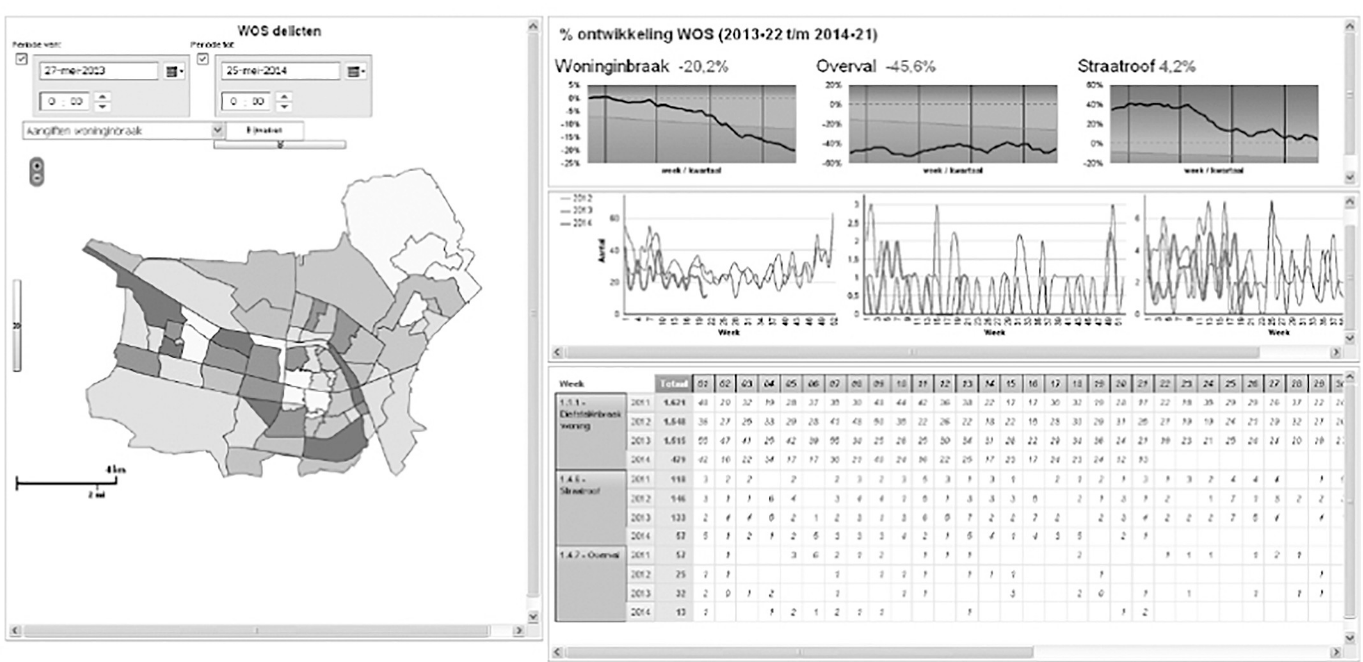

Fig. 1. Screenshot from the burglary, robbery and theft dashboard of Tilburg, the Netherlands.

was used to provide three-monthly updates to the members of the public safety 'triangle' (the mayor, police chief and district attorney) on the safety situation in the city. In addition, information specialists used the dashboard as an early warning system to identify trends. Civil servants working in the neighbourhoods used the dashboard to both monitor developments and inform citizens about sensitive issues, for example, at information sessions regarding shelters for asylum seekers and drug addicts. Finally, public safety specialists used the dashboard to track crime trends. The data analyst indicated in an e-mail:

> "The security specialists use the dashboard to track their themes (such as violence, breaking into houses, youth, etc.). The dashboard is one of the sources that is used to form a full picture of the situation."

Respondents from the city emphasized that decisions were not made based only on the datasets. Usage of the data was embedded in decision-making procedures that also relied on the inputs of experts and citizens. Analyses of crime patterns and suggested measures were always discussed with safety experts, to glean their opinions about the problems and possible solutions. Neighbourhood visits were an additional invaluable information source (Tilburg respondent 2). These visits provided a deeper understanding of the locales and also helped to engage citizens and stakeholders in local public safety issues.

5.3. The interaction between analysts and decision makers

Introduction of the dashboard resulted in new structures in Tilburg's local government organization. While the body responsible for decisions on public safety remained the 'triangle' of mayor, police chief and district attorney, a working group on burglary, robbery and theft (BRT) was established to take operational decisions on the basis of big data. Created to speed up decision-making (Tilburg respondent 2), the working group was made up of employees from the police, the public prosecution service and the municipality. Participants from the municipality included both civil servants with substantive expertise, such as criminologists, and data scientists. Thus, data science seems to have been integrated into the working procedures of the organization in this case.

The working group was mandated to take operational decisions in line with the strategic framework formulated by the 'triangle'. An example of such a decision was the development of an operational approach focused on hotspots. The working group used the dashboard to inform the 'triangle' of its activities and results. In an e-mail, the data analyst indicated:

> "We use the system as an early warning: if we see important developments that require action, we immediately inform the mayor."

Tilburg respondent 2 spoke of the use of information to influence the behaviour of politicians:

> "[O]n the basis of this data I can just go to the triangle, turn on the dashboard and show [them], "Mayor, this is our situation."… I can indicate, "Dear mayor, if we want to score on street robberies, we need to push here." That means I need extra means and men to realize this."

The dashboard influences politicians' decisions, but politicians are also influenced by other sources. These may steer them in a direction that is not entirely in line with the data from the dashboard. According to Tilburg respondent 2, the mayor sometimes identified relevant issues that then influenced approaches at the operational level. In addition, Tilburg respondent 1 noted that safety is not only a matter of crime statistics; it also involves how residents feel about their environment. Therefore, measures sometimes had to be taken in response to signals from society. Sensitivity to those signals, she said, is crucial for a strong public safety policy.

6. Digital traces in Milan

6.1. Context of the case

The Urbanscope project was the brain child of a member of a non-profit association committed to value creation in Milan. Association members were politicians, journalists, managers and many others involved in life and policy in the city. The initial proposition was to study Milan's position as an international city. Ultimately, the project was implemented in the lead-up to election of a new city mayor, to use data-derived insights to stimulate public debate. The project concerned both a 'digital traces' layer, comprising the informational landscape, and an 'urban' layer, comprising the urban space, buildings and infrastructure. Observation, analysis and representation of these two layers combined was to provide valuable insights on the city's international position and how it is lived and used.

On the part of decision makers, the association was represented by a key leader with a background in management, henceforth termed simply 'leader'. Analysts came from various disciplines: computer engineering, statistics, design and management. An important contextual element during the project was preparation for and opening of Expo 2015. Being an international event, Expo 2015 had a role of its own, as it was expected to be a global catalyst for Milan.

6.2. Description of the big data use process

6.2.1. Data collection

The first step was activated by a general research question: "to what extent is Milan an international city?" The leader posed the problem to the data analysts, pressing them to search, collect and elaborate all possible information from traditional and digital sources. During the first meeting, the analysts posed questions on possible sources. But the leader, indicating a desire for a "big" dataset, wanted to keep the scope open and wide. The search encompassed census and traditional statistics on cities, social media data (from Twitter, Instagram and Foursquare) and phone data. During the search, analysts realized that they would need to narrow the scope. To do so they defined subcategories to better specify "internationalization". This was done at a meeting with the leader. Feasibility was then tested, leading to elimination of some of the sources. For example, Facebook was eliminated due to privacy issues, and some international census data was omitted due to lack of geographical and temporal homogeneity. These eliminations elicited disappointment in the leader, first because the traditional data turned out to be less rich than expected when explored in an innovative way; for example, data on Milan's districts were poor. Second, more information was expected from Facebook. Analysts had to prove objectively that their choice not to include certain data was not arbitrary but linked to actual constraints. A major ambiguity that arose was the definition of the spatial unit of analysis, which ranged from the entire Lombardy region to the metropolitan area, the city, districts and specific positions of individuals. This is a typical dilemma preliminary to data integration. Here, data analysts proposed algorithms to cross-analyse different data and make a statistical inference on different units of analysis.

6.2.2. Data processing

Processing encompassed decisions on data aggregation and analysis and on visualization. Regarding the analysis and aggregation step, the main problem was weighing the trade-off between the number of sources and the time resolution for data fusion. Interestingly, in this phase analysts wanted to keep all of the collected data, searching for algorithms to conciliate the temporal diversity. Analysts saw the problem as just technical, a matter of "dividing totals" into subunits. The leader, however, was sceptical of applying algorithms, as he saw this as "data manipulation", given that any assumed division (e.g., monthly data into daily data) would lead to constructed data, not real data. The

Fig. 2. An example of data visualization for the project.

solution was to use diverse lenses to represent the city, so as not to overstretch data fusion. For example, the initial decision to analyse census, phone and Twitter data together was abandoned.

This phase revealed the importance of visualization in translating data into information. The first visualization was made using standard statistical software, showing matrices and box plots with dots representing data. After the meetings using this visualization the leader admitted, "the data don't tell me anything". Designers (always present even in the previous phases) then took the lead in the analyst group and began to elaborate graphical solutions using city geography as the background and reference framework. Here, the designers can be seen as 'intermediaries' between the analysts and the leader. The new visualizations catalysed the leader's thinking towards the next step: how to use the data. Fig. 2 shows the final result; the central visualization is based on the map of Milan districts which have a darker colour where the number of Tweets is higher in the temporal window selected. On the left part the interactive visualization allows to see the diversity of the tweets' languages for each district (for example Loreto or Brera in the screenshot).

6.2.3. *Using digital traces for decision-making*

The visual narrative of the city data activated the most interactive phase between analysts and the leader. All sought to give sense to the data. The analysts wanted to see their data valued and the leader wanted "stories to tell", to feed public debate and highlight problems and opportunities for the city. The starting point, the question "to what extent is Milan an international city" remained, but some elements of benchmarking had to be abandoned. In some cases, this was due to analytical problems (incomparability). In other cases, the results were not aligned with the politics that the leader wanted to promote. For example in some areas Milan appeared 'less international' than other cities. After several brainstorming sessions, four lenses were defined, restricting the number of sources but favouring the communication of narratives.

Another major issue was the temporal resolution. Social media data and phone data allow real-time frequency; yet data analysis and the stories selected did not change at a real-time pace. In this regard, the leader perceived a risk of trivializing the value of the digital traces. As a consequence, a much larger resolution was chosen for reporting: monthly and weekly. Analysts objected this solution at first, given the technical feasibility of reconciling sources to a lower resolution. However, the analysts soon began to complement their technical vision with a decision-making view.

6.3. *The interaction between analysts and decision makers*

The interaction between analysts and decision makers was lively throughout the project. From the start, the leader had a genuine interest in learning about internationalization, with the intuition that Milan was becoming more international and that traditional data did not capture the pace of this phenomenon. During data generation, both analysts and the leader understood that the question was too broad and tensions arose.

In a traditional data context, analysts and decision makers compete in data elimination. Here the opposite happened. Neither wanted to make decisions on eliminations, even though it was obvious that the question to be addressed was too broad. The main reason was the risk of losing signals from the data. Analysts proposed splitting the problem into subcategories, such as 'business', 'education' and 'entertainment'. These were marginally used to reduce the problem. Ultimately, the analysts and leader found their roles: the leader selected categories based on political relevance and the analysts fed the categories.

Interactions during information generation were marked by scepticism due to the participants' different knowledge of algorithms and politics. The leader tended to challenge every proposal that the analysts made, for example, on the reconstruction of temporal resolutions. He feared any hint of subjectivity would be dangerous when put to a public audience. Analysts were sceptical of masking some numbers, because to them, "incomplete information is incorrect". The mediator in this phase was the visual representation. Adoption of a more refined presentation enabled different stories to be told while also granting different tiers of access: a basic level with key information and a more refined level for analysis.

The third phase was highly interactive but with fewer tensions. Here, the political view of the leader tended to prevail, notwithstanding the rational view brought by the analysts, based on the knowledge they acquired about the context of data use. For example, instead of using arbitrary categories for the lens called ‘City Magnets’, analysts suggested adopting categories already employed by a specific social media. This was considered a more rational choice and arguably a better way to engage citizens.

7. Analysis: four theses, four stories

The four research questions derived from our four quadrants shed different light on the big data use processes as described.

7.1. The information optimization thesis

Does big data provide data analysts opportunities to provide better information to decision makers?

In both of our cases the very idea of using big data for decision-making was based on the view that big data has added value over traditional data. This added value was very clear with regard to data collection. In Tilburg, the variety of data types and data sources enhanced the quality and functionality of the map that provided an overview of crime in the city. To this end, datasets were combined that had not been combined before. It was the job of data analysts to identify crime trends and present these to the concerned officials, not least, the high-level ‘triangle’ responsible for public safety. Furthermore, experts and residents were consulted to confirm the accuracy of the analyses using the dashboard. In Milan, too, the variety of data sources provided extra information, in this case, about Milan's position as an international city.

In both cases data collection was driven by data analysts rather than decision makers. In Milan, data analysts' big data knowledge was used in categorizing definitions, to manage the scope of the project, and in aligning with social media formats in order to engage citizens. Hence, the first thesis can be confirmed: big data does provide data analysts opportunities to better serve decision makers.

7.2. The decision optimization thesis

Does big data provide decision makers opportunities to better absorb information from data analysts?

In Tilburg the dashboard played an important role in enabling decision makers to absorb the information that data analysts provided. Thanks to the dashboard, a targeted, area-based approach could be used instead of a scattergun approach to combat crime. The dashboard offered a variety of information – where the top-100 offenders lived, types of houses most affected and characteristics of areas – leaving users the opportunity to combine them. Still, some preconditions were required for this to work. First, it required a good understanding of both the city and the nature of the criminals. Second, privacy issues had to be solved. Finally, the dashboard could not be considered the final answer. It provided and ordered data, but did not offer explanations. Additional district visits were therefore conducted. This suggests that while visualization and representation tools are important in transforming big data insights to use, they cannot close the gap entirely. Tacit knowledge and the professional expertise of street level bureaucrats is deeply embedded in policy and decision-making practices and not easily substituted by analysts' capacity to make sense out of big data.

In Milan, too, decision makers got a richer picture of the city, and this promoted data absorption. The main motive for this project was much more political: the data were used to frame Milan as an international city on the eve of a mayoral election. In this context, the project leader wanted “a story to tell” to politicians. This proved hard with the data. As the leader admitted, “the data don't tell me anything”. However, with mediation by designers, who created visuals, the connection between data analysts and decision makers was re-established. This again points to the importance of the presentation of insights, while showing how decision makers can aid analysts in translating data into insights that are of use for decisions (whether this is desirable is another matter).

Thus, we found that big data indeed provides decision makers opportunities to better absorb information from data analysts. However, the usefulness of information is not just a function of the sheer quantity or quality of data. Extra effort has to be taken to connect data to an envisioned use. Without this connection, data are hard for decision makers to digest to achieve their ends.

7.3. The politics of algorithms thesis

Does big data provide data analysts opportunities to pursue their own interests while providing information to decision makers?

In Milan, data analysts made politically significant decisions. Though the leader wanted a broad scope, analysts defined categories in order to facilitate analysis. Data analysts also developed algorithms to tackle the ambiguity of the spatial unit of analysis (Milan). Wittingly or not, these categorizations and definitions had considerable impact on the outcomes of the analyses, yet were outside the decision maker's reach.

A further observation is that data analysts explicitly anticipated on decision makers' political needs. In Milan, this was the “strong story to tell”. In Tilburg, it was the priority given to certain areas or crimes. This can be viewed as rational, since analysts' role is to serve decision makers, who in turn have their own political ends. However, the velocity of big data puts data analysts in a rather firm position, since they are quicker in digesting data than decision makers. A nice illustration of this is the comment of the analyst in Tilburg who went to the ‘triangle’ stating, “if we want to score on street robberies, we need to push here”. This example underlines the added value of big data for decision makers, but it also demonstrates that decision makers can be put in a reactive role, reduced to pushing a button at the bidding of an analyst.

The cases suggest that data analysts have considerable power in constructing information out of data and in framing it to political needs. Of course it is difficult to detect analysts' intentions. Data analysts are hired to serve decision makers, regardless of the decision makers' aims. However, if they had other interests, there would be ample opportunities to follow them. In this regard, the politics of algorithm thesis is confirmed.

7.4. The information market thesis

Does big data provide decision makers opportunities to pursue their own interests while absorbing information from data analysts?

Big data's potential for decision making is much stressed in the literature and confirmed by our two cases. However, this does not necessarily mean that decision makers prefer information derived from big data, even if it is digestible. In Tilburg, the usefulness of the dashboard depended in part on the definition of public safety being applied. The data concerned objective safety. However, decision makers sometimes received signals of a subjective feeling of danger or threat among residents. These subjective signals, though difficult to capture with big data, were sometimes considered even more important than the objective dashboard data. The mayor, as a result, at times bypassed the dashboard to prioritize activities responding to signals from society or personal preferences.

In Milan the project leader went further. He interfered in both data collection and in the construction of information. First, he excluded data sources for privacy reasons. Second, he abandoned a benchmark that did not show his city in the desired light. Finally, he insisted upon relaxing the temporal data resolution, because he perceived a risk in presenting real-time information to a public audience. All this was against the wishes of the data analysts. This extends beyond the selective absorption of information from data analysts. The leader interfered in methodological considerations and even made data analysts compromise on their professional values.

This confirms the information market thesis. Decision makers are customers for the information provided to them, and sometimes they even proactively interfere in methods of data analysts, as we saw in the Milan case.

8. Conclusions: four legitimate theses

8.1. The views and logics revisited

Big data promises to transform public decision-making for the better, by making it more responsive to actual needs and policy effects. However, this promise is informed by a rational view of decision-making, which has been much criticized in the public administration debate. This paper applied this view, and a more political one, to the context of big data and provided a qualitative study. We questioned the impact of big data on decision-making, realizing that big data – including its new methods and functions – must inevitably encounter existing political and managerial institutions. Specifically, we looked at the interaction between data analysts and decision makers. We distinguished a rational view and a political view and applied an information logic and a decision logic. Combining the categories produced a framework consisting of four theses on the impact of big data on public decision-making.

- The 'information optimization' thesis suggests that big data impacts decision-making by enabling data analysts to provide better information to decision makers.
- The 'decision optimization' thesis suggests that big data impacts decision-making by enabling decision makers to better absorb information from data analysts.
- The 'politics of algorithms' thesis suggests that big data impacts decision-making by enabling data analysts to pursue their own interests while providing information to decision makers.
- The 'information market' thesis suggests that big data impacts decision-making by enabling decision makers to pursue their own interests while absorbing information from data analysts.

It is not possible to establish hard conclusions based on two case studies. The purpose of the case studies was learning rather than establishing facts (Yin, 2009). The following observations shouldn't therefore be interpreted as evidence. They however point us towards interesting reflections on literature and research directions.

The cases showed that big data does provide data analysts opportunities to serve decision makers while also pursuing their own interests. The same holds true for decision makers. Big data provides them opportunities to pursue their own interests, though under certain conditions, while absorbing information from data analysts. This means that all of the theses legitimately but partially answer the question of how big data impacts public decision-making processes. In other words, they do not exclude one another, and to fully understand the relationship between data analysts and decision makers, all four views are needed. The rational view helps to identify how big data enables decision makers to do their jobs better. The political view helps reveal how big data empowers data analysts and decision makers to influence one another.

Our observation that both data analysts and decision makers are enabled and empowered by big data underlines the importance of distinguishing the decision logic from the information logic. The four theses, as defined here, are all fruitful and point to distinct issues regarding the impact of big data on decision making. This suggests that the findings of scholars on big data depend on the perspective chosen. Our theoretical framework helps to identify these perspectives, which will facilitate debate on big data.

8.2. Data analysts and decision makers as autonomous agents

We looked at decision-making as a process that eventually results in some choice. By focusing on the process leading up to that choice, we demonstrated that big data's impact cannot be interpreted merely as functional, as often argued (e.g., Chen et al., 2012; Choi, 2012; Omalley, 2014). Big data provides both decision makers and data analysts expanded means to add value to the decisions made. However, before a choice is made, it also provides them more means to interpret this added value; and their interpretations are not necessarily aligned. Thus, Simon's (1960) famed observation, that within decision-making processes the phase of design is significant for the choices made, could well be particularly valid in the data revolution era. In other words, big data enables both data analysts and decision makers to choose before a final choice is made.

We found, indeed, politics before, during and after the rational analysis. Political analysis constructs the framework for the rational analysis (before), and processing data into information raises political considerations (during), and decision makers need to be able to attribute projects as a success (afterwards). Big data enables data analysts to choose, as they can construct information from data and frame it to fit political needs. The velocity of big data reinforces this quality, because big data analysts are quicker in handling data than decision makers are. Yet, decision makers can still interfere in the professional domain of data analysts. For example, they might impose preconditions for data collection, such as privacy considerations, and add a desirable slant to information construction.

This implies that big data enables both data analysts and decision makers to be autonomous agents rather than links in a functional chain. This calls into question traditional representations of decision-making processes, as these are commonly depicted as a sequence of activities. Our cases revealed that the interfaces between links – here, between data analysis and decision-making – become more important and deserve an explicit position in these representations. It is at these interfaces that data are translated into relevant information that decision makers can digest. Yet, political and legal sensitivities also find their way to data analysts. Neither data analysts nor decision makers are trained for the intensive interaction on these interfaces that we found in our cases. Moreover, a significant role was played by intermediaries, such as the designers who could visualize the data.

9. Future studies: a new but still developing power balance

We found that big data has the potential to enable and empower both data scientists and decision makers for public decision making. As a consequence, those viewing big-data-informed decision making as merely a sequence of activities will probably miss the institutional dimension of big data. In this section we will discuss some implications for an institutional research agenda. It includes a discussion of both the results and the conclusions, and is therefore positioned after the conclusion section.

9.1. Professional norms and values

We found that data analysts were politically significant in all phases. They coordinated the data-generation process. They had a crucial role in constructing information from the data. Their definitions were politically significant. They facilitated the framing, even the adjustment, of information to political preferences. We also found that the powers of data analysts were not unlimited. Data analysts did their work in the context of predefined policies. Decision makers sometimes neglected or bypassed big data analysis. They used information selectively. Selection criteria were sometimes political. Decision makers and data analysts both anticipated on political requirements. This suggests that in the cases under study some kind of power balance had been achieved. Still, this power balance may shift. Decision makers may be urged to decide more quickly because of the availability of real-time data. For this they must trust these data and the data analysts processing them. In fact, the trust relation has changed considerably. Data analysts nowadays are different from their forerunner policy analysts, who used to be the main integrators of data. In the big data era the computer is the integrator, run by data analysts. Transparency, too, is at issue, because of the volatility of data and the complexity of data processing.

To really understand how the power balance is shifting, we must look at the way data analysts and decision makers compromise on their professional norms. The Milan case revealed that data analysts and decision makers have different aesthetics and values regarding the big data use process. The decision maker saw the use of algorithms as manipulation, while the data analysts considered the non-use of available data sources unprofessional. In the Milan case, data analysts compromised on their professional values multiple times to make the big data project successful. Other cases may show the same trade-offs, albeit perhaps with different outcomes, as the push to use big data may force decision makers to compromise on their professional position as well.

9.2. An impact on legitimacy?

Looking further, the increasing importance of big data may evoke changes in whose values get institutionalized. We found that decision makers had discretionary freedom to include, exclude, use or ignore information derived from big data projects. As such, big data does not seem to have had a big impact on the way decision makers legitimized their decisions. One could argue that this is a good development, and that it strengthens the role of evidence in policymaking. Politicians and managers are required to weigh values – such as public safety, efficiency and sustainability – and make judgements aligned to both hard information and softer considerations. This constitutes the core of their legitimacy.

Up until now, big data projects have been initiated to serve specific values, and so have been unsuited to weigh values against one another. Moreover, as suggested by the Tilburg case, big data projects shed light on only parts of a problem, as public safety has both an objective and a subjective side, and big data illuminates the objective side only. This does suggest a reason for concern; that is, techno-optimism could promote big-data-informed decisions over other decisions.

9.3. Follow up: the dynamics of interaction

These findings are the result of a limited, qualitative study meant to explore the field of big data and decision-making and provide structure for the debate. One limitation is our assumed uniformity of 'decisions'. Decisions may be descriptive (like in our Tilburg case), explorative (like in our Milan case), or predictive or prescriptive. Decision-making processes and the relations within them may differ per decision type. This suggests that there is still much work to do. As a follow up, we suggest a more extensive investigation of the interaction dynamics between data analysts and decision makers. A special topic of interest is the way data analysts and decision makers anticipate on each other's preferences. They did so quite explicitly in our cases, though sometimes unwittingly. This tells us much about information asymmetries and transparency in big data processes, and may be better understood with insights from behavioural sciences and by studying the phenomenon at multiple levels – such as the organizational, process and individual levels.

Acknowledgements

We would like to thank Tom Daalhuijsen for collecting most of the data for the Tilburg case as part of his MSc Thesis. Furthermore, this paper is partially supported by the research project "Governing Public–Private Information Infrastructures", financed by the Netherlands Organisation for Scientific Research (NWO) as Veni grant 451-13-020.

References

Ackoff, R. L. (1989). From data to wisdom. *Journal of Applied Systems Analysis, 16*, 3–9.
Adrian, M. (2011). *Information management goes' extreme: the biggest challenges for 21st century CIOs.* SAS: Business Analytics and Intelligence Software.
Aldrich, H. (1999). *Organizations evolving.* London: SAGE Publications Ltd.
Arnaboldi, M., Busco, C., & Cuganesan, S. (2017). Accounting, accountability, social media and big data: revolution or hype? *Accounting, Auditing & Accountability Journal, 30*(4), 762–776. https://doi.org/10.1108/AAAJ-03-2017-2880.
Asselt, M. B. A.v. (2000). Perspectives on uncertainty and risk - the PRIMA approach to decision support. *Perspectives on Uncertainty and Risk, 443*. https://doi.org/10.1007/978-94-017-2583-5_10.
Axelrod, R., & Hamilton, W. D. (1981). The evolution of cooperation. *Science, 211*(4489), 1390–1396. https://doi.org/10.1126/science.7466396.
de Bruijn, H., & ten Heuvelhof, E. (2008). *Management in networks.* https://doi.org/10.4324/9780203885666.
Chatfield, A. T., & Reddick, C. G. (2017). Customer agility and responsiveness through big data analytics for public value creation: A case study of Houston 311 on-demand services. *Government Information Quarterly*. https://doi.org/10.1016/j.giq.2017.11.002.
Chen, H., Chiang, R. H. L., & Storey, V. C. (2012). Business intelligence and analytics: From big data to big impact. *MIS Quarterly, 36*(4), 1165–1188. https://doi.org/10.1145/2463676.2463712.
Chen, M., Mao, S., & Liu, Y. (2014). Big data: A survey. *Mobile Networks and Applications, 19*(2), 171–209.
Chen, Y.-C., & Hsieh, T.-C. (2014a). Big data for digital government. *International Journal of Public Administration in the Digital Age, 1*(1), 1–14. https://doi.org/10.4018/ijpada.2014010101.
Chen, Y.-C., & Hsieh, T.-C. (2014b). Big data for digital government: Opportunities, challenges, and strategies. *International Journal of Public Administration in the Digital Age, 1*(1), 1–14. https://doi.org/10.4018/ijpada.2014010101.
Choi, H. (2012). Predicting the present with Google Trends. *The Economic Record, 88*(s1), 2.
Cohen, M. D., March, J. G., & Olsen, J. P. (1972). A garbage can model of organizational choice. *Administrative Science Quarterly, 17*(1), 1. https://doi.org/10.2307/2392088.
Crawford, K., Miltner, K., & Gray, M. (2014). Critiquing big data: Politics, ethics, epistomology. *International Journal of Communication, 8*, 1663–1672.
Davenport, T. H., Barth, P., & Bean, R. (2012). How big data is different. *MIT Sloan Management Review, 54*(1), 43–46.
Feldman, M. S., & March, J. G. (1981). Information in organizations as signal and symbol. *Administrative Science Quarterly, 26*(2), 171. https://doi.org/10.2307/2392467.
Fidel, R. (2004). A multidimensional approach to the study of humaninformation interaction: A case study of collaborative information retrieval. *Journal of the Association for Information Science and Technology, 55*(11), 939.
Gantz, J., & Reinsel, D. (2011). Extracting value from chaos. *IDC Iview, 1142*, 1–12.
Gartner. (n.d.). Gartner IT glossary. Retrieved December 28, 2017, from http://www.gartner.com/it-glossary/big-data/
Goel, S. (2010). Predicting consumer behavior with Web search. *Proceedings of the National Academy of Sciences, 107*(41), 17486–17490.
Bruijn, H., & Heuvelhof, E. (2008). *Management in Networks.* https://doi.org/10.4324/9780203885666.
Helbig, N., Dawes, S., Dzhusupova, Z., Klievink, B., & Mkude, C. G. (2015). Stakeholder engagement in policy development: Observations and lessons from international experience. *Policy Practice and Digital Science*. https://doi.org/10.1007/978-3-319-12784-2_9.
Höchtl, J., Parycek, & Schollhammer (2016). Big data in the policy cycle: Policy decision making in the digital era. *Journal of Organizational Computing, 26*(1–2), 147.
Hota, C., Upadhyaya, S., & Al-Karaki, J. N. (2015). Advances in secure knowledge management in the big data era. *Information Systems Frontiers, 17*(5), 983–986. https://doi.org/10.1007/s10796-015-9593-y.
Jann, W., & Wegrich, K. (2007). Theories of the policy cycle. In F. Fischer, & G. J. Miller (Eds.). *Handbook of public policy analysis: Theory, politics and methods* (pp. 43–62). Boca Raton: CRC Press.
Jansen, M., & Meckling, F. (1976). Theory of the firm: Managerial behavior, agency, and ownership structure. *Journal of Financial Economics, 4*(4), 305–360.
Janssen, M., & Kuk, G. (2016). Big and Open Linked Data (BOLD) in research, policy, and practice. *Journal of Organizational Computing and Electronic Commerce, 26*(1–2), 3–13. https://doi.org/10.1080/10919392.2015.1124005.
Janssen, M., van der Voort, H., & Wahyudi, A. (2017). Factors influencing big data decision-making quality. *Journal of Business Research, 70*. https://doi.org/10.1016/j.jbusres.2016.08.007.
KING (2014). Big data op kleine schaal: data en gemeenten. Retrieved from http://magazine.kinggemeenten.nl/data-en-gemeenten.
Kitchin, R. (2014). *The data revolution: Big data, open data, data infrastructures and their consequences.* London: SAGE Publications Ltd.
Klievink, B., Romijn, B., Cunningham, S., & Bruijn, H.d. (2017). Big data in the public sector: Uncertainties and readiness. *Information Systems Frontiers, 19*(2), 267.
Klijn, E., & Koppenjan, J. (2015). *Governance networks in the public sector.* London: Routledge.

Kogan, M. (1999). The impact of research on policy. In F. Coffield (Ed.). *Research and policy in lifelong learning* (pp. 11–18). Bristol: Policy Press.
Koppenjan, J. M. F., & Klijn, E. H. (2004). Managing uncertainties in networks: A network approach to problem solving and decision making. Retrieved from http://scholar.google.com/scholar?q=Koppenjan+and+Klijn,+2004#0.
Kuit, M. (2002). *Strategic behavior and regulatory styles in the Netherlands energy industry.* Delft: Eburon.
Lasswell, H. (1956). *The decision process: Seven categories of functional analysis*.
Lindblom, C. (1959). The science of "muddling through". *Public Administration Review,* 79.
Lueg, C. (2001). Information, knowledge, and networked minds. *Journal of Knowledge Management, 5*(2), 151–160. https://doi.org/10.1108/13673270110393194.
Marx, V. (2013). Biology: The big challenges of big data. *Nature, 498*(7453), 255–260.
Mayer-Schönberger, V., & Cukier, K. (2013). *Big data. A revolution that will transform how we live, work and think.* London: John Murray Publishers.
Milakovich, M. E. (2012). Anticipatory government: Integrating big data for smaller government. *Internet, politics, policy 2012: Big data, big challenges*.
Nachmias, D., & Felbinger, C. (1982). Utilization in the policy cycle: Directions for research*. *Review of Policy Research, 2*(2), 300–308. https://doi.org/10.1111/j.1541-1338.1982.tb00676.x.
Nardi, B., & O'Day, V. (1999). *Information ecologies: Using technology with heart*.
Omalley, M. (2014). Doing what works: Governing in the age of big data. *Public Administration Review, 74*(5), 555.
Quinn, R., Bright, D., Faerman, S., Thompson, M., & McGrath, M. (2014). *Becoming a master manager: A competing values approach*.
Scharpf, F. (1993). Coordination in hierarchies and networks. *Games in hierarchies and networks. Analytical and empirical approaches to the study of governance institutions* (pp. 125–165). Boulder: Westview Press.
Scharpf, F. W. (1997). *Games real actors play. Actor-centered institutionalism in policy research.* Oxford: Westview Press.
Schintler, L. A. (2014). Big data for policy analysis: The good, the bad, and the ugly. *Review of Policy Research, 31*(4), 343.
Simon, H. A. (1960). *The new science of management decision.* New York: Harper Brothers.
Simon, Herbert (1997). *Models of bounded rationality: Empirically grounded economic reason. Vol.3*. Cambridge: MIT Press.
Simon, P. (2013). *Too big to ignore: The business case for big data.* Hoboken, New Jersey: John Wiley & Sons Inc.
Steenhuisen, B., & van Eeten, M. (2008). Invisible trade-offs of public values: Inside dutch railways. *Public Money & Management, 28*(3), 147–152. https://doi.org/10.1111/j.1467-9302.2008.00636.x.
Stewart, J. (2006). Value conflict and policy change. *Review of Policy Research, 23*(1), 183–195. https://doi.org/10.1111/j.1541-1338.2006.00192.x.
Stonier, T. (2012). *Information and meaning: An evolutionary perspective.* New York: Springer.
Teisman, G. (2000). Models for research into decisionmakingprocesses: On phases, streams and decision making rounds. *Public Administration, 78*(4), 937.
Thacher, D., & Rein, M. (2004). Managing value conflict in public policy. *Governance, 17*(4), 457–486. https://doi.org/10.1111/j.0952-1895.2004.00254.x.
Tversky, A. A., & Kahneman, D. (1981). The framing of decisions and the psychology of choice. *Science, 211*(4481), 453–458.
Uprichard, E. (2015). Most big data is social data. Retrieved December 28, 2017, from http://blogs.lse.ac.uk/impactofsocialsciences/2015/02/12/philosophy-of-data-science-emma-uprichard/.
Wildavsky, A. (1978). A budget for all seasons? Why the traditional budget lasts. *Public Administration Review, 38*(6), 501.
Yin, R. K. (2009). *Case study research: Design & methods.* https://doi.org/10.1007/BF01103312.

Haiko van der Voort is an assistant professor at the department of Policy, Organization, Law and Gaming, Faculty Technology, Policy and Management of Delft University of Technology. He earned his Phd with an extensive study on co-regulation, which is a collaboration effort among public regulators, regulated industries, and standardization organizations. He has developed an assessment framework for co-regulation, that respects both public and private interests, as well as the dynamics of co-regulatory efforts. His recent work concerns ict-governance, including the governance of data sharing standards, decision making on ict-projects by the Dutch government, and the governance of data quality in the big data era.

Bram Klievink is Associate Professor of Collaborative Digital Governance and Head the Organization & Governance section in the Multi-Actor Systems (MAS) department at the Faculty of Technology, Policy and Management. He studies governance in digital-induced institutional voids and specialise in collaborative governance arrangements for the digital age. He studies this topic in the domains of electronic public services, international trade and logistics, financial cybersecurity, infomobility, smart cities and big data.

He has been and still is involved in multiple European FP7 and NWO projects, in various capacities, including as (co)-PI, project leader, scientific coordination, supervisor and researcher. Furthermore, he was a board member of the Dutch Alliance for Vital Governance, a strategic research alliance for collaboration and knowledge transfer between government organizations and research institutes in the Netherlands.

Michela Arnaboldi is full professor at Politecnico di Milano, Department of Management, Economics and Industrial Engineering. In 2008 she has been visiting professor at the Centre for Analysis of Risk and Regulation (CARR) at the London School of Economics. She is member of the scientific committee of the l'Institute of Public Sector Accounting Research of the University of Edinburgh. Her research activity is mainly focused on two areas: (1) Managing performance in the Public Sector; (2) the evolution of Management Control Systems: risk and networks.

Regarding the first area - Managing performance in the Public Sector – three issues have been developed: (i) analysis of performance management system in higher education, (ii) exploration of local authorities reforms in Local authorities, (iii) development of management models in Cultural Heritage.

The second area - the evolution of Management Control Systems: risk and networks has been carried out along two dimensions: (i) analysing of the relationship between budgeting system and enterprise risk management, (ii) investigation of the role of management control system in private-public networks.

Albert Meijer is Professor of Public Innovation at the Utrecht School of Governance (Utrecht University). He is the chair of the permanent studygroup on e-government of the European Group for Public Administration and on the editorial board of Information Polity and the Electronic Journal of E-Government Studies. His current research program on public innovation revolves around the transformation of government in an information age and ranges from virtual organizations to electronic democracy. He has published extensively in top journals in the field of public administration and government information studies. He collaborates with researchers from various countries on issues of transparency, e-government and innovation.