Future Scenarios of Data-Driven Smart Cities

GEORGIANA NENCIU, PATRICK VAN DER DUIN, MARIJN JANSSEN, DOYA SNIJDERS, JORIS HULSTIJN

Georgiana Nenciu, Delft University of Technology, Patrick van der Duin, Dr., Delft University of Technology/Fontys Academy for Creative Industries, Marijn Janssen, Prof. dr., Delft University of Technology, Dhoya Snijders, Dr., The Netherlands Study Centre for Technological Trends, Joris Hulstijn, Dr., Tilburg University.

Abstract
This paper aims to investigate the role of Big Open and Linked Data (BOLD) in the development of smart cities. Future scenarios of data-driven smart cities were created, and defining characteristics of smart cities were identified based on their approach to data governance (top-down/bottom-up) and openness of data (open/closed data environment).

BOLD can increase the “smartness” of cities. However, until now, there has not been a research that captures the diverse and long-term contribution of BOLD to smart cities. BOLD is characterized by its large size and unstructured features (big data), availability to the public without any predefined restrictions (open data), and integration of different raw data sources to obtain new insights (linked data). In today’s connected world, big data is seen as “an essential element of economic growth” (Bulger, Taylor, & Schroeder, 2014, p. 12), and open data is perceived as having a lot of potential for development - “the new raw material of the twenty-first century” (Deloitte, 2012, p. 1). Smart cities promise to improve the quality of life by using technological solutions that generate huge amounts of data. Given the complexity and uncertainty of urban environments, BOLD can contribute to solving problems of cities only if it overcomes certain associated challenges such as the politics behind the use of data, and the privacy and security issues.

The present study takes an outside-in perspective to data-driven smart cities, starting from the current trends of two Dutch cities that aim to be smart cities, Rotterdam and The Hague. The societal developments were identified by conducting interviews with stakeholders from Rotterdam and The Hague from the public sector, business, and academia. A workshop was organized in collaboration with The Netherlands Study Centre for Technological Trends, with participants from public sector, business, and academia, to develop four potential scenarios of data-driven smart cities for 2040. The four scenarios were used as input to create a smart city governance model, based on characterizing elements of the urban data environment. The model can be used by policy-makers to understand the implications of data decisions for the whole urban ecosystem. The study has the following relevant scientific outcomes: (1) development of a smart city governance model by considering data decisions as defining elements of the city; (2) development of scenarios for smart cities with data as central focus element; and (3) overview of the current situation for the smart cities Rotterdam and The Hague.

Keywords: big data, open data, smart city, foresight, scenario development
Introduction

Big Open and Linked Data (BOLD) is a research field that deals with data having the following characteristics, as suggested by its name: big - defined as a very large set of structured and unstructured data, that exceeds the current architectures and databases (Khan, Uddin, & Gupta, 2014), open - defined as “data that is made available by government, businesses and individuals for anyone to access, use and share” (The Open Data Institute, 2015, p. 3), and linked - the result of aggregating different datasets from multiple sources, scattered across organizations or the Internet (Janssen, Estevez, & Janowski, 2014).

Smart cities aim to improve the quality of life by using technologies that generate large amounts of data, and BOLD can be an important enabler for smart city applications. The promise of BOLD in smart cities is that it will reveal unexpected patterns, improve decision making of governmental bodies, increase the efficiency of urban services, enhance transparency of public agencies, and dismiss the communication gap between the local government and its citizens (Rabari & Storper, 2015; The European Data Portal, 2016). Examples of such initiatives are: visualization of how taxes are spent by governments1, transport-related applications (including apps for visually impaired people2), and prediction of places where fire might spark3. Data-driven applications help citizens and governmental bodies better understand their city and the problems that the city faces.

While studies have discussed the combination of big data and open data (e.g. Jaakola et al. (2015), Janssen and Van den Hoven (2015)), the academic research on the implications of BOLD on smart cities is limited. Only a paper by Janssen et al. (2015) investigated the relationship between smart cities and BOLD, and they concluded that BOLD can contribute to increasing the “smartness” of cities. However, their research does not capture the impact that BOLD can have on cities, as their unit of analysis was a smart city initiative (smart energy, smart mobility), and not the city as a whole. Hence, there is a need for research on the impact of BOLD in smart cities, as suggested by Janssen et al. (2015). Leszczynski (2016) encouraged future research about “anticipating particular kinds of cities-to-come” (Leszczynski, 2016, p. 1) and big data in the urban context. Lastly, Meijer et al. (2015) identified the lack of research regarding governance models in smart cities.

The present study addresses these gaps by conducting a foresight study that investigates the impact of BOLD in the development and governance of smart cities. Using the combined case of Rotterdam and The Hague as smart cities, scenarios of data-driven smart cities were developed during an interactive workshop and characterizing elements of data-driven smart cities were identified. A framework for data-driven smart city is developed, which can be used by policy-makers to understand the implications of data governance decisions for the whole urban ecosystem.

The paper has the following structure: first the scenario development process is explained, and then the four scenarios that resulted from an interactive workshop are presented. The next sections discuss the resulting smart city governance model and provide practical recommendations for smart city professionals.

---

1 http://app.wheredoesmymoneygo.org/
2 http://blindsquare.com/
Research approach

The study conducted about BOLD in smart cities is a futures research study. According to Berkhout et al. (2007, p. 74), futures research is “the ability, the competence and the art of describing, explaining, predicting, exploring and/or interpreting future developments as well as its consequences for actions and decisions in the present”. The purpose of futures research is not to predict the future, but to help people and organizations anticipate future threats and opportunities and make better decisions today (Giaoutzi & Sapio, 2013). There are different approaches to futures research: predictive (what will happen), explorative (what could happen), or normative (what should happen) (Vergragt & Quist, 2011). Examples of futures research methodologies are depicted in Figure 1 below.

![Futures research methodologies](Van der Duin, 2016)

The potential evolution of future smart cities was explored by using scenario development, which is an explorative futures research method. Developing scenarios is a method for imagining a (limited) number of possible futures, by taking into consideration external trends and uncertainties (Schoemaker, 1995). The seven-steps trajectory proposed by Nekkers (2016) was followed to have a structured approach towards the scenario development. The seven steps are:

1. **Preparation** - identify the need and purpose of the scenario project
2. **Orientation** - determine the question to be answered and the time horizon
3. **Exploring the environment** - pinpoint the relevant trends or developments
4. **Determining scenario structure** - identify the framework for building the scenarios
5. **Building scenarios** - identify the most important elements and the scenarios’ titles
6. **Using scenarios** - draw conclusions from the scenarios
7. **Monitoring scenarios** – observe which scenario(s) is going to happen (this last step goes beyond scope of this research)

During the **preparation** step, the need for a scenario project was defined because of two reasons: (1) discussing the impact of BOLD in cities needs a dynamic perspective, as it is related to how cities develop over time; and (2) big data is inherently future oriented, as it enables “focus on the future more than on the present and the past” (Lyon, 2014, p. 6). The purpose of developing scenarios is to explore how future smart cities could look like, in order to use them as foundation for further analysis. The decision of creating scenarios was reinforced by the current lack of scenarios that emphasize the role of data in smart cities, in spite of the increased popularity of these concepts nowadays.
During the second step, **orientation**, the question that the scenarios would answer was defined: “what will data-driven smart cities look like in 2040?”. The time horizon was 2040 – 24 years were considered a period long enough to allow changes in smart cities, but still short enough to create foreseeable scenarios.

The third step, **exploring the environment**, was performed by investigating smart city and data trends for Rotterdam and The Hague. The two cities were considered a combined case study because of the following two reasons: (1) their geographical proximity eased the interactions with smart city stakeholders, and (2) considering developments of two cities that aim to become smart cities enhanced the holistic approach of the research.

The information was collected in two ways: semi-structured interviews and desk research. Nine semi-structured interviews were conducted with smart city stakeholders from the triple helix areas: public sector, academia, and business. Desk research complemented the interviews, by gathering online information about the main trends in the Netherlands and Europe. In addition to interviews and desk research, participation to smart city and big data events contributed to a better understanding of Dutch urban data trends and to discovering new directions of investigation.

The identified smart city and urban data trends from Rotterdam and The Hague were clustered using the PESTLE tool, which allows the analysis of the external environment in a systematic manner (Jackson, 2013). PESTLE stands for Political, Economic, Social, Technological, Legislative, and Environmental. Figure 2 summarizes the trends identified in the two Dutch cities, trends used as input for the next steps of the scenario process.

The **scenario structure** was created by using an outside-in perspective, which means that societal and market developments were the starting point for the foresight study (Cramer, Duin, & Heselmans, 2016). The PESTLE developments identified during the previous step were used as input to determine the scenarios’ axes of uncertainty and also to develop the process of creating scenarios. For each trend identified, there were one or more corresponding elements in the workshop. During the process of establishing the scenarios’ structure, valuable feedback was received from foresight experts.

The following step, **building scenarios**, was performed by organizing an interactive workshop in collaboration with The Netherlands Study Centre for Technological Trends (STT). Since the participatory approach is an important aspect of futures research (Baerten, 2016), people from various fields (public sector, academia, and business) worked together in teams and shared their perspectives on future smart cities. The structure of the workshop was ensured by a three-step process: **Step 1 - City foundations**: choosing the main characteristics of the city; **Step 2 - City life**: detailing the implications of the chosen foundations; and **Step 3 - State of the city address**: presenting the future smart city. Instead of providing a 2x2 matrix and asking participants to imagine that type of smart city, teams were free to make decisions on four axes of uncertainty. This proved to be an engaging process, as participants felt more empowered to express their own opinions. The four axes of uncertainty were:

- **governance of the city** – does the smart city have a top-down approach (centralized system with a central entity) or a bottom-up approach (decentralized system with a collaborative environment)?
- **openness of data** – does the smart city have open data (anyone can access and use data) or closed data (data exists, but it is not available to everyone)?
- **acceptance of technology** – do citizens (gladly) accept the intensive use of technology (high acceptance) or do they (strongly) reject the use of technology (low acceptance)?
- **main drivers of the city** – what drivers are prioritized: environmental (related to climate change, pollution, etc.) or economic (related to financial gains, etc.)?
Figure 2. Summary of trends and developments using PESTLE
The sixth step, using scenarios, was performed in two ways. Firstly, the scenarios served as input to develop a model of data governance in smart cities, and recommendations for the smart city team of the municipality were discussed. Secondly, the knowledge was disseminated to a broader audience by publishing a blog post on STT’s website, with the purpose of raising awareness about issues that concern smart cities.

The final step, monitoring scenarios, can be part of a next research, as time has to go by to see which of the scenarios will actually happen.

Scenarios of future smart cities

Four potential scenarios of data-driven smart cities emerged from the interactive workshop. The characterizing elements of each scenario are presented in Table 1 below. It can be observed that two of the four scenarios, namely Greenville and Circular City, have identical characteristics: bottom-up governance, open data, high acceptance of technology, and environmental drivers. However, the participants presented different sides of the smart city, as detailed in their descriptions below. This might mean that another variable, not considered during the scenario development, differentiates the two scenarios.

1) **Dataflex** is a smart city with top-down governance, the data environment being governed by the municipality. The smart city has a ubiquitous sensors network that generates data about every aspect of the city. The city uses data analytics to interact in real time with citizens, for example when sensors feel a person is very hungry, the closest shop can suggest buying a meal. Every citizen can opt out of the system, if (s)he wants to. Data is used by the municipality to find out what is best for its citizens, to prevent free-riding behaviour and to manage the use of resources in a balanced manner. Laws and regulations are replaced by data-driven systems that encourage citizens to do what is best for them. For example, when a citizen is prone to an addiction, the system can prevent him/her to get in contact with its temptation.

2) **Incorpolis** is a smart city with top-down governance, with the data environment governed by private entities. The city takes its name from the two most important aspects of the city: income and corporations. Corporations collect and own all data generated by citizens and urban services, and the results of the data analysis are traded on the data market. These companies have the intelligence to create useful and personalized services. Citizens are the

<table>
<thead>
<tr>
<th>Smart city characteristics</th>
<th>Dataflex (The city ruled by AI)</th>
<th>Incorpolis (The corporation-city)</th>
<th>Greenville (The open-source city)</th>
<th>Circular City (The local community)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Data governance</td>
<td>Top-down</td>
<td>Top-down</td>
<td>Bottom-up</td>
<td>Bottom-up</td>
</tr>
<tr>
<td>Openness of data</td>
<td>Open data</td>
<td>Closed data</td>
<td>Open data</td>
<td>Open data</td>
</tr>
<tr>
<td>Acceptance of technology</td>
<td>High</td>
<td>High</td>
<td>High</td>
<td>High</td>
</tr>
<tr>
<td>Economic drivers</td>
<td>Environmental</td>
<td>Economic</td>
<td>Environmental</td>
<td>Environmental</td>
</tr>
</tbody>
</table>

Table 1. The four smart city scenarios that emerged from the foresight workshop
data creators, they are “data-enslaved” by the data companies – the more and diverse data one generates, the better services it receives. Citizens communicate more and more by using technology, the amount of live interactions decreases, and data-profiling helps people adapt the way they communicate with others. Every month there is a data free day, when citizens do not generate any data, and every first of January the new data milestones are celebrated.

3) **Greenville** represents a bottom-up utopian city with a horizontal hierarchy - there are no elections, and there is no official entity that governs the city. Advanced algorithms aggregate all data generated in the city, and they make only data-driven decisions. Data-driven technology knows what is best for people, and people are very happy in this city. Data is used for predicting the behaviour of Greenville’s citizens, and preventing them from taking harmful actions. Greenville is a self-fulfilling city, where all jobs are automatized. By using the advanced virtual experience system, you can do almost anything you want without harming other people.

4) **Circular City** is a smart city driven by the aspiration of enabling the local potential of the city. The governance of the city is a mix of top-down and bottom-up approaches - while having bottom-up roots, starting from a decentralized system, it also has top-down influences, because there was the need of a higher authority to enable the local perspective. The government takes into account people’s opinions by using the Democracy Platform. The government uses data-driven technologies to facilitate the fulfilment of necessities. The open data approach enables a transparent system, where all costs (including environmental ones) and benefits are visible to everyone. People can make conscious decisions about what they buy and from where.

The four smart city scenarios have certain things in common, related to the choice of uncertainties and elements incorporated in the scenario description. Firstly, all scenarios have “high acceptance of technology” as foundations, suggesting that a future where citizens are reluctant to technological advancements seems unlikely to happen. Secondly, three out of four scenarios incorporate top-down elements. This might indicate a need for higher coordination of smart cities and data environments, as a pure decentralized approach is difficult to imagine and achieve. Thirdly, three out of four scenarios have open access to data and environmental drivers prioritized over economic drivers. The reason behind these choices might be the intense debates the two topics received. Open data is a “hot topic” nowadays in the Netherlands, with many datasets opened by governments. The three scenarios show mostly optimistic approaches to open data, when in reality there are many debates regarding the lack of privacy and security. Fourthly, all four scenarios have aspects that might pose ethical questions. In Dataflex and Greenville, data is used for deciding strategic issues at city level. But how can algorithms distinguish what is desirable for citizens? In Incorpolis, citizens are “data-enslaved” and the quality of the services offered to them depends on the data they generate. Circular City offers transparency for everyone, raising the question: what are the boundaries that should not be crossed when it comes to data openness?

From the scenarios description, as well as from the interviews conducted as preparation for the workshop, it was observed that BOLD can enable value creation by facilitating the emergence of citizen-centric services and by providing the prerequisites for an urban digital platform. BOLD can offer insights into the urban environment and citizens’ behaviours, being the first step towards providing personalized services and solutions. Also, BOLD emerged multiple times as the element that empowers different innovative applications in the city, provided either by public or private actors.
Smart city governance model

From the descriptions of the four future data-driven smart city scenarios, it can be observed that two of the four elements proposed to build future cities had a great impact in characterizing the future scenarios, whereas the other two did not make a big difference. Governance of data (top-down/bottom-up) and openness of data (open/closed) contributed the most to defining the data environment in smart cities, while the drivers (economic/environmental) and the acceptance of technology (high/low) are not differentiating elements for smart cities, as they contributed to a smaller extent.

Building upon the two differentiating elements, governance of data and openness of data, a tentative 2x2 governance model was constructed (Figure 3). Each quadrant represents a potential data-driven smart city typology.

![Data-driven smart cities typologies](image)

The See-through City is the smart city with top-down governance and open data environment. It values data transparency, having data-driven policies and open data platforms where multiple up-to-date datasets are open and can be accessed by anyone. Data regulations are in place, as the municipality of the city is in charge of governing the data environment.

Data Masters has top-down governance and closed data environment. The data ecosystem is governed by technology or data companies, and the data power is concentrated to few private entities. Citizens generate and share their data in exchange for services.

Grass-roots Town is the smart city with bottom-up governance and open data environment. The city that encourages citizen involvement and has developed a data-driven entrepreneurial culture. However, the open and decentralized approach comes with the cost of low privacy and less security.

Data Bazaar has bottom-up governance and closed data environment, which fosters the development of data infomediaries. Data is highly valued, and citizens have data vaults where they safely store their private data.
These city typologies can be used by smart city policy-makers to understand how decisions related to data environment that are taken now can impact the smart city on long-term. For example, in a city with top-down approach and open data, the municipality can centralize the data and make data-driven policies, whereas in a city with top down approach and closed data, a company can benefit from all the data generated in the city. The four types of data-driven smart cities can be used by policy-makers to develop scenarios for their own city, and understand the implications that data environment decisions can have for the whole urban ecosystem.

**Recommendations for smart city policy-makers**

Based on the foresight study regarding data-driven smart cities, practical recommendations for smart city policy-makers were derived by the researchers. These recommendations aim to help policy-makers develop the urban data ecosystem of their smart city and mitigate the potential risks associated with BOLD in smart cities. The recommendations are the following:

- *Create a multi-disciplinary smart city team* with members from technological, social and policy backgrounds, to fully understand the technological developments and the subsequent social consequences. A multi-disciplinary team would also decrease the dependence on the vendors’ advice and capabilities.
- *Ensure the interoperability of data systems in a smart city*, in order to mitigate the risk of depending 100% on a single company when it comes to its data systems (as it is depicted by the Incorpolis scenario). The smart city team should ensure that other companies are able to provide complementary smart city services, to prevent monopoly regarding urban data.
- *Clarify ownership of projects and data from early stages*, to encourage a preventive approach towards data gathering. Accountability is also an aspect that needs to be discussed – who is accountable for the potential undesired consequences? Should one entity be 100% accountable or should the accountability fall on both entities?
- *Involve in best practices communities* that deal with smart city standards and data quality indicators (e.g. World Council of City Data, Open and Agile Smart Cities). Smart city teams can benefit from the global knowledge and working methods shared by these communities.
- *Create a long-term BOLD strategy*, to guide the process of implementing data-driven initiatives in the smart city. The strategy can be also a communication tool for the smart city team in relation with other teams of the municipality, as well as other stakeholders such as companies and citizens.
- *Proactively encourage the use of open data* to overcome potential hindrance factors for open data initiatives (e.g. lack of information, bad quality of data). Some measures that can be taken by the smart city team are: make visible the open data platform in relevant environments by organizing data competitions or hackatons, promote the open data website on social media, involve data organizations (companies or start-ups), and put in place feedback mechanisms regarding the data quality.
- *Discuss the active responsibility of engineers and policy-makers*, as data systems might involve citywide projects, where all citizens are involved without their individual approval. When it comes to cities and urban data, it is important to enable the active responsibility, which is forward looking and considers future consequences.
Conclusion

Data is an intrinsic element of smart cities, because the generation and use of data in a smart city is inevitable. Therefore, decisions regarding data and how data is tackled say a lot about the type of smart city envisioned.

Following the seven steps method proposed by Nekkers (2016), future scenarios of data-driven cities were developed. The future scenarios were used to develop a framework for data-driven smart cities, with two axes of uncertainty: governance of data (open/closed) and data openness (open/closed). The resulting four types of data-driven smart cities are: See-through City, Data Masters, Grassroots Town, and Data Bazaar. The framework can serve as a starting point for policy makers in the development of smart city scenarios, and as a lens towards understanding the implications of data decisions on the urban ecosystem.

Smart city professionals can benefit from the use of the practical recommendations concerning the setting-up of a multi-disciplinary smart city team, interoperability of data systems in a city, creation of a long-term BOLD strategy, and others.

Next to that, the study advances the academic literature regarding data-driven smart cities with the following relevant outcomes: (1) development of a smart city governance model by considering data decisions as defining elements of the city; (2) development of scenarios for smart cities with data as central focus element; and (3) overview of the current situation of the smart cities Rotterdam and The Hague.
References


