Data science empowering the public: Data-driven dashboards for transparent and accountable decision-making in smart cities

Ricardo Matheus*, Marijn Janssen, Devender Maheshwari

Delft University of Technology, Faculty of Technology, Policy and Management, Jaffalaan 5, 2628 BX Delft, The Netherlands

ARTICLE INFO

ABSTRACT

Dashboards visualize a consolidated set data for a certain purpose which enables users to see what is happening and to initiate actions. Dashboards can be used by governments to support their decision-making and policy processes or to communicate and interact with the public. The objective of this paper is to understand and to support the design of dashboards for creating transparency and accountability. Two smart city cases are investigated showing that dashboards can improve transparency and accountability, however, realizing these benefits was cumbersome and encountered various risks and challenges. Challenges include insufficient data quality, lack of understanding of data, poor analysis, wrong interpretation, confusion about the outcomes, and imposing a pre-defined view. These challenges can easily result in misconceptions, wrong decision-making, creating a blurred picture resulting in less transparency and accountability, and ultimately in even less trust in the government. Principles guiding the design of dashboards are presented. Dashboards need to be complemented by mechanisms supporting citizens' engagement, data interpretation, governance and institutional arrangements.

1. Introduction

Governments are more and more utilizing data in all aspects of their functioning. *Data science* in government deals with the extraction, interpretation and presentation of insights from unstructured and structured data that can be either closed or opened. An important area of data science is to visualize the data in dashboards. Cleveland (2001) argues that data science consists of multidisciplinary investigations, models and methods for data, computing with data, pedagogy, tool evaluation, and theory. Government data scientists need in-depth knowledge of statistics and data analytics for analyzing data, as well as knowledge on the use of techniques and instruments for predictive purposes and to visualize the results. By combining disciplines, new insights and applications can be created and communicated using dashboards. Nevertheless, data scientist also need to have an understanding of other elements like the policy-making, organization, legislation and public values. This knowledge allows them to positioning the data in the context and to understand its use and implications.

Data science is an essential area for governments, as they collect a lot of data in various areas (geographical, traffic, social security, energy, etc.) that can be combined or enriched with data from smart devices and other sources such as discussion forums, social media, and private sector data (Janssen, Matheus, & Zuiderwijk, 2015). The making of sound decisions depends on the use of high-quality data (Chengalur-Smith, Ballou, & Pazer, 1999). Data might be an enabler for creating new innovative applications (Marsh, Pane, & Hamilton, 2006) to improve public values like security, safety, transparency and accountability.

In data science, the sharing, use and interpretation of data are key aspects in bridging the gap between the government and the public. Platforms can be created to share data (Brown, Fishenden, Thompson, & Venters, 2017). The use of data and the accompanying instruments will likely influence government policy-making, resulting in new applications, but can also impact the interaction with the public (Ganapati, 2011a). Dashboards can be used to release information for governmental decision-makers (Maheshwari & Janssen, 2014), but also for the public to scrutinize government actions, to engage in the decision-making processes and to improve decision-making. Dashboards should help to facilitate transparency, governance, trustworthiness and enable citizens' to participate in decision-making in smart cities (Allio, 2012).

Recently, data science and dashboards have gained more and more attention in the public sector. A dashboard is often created by having a webpage which visualized all kinds of data given a certain purpose. For example, in 2009, the US federal government developed dashboards with federal stimulus funding (www.recovery.gov) aiming transparency

and accountability of national economic recovery policy (Ganapati, 2011b). The Dutch government uses dashboards to enable the monitoring of large IT-projects by the public (www.rijksictdashboard.nl). As such, dashboards are becoming an important means of communicating and interacting with the public to create transparency and to achieve accountability. The latter refers to the answerability for one's actions or inactions and to be responsible for their consequences (Roberts, 2002).

In daily life, a dashboard is a control panel of the driver of a car in which the driver can see what happens and has controls to intervene. Few (2006) defined dashboards as "*a visual display of the most important information needed to achieve one or more objectives, consolidated and arranged on a single screen so the information can be monitored at a glance*" (p. 34). This definition might be challenged, as governments might have multiple sometimes conflicting objectives, and it does not acknowledge the ability to intervene. Furthermore, dashboards can enable the zooming in on more detailed information and might not be limited to a single screen. Therefore, in this paper we define dashboards as "*the visualization of a consolidated set data for a certain purpose, which enables to see what is happening and to initiate actions*". The purpose varies depending on the focus on the public or for use for policy-making. In our definition the possibility to intervene is a key aspects, as only viewing data without having the possibilities to take actions based on the result has limited use and might result in the abandon of the dashboards. Dashboards are often part of public organizations' '*open government*' efforts, which aim at creating transparency and stimulating engagement with citizens and business. The efforts of open government can result in a more democracy (Cuadrado-Ballesteros, 2014), efficiency (Navarro-Galera et al., 2016), and transparency, accountability, collaboration and engagement (Bertot, Jaeger, & Grimes, 2010; Dawes & Helbig, 2010) and trust into the government (Lourenço, 2015). Despite the promises, employing dashboards is often a difficult endeavor. Data is often context-specific, and without in-depth knowledge of the context in which the data is collected, interpretation will likely be wrong (Matheus & Janssen, 2013). As such, data science in government requires in-depth skills and knowledge about the inner workings of the governments and its environment (McAfee, Brynjolfsson, Davenport, Patil, & Barton, 2012).

The objective of this paper is to understand and to support the design of dashboards for creating transparency and accountability. Literature and two case studies are investigated to identify benefits, risks and principles for designing dashboards in the public sector. Following these principles can lead in realizing the benefits and overcoming the risks. A cycle describing the data cycle of open and private data for dashboards is presented to show how value can be created from data.

This research is conducted by reviewing literature about dashboards and complementing the literature by investigating two case studies in detail. The literature review was performed using the top 20 journals of 2015 in the Scimago Journal Rank (SJR) in the Information System (IS), Information Systems and Management, and, Library and Information Science. All the journals were surveyed, using the keyword "dashboard". Around 130 papers were found, considering title, abstract, keywords and citations. The major part of the papers mentioned only dashboard in the text, whereas the research was not focused on dashboards. Only 9 of these papers were related to smart cities. The scant literature shows that dashboards suggests that dashboards are underexplored, whereas they are essential in data science.

This paper is structured as follows. In Section 2, two different cases are presented in which both the government and the public plays a major role. The cases and literature are used to understand the value creation mechanisms, benefits, risks and challenges when designing and using dashboards, which is presented in the Section 3. Section 4 presents design support for developing dashboard in the form of design principles and a data cycle for dashboard. Finally, conclusions are drawn in Section 5.

2. Dashboards in Smart City practice

Due to the economic boom and access to credit in Brazil, there has been an increase in the number of cars (3 million cars for 6.5 million inhabitants) resulting in huge traffic jams, even outside of rush hours. The Smart City dashboards in Rio de Janeiro, Brazil, were created to solve problems related to public transportation and traffic. For this an infrastructure, a dashboard, and a data portal with more than three thousand datasets and seven APIs for real-time data use (www.data.rio) were developed and used by the Center of Operations Rio (COR). COR is situated in a four floor building that reunites almost 30 local secretariats, public and private enterprises to identify and solve in real-time issues on the city. COR collects around 4 Gigabytes (GB) of data on transit every day. This includes data about bus stops, car accidents, constructions works, and accidents like tree falling down on the streets. This excludes real-time GPS (geographical location collected using the Global Positioning System – GPS) data coming from buses driving their routes. Every 2 min, data is collected from the position of over 8000 buses. This result in another dataset which amounts to some 12 GB per day. With data from COR and the work of the data scientist group Pensa dashboards are created with the objective of putting key organizational elements into a consolidated format using several visualization tools, gauges, graphs, charts, and pictograms. The Pensa (www.pensa.rio), a group of big data scientists at Rio de Janeiro City Hall, enabled COR to visualize the data in a structured, integrated, and organized manner at a glance. The process of data analyses used by Pensa in Rio de Janeiro is based on questions made by the political decision-makers (mayor and secretariats) to create dashboards.

2.1. Case 1: traffic dashboards

The first case represents a partnership between the Smart City of Rio de Janeiro and the Social GPS Smartphone application called 'Waze' (www.waze.com) to employ citizen-generated data combined with government data as shown in Fig. 1. This app allows citizens to send real-time information about traffic conditions and accidents to the city. The COR uses the Waze application to send real-time information about route changes, flood routes, traffic jams, and car accidents to its citizens. The result is a combination of open data from the Rio de Janeiro City Hall and user-generated content collected through Waze about Rio de Janeiro's 7 million inhabitants. This big data is presented in real time using electronic panels positioned all over the city.

Two dashboards were created based in the Big Data Analytics. The first dashboard is showed in a big video-wall remembering the North American Space Agency (NASA) Lyndon Johnson Center, in Houston, as an interactive map version at COR called GeoPortal (Dashboard A in Fig. 1). This dashboard shows in real time, 24 h per day, 7 days per week, where the traffic jams and accidents are. Combined with other data sources, including 900 cameras that can turn 360°, spread over the city recording 24 h of the traffic, allows public servants at the operational level to make decisions to solve traffic problems. Accidents street holes and other issues are easily identified with traffic jams and prescanned using video cameras. The most-suitable team to deliver the service and solve the problem will be called. For example, one team might be able to arrive faster, but has no health equipment or technical capacity to solve the issue. This helped also the local government to plan the long-term public service delivery based in frequency and type of events, positioning proper teams to solve long-term issues.

The second dashboard is based on the same data, but it is presented in different format to the public. Ten dashboards were built in the top-10 most congested parts and directions of avenues in Rio de Janeiro (e.g. Dashboard B in Fig. 1). The top-10 most congested places were identified with Big Data analysis from the data scientist group Pensa, taking into consideration data from several internal databases (GPS data from buses, traffic jams reported by civil servants, speed of traffic from speed traps, car accidents, and others) and the Waze application
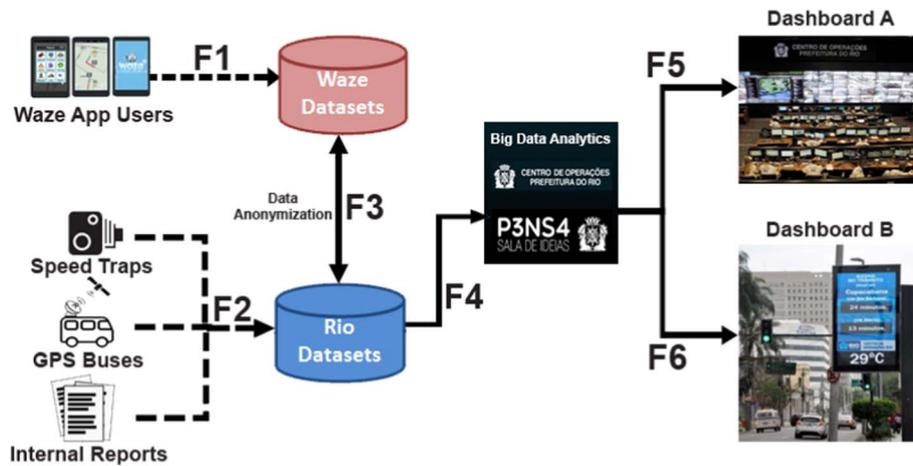
Fig. 1. Traffic dashboards.

(level of traffic jam per hour and accidents reports). These dashboards do not have any interactive map such as COR has in the video-wall, but just present to the citizens textual information about the two most suitable routes between peripheral and downtown regions (during the morning and the opposite in the afternoon), in order to avoid delays due to traffic jams. This enables the public to make a decision concerning the best route.

2.2. Case 2: public transport dashboard

The second case studied was a partnership with the Moovit social application for buses (www.moovitapp.com). This partnership had the objective of improving the quality of public transportation and transparency by showing in real-time the city's bus system. User-generated data is collected and combined with location data about buses. This data is integrated in a dashboard as shown in Fig. 2. Moovit empowers citizens to choose the best route (faster, closer, etc.) in accordance with real-time information on events (traffic jams, accidents, broken down buses, etc.) presented via the duration and distance to get from their starting to their finishing location. For commuter traffic this is normally between peripheral regions and downtown during the morning and in the afternoon the opposite direction, as well it happens with private cars and the first scenario "Traffic Dashboard" described above. This application used the public Application Programming Interface (API) from Rio de Janeiro city Hall at Open Data Portal www.data.rio to access the GPS data from buses. Based on this data, Moovit calculates the estimated time of arrival (ETA) of buses at a certain bus stop, as well as the distance between the starting and finishing location. The COR, the data scientist group Pensa, and the local government transport department have access to this data sets and are constantly using them

to monitor the public transportation, and also using for predictive purposes and planning future changes needed. Decision-making include the re-distribution of bus lines for big events such as World Cup in 2014 and Olympics games 2016.

COR has a dashboard that can create a "high level of data visualization", given to people, but also create dashboards drilling into details needed in accordance with each department and scenario. For example, a shortage of buses in one line can be used to reallocated buses from another line to this line. By checking in real-time which buses have low utilization, it is possible to reduce the numbers of buses in idle lines and use them other line, in his way balancing the overall system. In the long-term, a machine-learning algorithm to automatically identify how to balance the system based in several characteristics (i.e. bus utilization, speed, average fleet age, type of buses) can be created.

3. Understanding dashboards

The cases show that the design of a dashboard is dependent on many factors including the information available and its purpose. Although there are other purposes, dashboards are often aimed at empowering the public by creating transparency and accountability (Janssen & van den Hoven, 2015). Dashboards are an instrument for reducing information asymmetry (e.g. Bugaric, 2004). *Information asymmetry* is the situation in which one party has more information than another party (Jensen & Meckling, 1976). By making data available dashboards can help to reduce the information asymmetry by providing more insight into a certain situation. The overcoming of information asymmetry should result in higher levels of transparency. In the next subsection, we discuss the traditional view on value creation by employing dashboards. In the next subsections benefits, risks and challenges are
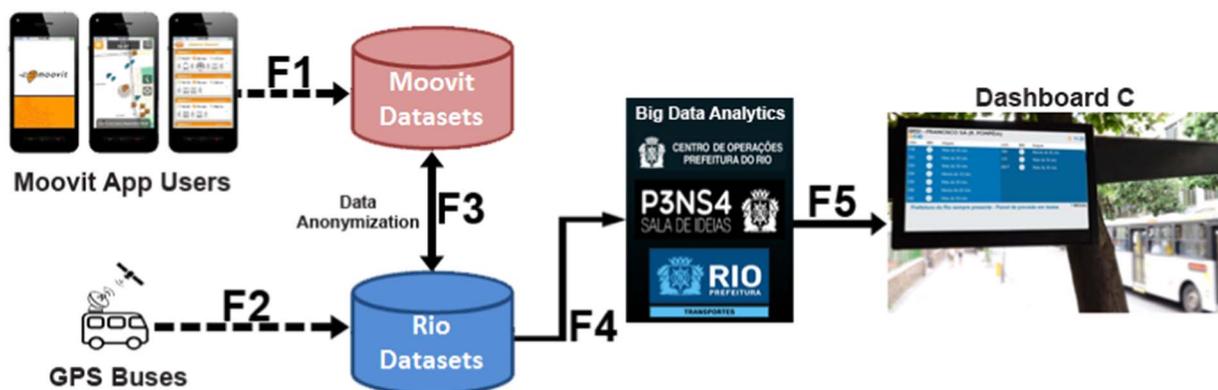

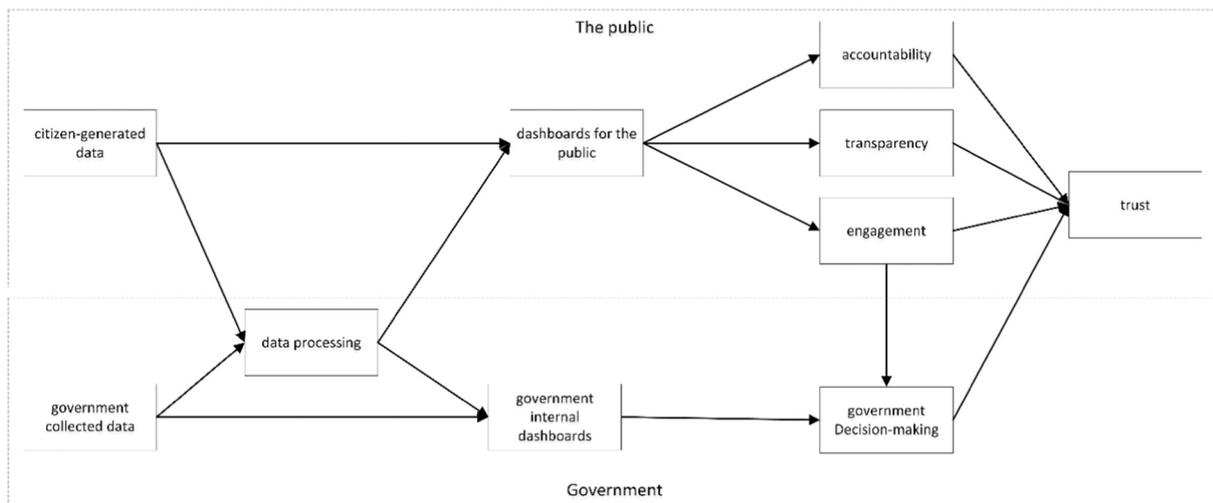**Fig. 2.** Moovit public transportation dashboards.

Fig. 3. Value creation logic behind dashboard.

discussed with challenge this view. Finally, principles guiding the design of dashboards overcoming the risks and challenges and capturing the benefits are discussed.

3.1. Value creation by dashboards

Dashboard can be used for decision-making by the government and the public (Allio, 2012). Data can be generated by citizens or collected by governments. Data might have varying data qualities and collected in different manners. Thereafter data need to be processed and visualized in dashboards. The use of dashboard should resulted in transparency and accountability and ultimately in more trust in the government (Harrison & Sayogo, 2014; Villeneuve, 2014). Also, the dashboard can stimulate citizen engagements. Governments might also develop dashboards for their own decision-making based on the input from citizen-engagement. Sound and evidence-based decisions and policies should result in more trust by the public. Fig. 3 visualizes the idea that is often assumed behind the creation of dashboards for smart cities and open government efforts in general. Yet, developing a proper dashboard is challenging.

Reality is often more cumbersome than the simple conceptualization in Fig. 3, as can be derived from the cases in the previous section. Only a blurred picture might be created. Furthermore, also distrust can be created by making mistakes or by making the problems of the government transparent to the public. For example, if the data in the dashboards for private cars to reduce traffic does not result in a reduction of the traffic, people might distrust the government and might avoid using the dashboard. If the dashboards show 15 min but the actual time spent was 30 min. Hence, there are many risk and challenges that will be discussed after having presented the benefits.

3.2. Benefits of public sector dashboards

Public organizations can use dashboards for a variety of purposes, including transparency, performance monitoring, reporting, planning, and policy-making. Dashboards can be designed for use by governments (internal) or the public (external). Internal objectives can be related to monitoring and analysis for faster and more accurate decision-making, resulting in increased efficiency and effectiveness of operations. External objectives are often related to creating transparency and accountability, mobilizing external capacity to gain feedback, and facilitating participation by the society.

There are many pitfalls that can prevent the achievement of the intended benefits. If data are not properly cleansed in advance, their usage can lead to inappropriate analysis. Higher data quality comes at a price and might not always be necessary. Data improvement techniques can solve part of the problem of low data quality. Proper vocabularies for metadata (ontology, semantic web, etc.) can solve another issue related to understanding and interpretation. The datasets must also have proper information quality like, timeliness and granularity, and system quality, like being able to be accessed based in different formats. The possibility to drill into the details of time, date, place, and description is as important as having the high-level statistics of, for example, car accidents. Both data can be used for the same issue, but with different objectives. Whereas high-level statistics can show the trends of a city traffic jam and identify possible bottlenecks, detailed data can give insight for how to overcome traffic congestion for a driver. Finally, data can be visualized in the form of tables, graphs, or at a glance in heat maps on the dashboards.

Dashboards can help speed up decision-making for both policymakers and the public. In this way dashboards can improve the effectiveness and efficiency of public policies and operations, like reducing traffic jams and car accidents. This can mobilize external knowledge to use the open government data and combine these with other datasets to create dashboards. The dashboards also enable public participation in decision-making and the improvement of services. Using electronic panels to show the average time need to get from one place to another one utilizing two different routes results in better decision-making for citizens to select the most appropriate route. Such transparency is better than government enforcing an option. In addition, the public can use this information to evaluate, provide feedback, and suggest improvements. Similarly, government can benefit by understanding the daily decision-making (planning) by people and influencing their behavior by displaying information on smart located electronic panels spread over the city. The use of these panels resulted in more transparency by reducing the information asymmetry between citizens and public agencies. As a result of this transparency the of transit time of citizens was reduced up to 15% time savings by travelers in rush hours. The list and descriptions of public dashboard benefits as found in the 2 cases are presented in Table 1.

3.3. Risks and challenges

Dashboards can provide both an overview and detailed views at the same time and might create transparency and can be used for accountability and to stimulate engagement. The usage of data on dashboards by the public does not result in transparency and accountability per se. Benefits can only be gained if dashboards are properly designed. For this it is essential to understand the risks and challenges encountered when designing dashboards.

Table 1
Overview of the main benefits.

Benefits	Description
Strategic and political benefits	
1. Overview at a glance	Dashboards permit flexibility to create various types of overviews at glance. This enables different stakeholders to gain insight easily.
2. Drill into detail	The same flexibility that allows a broad overview also can be used to drill into detail for any event. During traffic monitoring, it is possible to zoom in on accidents.
3. Transparency	Creating a view at a glance can provide transparency about a situation, combined with the ability to drill down to details to really understand the data and conclusions that might be inferred from it.
4. Customer/user-orientated presentation/ visualization oriented	Dashboards enable the creation of customized views. In this way dashboards provide information to people at the time and place they really need it.
5. Identification of fraud and corruption	Anomalies and patterns in data can be used to detect corruption by showing them in dashboards combined with data analytics.
6. Visualizing trends	The information presented in dashboards concerns the trends or issues in the city. It is not necessary to know what is still good, but rather what is needed to avoid the worst-case scenario.
7. Accountability	The insights can be used to help hold organizations or people accountable for their actions (or inactions). This requires the possibility of following up the results from the data analysis at the institutional level.
Operational benefits	
8. Better and faster decisions and policy-making	Real-time data can enable operational strategies based on specific situations and immediate decisions.
9. Mobilize external knowledge	External partners such as smartphone applications and media (radio, TV) can use and spread the same information used on dashboards, collaborating for better effectiveness and efficiency (for instance, operational traffic conduction).
10. Improving the effectiveness and efficiency of operations	Transparency of relevant information on dashboards in real time can improve the efficiency of operations with better management of many situations such as reduction of traffic congestion. The efficiency makes people trust the dashboard, improving the effectiveness of operations in the short and long term.
11. Disclose relevant information to people	Dashboards give people relevant information for real-time decision-making. If not used, they can be just an expensive data storage system.
12. Enable participation	Relevant information on dashboards gives people the opportunity to participate in public decision-making and policy-making.
13. Public participation in service improvement	Open data and accountable systems give people and enterprises the opportunity to create new ideas on data usage. Co-creation allows to combine the data to develop new applications.

Data quality varies and the data quality should have the quality required for the dashboards. However, quality varies as is demonstrated by Vetrò et al. (2016) by investigating two case studies. Furthermore, it is not only about the technology, but also about how the results created by data scientists will be used. For example, Welch, Feeney, and Park (2016) found that data sharing is more strongly determined by persuasive mechanisms and technical engagement. Even if a dashboard is properly designed from a technical perspective, the detection of fraud or identification of strange patterns is useless if there are no means to further investigate these patterns by legitimate agencies. This requires that institutions should be in place which are able to follow up results originating from the data science analysis. The use of data outcomes requires institutional changes, also to ensure accountability. Only formal authorities, such as ombudsmen, ministries, politicians, and judges, can hold organizations accountable. Lourenço (2015) analyzed open data portals and found that they often have no organizational and structural elements to support public accountability.

In summary, the development and use of dashboards may involve many risks and challenges as presented in Table 2. One of the main risks is the misunderstanding of information, which could lead to incorrect conclusions about the data. This also raises the question of whether dashboards can result in transparency and accountability. Dashboards might help to improve the level of transparency, but much is dependent on the proper design. An information overload or showing the incorrect or incomplete information might result in less transparency (Matheus & Janssen, 2013). As long as there is information asymmetry, there will be no complete transparency.

Developing dashboards can be expensive. Their development and operation might consume a lot of resources and data scientists are scarce. This becomes even worse when considering that specialized human resources are expense and rare. This can lead to a lack of new development and a standardized predefined view in the long term. If data originates from different departments or the dashboards have many owners, it may lead a decrease in quality. There is also a risk of politicians and civil servants boycotting or not using the dashboard due to the difficulty of accepting new technologies, or due to the possible

loss of political power in decision-making processes.

Data cleansing and processing is an essential part of dashboard development. Implementing the legal requirements to anonymize public data in order to protect people's privacy is both a risk and a challenge (Janssen & van den Hoven, 2015). It is risk because there is possibility to identify people, even when using anonymized data sets. It is challenge as anonymizing data sets requires specific knowledge and skills, and comes at a price. Statistical analysis also helps to reduce outliers and 'bad data', which leads to bad results. The use of low quality data can lead to incorrect or inaccurate decision-making (Lazer, Kennedy, King, & Vespignani, 2014). Even correct data can be misinterpreted by people.

4. Designing dashboards

The creation of value, overview of benefits, risks and challenges helps to understand the concept of dashboards. In this section support for the design of dashboards will be derived. First, design principles will be presented, followed by the data cycle.

4.1. Design principles

The design of dashboards should accomplish the goal of creating transparency and accountability, but is encountering many risks challenges. Our literature review of dashboards shows that there are only a few papers about dashboards in the public sector, whereas much more information is available about private sector dashboards. However, there exist no support for creating dashboards in the literature. Therefore, we opted to identify *design principles* as these are useful for guiding solving ill-structured or 'complex' problems (Simon, 1996). Gibb (1997) views principles as rules of thumb that can be used to guide the designers. Informed by our 2 cases and using both public and private sector literature the design principles presented in Table 3 were derived. The principles can guide the designing of more effective public sector dashboards.

These principles presented in table 3were not always followed in the

Table 2
Description of the main risks and challenges.

Risks and challenges	Description
1. Low data quality	If data is not properly cleaned via statistics processes, for example, bad data can lead to bad analytics and give people the wrong message and information. Wrong information is worse than having no information!
2. Fragmented responsibility for data quality	Many owners can rapidly decrease the quality of data and the information delivery on the dashboards.
3. Limited readiness to adapt to new changes	Politicians and public managers can boycott the new system or new technologies for a variety of reasons, such as the elderly having difficulty using the technology or civil servants avoiding loss of their political power for decision-making on public policies.
4. Limited knowledge and no suitable staff	There are not many people having the capabilities needed to design and operate dashboards. As the number of qualified people is low, they are expensive and in high demand by the job market.
5. Wrong interpreting information	Incorrect or inaccurate data and information visualized using dashboards can reduce citizens' trust in government.
6. Wrong decision-making and mistrust in government	Governments should take care when disclose data in dashboards. Part of those public dashboards will lead citizens' decision-making. If the public organization cannot deliver a proper level of public services, it will lead to mistrust in the government.
7. Reducing transparency	Although dashboards can enhance transparency, dashboards might only blur the picture by providing pre-defined views defined by those in power.
8. Lack of resources	The development of a suitable dashboard can be expensive as numerous factors must be taken into account during development. This might require many interpretations and continuous updating after completion.
9. No maintenance and updating	Data in a dashboard is not updated or the dashboard interface is no longer current. If dashboards do not perform well, citizens will abandon them and less likely trust them.
10. Pre-defined view (level of dashboard customization)	Showing to people a predefined way without being able to delve into details and to take complementary views might result in bias, not being able to see the complete picture and in distrust. Offering alternatives to customize visualization and showing alternative views can increase insights can improve governance and trustworthiness, avoids bias from governmental perspective.
11. Not being able to adapt to new developments	The use of dashboards can result in new detailed questions, which might require additional data collection, or new plans and strategies to reduce the impact of public decision-making.
12. Data privacy	If data is not properly anonymized, private data can be displayed erroneously.
13. No participation support	Our definition of dashboard suggests that there is some level of control or intervention possible. Engagement support, providing comments and suggestion is essential.
14. No institutional mechanisms	The results of the data science efforts can result in the need for taking actions, for example, to hold organizations accountable. This requires having institutional mechanisms in place.
15. Scarce usage	Scarce usage of dashboards by citizens may not bring the efficiency and effectiveness expected by government and the public. Traffic congestion remained the same and solutions were not working properly in our case study.

two case studies. This resulted in a number of risks and challenges as reported in the previous section. Although these principles do not necessarily need to be followed, they can help guide the design of a dashboard to improve the creation of transparency and accountability. For example, the lack of updated, timely, and precise data contributed to misunderstanding and misinformation among citizens and decision-making managers (Ballou, Heitger, & Donnell, 2010).

5. Data cycle for dashboards

We abstracted the information flow and information processes from the cases. The main phases often followed by data scientists are data capturing, data storage, data searching, data sharing, data analysis, and data visualization (Bizer, Boncz, Brodie, & Erling, 2012), whereas in our cases different phases collection, processing, analysis, visualization, and usage phases were followed. The step 'data usage' is introduced, as e value can only be created from data when it is actually used (Janssen, Estevez, & Janowski, 2014). Fig. 1 shows the steps and the two main information flows in the cycle. One flow is used for creating a dashboard for internal use by the government to support policy-making, whereas the other dashboards is aimed at creating transparency and accountability for the public.

The first flow is labelled as F1 and depicted using a dashed line. This F1 shows data from sensors, forms by the government or generated by citizens.

In Stage B, a division between the Database with Public Private Data (PPD) and the Database with Public Open Data (POD) is made. PPD cannot be shared with external parties, due legal reasons such as privacy of personal data (names, nicknames, geographical positions etc.) and confidentiality of strategic governmental data (strategic plans anti-terrorism, emergency plans, etc.). For these reasons, a normalization and standardization is conducted on the PPD to create the POD, labelled as "anonymization" (F2). In the two cases studies presented in this paper, the GPS of buses is considered POD, which is freely

accessible in open format at the Open Data Portal Data.rio. The GPS data can be accessed by downloading a dataset in the Comma Separated Values (CSV) file format, or by invoking the Application Programming Interface (API). Furthermore, the dashboard using external data from Waze was accessible via a special API created by Waze in 2013.

The two databases POD and PPD enable the creation of two different flows with different objectives. While the PPD helps government to create internal strategies for public policy, the POD can be freely used by individuals and enterprises by accessing the Open Data Portal. The governmental flow of information is shown in red by a continuous solid line (F3, F5, F7, F9) and the public flow of information is shown in green by a dashed line (F4, F6, F8, F10).

After Stage B, both flows (F3 and F4) go to Stage C (Data Analysis). Based on the case studies, Stage C includes the Big Data Analytics Processes either manually performed by data scientists or automated. The statistical analysis and geographical analysis enable Stage D (Data Visualization). The stage D combines results from statistical and geographical analysis into a dashboard to present for internal and external audiences. The F5 flow of information enables visualization of dashboards for civil servants and politicians, whereas, the F6 flow enables the creation of public dashboards for private cars avoid traffic jams (case 1) and users of public transportation (case 2).

After the Stage D, this data is used by internal audience of Rio de Janeiro City Hall (i.e. COR) represented by the flow F8 and externally by the public, e.g. people in the traffic jams or using public transportation. Both flows of information end at Stage A, being collected again by sensors in the city (speed traps and GPS of buses) or via Big Data using Waze and Moovit API, represented by flows F9 and F10 (Fig. 4).

The cycles in Fig. 4 suggest that dashboards need to be continuously updated. One cycle opens data for citizens, which should result in higher levels of transparency. For citizens, various views and visualizations can be created. Due to privacy issues and for other reasons not all data might be indiscriminately shared. The second cycle opens data for a limited group of policy-makers. In this way this data can be used to

Table 3
Overview design principles for dashboards.

Principle	Description	Source
1. Collect accuracy and precise data	Governments must give the most correct and precise information, to prevent users from being unable to understand the data and being misled. Incorrect information in the dashboard can result in bad decisions.	(Abelson, Gauvin, MacKinnon, & Watling, 2004; Obama, 2009), case studies
2. Customize views	Dashboards should not be merely simple or generic visualizations; dashboards should contain customized views for showing the problem at hand. In this way decision makers and users can gain insight. Customized views can help them understand the situation. The design requires understanding of organizational strategies, viewpoints, business processes, indirect effects, decision support systems, and priorities. In the cases found in the literature review, separate apps are developed for each purpose to enable a clear view of the problem at hand.	(Eckerson, 2010; Johnston & Pongatichat, 2008; Kaplan & Norton, 2001; Little, 2004), case studies
3. Support different view	A single view might result in a limited picture on the situation. Different views can avoid bias and improve the understanding. By providing raw data, others can create new views which can result in updating the dashboard and improving usage.	(Eckerson, 2010; Johnston & Pongatichat, 2008; Kaplan & Norton, 2001; Little, 2004), case studies
4. Clear presentation	Dashboards enable the use of charts, graphs, pictograms, bars, and numbers, etc., to visualize information for monitoring and analyzing performance. Dashboards should visualize data in and easy-to-understand manner. In our cases, the simplicity of the dashboards enabled their use by a broad public.	(Baskett, LeRouge, & Tremblay, 2008; Few, 2006; Velcu-Laitinen & Yigitbasioglu, 2012), case studies
5. Offer decision-making support	Relationships between performance metrics and organizational desires must be clear. Dashboards can provide decision-support to evaluate 'what if' scenarios and to use predictive analytics. This can help provide more insight into the situation and help decision makers. By providing insight into possible alternatives, the effect of choosing an alternative can be predicted in our cases.	(Ganapati, 2011b; Velcu-Laitinen & Yigitbasioglu, 2012), case studies
6. Interaction support	Static dashboards often provide limited insight. More insight can be gained by providing interaction features, which enables users to view the data from various perspective, to suggest recommendation based on the data but also to provide feedback to improve the use. Real-time information was a key element for supporting the decisions in the cases.	(Chen, Chiang, & Storey, 2012; Mayer-Schönberger & Cukier, 2013; McAfee et al., 2012; R. M. Peters, Janssen, & Engers, 2004; Svensson, Saeverhagen, & Bouillouta, 2015), case studies
7. Provide overview and details	Dashboards should be able to deal with an enormous volume of big and open data. By providing an overview and the opportunity to zoom in on details, the high-volume of big data analyzed was presented in a simple visual display.	(Lohr, 2012; Marz & Warren, 2015), case studies
8. Focus on creating public values	Merely visualizing data has limited use if this is not suitable for creating added value. Dashboards are difficult to develop, especially for big and open data. Often a business case is required to determine the added value. Dashboards should be designed to create public values like engagement, transparency and accountability and adhere to public values like privacy.	(Chen et al., 2012; Dietrich, Plachy, & Norton, 2014; Schroeck, Shockley, Smart, Romero-Morales, & Tufano, 2012), case studies
9. Ensure real-time updates of data	The majority of dashboards are not based on real-time data. Governments present dashboards of what has happened in the past and can use this to predict future events.	Case studies
10. Ensure institutional support	The creation of transparency results in the detection of incorrect data or behavior. For citizens, it is hard to take any action if there is a suspicion of fraud or corruption. It should be possible for citizens to report to an independent and trusted agency. There should be institutions to deal with such matters further.	Case studies

inform policy-making and other decisions, however, this does not result into transparency for citizens. There is still information asymmetry, as citizens still have less information than governments have. As a result, not being able to open all data might result in less transparency and trust in government.

Our case study findings show also that, although dashboards are often used for policy evaluation, dashboards can support the complete policy-making cycle including policy formulation, implementation, and evaluation. Dashboards should help citizens to create an understanding of the situation at hand, avoid long-lasting search processes and an information overload. Due to this reason, the creation of dashboards for citizens need a design focused on relevance and take the citizens' situation into account. Balancing issues such as privacy, information overload and designing an overview at glance is challenging. A good dashboard is user-centric and provides insight, however, for some citizens this might not be sufficient. Although dashboards in different areas can have different shapes and forms, there might be users that want to gain access to the raw data. Due to the diversity of possible users of data, it is paramount that both the societal issues that need to be addressed and the users' needs are elicited.

6. Conclusions

Dashboards can be used by governments to interact with the pubic. Smart cities can use dashboards to empower citizens to create 'smarter' citizens and can be viewed as the missing link between Smart Cities and Smart Citizens. Developing dashboards needs data science activities to extract data from various sources, integrate them to find new insights through the use of data and to design visualizations. Dashboards are instruments for presenting data into an integrated visual display and to initiate actions. Dashboards are likely to become more important with the availability of more and diverse data. Dashboards can play a crucial role in providing insights into a situation and helping the situation to improve and evolve. The identified benefits of dashboards are related to creating an overview and being able to zoom in on the details. Transparency can be created by overcoming information asymmetry between public organizations and the public. However, the benefits can only be gained when the dashboards are properly designed. The cases show that dashboards can provide many benefits, but also that the design of dashboards is a non-trivial activity encountering many risks and challenges. The design principles presented in this paper can help
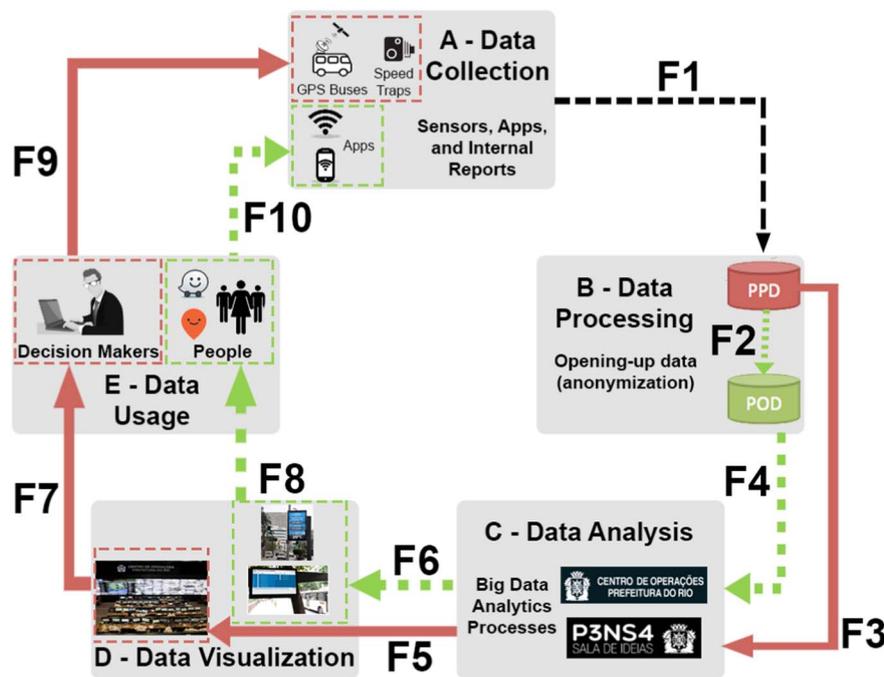
Fig. 4. Data cycle for dashboards.

to design dashboards.

Our findings show that the introduction of dashboards might be useless if their introduction is not accompanied with organizational changes. Dashboards should not only be used to communicate with the public, but also to gain feedback from them and to stimulate interaction. Finding new insights or detecting corruption is useless if there are no means to deal with the feedback or to further investigate by legitimate agencies. Formal authorities that can hold organizations accountable need to be involved or created. Engagement in dashboards, with citizens having the opportunity to provide data and discuss results, plays a crucial role in achieving the benefits. Furthermore, a bad-designed dashboard might result in misunderstanding of data and can affect the public's trust in the government.

Acknowledgement

References

Abelson, J., Gauvin, F.-P., MacKinnon, M. P., & Watling, J. (2004). *Transparency, trust and citizen engagement.* Ottawa: Canadian Policy Research Networks Inc.

Allio, M. K. (2012). Strategic dashboards: Designing and deploying them to improve implementation. *Strategy & Leadership, 40*(5), 24–31.

Ballou, B., Heitger, D. L., & Donnell, L. (2010). Creating effective dashboards: How companies can improve executive decision making and board oversight. *Strategic Finance, 91*(9), 27–33.

Baskett, L., LeRouge, C., & Tremblay, M. C. (2008). Using the dashboard technology properly. *Health Progress, 89*(5), 16.

Bertot, J. C., Jaeger, P. T., & Grimes, J. M. (2010). Using ICTs to create a culture of transparency: E-government and social media as openness and anti-corruption tools for societies. *Government Information Quarterly, 27*(3), 264–271. https://doi.org/10.1016/j.giq.2010.03.001.

Bizer, C., Boncz, P., Brodie, M. L., & Erling, O. (2012). The meaningful use of big data: Four perspectives — four challenges. *SIGMOD Rec. 40*(4), 56–60. http://dx.doi.org/10.1145/2094114.2094129.

Brown, A., Fishenden, J., Thompson, M., & Venters, W. (2017). Appraising the impact and role of platform models and government as a platform (GaaP) in UK government public service reform: Towards a platform assessment framework (PAF). *Government Information Quarterly, 34*(2), 167–182. http://dx.doi.org/10.1016/j.giq.2017.03.003.

Bugaric, B. (2004). Openness and transparency in public administration: Challenges for public law. *Wis. Int'l LJ, 22,* 483.

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

Chengalur-Smith, I. N., Ballou, D. P., & Pazer, H. L. (1999). The impact of data quality information on decision making: An exploratory analysis. *Knowledge and Data Engineering, IEEE Transactions on, 11*(6), 853–864.

Cleveland, W. S. (2001). Data science: An action plan for expanding the technical areas of the field of statistics. *International Statistical Review, 69*(1), 21–26.

Cuadrado-Ballesteros, B. (2014). The impact of functional decentralization and externalization on local government transparency. *Government Information Quarterly, 31*(2), 265–277.

Dawes, S. S., & Helbig, N. (2010). Information strategies for open government: Challenges and prospects for deriving public value from government transparency. *Electronic government* (pp. 50–60). Springer.

Dietrich, B. L., Plachy, E. C., & Norton, M. F. (2014). *Analytics across the enterprise: How ibm realizes business value from big data and analytics.* IBM Press.

Eckerson, W. W. (2010). *Performance dashboards: Measuring, monitoring, and managing your business.* Wiley.

Few, S. (2006). *Information dashboard design.* O'Reilly.

Ganapati, S. (2011a). *Key features for designing a dashboard.* Government Finance Review.

Ganapati, S. (2011b). *Use of dashboards in government-been there and done that modified 12: 00, 11 Nov 2011 by admin|page history.*

Gibb, T. (1997). Towards the engineering of requirements. *Requirements Engineering, 2,* 165–169.

Harrison, T. M., & Sayogo, D. S. (2014). Transparency, participation, and accountability practices in open government: A comparative study. *Government Information Quarterly, 31*(4), 513–525.

Janssen, M., & van den Hoven, J. (2015). Big and open linked data (BOLD) in government: A challenge to transparency and privacy? *Government Information Quarterly, 32*(4), 363–368. http://dx.doi.org/10.1016/j.giq.2015.11.007.

Janssen, M., Estevez, E., & Janowski, T. (2014). *Interoperability in big, open, and linked data—Organizational maturity, capabilities, and data portfolios.* IEEE Computer Society.

Janssen, M., Matheus, R., & Zuiderwijk, A. (2015). Big and open linked data (BOLD) to create smart cities and citizens: Insights from smart energy and mobility cases. *Electronic Government* (pp. 79–90). Springer.

Jensen, M. C., & Meckling, W. H. (1976). Theory of the firm: Managerial behavior, agency costs and ownership structure. *Journal of Financial Economics, 3*(4), 305–360.

Johnston, R., & Pongatichat, P. (2008). Managing the tension between performance measurement and strategy: Coping strategies. *International Journal of Operations & Production Management, 28*(10), 941–967.

Kaplan, R. S., & Norton, D. P. (2001). Transforming the balanced scorecard from performance measurement to strategic management: Part II. *Accounting Horizons, 15*(2), 147–160.

Lazer, D., Kennedy, R., King, G., & Vespignani, A. (2014). The parable of Google flu: Traps in big data analysis. *Science, 343* (14 march).

Little, J. D. (2004). Models and managers: The concept of a decision calculus. *Management Science, 50*(Suppl. 12), 1841–1853.

Lohr, S. (2012). The age of big data. *New York Times, 11.*

Lourenço, R. P. (2015). An analysis of open government portals: A perspective of transparency for accountability. *Government Information Quarterly, 32*(3), 323–332. http://dx.doi.org/10.1016/j.giq.2015.05.006.

Maheshwari, D., & Janssen, M. (2014). Dashboards for supporting organizational

development: principles for the design and development of public sector performance dashboards. *Paper presented at the proceedings of the 8th international conference on theory and practice of electronic governance* (pp. 178–185). ACM.

Marsh, J. A., Pane, J. F., & Hamilton, L. S. (2006). *Making sense of data-driven decision making in education*.

Marz, N., & Warren, J. (2015). *Big data: Principles and best practices of scalable realtime data systems.* Manning Publications Co.

Matheus, R., & Janssen, M. (2013). Transparency of civil society websites: towards a model for evaluation websites transparency. *Paper presented at the proceedings of the 7th international conference on theory and practice of electronic governance, Seoul, Republic of Korea* (pp. 166–169). ACM.

Mayer-Schönberger, V., & Cukier, K. (2013). *Big data: A revolution that will transform how we live, work, and think.* Houghton Mifflin Harcourt.

McAfee, A., Brynjolfsson, E., Davenport, T. H., Patil, D., & Barton, D. (2012). Big data. The management revolution. *Harvard Business Review, 90*(10), 61–67.

Navarro-Galera, A., et al. (2016). Online dissemination of information on sustainability in regional governments. Effects of technological factors. *Government Information Quarterly, 33*(1), 53–66.

Obama, B. (2009). *Transparency and open government.* (Memorandum for the heads of executive departments and agencies).

Peters, R. M., Janssen, M., & Engers, T. M. V. (2004). Measuring e-government impact: Existing practices and shortcomings. *Paper presented at the proceedings of the 6th international conference on electronic commerce* (Delft, The Netherlands).

Roberts, N. C. (2002). Keeping public officials accountable through dialogue: Resolving the accountability paradox. *Public Administration Review, 62*(6), 658–669.

Schroeck, M., Shockley, R., Smart, J., Romero-Morales, D., & Tufano, P. (2012). *Analytics: The real-world use of big data: How innovative enterprises extract value from uncertain data.* IBM Institute for Business Value.

Simon, H. A. (1996). *The sciences of the artificial* (3rd ed.). Cambridge: MA: MIT Press.

Svensson, I., Saeverhagen, E., & Bouillouta, F. M. (2015). Driving rig performance through real-time data analysis, benchmarking. *Paper presented at the SPE digital energy conference and exhibition.* .

Velcu-Laitinen, O., & Yigitbasioglu, O. M. (2012). The use of dashboards in performance management: Evidence from sales managers. *The International Journal of Digital Accounting Research, 12*(18), 1–20.

Vetrò, A., Canova, L., Torchiano, M., Minotas, C. O., Iemma, R., & Morando, F. (2016). Open data quality measurement framework: Definition and application to open government data. *Government Information Quarterly, 33*(2), 325–337. http://dx.doi.org/10.1016/j.giq.2016.02.001.

Villeneuve, J.-P. (2014). Transparency of transparency: The pro-active disclosure of the rules governing access to information as a gauge of organisational cultural transformation. The case of the Swiss transparency regime. *Government Information Quarterly, 31*(4), 556–562.

Welch, E. W., Feeney, M. K., & Park, C. H. (2016). Determinants of data sharing in U.S. city governments. *Government Information Quarterly, 33*(3), 393–403. http://dx.doi.org/10.1016/j.giq.2016.07.002.

Ricardo Matheus is a PhD Candidate at TUDelft. Currently works with Big and Open Linked Data (BOLD) being part of OpenGovIntelligence (www.opengovintelligence.eu/).

Marijn Janssen is full professor in ICT and Governance and head of the Information and Communication Technology section of the Technology, Policy and Management Faculty of Delft University of Technology. More info: www.tbm.tudelft.nl/marijnj.

Devender Maheshwari holds a PhD from Delft University of Technology at the faculty of Technology, Policy and Management. He is a telecommunications engineer and his research specializes in measurement and benchmarking.