Ethics of Public Use of AI and Big Data

The Case of Amsterdam's Crowdedness Project

Ryan, Mark

University of Twente, Deptartment of Philosophy

Corresponding Author: Mark Ryan, m.j.ryan@utwente.nl

Abstract:

Smart information systems (Big Data and artificial intelligence) are used by governments to improve mobility, reduce over-crowdedness in hotspots, and provide more effective management of crowds. I looked at how Amsterdam municipality is using smart information systems (SIS) in their DrukteRadar Project to identify, report, and tackle issues surrounding crowdedness levels in the city.

SIS are becoming popular amongst governmental officials to automate activities more effectively. SIS provide the opportunity to improve mobility, increase economic growth, reduce energy outputs, improve management decisions, respond to disasters quicker, and improve citizens' quality of life. They offer governments the possibility of improving services, while reducing costs. The use and implementation of SIS is becoming widespread and governments are observing the benefits posed by SIS, particularly in relation to urban management.

80% of Europe's population will live in cities by 2020 and governments face a huge strain on resources and infrastructure. The use of SIS is being pioneered to help governments meet these needs and to provide a sustainable future for urban citizens. Ethical issues in this context can include that data may not be accurate, faithful or representative of a city and its citizens, which may cause bias, prejudice and harm to a population, by leading to unfair service provision. ICT companies' involvement in governmental SIS projects may also lead to technological lock-in and dependency on corporations. Instantaneous and ubiquitous retrieval and analysis of data may infringe

upon citizens' privacy and may lead vulnerabilities of malicious hacking, stolen data and a city's security.

To uncover if these issues correlate with the experience of those working in the field, I interviewed the Project Owner of Amsterdam's DrukteRadar project (translated as crowdedness project). This project implements SIS to anticipate and prevent overcrowding in Amsterdam, and was created in response to growing pressures on the city's amenities. The DrukteRadar Project collates a wide array of datasets to predict crowd levels and potential problem hotspots, visualised through a digital dashboard. The project aims to improve municipality management, provide help to tourists planning their trips, and assisting citizens' navigation through the city.

Through my discussions with the Project Owner of the DrukteRadar, I uncovered two additional issues not found in the literature: access to SIS and data ownership. The DrukteRadar team is concerned about access to SIS to promote fairness, equality, and provision of services amongst citizens. It aims to make its dashboard user-friendly and available to as many people as possible to promote inclusion. Data ownership is a concern for the project – who owns the data and what can be done with it. The DrukteRadar Project ensures they have data sovereignty, so that they do not become technologically locked-in to relationships with private organisations.

The Project Owner was aware that inaccurate of data may lead to discriminatory recommendations and harmful consequences. The DrukteRadar Project tries to minimise their algorithmic inaccuracy through extensive monitoring; secure technical infrastructure; and stakeholder review sessions. Another interesting finding was identifying how projects ensure privacy protection of its citizens. The DrukteRadar ensures that data is not traceable to individuals and the use of datasets follow privacy-by-design principles. The project also has strong security protocols, cyber-security measures, anonymization techniques, and repeated vulnerability tests. Overall, my report was able to evaluate how ethical issues found within the SIS literature correlate to those identified, and tackled, in practice. as well as highlighting the two additional concerns not explicitly mentioned in the literature.

Acknowledgements: This SHERPA project has received funding from the European Union's Horizon 2020 Framework Programme for Research and Innovation under the Specific Grant Agreement No. 786641. The author(s) acknowledge the contribution of the consortium to the development and design of the case study approach. They are grateful to UCLan Cyprus for their contribution to quality assurance of the case study.

Citation: Ryan, M. (2018). Ethics of Public Use of AI and Big Data. ORBIT Journal, 2(1). Retrieved from https://www.orbit-rri.org/ojs/index.php/orbit/article/view/101

Government, Smart Information Systems, and Ethics

In 2050, 70% of the world's population will live in cities by 2050 accounting for 80% of the world's CO2 emissions (Nigon et al. 2016). 80% of Europe's population will live in cities by 2020 (Albino, Berardi, and Dangelico 2015).

Due to the large and growing populations, many cities currently face a huge strain on resources, infrastructure and transportation whilst simultaneously battling with pollution. The use of smart information systems (Artificial Intelligence and Big Data) are being pioneered to help governments meet these needs and to provide a sustainable future for their citizens.

This case study will evaluate a practical example of such efforts, namely the DrukteRadar Project in Amsterdam. One of the aims of the project is 24/7 digital congestion management.

Based on the literature and the DrukteRadar Project, this case study will analyse the ethical and social implications of using smart information systems (SIS) in the governmental domain, asking the primary research question: Which ethical issues arise in the use of SIS in governmental domains and how can they be addressed.

Section 1 will give an overview of the benefits of using SIS in government contexts and provide examples of how they are used in practice. Section 2 will analyse the current literature on ethical and social implications of using SIS in government. Section 3 comprises background research on the DrukteRadar Project in Amsterdam. Section 4 focuses on ethical and social issues in the DrukteRadar Project. Interview excerpts and results from talking to the Project Owner of the DrukteRadar are provided in both Sections 3 and 4.

1. The Use of SIS in Government

Smart information systems (SIS) offer the promise of improving services provided to citizens while reducing costs for city administrations (Zanella et al., p. 23). Big Data will

underpin the future of urban data analytics, and will be an important component within governmental agendas (Bibri 2018, p. 193). The use and implementation of Artificial Intelligence (AI) and Big Data are starting to be widespread and governments are observing the benefits of SIS through improved infrastructure, mobility, and healthcare management.

According to some authors, SIS offers the potential to make governments more efficient, citizens happier, businesses more prosperous, and the environment more sustainable (Yin et al. 2015). SIS allow governments to: improve facilities, improve mobility, develop new services, increase economic growth, increase productivity, reduce energy outputs, improve air and water quality, improve management decisions, respond to disasters quicker, create new business opportunities, and improve citizens' quality of life (Kitchin 2014; Kitchin 2016a; Nam and Pardo 2011; Pan et al. 2016).

SIS are being used in a wide number of different areas within the governmental domain, namely: healthcare, homes, governments, offices, transport, decision-making, security, e-service, and agriculture (Rjab and Mellouli 2018). Some examples are given below.

Yokohama, are implementing AI to aid them with their growing elderly population (Boenig-Liptsin 2017, p. 18)

Picture Freeimages, fabel nard

San Francisco police use the ShotSpotter AI tool, which uses neural networks to detect gunfire by listening for 'sound signatures' (Srivastava, Bisht, and Narayan 2017).

Picture Freeimages, José A. Warletta

50 AI cameras are being trained to prevent drunk passengers from boarding trains in Kyobashi (Osaka) (ibid.).

Picture Freeimages, Midori Sakurai

Drones are being used in fire control in Kansas, US by gathering data about a fire before firefighters are deployed (ibid.).

Picture Freeimages, Sias van Schalkwyk

AI has the potential to assess complex data streams, such as traffic congestion and mobility (Devi and Neetha 2017). The E-Taoyuan and U-Taoyuan projects in Taiwan are implementing e-governance strategies and options for citizens (Albino, Berardi, and Dangelico 2015). SIS may also be used to detect poverty and whether certain actions are improving wealth distribution levels (Glaeser et al. 2018).

There is an abundance of data being retrieved from cities in the hope of applying this information in effective and productive ways for its citizens. The main purpose of using and implementing SIS is to adequately accommodate citizens' needs (Chin, Callaghan, and Lam 2017, p. 2055). However, the different categories of stakeholders involved is quite diverse: private companies, philanthropy organisations, research agencies, universities, institutions, governmental bodies (regional, national, and international), city administrators, and NGOs. For the use and implementation of SIS, SIS specialist bodies are being created: regional bodies, national bodies, specialist units within existing bodies, international specialist institutions and lobby groups, as well as international standards bodies (Kitchin et al. 2017; Yin et al. 2015).

One of the key challenges facing governments is mobility planning and it is hoped that SIS will enable viable solutions in this area (Vázquez Salceda et al. 2014). Big Data can be used to change traffic lights based on the most conducive way for traffic to flow, which in turn is based on patterns and real-time traffic flow. Big data can also be used to map and correlate the consequences of implementing particular changes in traffic management, i.e. 'by closing a road or siting a new hospital on the network' (Kitchin 2016c, p. 3). It is hoped that SIS will be able to provide effective and efficient methods to initiate appropriate changes to these structures. As European populations living in cities will increase to 80% by 2020, there is a growing concern about crowdedness levels within cities. One approach being implemented to combat this is Amsterdam's DrukteRadar project, which is the focus of this case study. However, firstly, it is important to analyse ethical issues found in the literature relating to governmental SIS to be able to contrast this with issues found within the Dutch project.

2. Ethical Issues of Using SIS in Government

The extensive literature review undertaken for this case study used the following bibliographical databases: Google Scholar, ScienceDirect, Web of Science and Scopus. An array of keyword combinations was used to find articles about ethical issues arising from the use of Big Data and AI in governmental contexts. Because of the vagueness of the term 'governmental', several different keyword compositions were needed to allocate appropriate articles, such as city, municipality, government, state, national, and the public. As a result of the literature search, four key ethical issues emerged: accuracy of data; the accuracy of recommendations; power asymmetries; and privacy.

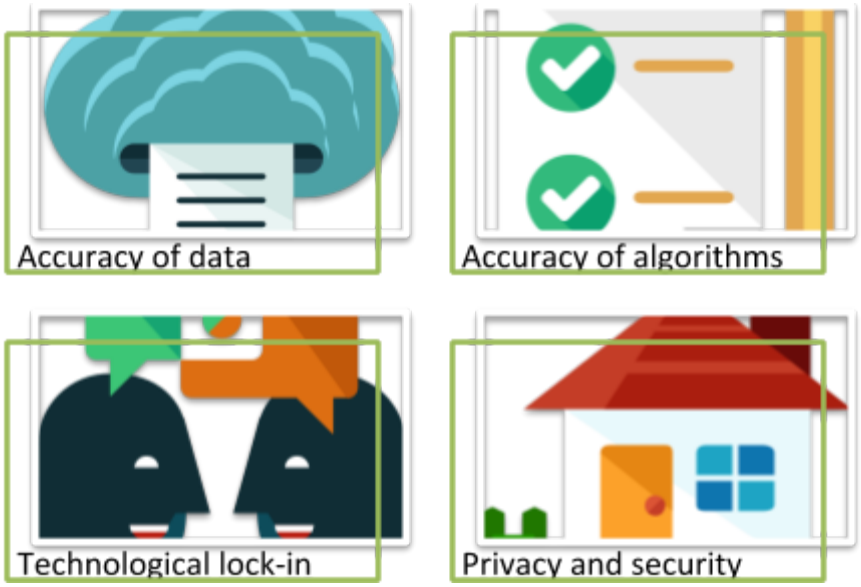

Figure 1 – Key Ethical Issues in the Literature– Using SIS in Government

2.1. Accuracy of Data

There is a shift from the old way of correlating and analysing datasets, which involved physically gathering data by conducting surveys, census, interview data and so forth. These required extensive labour-resources and the findings were highly irregular; monthly, bimonthly, or yearly. With many of the real-time, instantaneous ways of gathering and analysing data, one can instantly create reports and analytics. However, with the ubiquitous drive to find valuable information from Big Data, there is the

possibility that important information is lost from traditional small datasets (Kitchin 2013, p. 265). Not only is there the possibility that one may lose important factors contained within small datasets if only Big Data analytics are used, but one may also be limited by the ability to create integrated models that can effectively analyse this data (Batty et al. 2012).

The data derived and used to make decisions for cities are not value-neutral or impartial (Kitchin 2015a, p. 15). Data is not objective, it is always contextually loaded: established, derived, related and integrated within a wider information and logistical system. Therefore, data may not be accurate, faithful or representative of a city. There is the possibility that the data has errors within it, it may be inconsistent or unreliable (Kitchin, Lauriault, and McArdle 2015, p. 28). Analysing particular city indicators to find particular details may lead to overfitting and 'generalize beyond the data an urban analyst is looking at in a particular context' (Bibri 2018, p. 197). Therefore, it is important to identify potential misclassifications of data and applying data models to potentially incompatible urban data frameworks (Bibri 2018; Glaeser et al. 2018).

2.2. Accuracy of Algorithms

The philosophy that underpins governmental SIS projects is often criticised as being limited in scope and 'not provid[ing] a full and multidimensional picture of the city' (Kitchin 2016b). Analysing Big Data may reduce cities to specific, definable and operable dimensions, conflating many important factors that define a city. This reductionist approach may fail to acknowledge the richness of how a city has functioned 'socially, culturally, politically and economically' (Kitchin 2013, p. 264). There is the belief that cities, guided by SIS, are 'largely rational, mechanical, linear and hierarchical … and can be steered and controlled' (Kitchin 2015a, p. 9; see also Creemers, 2018). In SIS projects, cities may be treated as these knowable, predictable, and guidable entities, when in fact, they are a complex interweave of issues, problems, interests and uncertainties (Kitchin 2016a).

While AI is used to calculate and predict certain behaviours and patterns within the city, there are limitations on its effectiveness. Even though AI can help us with 'highly routinized' patterns, 'there are limits on the extent to which we can explain them and reproduce them' (Batty 2018). There is the possibility that SIS will cause bias, prejudice and harm to a population (Sholla, Naaz, and Chishti 2017). For example, predictive policing is criticised for reinforcing prejudices and racial profiling (van Zoonen 2016, p.

476). When human bias is put into AI algorithms, city services may not be provided equally or fairly (Capgemini Consulting 2017).

A potential issue along these lines was experienced in Boston, where a smart phone app was proposed to uncover potholes on public roads through automated recognition by the phone's accelerometer[1]. "People in lower income groups in the US are less likely to have smartphones, and this is particularly true of older residents, where smartphone penetration can be as low as 16%. For cities like Boston, this means that smartphone data sets are missing inputs from significant parts of the population — often those who have the fewest resources" (Crawford, 2013). In the event, this was recognized before the app was used, and so the issue avoided. However, similar cases may have less foresight in this regard and recommendations based on such efforts would be biased.

For data analytics to have a prescriptive component, there need to be specific benchmarks established prior to data analysis and recommendations. 'Benchmarking and dashboard initiatives thus inherently express a normative notion about what should be measured and how it should be measured' (Kitchin, Lauriault, and McArdle 2015, p. 29). As a result, these benchmarks have the potential to guide discussions, set research agendas, influence governance, and ensure effective decision-making. Establishing unfair or inaccurate benchmarks may have a dramatic effect on issues being evaluated, citizen welfare, and the city as a whole.

Furthermore, there is the possibility that data retrieved from citizens will be used to 'nudge' them in certain directions and conduct activities they may not have performed otherwise (Cardullo and Kitchin 2017). Citizens may be manipulated in a socially-controlled manner, nudging individuals and groups in a particular direction (Kitchin 2018, p. 25). There is the possibility that SIS will be used to socially engineer a population, either to conform to particular practices and behaviours instituted by governmental bodies, which may also lead to power and control issues.

2.3. Technological Lock-in

There is a heavy financial burden to adopt SIS in the short-term, despite holding the potential to save more money in the long-term, so it is debatable if this is the best means of governmental expenditure (Glasmeier and Christopherson 2015, p. 7). Big Data storage facilities, research and innovation, data analysis and implementation, all incur high costs (Hashem et al. 2016, p. 749). Government use of SIS are not guaranteed to become successful and there is a great deal of uncertainty about return on investment.

[1] An instrument that measures the acceleration of a moving body.

Cities implementing SIS are aiming to propel the development of their city, but this is not always possible; such as the Assen Smart City Project, which lost €50 million by installing over 200 sensors around the city (Cloin 2017). This ready-to-go 'Sensor City', it was hoped, would attract ICT investment. This did not happen. The municipality now has to find a private company to purchase the outdated sensors at a fraction of the investment (Naafs 2018). This money could have been much better spent elsewhere.

If corporations are heavily involved with any SIS government project, the city may become overly dependent on those corporations, putting public decision-making and governance in jeopardy (Kitchin 2015b, p. 2). Another concern about the adoption of SIS is that they will increase levels of privatisation of public goods (Kitchin 2016a, p. 23). SIS may drive the privatisation of state services and facilities. Public space, state services and facilities may become commodified, privatised, and used for advertising purposes (Hollands 2015, p. 68). As future cities will be run and controlled by electronic means, how a city functions and operates may be open to attack from to third-parties (Batty et al. 2012). This may cause a wide range of problems for governments, particularly relating to the privacy and security of its citizens.

2.4. Privacy and Security

Individuals are being traced at multiple and varied routes along their journeys within cities and 'are increasingly becoming open to geo-targeted profiling and social sorting' (Kitchin 2016c, p. 8). One's location may be used for stalking, burglary, surveillance, or marketing purposes (Elmaghraby and Losavio 2014; Li et al. 2016). One such example of this is how Renew installed sensors on 200 London bins in 2014, in order to track over four million individuals through their smartphone MAC addresses: 'The company reported that they could measure the proximity, speed and manufacturer of a device and track the stores individuals visited, how long they stayed there, and how loyal customers are to particular shops' (Kitchin 2016c, p. 7). While government officials state that they did not use this data for purposes other than to improve the functionality of the city, it does not mean that such means will not be used for surveillance in the future.

Another key concern relating to privacy protection and governmental implementation of SIS is 'identity privacy'. Identity privacy is the protection of personal and confidential data about individuals' identity (Kitchin 2016c, p. 5). For example, identifying who is using smart parking services or accessing smart buildings (Martínez-Ballesté, Pérez-Martínez, and Solanas 2013). Cities are anonymising and repurposing data in order to be able to use it. However, this data may be de-anonymised in the future. There are also a growing number of unmanned aerial vehicles (UAVs) using cameras, image

repositories, and AI-technology to identify individuals; even if those individuals are not doing anything illegal or untoward. These technologies have the potential to harm identity privacy through SIS, such as AI-generated facial recognition.

Real-time analysis is fundamentally important for governmental SIS because it allows realistic planning of issues that require constant monitoring, updating and adjusting. However, having persistent, instantaneous and ubiquitous retrieval and analysis of data may have a seriously negative effect on citizens' way of life, with the feeling of constantly being monitored. The adoption of SIS may lead to 'dataveillance'[2] and extensive geosurveillance, social and spatial sorting, and anticipatory governance' (Kitchin 2015a, p. 9). In order to counter this effect, citizens should be informed about how their data is being handled, what it is being used for, how it will be stored, and who will have access to it (Galdon-Clavell 2013, p. 720).

There is the possibility that sensitive information may be retrieved, jeopardising the city's functionality if this information is used nefariously. The greater digitalisation of the city infrastructure, the greater the vulnerability to malicious hacking, stolen data, or disruption of systems within the city (Kitchin, Lauriault, and McArdle 2015, p. 20). The costs of security may prevent cities from implementing SIS, causing a trade-off between costs and secure systems (Sen et al. 2013). There is a tension between cost-effective data retention, resulting in shorter storage times and diminished data usability, and incurring higher costs for improved data storage (Li, Cao, and Yao 2015). If there are inadequate security protections, citizens may reject SIS implementation (Zhang et al. 2017).

Having now discussed the ethical issues that the literature has identified with regards to the governmental use of SIS, it is clear that these cover a broad array of issues, most of which are not specific to governmental use. In order to better understand the way in which these issues arise practical contexts, the next section describes a case of a governmental SIS. The case was chosen because this project is using a large abundance of different datasets, obtaining public and private data to make predictions for the benefit of citizens and the departments within the municipality. The project is a unique way of integrating and using SIS in a governmental context, which could prove to be an innovative application that would be beneficial for all busy cities around the world. However, pioneering projects also come with their disadvantages, namely, that there are no similar projects to compare and contrast them with, and often ethical evaluations have not been carried out on them. Therefore, it is important that they are critically evaluated by third-party organisations (such as SHERPA) to determine their ethical viability and societal impact.

[2] The practice of monitoring digital data relating to personal details or online activities.

3. DrukteRadar: Governmental Use of SIS

The focus of this section will be on a governmental use of SIS in practice, namely, Amsterdam's DrukteRadar Project, which loosely translates as 'crowdedness project'. While the project is in its early stages of development, it can offer some insights about SIS projects being implemented in Europe in a governmental context. Interviews were conducted with the Project Owner of the DrukteRadar, a computer scientist working on it, and a further colleague working on the project. The computer scientist and the other colleague working on the project were not formally interviewed, thus not quoted in the case study. The interviews were analysed using a qualitative analysis software tool (NVIVO).

3.1. Description of Amsterdam's DrukteRadar

Amsterdam is the capital, and largest city, of the Netherlands, with the municipality comprising of over 200 km². The 850,000 Amsterdammers are proud of their city and the municipality is always making improvements for the 180 different nationalities that inhabit it. Along with a highly condensed population, the city receives a staggering 18 million visitors a year and is the 8th most visited city in Europe. This number is set to jump to 25 million by 2025 (Hein 2016). Tourism accounts for over €75 billion in the Netherlands and consists of 3.9% of GDP, employing 641,000 people (Pieters 2018). The municipality realises the benefits tourism brings to the city and its dependence on ensuring that all significant sites are well-preserved, amenable, and accessible to visitors and locals alike. Tourism is very important for Amsterdam, but it comes with the cost of overcrowding. In recent years, Amsterdam municipality has attempted to take a proactive approach, initiating a number of innovative approaches, through their Chief Technology Office (CTO).

Amsterdam, photo: Peter Hellbrand, free images

One of the goals of the CTO is to create a roadmap to ensure the city is futureproofed. Its aims are to enable the city and its authorities to connect, accelerate and strengthen new projects and to create solutions to problems within the city. The CTO and the municipality have been involved in a number of projects in recent years such as:

- sharing traffic data with a technology company in order to help them tackle traffic congestion in the city;
- using data to identify depression hotspots to provide better care and services to those areas;
- "Beautiful Noise", which assesses social media comments from museum queues to improve services;
- "Rain Sense", which identifies places with bad rainfall to prevent flooding;
- smart street lighting; and
- pay-by-phone parking apps (Fitzgerald 2016).

The CTO is responsible for some of the most innovative approaches being implemented within the municipality, such as the DrukteRadar Project, the planning for which started in early 2017. The aim of the DrukteRadar Project is to use data analytics to anticipate and prevent overcrowding in Amsterdam. The output of the project was supposed to be launched in the summer of 2018, but it was postponed until the second quarter of 2019 to ensure functionality and that it was fit-for-purpose. The interviewee for this case study is the Project Owner of the DrukteRadar. He is the person who has to make strategic decisions about the project's direction. He said that the project was funded for the next four years and had the full endorsement of the municipality. It is an innovative platform that is set to greatly benefit the city of Amsterdam.

3.2. Description of SIS in the DrukteRadar Project

The aim of the DrukteRadar Project, as noted earlier, is to use data science to anticipate and prevent overcrowding in Amsterdam. The project uses technological tools to create data about crowd levels in the city for effective use by locals, tourists, and city managers. [3] It was created by the CTO and City in Balance[4] group in response to growing pressures

[3] The DrukteRadar operates through a number of docker containers, which are run on their server. There is a mixture of front-end, back-end, and databases connected through this network to run the dashboard. For the front-end, there is standard HTML, CSS and JavaScript, and Application Programming Interfaces

on Amsterdam's amenities, navigation, and tourist and citizen experience. The aim was to create a dashboard so that both the municipality and Amsterdammers could identify when areas of the city are crowded.

The project aims to identify potential crowd numbers from hotel stays, events, and crowd-size projections. This involves retrieving very large amounts of data from inside and outside of the municipality, while effectively analysing and visualising those data. The project first identified that in order to have a fully functional and successful crowd prediction model, they needed to collect and analyse large datasets. The data retrieved focuses on the number of people in particular areas, length of time they stay there, and problem areas. The CTO wants a wide array of Big Data from:

i. Public transport: GVB, Connexxion, NS and Translink.

ii. Private transport: TomTom, HERE, tech companies, TCA, the Municipality itself.

iii. Pressure on locations: Google.

iv. Telephone: Vodafone, T-Mobile and KPN.

v. Hotel occupation: Booking.com, Airbnb.

vi. Economic activity: Mastercard, ABN AMRO, ING, Rabobank.

vii. Tourism & recreation: Tours & Tickets, Booking, AirBnB, Schiphol.

viii. Social media: Facebook and Twitter.

All of the data retrieved from these different sources is stored on the municipality's data storage facility, DataPunt. This is a low-cost open-source storage space and the municipality places a great deal of emphasis on ensuring its security. DataPunt was created in early 2016 and is the data store function of Amsterdam Municipality. It has a number of modules that retrieve, link, and receive data as well as making this data accessible to other ICT systems. There is a need to develop the data landscape of the municipality and to effectively use and integrate data into helpful, manageable ICT processes (Moerman 2017, p. 7). Some of the data types are: 'Information about public space, buildings and lots, traffic, care, environment, quality of life, permits, subsidies and numerous other data' (Municipality of Amsterdam 2018). DataPunt is the municipality's storage facility and functions as the central data portal between the municipality and its

(APIs) running in separate containers. The database is in a separate container and the importer transforms the data into the right format. They use weak patterns as complex machine-learning methods would over-fit the data.

[4] It experimented with ways to find a better distribution of tourists in the city (City of Amsterdam 2019).

partners. The large datasets stored in DataPunt are ready to be reused in Amsterdam projects, such as the DrukteRadar.[5]

The project's dashboard is an interactive map of the city, with specific 'hotspots areas' map and a 'district level' map of the city. The hotspot map identifies how crowded specific locations within the city are. It is meant to inform Amsterdammers, tourists, and also those working in the municipality about particular locations in the city. There are also layers for automobile traffic, bicycle traffic, and pedestrian traffic in the city.

The district level dashboard identifies how crowded entire districts are and is designed for the municipality, rather than locals or tourists. The hotspot and district maps distinguish the crowdedness by colour coding: the lowest levels are green, with mid-levels shades of orange, and problematic areas are red. However, the way of measuring district and hotspots is different. The district level is measured on an objective person-per-square-meter metric, while the hotspot level is measured on a historical relative account of those areas (however, this may change to the more objective approach in the future).

3.3. The Aims of the Organisation Using These Technologies

There is a pressure on the city of Amsterdam to accommodate increasing numbers of people. Everyone wants to visit the same landmarks, which are centrally located or are within certain areas of the city.[6] There is a concern about how the municipality will cope with the pressure on amenities in the city. As a result of these issues, there were a number of different approaches suggested in order to ensure that residents and visitors can enjoy the beautiful city with comfort and ease. One of the main issues raised was the need to identify problem areas, times of peak busyness, and ways to ameliorate these issues. In conjunction with different organisations and departments within the municipality, it was proposed that crowd management could be tackled by the program City in Balance, which created the DrukteRadar Project.

[5] The data required by DrukteRadar is sent from DataPunt, and is processed in the back-end docker container. The analyser container uses machine-learning and works with the database container, to find weak patterns emerging within the datasets. The developers have created an Application Programming Interface (API) that is used by the visualisation docker container to illustrate the crowdedness levels on a dashboard.

[6] Tourists mainly congregate in the following areas: The Old Centre; the Museum District & Vondelpark; Oud-West; the Canal Belt and Jordaan. While other areas in Amsterdam are also witnessing increased volumes of tourists in recent years: De Pijp; Plantage; North Amsterdam; the Docklands; and Amsterdam Zuid.

The DrukteRadar Project will collate a wide array of datasets to predict crowd levels, movement, and potential problem hotspots. The data will be used for better municipality management, for greater visibility for tourists planning their trips, and for citizens to navigate through crowds during their daily activities. One of the main benefits of this project is anticipated to be that the Council can implement better management decisions and strategies, as a result of increased knowledge of crowd levels. Different departments will need data on crowd levels for different reasons and the Project Owner's responsibility is to determine what they need from the data. So, for instance, the Waste Department asks the Project Owner to identify crowd levels so that they can ensure the city remains clean; the Transportation Department needs data to alleviate traffic congestion, and the Police Department needs data to identify potential problematic areas.

3.4. The Effectiveness of the SIS Technology during Use

The Project outlines six specific steps to ensure the effectiveness of the DrukteRadar project:

Steps	Description	Details
1	Acquiring data	Individuals tasked with retrieving data within the city and externally.
2	Estimating and planning potential	Prioritisation of data retrieval is agreed upon, depending on the high level of impact and necessity.
3	Make good agreements	Identifying who manages the data and how it is used.
4	Database storage and API creation	Data is stored at DataPunt, which is under supervision of Amsterdam municipality.
5	Data processing and front-end development	Data is evaluated to determine specific issues and is made available to front-end developers for visualisation purposes.
6	Testing and development	The software is tested to determine if it is fit-for-purpose. End-users are consulted to identify issues with the tool.

The CTO will establish a number of different Key Performance Indicators (KPIs) and metrics to determine the success and worthwhileness of the project. While the project is still in its early stages, they want to identify a number of different metrics, such as:

- how many stakeholders will the project help to resolve issues;

- will the project help optimise garbage collection routes, also leading to a reduction of Co2 emissions and lower fuel costs;
- how many people use the website or the project and how do they use it; media coverage received; positive or negative feedback; and reductions in busy hotspots.

3.5. Stakeholders Involved in the DrukteRadar Project

The DrukteRadar aims to incorporate a wide range of stakeholders in its approach, and suggestions for further stakeholders are considered necessary to improve the project. The end-users are incorporated into the stakeholder group and are involved throughout the design and implementation of the project. The DrukteRadar Project Owner receives regular feedback from users and test panels, which are incorporated into dashboard developments.

The project team works in an Agile format, according to SCRUM principles,[7] and the Project Owner is responsible for the overall project. His main responsibility is transforming stakeholder input and feedback into concrete tasks for the development team to implement. The project also has an advisory group, consisting of scientific experts, where they provide advice on new priorities every six months. There is also a tactical management session every three months to identify potential issues. The stakeholder group consists of members from the following groups: V&OR; OOV; Smart Mobility; EZ; CTO; R&D; Parkeren; City Works; Clean Stadsdeel Centrum; Schoon; Afval; and OIS. This stakeholder group assists the project in developing the agenda and ensuring agreed standards are met.

In order for the project to work successfully it requires internal and external data. The municipality wants to pay as little as possible for data. Currently, data providers only have data that is a week old, or even a month old, which is too old for the project's requirements. External data retrievers should therefore coordinate partnerships on behalf of the city of Amsterdam to obtain newer data. There needs to be a bridge between data providers and the municipal government. There is a legal memorandum of understanding (MoU) between the municipality and data providers to ensure agreement about their partnership. The municipality aims to incorporate prevalent public figures in the discussion of crowd management to retrieve insights into the development of the project. The Project Owner said that it is a hot topic in the media and the public is aware of the

[7] The SCRUM principles are: focus, openness, respect, courage and commitment ("Scrum Values & Agile Methodologies for Project Management," 2018).

need for such intervention in the city. The project is considering the integration of a feedback option for citizens to provide recommendations. They want to make clear that there will be many iterations throughout the project's development.

4. Ethical Implications

After careful analysis of the DrukteRadar documentation, conversations with the team, and the interview with the Project Owner, a number of pertinent ethical issues were identified that may result from the project's implementation and use. The six ethical issues identified were: access to SIS; accuracy and availability of data; the accuracy of algorithms; ownership of data; technological lock-in; and privacy and security.

Figure 2 – Key Ethical Issues DrukteRadar project (some images from freeimages.com)

These topics were discussed during interviews and largely coincide with the ethical issues that are found within the literature on the topic of SIS implementation in a governmental context. Two ethical issues that did not specifically arise in the literature concerning the governmental use of SIS, but did arise in the interview, were access to SIS and ownership of data. Overall, there were a lot of similarities between what the project views as ethical concerns that need to be incorporated and what has been identified as concerns in the literature.

4.1. Access to SIS

The Amsterdam municipality is aware that most people are not technologically savvy, so it had to adopt a more bottom-up approach trying to make their actions as transparent as possible to citizens (Fitzgerald 2016; Grashoff 2017b, p. 3). One goal of the DrukteRadar Project was to ensure that it was user-friendly and to make it available to as many people as possible. The effectiveness of the DrukteRadar Project relies heavily on whether or not it is used, which relies on how it is disseminated. The DrukteRadar team have identified that social media and Google were obvious dissemination channels, but how exactly this data would be distributed was a little unclear. The Project Owner noted that they were going to do two things to ensure availability and access to the DrukteRadar: use the project for citizens and the public, but also provide it to other departments within the municipality to use.

He also mentioned that the project will disseminate its services through the dashboard, which was explained in section 3.2. of this case study. This visualisation will clearly show citizens where it is busy and what levels of crowdedness they can expect at a particular time. In order to ensure that citizens are made aware and have access to the dashboard, the municipality is going to begin a social media advertising campaign in 2019, through the city of Amsterdam website, and possibly through newspapers and other outlets. However, at the moment, this is not a top priority as they are more focused on having the project successfully deployed. Most of the data and documents are available on the municipality's open data platform, and for the success of the project, it is important that it has accurate and available data for its recommendations.

4.2. Accuracy and Availability of Data

One of the chief problems with implementing SIS is the lack of available data repositories. In itself, it is not an ethical issue, but as a result of poor-quality data repositories and data availability, algorithmic predictions may lead to discriminatory recommendations, inaccurate predictions, and harmful consequences. Data accuracy and availability is therefore an important concern for Amsterdam Municipality.[8]

[8] Amsterdam's Chief Technology Officer mentioned that they found it difficult to calculate how many bridges they have in the municipality (Baron 2016). When Ger Baron took on the role of Chief Technology Officer he grew their inventory to 12,000 datasets, which allowed them to get a wider picture of problems facing citizens (Fitzgerald 2016). He noted that they also use analytics frameworks to assess small data so that they do not miss crucial data in these sets (Baron 2016). For example, the AMS Institute used Wi-Fi hotspots, cameras, GPS tracking and social media posts to determine crowd sizes, movement of crowds, and congestion areas during the 2015 SAIL Amsterdam Festival (Fitzgerald 2016). During this event, Baron noted that there was something delaying people at a point on the map and researchers made guesses about what this could be. In reality, there was a trash-can in that location, which was not on their map, and people were slowing down to throw rubbish into it (Fitzgerald 2016).

Accuracy and availability of data issues are prevalent in the integration and use of emerging technologies, and one of the limitations for their effectiveness is the lack of available data to train their models. For example, the DrukteRadar project requires more real-time data to improve their models. Data is only compiled from a few months previously, so the project is unable to give a cohesive evaluation and comparison of the levels of crowdedness of the previous year. The project needs to have larger, longer, and more detailed datasets so that they can produce more effective models. Otherwise, if they apply their current datasets to these models, they would be over-fitted[9], which would be problematic for accuracy and effectiveness.

So far, the project has incorporated data from various sources in the municipality. Internal sources are a wide range of different city departments and groups, whereas external data is retrieved from different partner organisations of the municipality. It is often difficult to obtain the most effective and useful data for the project and as a result of the General Data Protection Regulation (GDPR) and other legislation, many departments are apprehensive about providing data to the project. In addition, the Project Owner highlights that sometimes it is difficult to get data from different departments because they want to see a benefit for providing this data. Even within the municipality, there needs to be a mutual benefit for data provision. The DrukteRadar team needs to identify how departments can benefit from the use of this tool if they provide their data repositories.

The Project Owner acknowledges that internal data has its limitations. He stated that when analysing the busyness of the city, less populated and more secluded areas have little data on how busy they are. Private organisations, such as telecom providers or Google Maps, have a greater availability of datasets covering wider ranges of the city. Another limitation with acquiring internal datasets is that it requires a great deal of time. He pointed out that it is a challenge to acquire these datasets because there needs to be a mutually beneficial relationship to acquire them.

'So, if you can say that, "Give us the data, and we'll make sure that you can run your external operations more efficiently", then that helps. That's the kind of win-win situation we want to create' (Interviewee 1).

In relation to external data, the project has 'data hunters' that deal with semi-public or private organisations to obtain appropriate data for their algorithms. The data hunter tries to negotiate partnerships with private companies to establish a mutually beneficial relationship. The private organisations benefit from having a big client like the

[9] Overfitting a model is when there are more parameters inputted than are justified by the data provided. It is attempting to extract richer, more nuanced information from limited datasets.

Amsterdam municipality, and the project benefits from the provision of data. The Project Owner says that

'they can fine-tune their business case, and together we can develop a project that's also interesting for other similar cities and based on that partnership, we get the data for free, and they can use their insights they get from us for other cities' (Interviewee 1).

The DrukteRadar Project aims to minimise the number of errors occurring, but it is important to have preventative and ameliorative steps in place for when they happen. If there are errors or discrepancies within city data, Amsterdam has a number of contingency plans in place: it has extensive monitoring facilities, and a secure technical infrastructure (Gemeente Amsterdam 2018b). A range of questions have to be asked in terms of the data derived, such as:

- is the data suitable,
- where does it come from,
- who has access to it,
- is it current,
- and how is it managed (Moerman 2017, p. 5)

Furthermore, to ensure that their data is accurate and effective, the DrukteRadar Project provide the general public with the option to contact them if they identify any problems with the product's recommendations (Gemeente Amsterdam 2018a). It is important to establish accurate and effective data so that the municipality can make accurate and appropriate recommendations and predictions, thus minimising any harmful effects as a result of improper measures and policy.

4.3. Accuracy of Algorithms

The DrukteRadar Project, initially implemented relative data measurements for the hotspot levels in cities and objective data measurements for the district-level dashboard. However, there is a concern about using different methodologies to depict the crowdedness level in the city, with the district level applying persons-per-square-meter, whereas the hotspot level applies the historical relative account of crowdedness.

The hotspots are compared against historical data, and against what a 'typical Thursday' looks like; while seasons or other variables are not accounted for, yet. The amount of historical data only extends back a few months, making it impossible to give an annual

forecast or projection. There are options to increase and decrease the approximate levels depending on seasons, but the methods used might be inaccurate. The hotspot measurement identifies a particular location. For example, the level of busyness of Dam Square at 13:00 on a particular Saturday is assessed against what the standard level of busyness for Dam Square on a Saturday at 13:00 (relative data), rather than how busy Dam Square is compared to other locations in the city (objective data). This is to cater to locals so that they can identify areas that may be problematic for their commutes or activities, while also acknowledging that there are certain areas that will always be quite busy, such as Dam Square.

However, one would need some previous knowledge and understanding of areas within Amsterdam to unpack this information, making it difficult to cater to both tourists and locals alike. The tourist may not understand the relative index being used in the project; for example, knowing that Dam Square at a medium level of business is still very busy. This may make it difficult and confusing for the non-local user. If the project uses two different maps, one for the local and another for the tourist user, it may also be problematic because the local may not be able to interpret what a standard level of busyness is for an area. For example, what if they are a local but have never been to that particular landmark in the city so are unaware how busy it is regularly, or perhaps they are local, but new to the city.

The Project Owner mentioned that an important issue for the project was to identify who the end-user will be and then design the platform based on this understanding. The needs of a citizen of Amsterdam would be different from a tourist, which would be different from departments within the municipality using it. Once the team finalise who the end-users will be, they will be able to create different landing-pages, structuring the data visualisation according to their needs. He noted that they recently veered towards using more objective data, rather than relative data, for identifying the hotspot busyness level of the city.

4.4. Ownership of Data

One of the explicit goals of the DrukteRadar Project is to ensure data sovereignty. While Ger Baron agrees with the need to incorporate SIS in Amsterdam, he claims that tech companies do not always understand cities and how they work, emphasising that what works in theory does not always work in practice (Baron 2016). Models and technological propositions may confine and reduce the complexity of a city such as Amsterdam. Baron states that they would have left a great deal of technological innovation to the companies in the past, but he has been looking closely for governmental

involvement in recent years (Baron 2018). He wants to take a stronger role in SIS development, as he believes that they can 'deal with this much more effectively and efficiently' (Baron 2018).

When the municipality works with private organisations to evaluate data, and use it for the DrukteRadar's functionality, sometimes the issue of data ownership is raised. The Project Owner indicated that the ownership of data depended on where the data was collected from. If the municipality creates the data, then it is the owner of this data; whereas, if the partner organisation creates the data, then they claim ownership of it. If the data is collected in a public space, then the city of Amsterdam tries to ensure that they obtain ownership of that data. This is achieved through contracts and partnerships with private organisations.

The relationship between the municipality and internal and external data providers is a data partnership. The Project Owner stated that the organisations' role is a data provider and they have no control over the running of the project. Both internal and external partners do not have any involvement or say in how the project is structured, organised, or how they create their models. They have a vested interest in the success of the project, either for personal use (internal partners) or having Amsterdam municipality as a customer (external partners). The DrukteRadar Project aim to ensure a mutually beneficial relationship is achieved between public and private entities with a high degree of data sovereignty.

4.5. Technological Lock-in

There will always be tensions between public and private interests in the use and implementation of SIS in governmental applications. Ger Baron stated that the city of Amsterdam pushed back on companies from installing transmitters and sensors on street lamps around the city because it was not in the public's interests (Baron 2016). Amsterdam rejected the idea to install smart lighting because of the extremely high costs of investment (Fitzgerald 2016). The municipality tries to succeed with an independent approach from third-party companies, but often this is not possible.

The DrukteRadar project has already come up against problems attempting to establish a data-sharing partnership with a bank, but the bank's lawyers opposed such a relationship. While some telecommunications companies will not even discuss this topic with the municipality of Amsterdam. The CTO has realised that obtaining data from third-party organisations is not easy and often requires considerable costs. Often, these third-party organisations want a collaborative arrangement in order to benefit both parties.

The municipality is in the process of setting up arrangements with NS and Schiphol airport to create data sharing partnerships. There is the hope that these agreements will be mutually beneficial, rather than Amsterdam purchasing or becoming a customer of these organisations, rather than Amsterdam municipality necessarily being dependent or locked-in to relationships with these organisations. The Project Owner mentioned that one way of creating this partnership is data-sharing.

4.6. Privacy and Security

Because data is being retrieved about citizens, their location, their time spent there, and because these data may be linked to what they are doing there, there is a concern over privacy for the DrukteRadar Project and how to ensure that it is protected. The project has a mixture of objective data, but only uses aggregated data from a minimum of 15 people so that it is impossible to identify individuals from the group. In instances where the number of people in an area is below 15, the system would produce a result of zero, in order to ensure that individuals are not identifiable to ensure privacy. Fundamentally, the project does not want

'any personal data of any sort. And we actually want as little data as possible; just the necessary data that we can use to run our models, and make it as efficient as possible' (Interviewee 1).

There are a number of different data sources being retrieved, ranging from low to medium privacy risk. For example, weather data has no privacy concerns, whereas, park and ride data (which shares how many spots are being used) and NDW traffic data (such as speeds on the road and average speeds) have a low privacy risk, because of the lack of personally-identifiable information. The data retrieved from bike rental has no personalized data, just the quantity of bikes; and public transportation data is protected by having specific cut-off values as described earlier.

The city aims to protect citizens' privacy as much as possible and claims to ensure that the open data that is available on their website is not traceable back to any individual person (City of Amsterdam 2018). They do this by ensuring that no individual's data is input in the first place and any datasets that are used follow their privacy-by-design principles. The Municipality explicitly state that an individual's data will never be given to commercial institutions or private individuals unless written permission is granted. In addition, one can request that one's data is not shared with non-mandatory administrative authorities. The DrukteRadar team ensure that

'every data source we use has the privacy-by-design incorporated in the source. So that means that privacy is designed within the data collection mechanism' (Interviewee 1).

Partners providing data must follow strict adherence to the organisation's privacy-by-design principles. This is carried out at the start of any collaboration with different internal or external partners. Internally, it is achieved by the privacy officers within the city itself,

'who advise on these projects, and they always advise to collect no personal data, or at least as little as possible' (Interviewee 1).

The privacy officers ensure that no personal information is passed on in the datasets and that all the privacy concerns are adhered to. The project only wants to use and have access to non-private information. The Project Owner mentioned that sometimes this can take a lot of time with back-and-forth communications between different departments and privacy officers, but that they wanted to ensure they meet current privacy protection standards.

'Normally all the data that's open, or that's not privacy related, we can share' (Interviewee 1).

He elaborates that the open data does not contain any private or personal information of any of the citizens of Amsterdam, as they do not store or retrieve any personal data or personally identifiable information. They place a strong focus on ensuring the privacy of its citizens is protected. In relation to external datasets, the Project Owner pointed out that the level of aggregations of these datasets makes it impossible to extract personally identifiable information. However, the department still ensures that external partners also abide by the seven privacy-by-design principles.[10] In addition to implementing privacy protocols, the project also protects privacy by ensuring that all data is safely secured on the municipality's data storage system.

The DrukteRadar uses data from DataPunt, which is a heavily secured cloud system, used to store and collect data. DataPunt is the Amsterdam data node facility that makes information about the city open and available (Grashoff 2017b). They emphasise that distribution of data is only provided to those that need it and respecting legal requirements (i.e. in accordance with national, EU, and municipal legal regulations). Amsterdam Municipality protects this data from being hacked and compromised. DataPunt minimises privacy risks by reducing the amount of personal data permanently

[10] The seven principles are: proactive not reactive; privacy as the default setting; privacy embedded into design; full functionality: positive-sum, not zero sum; end-to-end security: full lifecycle protection; visibility and transparency; and respect for user privacy ("The Seven Foundational Principles," 2018).

stored, strengthened cyber-security measures, anonymization techniques, and repeated vulnerability tests (Grashoff 2017b). The DrukteRadar follows the principles outlined in DataPunt to ensure security levels are met. The Project Owner claims that it is very secure and reliable.

5. Conclusion

SIS may be used by governments in a wide diversity of applications, which can result in a wide range of social and ethical issues. This case study evaluated ethical and social issues found in the literature, how SIS is being used, and what actions are being taken to counter harmful impacts on society. The case study evaluated a specific application of governmental SIS, the DrukteRadar Project, to provide context for how municipalities are dealing with ethical issues. Focusing on the use of SIS in Amsterdam, provided key insights into ethical issues and how they contrast with those found in the literature.

The interview with the Project Owner provided information about the real-life application of governmental SIS while addressing a number of ethical issues as a result: access to SIS; accuracy and availability of data; accuracy of algorithms; ownership of data; technological lock-in; and privacy and security, which are being addressed by the DrukteRadar team.

The DrukteRadar team are ensuring that stakeholders are made aware of the dashboard developed for the project through a public awareness campaign on social media, newspapers, and their website. The DrukteRadar team realise that accuracy and availability of data underpin the success of their project, so they have been establishing partnerships with internal and external data providers. The municipality also wants to maintain its data sovereignty, so the project tries to concentrate on using internal data to guarantee data ownership and avoid technological lock-in. The project also ensures that datasets follow privacy-by-design principles and are GDPR compliant.

5.1. Limitations

A notable limitation of this case study is that it is analysing SIS that have not been fully implemented and disseminated to its target audiences. When I first contacted the organisation, the project was supposed to be launched by October 2018, but this date was pushed back until Spring 2019 because the team wanted to refine it before launch. While many of the issues raised in this case study are pertinent, there may be additional issues

that are only visible once it is implemented and used in practice. This could have been valuable exploratory terrain for the case study.

One of the main limitations of this case study is that it was based on only three interviews, only one of which was analysed in a structured manner. While the primary interviewee was knowledgeable about SIS and their societal and ethical impacts, the case study would have been strengthened by additional interviews. Furthermore, as the case study only examined one example of SIS in Amsterdam, it does not allow for generalisations about their use of SIS, or broader conclusions about municipalities' use of SIS, generally. The project also focused on a specific application of SIS (i.e. for crowdedness detection and prevention), which would make it impossible to difficult to make broader deductions about the use of SIS by governments.

5.2. Contribution to Knowledge

This case study offers the discipline a fresh look at the ethical concerns associated with using SIS in a governmental setting. While some of the topics and issues discussed in this case study have been evaluated elsewhere, they have rarely been applied to specific projects in this way. Therefore, the case study has specific implications for SIS ethical theory and analysis. Furthermore, there has been very little written about Amsterdam's ICT projects, besides self-published articles on their websites. This case study offers the municipality an objective evaluation of one of their latest projects, prior to its full-scale implementation, and ethical issues that may arise in the project's development.

The analysis of the DrukteRadar Project contributes an empirical evaluation of SIS being used in practice, uncovering two specific issues that were not found in the literature: access to SIS and data ownership. It was shown that a municipality can increase access to SIS, promoting fairness, equality, and provision of services amongst citizens. The DrukteRadar project achieves this by ensuring that its dashboard is user-friendly and available to as many people as possible.

While there are general policy guidelines and frameworks for the ethical use of SIS, there are few recommendations specific applications of SIS, particularly around data ownership and relationships between public authorities and private companies working on SIS. The case study offers insights into how municipalities can tackle the issue of data ownership, namely, if the data is collected from a public space, then the city should claim ownership, rather than private companies.

5.3. Implications of this Report

This report has policy implications for the use of SIS within the governmental domain, particularly municipalities. When cities integrate SIS, there should be ethical considerations about how they will affect citizens, the city's relationships with private companies, and their impact on society.

One way to do this is by identifying who the end-user is during the design phase of SIS. Cities need to acknowledge that a lot of people are not technologically-savvy. The effectiveness of public SIS depends on whether or not it is usable and available to most people. Policymakers should have clear dissemination plans to reach their end-user, in this case, the citizens of Amsterdam. Municipalities may use this case study to understand important concerns related to the implementation of SIS. This report showed that there needs to be a win-win scenario in data partnerships, whether they are public or private.

However, what became evident in this case study is that private companies may not always understand how cities work, so city officials need to be careful when collaborating with them on SIS projects. Departments within the municipality may deal with tasks more efficiently and effectively than private companies. Investments into training internal staff, instead of outsourcing projects to costly ICT companies, is something that should be considered. This case study also demonstrated how ensuring data sovereignty could help avoid becoming technologically locked-in and dependant on private organisations.

This report provides policymakers with fresh insights into how they can reduce discriminatory recommendations and harmful consequences resulting from SIS – through extensive monitoring, secure technical infrastructure, and stakeholder review sessions. Technical insights were provided to show how cities can ensure citizens' privacy is protected: data is not traceable to individuals (aggregated data from a minimum of 15 people), internal and external datasets follow privacy-by-design principles, and try to obtain as little personal data as possible. Finally, this report provides ways municipalities can ensure data security: use a heavily secured cloud system, anonymization techniques and repeated vulnerability tests.

5.4. Further Research

The literature review aimed to provide an overview of the main ethical issues being discussed in governmental use and implementation of SIS. Additional research may need

to be completed on ethical issues that arise in the future relating to governmental applications of SIS. There is a need for further empirical investigations into the use of SIS by governments, through discussions about how cities are using SIS, as well as how different international institutions, or supra-international bodies, are using SIS. Furthermore, additional case studies may allow for cross-case analysis with other empirical research on SIS use. Overall, the hope is that this report encourages more academic evaluations into the ethical use and implementation of governmental SIS.

6. References

Albino, Vito, Umberto Berardi, and Rosa Maria Dangelico, "Smart Cities: Definitions, Dimensions, Performance, and Initiatives", *Journal of Urban Technology*, Vol. 22, Issue 1, 2015, pp. 3-21.

Amsterdam Smart City, "Perspectives on the Future of Mobility in Amsterdam", *Smart Mobility in Amsterdam*, 26th March 2018.

Baron, Ger, [Interview], in Fitzgerald, Michael, "Data-Driven City Management: A Close Look at Amsterdam's Smart City Initiative", *Sloan Review*, May 19th, 2016, https://sloanreview.mit.edu/case-study/data-driven-city-management/

Baron, Ger, [interview] in Daalder, Leonieke, "Ger Baron (CTO of the Municipality of Amsterdam): 'How do we go from Government to GovTech?'", Marketing Facts, 6th March 2018: https://www.marketingfacts.nl/berichten/ger-baron-cto-gemeente-amsterdam-overheid-govtech

Batty, Michael, "Artificial Intelligence and Smart Cities", *Environment and Planning B: Urban Analytics and City Science*, Vol. 45, Issue 1, 2018, pp. 3-6.

Batty, Michael, Kay W. Axhausen, Fosca Giannotti, Alexei Pozdnoukhov, Armando Bazzani, Monica Wachowicz, Georgios Ouzounis, and Yuval Portugali, "Smart Cities of the Future", *The European Physical Journal Special Topics*, Vol. 214, Issue 1, 2012, pp. 481-518.

Bibri, Simon Elias, "Data Science for Urban Sustainability: Data Mining and Data-Analytic Thinking in the Next Wave of City Analytics", *Smart Sustainable Cities of the Future*, Springer, Cham, 2018, pp. 189-246.

Boenig-Liptsin, Margarita, "AI and Robotics for the City: Imagining and Transforming Social Infrastructure in San Francisco, Yokohama, and Lviv", Artificial Intelligence and Robotics in the City, Issue 17, 2017, pp. 16-21.

Capgemini Consulting, "Unleashing the Potential of Artificial Intelligence in the Public Sector", Capgemini website, 2017, retrieved 27th July 2018: https://www.capgemini.com/consulting/wp-content/uploads/sites/30/2017/10/ai-in-public-sector.pdf

Cardullo, Paulo, and Rob Kitchin, "Being a 'Citizen' in the Smart City: Up and Down the Scaffold of Smart Citizen Participation", *The Programmable City Working Paper 30*, SocArXiv Website, 15th May 2017.

Chin, Jeannette, Vic Callaghan, and Ivan Lam, "Understanding and Personalising Smart City Services Using Machine Learning, the Internet-of-Things and Big Data", *IEEE 26th International Symposium on Industrial Electronics*, 2017, pp. 2050-2055.

City of Amsterdam, "City Data", *Gemeente Amsterdam* [website], retrieved 30th July 2018, https://data.amsterdam.nl

City of Amsterdam, "Policy: City in Balance", City of Amsterdam [website], retrieved 10th January 2019, available here: https://www.amsterdam.nl/en/policy/policy-city-balance/

Cloin, Cindy, "Assen moves to utopia 'Sensor City'", *Trouw*, January 7th, 2017, https://www.trouw.nl/home/assen-vertilt-zich-aan-utopie-sensor-city-~a2638a3b/

Crawford, K., "The Hidden Biases in Big Data" [WWW Document], *Harvard Business Review*, 2013, available at: https://hbr.org/2013/04/the-hidden-biases-in-big-data (accessed 1.4.17).

Devi, Suguna, T. Neetha, "Machine Learning Based Traffic Congestion Prediction in a IoT Based Smart City", *IRJET*, Vol. 4, Issue 5, 2017, pp. 3442-3445.

Elmaghraby, Adel S., and Michael M. Losavio, "Cyber Security Challenges in Smart Cities: Safety, Security and Privacy", *Journal of Advanced Research*, Vol. 5, Issue 4, 2014, pp. 491-497.

Fitzgerald, Michael, "Data-Driven City Management: A Close Look at Amsterdam's Smart City Initiative", *Sloan Review*, May 19th, 2016, https://sloanreview.mit.edu/case-study/data-driven-city-management/

Galdon-Clavell, Gemma, "(Not So) Smart Cities? The Drivers, Impact and Risks of Surveillance-Enabled Smart Environments", *Science and Public Policy*, Vol. 40, 2013, p. pp. 717-723.

Gemeente Amsterdam, "Availability and Quality of Data", *Gemeente Amsterdam* [website], retrieved 30th July 2018, 2018a, https://data.amsterdam.nl/#?mpb=topografie&mpz=11&mpv=52.3731081:4.8932945&pgn=content-detail&pgi=item1&pgt=beleid

Gemeente Amsterdam, "Vragen en antwoorden over Amsterdam City Data", *Gemeente Amsterdam* [website], retrieved 30th July 2018, 2018b, https://data.amsterdam.nl

Glaeser, Edward L., Scott Duke Kominers, Michael Luca, and Nikhil Naik, "Big Data and Big Cities: The Promises and Limitations of Improved Measures of Urban Life", *Economic Inquiry*, Vol. 56, Issue 1, 2018, pp. 114-137.

Glasmeier, Amy, and Susan Christopherson, "Thinking About Smart Cities", *Cambridge Journal of Regions, Economy and Society*, Vol. 8, 2015, pp. 3-12.

Grashoff, Wimfred, "Businesscase DataPunt", *Gemeente Amsterdam* [website], 21st June 2017, 2017a, https://assets.amsterdam.nl/publish/pages/841982/buca_DataPunt_v1_0_20170621_voor_publicatie.pdf

Grashoff, Wimfred, "Privacy en Beveiliging in DataPunt", *Gemeente Amsterdam*, 21st June 2017, 2017b, https://assets.amsterdam.nl/publish/pages/841982/20170621_privacy_en_informatiebeveiliging_DataPunt_pub.pdf

Hashem, Ibrahim Abaker Targio, Victor Chang, Nor Badrul Anuar, Kayode Adewole, Ibrar Yaqoob, Abdullah Gani, Ejaz Ahmed, and Haruna Chiroma, "The Role of Big Data in Smart City", *International Journal of Information Management*, Vol. 36, Issue 5, 2016, pp. 748-758.

Hein, Anton, "Amsterdam Europe's Eight Most Popular City Destination", *Dutch Amsterdam* [website], February 2nd, 2016, http://www.dutchamsterdam.nl/4456-amsterdam-europe-eight-most-popular-city-destination

Hollands, Robert G. "Critical Interventions into the Corporate Smart City", *Cambridge Journal of Regions, Economy and Society*, Vol. 8, Issue 1, 2015, pp. 61-77.

Kitchin, Rob, "Big Data and Human Geography: Opportunities, Challenges and Risks", *Dialogues in Human Geography*, Vol. 3, Issue 3, 2013, pp. 262-267.

Kitchin, Rob, "Data-Driven Networked Urbanism", *The Programmable City Working Paper 14*, 10th August 2015, 2015a.

Kitchin, Rob, "Getting Smarter about Smart Cities: Improving Data Privacy and Data Security", *Data Protection Unit*, Department of the Taoiseach, Dublin, Ireland, 2016a.

Kitchin, Rob, "Making Sense of Smart Cities: Addressing Present Shortcomings", *Cambridge Journal of Regions, Economy and Society*, Vol. 8, 2014, pp. 131-136.

Kitchin, Rob, “Reframing, Reimagining and Remaking Smart Cities”, *The Programmable City Working Paper 20*, 16th August 2016, 2016b.

Kitchen, Rob, “The Ethics of Smart Cities and Urban Science”, *Phil. Trans. R. Soc. A*, Vol. 374, 2016c, pp. 1-15.

Kitchin, Rob, “The Promise and Perils of Smart Cities”, *Society for Computers & Law*, Vol. 26, Issue 2, 2015b, pp. 1-5.

Kitchin, Rob, “The Real-Time City? Big Data and Smart Urbanism”, GeoJournal, Vol. 79, 2014, pp. 1-14.

Kitchin, Rob, "The Realtimeness of Smart Cities", *TECNOSCIENZA: Italian Journal of Science & Technology Studies*, Vol. 8, Issue 2, 2018, pp. 19-42.

Kitchin, Rob, Claudio Coletta, Leighton Evans, Liam Heaphy and Darach Mac Donncha, “Smart Cities, Urban Technocrats, Epistemic Communities and Advocacy Coalitions”, *The Programmable City Working Paper 26*, 8th March 2017, 2017.

Kitchin, Rob, Tracey P. Lauriault, and Gavin McArdle, “Smart Cities and the Politics of Urban Data”, *Smart Urbanism: Utopian Vision or False Dawn?* Routledge, London, 2015, pp. 16-33. ISBN 9781138844223.

Interviewee 1, “Interview”, November 6th, 2018.

Li, DeRen, JianJun Cao, and Yuan Yao, "Big Data in Smart Cities", *Science China Information Sciences*, Vol. 58, Issue 10, 2015, pp. 1-12.

Li, Yibin, Wenyun Dai, Zhong Ming, and Meikang Qiu, "Privacy Protection for Preventing Data Over-Collection in Smart City", *IEEE Transactions on Computers*, Vol. 65, Issue 5, 2016, pp. 1339-1350.

Martínez-Ballesté, Antoni, Pablo A. Pérez-Martínez, and Agusti Solanas, "The Pursuit of Citizens' Privacy: A Privacy-Aware Smart City is Possible", *IEEE Communications Magazine*, Vol. 51, Issue 6, 2013, pp. 136-141.

Moerman, Marcel, “Projectstartarchitectuur: DataPunt”, *Gemeente Amsterdam* [website], 21st March 2017, https://assets.amsterdam.nl/publish/pages/841982/20170626_DataPunt_psa_v1_0.pdf

Municipality of Amsterdam, “Launch of Amsterdam City Data: A Wealth of Data about the City”, *Municipality of Amsterdam* [website], 30th July 2018, https://www.amsterdam.nl/bestuur-organisatie/organisatie/overige/stad-vol-data/actueel-data/lancering-acd/

Naafs, Saskia, “'Living Laboratories': The Dutch Cities Amassing Data on Oblivious Residents”, The Guardian, March 1st, 2018, 2018,

https://www.theguardian.com/cities/2018/mar/01/smart-cities-data-privacy-eindhoven-utrecht

Nam, Taewoo, and Theresa A. Pardo, "Smart City as Urban Innovation: Focusing on Management, Policy, and Context", *Proceedings of the 5th International Conference on Theory and Practice of Electronic Governance, ACM*, 2011, pp. 185-194.

Nigon, Julien, Estèle Glize, David Dupas, Fabrice Crasnier, Jérémy Boes, "Use Cases of Pervasive Artificial Intelligence for Smart Cities Challenges", *IEEE Workshop on Smart and Sustainable City (WSSC 2016) associated to the International Conference IEEE UIC 2016*, Toulouse, France, July 2016.

Pan, Yunhe, Yun Tian, Xiaolong Liu, Dedao Gu, and Gang Hua, "Urban Big Data and the Development of City Intelligence", *Engineering*, Vol. 2, Issue 2, 2016, pp. 171-178.

Pieters, Janene, "Tourism in the Netherlands up 9 Percent; Biggest Increase in 12 Years", April 4th, 2018, *NL Times* [website], https://nltimes.nl/2018/04/04/tourism-netherlands-9-percent-biggest-increase-12-years

Rjab, Amal Ben, and Sehl Mellouli, "Smart Cities in the Era of Artificial Intelligence and Internet of Things: Literature Review from 1990 to 2017", *Proceedings of the 19th Annual International Conference on Digital Government Research: Governance in the Data Age*, May 30-June 1, 2018, Delft, Netherlands, ACM, 2018.

Scrum Values & Agile Methodologies for Project Management [WWW Document], 2018, available here: https://www.scrumalliance.org/learn-about-scrum/scrum-values

Sen, Mourjo, Anuvabh Dutt, Shalabh Agarwal, and Asoke Nath, "Issues of Privacy and Security in the Role of Software in Smart Cities", *International Conference on Communication Systems and Network Technologies (CSNT)*, 2013, pp. 518-523.

The Seven Foundational Principles [WWW Document], Ryerson University. 2018, URL https://www.ryerson.ca/pbdce/certification/seven-foundational-principles-of-privacy-by-design/

Sholla, Sahil, Roohie Naaz, and Mohammad Ahsan Chishti. "Ethics Aware Object-Oriented Smart City Architecture", *China Communications*, Vol. 14, Issue 5, 2017, pp. 160-173.

Srivastava, Shweta, Aditya Bisht, and Neetu Narayan, "Safety and security in smart cities using artificial intelligence—A review", *Data Science & Engineering-Confluence,*

2017 7th International Conference on Cloud Computing, IEEE, 2017, pp. 130-133.

van Zoonen, Liesbet, "Privacy Concerns in Smart Cities", *Government Information Quarterly*, Vol. 33, Issue 3, 2016, pp. 472-480.

Vázquez Salceda, Javier, Sergio Álvarez Napagao, José Arturo Tejeda Gómez, Luis Javier Oliva Felipe, Dario Garcia Gasulla, Ignasi Gómez Sebastià, and Víctor Codina Busquet, "Making Smart Cities Smarter Using Artificial Intelligence Techniques for Smarter Mobility", *SMARTGREENS 2014: proceedings of the 3rd International Conference on Smart Grids and Green IT Systems*, SciTePress, 2014, pp. IS7-IS11.

Yin, ChuanTao, Zhang Xiong, Hui Chen, JingYuan Wang, Daven Cooper, and Bertrand David, "A Literature Survey on Smart Cities", *Science China Information Sciences*, Vol. 58, Issue 10, 2015, pp. 1-18.

Zanella, Andrea, Nicola Bui, Angelo Castellani, Lorenzo Vangelista, and Michele Zorzi, "Internet of Things for Smart Cities", *IEEE Internet of Things Journal*, Vol. 1, Issue 1, 2014, pp. 22-32.

Zhang, Kuan, Jianbing Ni, Kan Yang, Xiaohui Liang, Ju Ren, and Xuemin Sherman Shen, "Security and Privacy in Smart City Applications: Challenges and Solutions", *IEEE Communications Magazine*, Vol. 55, Issue 1, 2017, pp. 122-129.