Expertise
and insight
for the future

Amanda Kauppinen

Smart Cities: Mass Data Collection and Processing

Metropolia University of Applied Sciences

Bachelor of Engineering

Information Technology

Bachelor's Thesis

10 March 2021

Author(s) Title Number of Pages Date	Amanda Kauppinen Smart Cities: Mass Data Collection and Processing 43 pages 3 April 2021
Degree	Bachelor of Engineering
Degree Programme	Information Technology
Specialisation option	Smart Systems
Instructor(s)	Sami Sainio, Senior Lecturer

The purpose of this thesis was to explore modern Smart City implementations and how they manage big data in terms of both storage and the process by which it is analysed. The thesis takes a closer look into what is required for a city to be considered Smart, and how modern cities are on the forefront of handling this transition. The main goal of this research was to provide detailed insight on the how large amounts of data are to be collected and stored in a Smart City setting. Additionally, the study includes a deeper evaluation of the analytical process in terms of both the requirements for success and various techniques by which the analysis is conducted. Furthermore, it incorporates the benefits and also the risks associated with city-wide implementation of Smart technologies. The thesis concludes with the future of Smart Cities and challenges they may face moving forward.

The research behind this thesis was conducted using varying journals, articles and e-books written within the past decade. These sources were comprised of topics such as Smart technologies, government-issued development plans, and case studies on modern Smart City implementations and piloting programs.

Based on this study, it can be concluded that there are both incredible benefits and substantial risks to the usage of big data analysis in Smart Cities. Smart City technologies have the ability to account for remarkable advancements in society. Ideally, they can accomplish a more connected, sustainable society, thus greatly improving the quality of life of inhabitants. However, the findings also indicate that it will be absolutely necessary moving forward that serious considerations are made in regard to cybersecurity and protection of both individual and government data.

Keywords	Smart Cities, Big Data, Privacy, Cybersecurity

Contents

List of Abbreviations

AI	Artificial Intelligence
AWS	Amazon Web Services
AMI	Advanced Metering Infrastructure
AMR	Automated Meter Reading
CCTV	Closed-circuit Television
GDPR	General Data Protection Regulation
GHG	Greenhouse Gas
GPS	Global Positioning System
ICT	Information and Communication Technology
ITS	Intelligent Transportation Systems
PII	Personally Identifiable Information
V2X	Vehicle-to-Everything

1 Introduction

As the world grows closer to global connectivity, the implementation of Smart Cities will have an astounding impact on how people and governments operate. As of 2020, the world has been confronted with rising populations in major urban areas, climate change, food shortages, and a mass of other major problems which require understanding and for which technology will be needed to offer solutions to. According to research done by Cisco, 3.9 billion people worldwide had access to the internet in 2018. This is projected to spread to 66% of the world’s population (5.3 billion) by 2023, with major increases in places like Asia Pacific (20%), Central and Eastern Europe (13%), and East Asia and Africa (11%). Additionally, the number of internet-connected devices and internet connections is growing at a rate faster than population growth. [1.] This is an indicator not only that the Internet is spreading at an incredibly quick rate on the global scale in regard to regions, but also that individuals are increasing their exposure and usage of IoT devices in their daily lives. The implications of this information point to a progressively more connected society, with potential to make the adoption of Smart City technologies more straightforward.

Smart Cities aim to simultaneously offer solutions to the growing issues facing major urban areas across the world, while also using data and statistics derived from its inhabitants to increase the quality and ease of life in these places. While this is most certainly beneficial to occupants and governments alike, mishandling of this process and the data derived from it can have potentially severe consequences. The purpose of this thesis is to take a closer look into how Smart Cities have developed from their first conceptions, and what they can offer the human population in terms of quality of life and optimisation of public resources. It will also discuss how the implementations of specific technologies used to achieve the goals of Smart Cities need to be handled with extreme caution in order to protect their citizens. Furthermore, the consequences of mishandling this process, citing current examples seen in today's society, will be used to demonstrate the dangers that threaten civilization while moving forward.

2 Smart City Conceptualization

The concept of a Smart City is a complicated idea. While there is no official definition for the term, there is a general consensus on the attributes required for a city to be considered Smart. These characteristics are a combination of both Smart objectives for a city, as well as the processes by which these objectives will be met. Firstly, there needs to be a standard for what makes a city Smart. Only after these specific guidelines have been set, can methods be defined on how to reach these goals.

2.1 Overview of Criteria

Because Smart City implementations and concepts have many different facets, there is no singular description that encompasses all necessary ideas. However, most entities can agree on a few basic goals that serve as a solid foundation:

- Quality of life improvement
- Sustainability
- Connectivity

These points, among others, are the establishment of any Smart City conceptualization. [2, p. 147; 3, pp. 5, 9-11.] Many of these terms are very broad, but each serve a vital purpose in the construction of a Smart City.

An improvement to a citizen's quality of life can of course be interpreted in many different ways. Emphasis on air quality improvements via pollution reduction, expanded healthcare systems, and emergency services are all areas in which modern cities have been focusing on for Smart implementations. [2, p. 147; 3, pp. 10-11, 38.] In London, using advanced analytics on city-wide sourced big data, the police have been able to expose trends in certain types of crimes, thus allowing the authorities to adjust their patrol routes for a safer London. They have also invested extensively in public health by promoting outdoor activity in London's Smartest region: Queen Elizabeth Park. [3, p. 11; 4.] Additionally, in New York, constant analysis on air-quality data in key parts of the city has helped identify major sources of pollution in the city. Because of the findings made by this evaluation, emissions in New York have plunged. [5, p. 703.]

Again, sustainability is another term that is fairly ambiguous in its meaning. For many cities, the idea of sustainability is centred on the growing, overwhelming demand of electricity and how cities must improve their power grids. There is a direct correlation between population size and energy usage, so as cities grow in size, there needs to be Smart infrastructure not only to handle the stress on the power grid, but to use clean energy where possible in order to reduce carbon emissions. [2, p. 293.] Water distribution and management directly impacts citizens through ensuring that a necessary resource is supplied to the masses, as well as safeguarding the agricultural system of a city. [2, p. 158.]

The challenge with sustainability is a city’s ability to intelligently manage the water, and in turn food supply. In established cities, the government must take into consideration the existing infrastructure and the difficulties that come with evolving it into a Smarter system. In New York, for instance, the government has put extensive effort into reworking the current water infrastructure to incorporate the idea of Smart metering. This has been done by implementing “Automated Meter Reading” systems, which includes the placement of measurement sensors throughout the city’s water distribution systems. The data provided by the placed sensors allows for statistics on usage, leaks, and water levels exposing areas of improvement and potential risks. Additionally, this helps address issues such as organised water distribution to the public as well as management of the different types of water sources so as to make use of them much more efficiently. [5, pp. 703-704.]

Connectivity is both a goal, and a method for reaching other objectives when working towards a Smarter city. The very foundation of a Smart City is rooted in its integration of technology, specifically information and communication technology (ICT) solutions. Various sources of valuable information, for example sensors or mobile applications, rely on widespread, working networks in a city. Both wired and wireless networks are used to retrieve and perform analysis on this data, therefore the more connectivity a city maintains, the more information it has at its disposal for problem solving. [6, p. 2.]

2.2 Components

While there are vast amounts of Smart City usages, the general process for how these are applied follow the same basic format. An example of this process is seen in Figure 1.

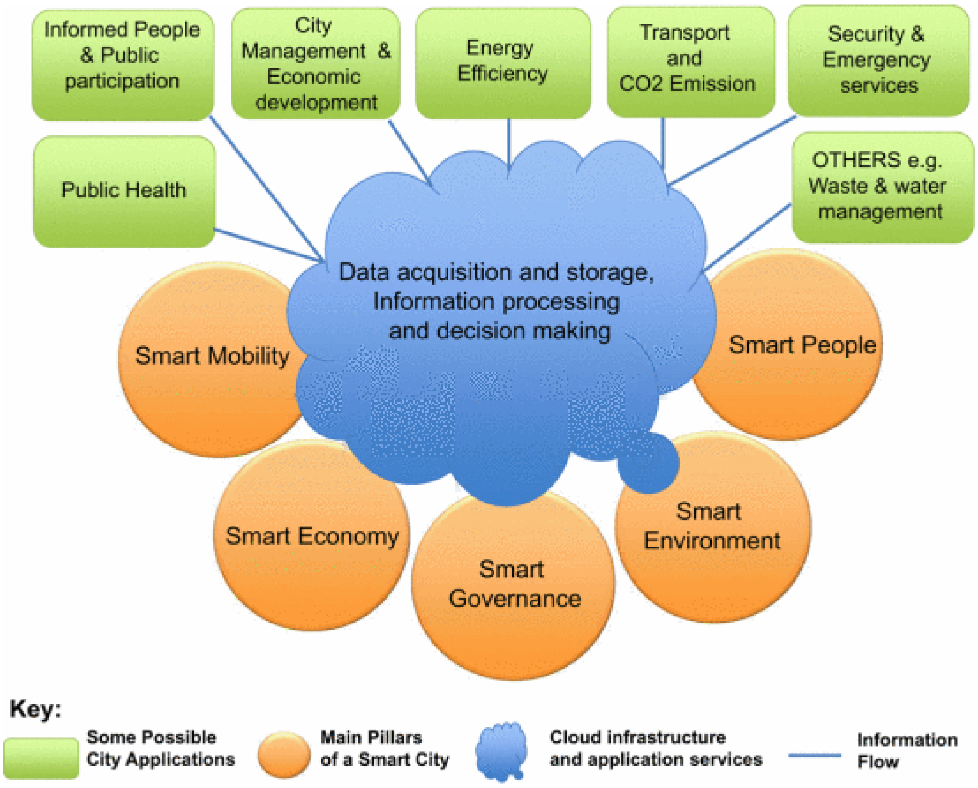

Figure 1. Cross-thematic Data Management and Analysis for a Variety of Smart City Applications in Cloud Environments. Copied from Khan et al. 2013 [7].

Here, it is noted that the handling of data is an integral part of this process. Firstly, this data needs to be generated from specific points in a city. These sites can range from widespread placement of water sensors to provide usage and leakage statistics, to traffic light sensors indicating how busy certain types of public transport are at key times of a day. The data is then acquired, stored, and examined with varying methods of advanced analysis utilized in regard to the sources. [7.] For instance, studying water or electricity usage in a city may benefit from different software or machine learning tactics than other processes such as predicting weather events to prepare a city for natural disasters. Thirdly, these evaluations are interpreted and applied to different sectors of a city. As seen in the green boxes in Figure 1, the conclusions drawn from the previous step can have a direct impact on concepts such as a city's healthcare division, transportation, and energy efficiency. [7.]

2.3 Transforming a City

Before discussing modern Smart Cities in the making, it is worth briefly examining the transformation of infrastructure that must occur in existing cities on their way to becoming

Smart. While there are no distinct definitions for what makes a city's infrastructure "smart", the key to advancement lies in the integration of ICT solutions. Individually, this will present differently in the varying sectors of a city. In terms of resources, this likely involves the deployment of different types of measurement sensors placed throughout the water distribution system in order to measure levels, usage, and leaks. Most modern cities already have a comprehensive water distribution network. However, intelligently transforming this sector allows for the analysis on the data reported by these sensors to address issues that otherwise would not be known without Smart metering. Regarding transportation, Smart infrastructure will be defined by the city's use of technology such as adding motion sensors, video surveillance devices, and air quality meters on existing roadways in order to determine patterns in traffic flow and to expose areas of risk in terms of accidents and GHG emissions. [2, pp. 155-158.] Instead of working to fully re-build a city's infrastructure which is costly in both time and finances, both of these examples instead focus on enhancing existing, proficient systems that have been successfully in use for a significant length of time [6, p. 2].

3 Modern Smart City Implementations

Within the last decade, Smart City global initiatives have been on an exponential, upward climb. In recent times, extraordinary factors such as population density, climate change, and the rapid advancement of technology have put global epicentres under extreme pressure. According to the United States Census Bureau, as of 2020 the top two populations by country (China and India) are believed to have over 1 billion inhabitants each. The top five most populous countries in the world have a combined total of roughly around 3.5 billion people. [8.] This means, the top five countries make up nearly half of the whole world's inhabitants. In Figure 2, the comparison can be seen between the current population in the year 2020 and the projected population of the year 2030.

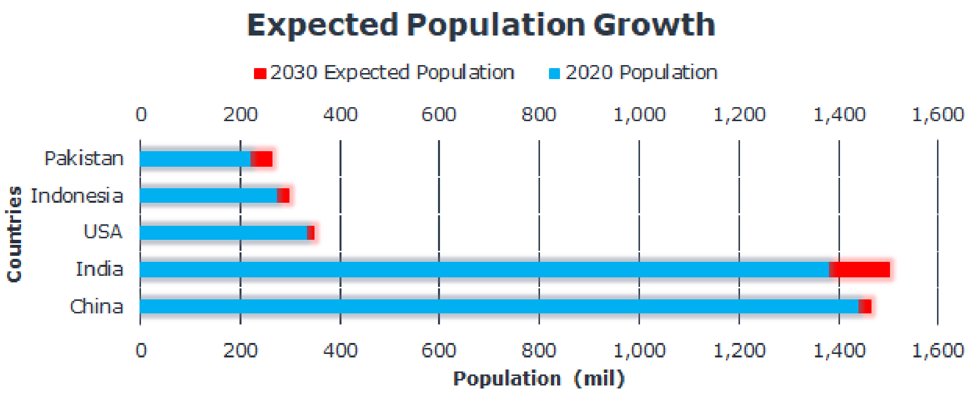

Figure 2. Expected Population Growths. Based on data acquired from U.S. Census Bureau Statistics [8].

As can be seen in the graph, just between these five countries, the world is expected to have over 235 million new inhabitants by 2030. India alone is projected to grow in over 100 million people within the decade. The focus by Smart Cities on sustainability and infrastructure will be key to not only providing the basic necessities to the increasing number of citizens, but also to make these cities much more optimised. [3, pp. 5-6.] In this section, the focus will be on Smart City implementations that have been developing in recent years.

3.1 London

London has spent several years working towards becoming a Smarter city. In 2019, the Mayor issued an extensive, revised roadmap to achieve this goal. Turning a city "smart" is a massive undertaking, affecting nearly all parts of a city and its people. London's plan reflects this notion, covering a wide variety of topics.

Smart City Focus Areas - London	
Citizen Involvement and Diversity	• Enhance leadership in design and common standards • Develop new approaches to digital inclusion • Introduce Civic Innovation Challenge • Renew Civic Platforms • Promote diversity in technology sectors
Data Protection and Management	• Launch London Office for Data Analytics program • Develop city-wide cybersecurity strategy • Strengthen data rights, accountability and trust • Increase transparency and innovation
Connectivity	• Launch the Connected London program • Consider planning powers • Enhance public WiFi availability • Support new generation of Smart infrastructure • Promote common standards
Digital Leadership	• Enhance digital and data leadership • Develop workforce digital capability • Support computing skills from early years • Recognise role of cultural institutions
Collaboration	• Establish the London Office of Technology and Innovation • Promote MedTech innovation in the NHS • Explore new technology sector partnerships • Support better Greater London Area digital delivery • Collaborate with other UK and global cities

Figure 3. Smart City Focus Areas – London. Based on London Government's Roadmap to a Smarter City [3, pp. 15-37].

In Figure 3, the most important focuses are outlined as citizen involvement and diversity, data protection and management, connectivity, digital leadership and collaboration. One of their main goals is centred around involving citizens as much as possible in the planning and use of key citizen services. For instance, they had found previous success in educating the population (especially older generations) in how to use key technologies and digital services offered by the government. This campaign's success indicates that the rollout of future public services such as Smart metering and the transition to onlineexclusive services may be more successful with the accompaniment of citizen education and their involvement during production. [3, p. 16.] The city has put great emphasis on diversifying technology education by offering specialty training programs and collaboration between various education facilities. Additionally, both encouraging the public to contribute their innovative ideas as well as increasing the diversity amongst tech-sector workers are goals that London's government believes will enhance the city's ability to foster more inventive ideas leading to a Smarter London overall. [3, pp. 17-20.]

In a Smart City setting, there are various risks to consider. This is why some of London's main priorities are data security and having informed, educated leadership in place as seen in Figure 3. The government recognises that moving forward, cyber security is going to be an absolutely necessary part of their Smart City vision and thus there have been significant investments in collaboration with the police and other local councils in order to develop a solid, renowned cybersecurity sector. Additionally, London has put great emphasis on making sure that the city's leaders are at least basically educated in tech-related matters so as to make knowledgeable determinations that impact the public positively. [3, pp. 23, 38.]

While building on existing infrastructure is detrimental for Smart City expansion, there are some cases in which cities will have the ability to plan areas intelligently from the conception stage. In London, Queen Elizabeth Olympic Park was built with Smart ideas in mind, acting as a technological testbed for both surrounding as well as international regions. The government has been able to design this park specifically to address key issues such as air quality and energy conservation. [3, p. 11.] Unlike using a broader service such as Google Maps, LivingMap (a mobile and desktop application developed specifically for the park) offers visitors a comprehensive guide to the park. In addition to the basic services such as showing the user, for example, nearby attractions and eateries, the app goes further in providing the ability for users to see well-lit paths for night-time activities and more unique features of the park such as biodiversity points of interest. Additionally, mapping is extended to indoor spaces as well, going beyond what more generic navigational apps tend to offer. The goal is to promote healthier lifestyles by promoting physical activity and interest in what the park has to offer. [4.]

3.2 Singapore

Singapore has produced several years of Smart City development. Starting at the end of 2014, the government of Singapore officially launched the Smart Nation Vision. Singapore's goal of becoming the first Smart Nation in the world is centred around the integration of widespread networks and data processing in order to create a stronger society. [9, p. 3.]

As outlined in the case study issued by the Republic of Singapore government, their plan is very ambitious and comprises a wide range of topics affecting citizens. These topics are covered in great detail, and are as follows:

1. Transportation and Urban Mobility
2. Safety and Citizen Security
3. Emergency and Response
4. Environment
5. Energy
6. Citizen Interaction and Communication Mechanisms

According to the case study, for the past decade Singapore has already been working on Smart transportation through the use of Intelligent Transportation Systems (ITS). Of the six main sections to its Smart Nation plan, transportation has been the most developed and prioritized sector. This is logical because, as of 2016, only about 12% of the population drives personal vehicles due to difficult measurements enacted on the obtaining and registering of this type of transportation. This has put much emphasis on the development of the nation's public transport sector. [9, pp. 8-25.]

Two examples of the government's action taken in the transportation sectors are the Transport.SG mobile application and the OneMotoring portal for both public transportation users as well as drivers. Transport.SG offers an advantage over traditional navigational applications in that it provides a tailored experience to the user. Citizens can set customizable journeys based on their most frequented routes, so that the app always offers up to date information on their daily commutes. Additionally, the app has the unique ability to inform drivers of nearby parking capacities to make their trips more efficient. [10.] The data provided by road surveillance devices has led to the creation of OneMotoring, a public-accessible portal with the purpose of not only delivering real-time road updates such as traffic, accidents, and detours, but also offering citizens general knowledge and advice on how to manage their own personal automobiles. [11.]

In 2015, Singapore put emphasis on investigating how to address the issues that come with a growing elderly population. Initiatives have been made in recent years with goals to reduce the use of care institutions and put more effort into assisting seniors with more independent life. [12.]

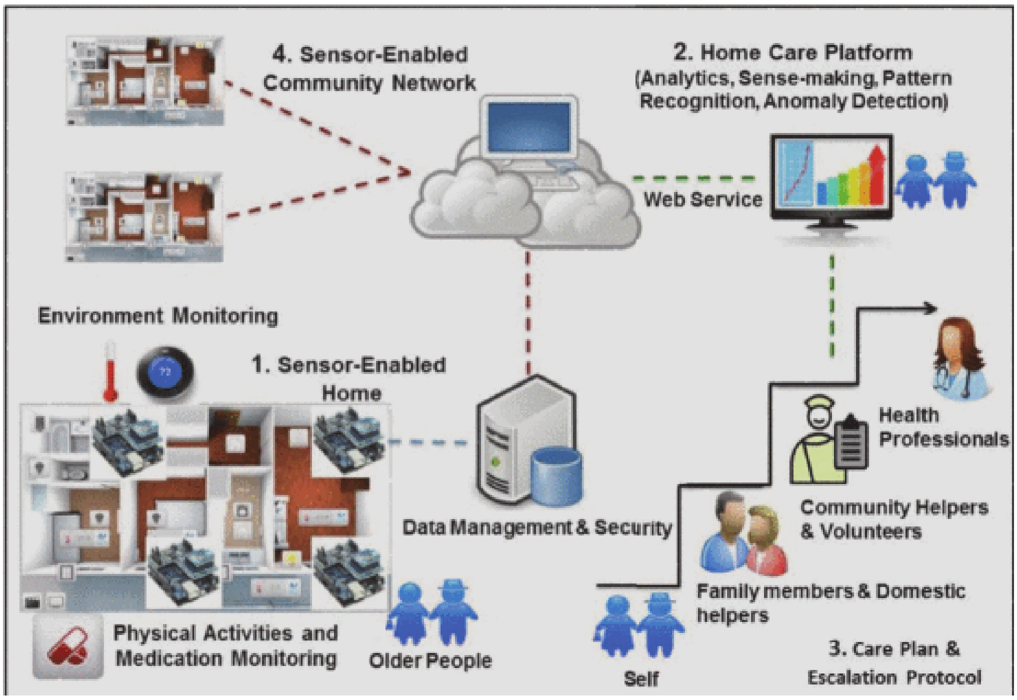

Figure 4. Design for Sensor-Enabled Care. Copied from Liming et al. 2015 [12].

In figure 4, the first step of this process can be seen as enabling the senior's home with a wide variety of sensors, monitoring everything from air quality to sound levels to motion. Step 2 is focused on data evaluation that could provide a mobility map of typical movements of inhabitants, so that anything out of the ordinary could indicate possible accidents or shed light on new, subconscious behaviours potentially indicative of undocumented ailments. Step 3 uses the conclusions from the analysis to make incredibly personalised plans for the seniors and intervening with further care where necessary. The final step is to expand this environment creating interconnected communities to support the aging populations. [12.]

3.3 New York

In 2015, the governing body of New York City (NYC) issued a plan to transition to a Smarter city with focus on sustainability and updates to infrastructure. Like both London and Singapore, it is long-established and when envisioning a Smart NYC, updating the existing infrastructure is absolutely necessary. NYC has spent the last decade investing in Smart metering to better manage the water distribution system. This elaborate system utilises over 800 thousand advanced water meters, each attached to a radio transmitter that relays current readings on water levels to the nearest operation centre. Readings are sent via a radio frequency specifically assigned for this communication at regular intervals, allowing a more complete, accurate picture on usage. While this is a complex system that requires much attention and planning, it has saved over three million dollars per year by removing the previous system in place of having to manually check the meters. Additionally, this is easier on citizens as the data is linked to their website where usage can be monitored, and bills can be paid. [13, pp. 3, 8.]

Another area NYC has been pushing for improvements is that of waste management. NYC manages an incredible amount of waste that varies greatly day-to-day based on things such as weather, seasons, or events. Because of this NYC has launched Smart trash bins not just to get a bigger idea on trends in waste disposal, but to track individual bins' contents in order to respond intelligently to trash pickup. The sensors used inside the bins keep track of the amount of trash inside and send alerts to management when full. By having many of these bins distributed throughout the city, the government gets a good idea of where and how frequently trash is generated and can then respond with more effective route planning of retrieval. [13; 14, p. 703.]

Furthermore, NYC has responded to the overwhelming need for internet connections by introducing LinkNYC. These "links" are spread throughout NYC and provide free, accessible WiFi to citizens. These hubs go beyond just delivering the internet, offering public device-charging stations, tablets for access to city information, and the ability to call anywhere in the United States for free. These hubs are built to withstand all manners of weather and collect information on the surrounding environment via internal sensors for various analytical purposes. [14, p. 704.]

4 Mass Data Collection and Storage

In order for a city to be Smart, there needs to be incredible emphasis placed on data collection and analysis. However, organising data collection and storage for a large Smart City is an arduous process. Firstly, the data needs to be considered in both type and number of sources. Secondly, the data needs to be stored with purpose, meaning that decisions should be made based on the location of the storage (for example in the cloud or at the end device) as well as the volume of the data, such as whether the data is sent to storage in small or large batches. [15.]

4.1 Data Acquisition

When considering the size of a Smart City in terms of both population and area, the process of collecting data is complex. In order to meet the criteria, cities need to rely on information from an immense number of sources. For example, sources can be:

- Smart Devices
- Environmental Sensors
- Urban Sensors
- Closed-circuit Television (CCTV)

These are all examples of potential points of data generation. [7, 16.] For instance, accessing location services of residents' cell phones can allow for statistics on how people tend to move around a city, what times of day are more popular for this movement, and the frequency at which they travel in order to provide real-time information on available parking in particularly congested areas [7]. Environmental sensors offer the ability to monitor a city's usage of valuable resources like water and electricity. When paired with advanced analytics, the data can even provide trends in weather related events. [2, p. 158; 17.] In South Korea, urban sensors such as CCTV and license plate detection cameras provided data on potential crime events in a Smart City piloting program with a goal of crime reduction [16]. The collection of the data supplied by the aforementioned sources relies on both wired and wireless networks for retrieval. [6, p. 2.]

4.2 Data Storage

Storage easily becomes an issue when handling such massive amounts of data. According to an article written in mid-2019, the predicted volume of generation on a global scale was projected to approach 2500 petabytes of data per day by the end of the year. This evaluation was based only on surveillance videos and not inclusive of the vast range of other Smart technologies to exist in such a city. [15.] Although this is a global prediction and not reflective of an individual city's needs, modern Smart Cities will still require a comprehensive and robust data storage process for even a fraction of that number.

At this point in time, there are a number of different processes, consisting of both physical and non-physical parts, by which a city can store data that need to be heavily contemplated. Smart Cities will likely have a complex approach to these requirements, making use of both the storage existing on the end device (referred to as "edge", for edge computing) as well as cloud storage, where the data can be collectively stored in much higher quantities on servers. Physically, the cloud for a Smart City with such large amounts of data would need to rely on a considerable number of servers, as well as a central location to host them for minimum latency. [15.]

4.2.1 Edge and Cloud Based Storage

In a modern Smart City, it is likely that storage will consist of both the edge and the cloud. Both of these resources provide a different set of advantages as well as limitations.

Parameters	Cloud Computing	Edge Computing
Service Location	Within the Internet	In edge network
Distance (Number of Hops)	Multiple hops	Single hop
Latency	High	Low
Jitter	High	Very low
Location awareness	No	Yes
Geo-distribution	Centralized	Distributed
Mobility Support	Limited	Supported
Data enroute attacks	High probability	Very low probability
Target user	General Internet users	Mobile users
Service Scope	Global	Limited
Hardware	Scalable capabilities	Limited capabilities

Table 1. Differences Between Cloud Computing and Edge Computing. Copied from Ahmed et al. 2017 [18, p. 2].

These differences between the two can be seen in Table 1, with the key difference being the location of the data. When edge storage is used, data is collected, stored, and analysed on the field device. This is a distributed system, as there would be many of these devices spread throughout a city. They would be used to monitor factors such as water distribution, and electricity usage. For cloud storage, the data is collected from the edge device, then sent to the cloud where it would be stored and eventually analysed. Storing some data at the edge allows for much faster and potentially more secure computing, but the hardware is very limiting in that the edge devices are not scalable and cannot offer nearly as much memory space as the cloud. [18, p. 2.]

Though analytics will be covered in further detail below, it is important to note that depending on how quickly the data needs to be analysed, varying storage processes may be followed for differing types of data. Storing data at the edge allows time-critical edge computing. For example, by storing and analysing the data at the edge, an emergency service may only receive the most critical information from the end device (post-analysis), thus allowing the workers to respond as quickly as possible. Alternatively, storing much larger amounts of data in the cloud can allow for city-wide studies of traffic trends that are not as reliant on the speed at which the data can be processed. [15.]

4.2.2 Physical Necessities

The physical aspect of mass data storage is just as important as cloud and edge technologies. When considering the physical needs of storage, there are many different requirements necessary for a data centre to operate optimally. Location is important, as the distance between the end devices and the data centre can affect the latency of data transmission from the collection point to the servers. [15.] In large, established cities, finding the space for a sizable data centre that is located centrally can be a great difficulty.

In addition to facilities, selecting the best hardware for the storage is an added challenge. With new technological advancements every day, there is a difficult task ahead of not only selecting a competent tool for the job, but also making sure that the tool will be able to accommodate future technologies. Furthermore, the chosen hardware needs to be able to tolerate physical conditions such as high temperatures and the stress of constant work for as long as possible. [15.]

5 Mass Data Analysis

As discussed above, a Smart city with multiple large data sources can quickly become overwhelming, with emphasis not only on storage but sophisticated analysis of this valuable information as well. In order to efficiently evaluate the data, some requirements must be met to ensure that the process not only scales for an ever-increasing number of data sources but is also built to analyse the data as effectively as possible. [7.] This section will also cover the varying techniques by which this analysis can occur, and how the data is communicated.

5.1 Requirements

There are several different needs that should be met when choosing an analysis architecture for a Smart city's level of data acquisition. While this is still a relatively new field, there are quite a number of suggestions for how the analytical process should go. This section will focus on just three aspects that are considered highly important to a well-structured method: scalability, security and privacy, and interoperability. [18, p. 7.]

5.1.1 Scalability

When a Smart City is ready for data analysis, it is to be assumed that a healthy number of data sources is already present. The amount of data sources will increase vastly as the city grows or more Smart solutions are implemented. This is where scalability needs to be considered not only in terms of how to scale the number of field devices collecting the data, but in the software/hardware combination that drive the analytical process. When considering edge devices, a city needs to heavily plan how an increase in this area will have an effect on the process in terms of network congestion and overall load on both storage and analytical software. [18, p. 7.]

Looking at the software and hardware that make up the data analysis sector of a Smart City, there are two options: vertical and horizontal scaling. Simply put, vertical scaling is focused on enhancing the processing platform so that it is more robust and can handle larger volumes of data. This means adding/improving components such as memory, the central processing unit, and storage. Horizontal scalability is focused on running the analytical process in several different places simultaneously, making use of numerous individual machines. [19.]

	Vertical Scaling	Horizontal Scaling
Scaling	Limited, compelled by technology manufactures.	Unlimited, unrestricted to specific technology
Management complexity	Easy, Most software applications can benefit additional resources easily.	Complex, management software has to handle the complexities of parallel processing on distributed data.
Financial investment	Considerable financial investment.	Affordable, added machinery is high-end servers.

Table 2. Comparison Between Vertical and Horizontal Scaling. Copied from Shahat et al. 2017 [19].

In Table 2, the differences as well as some advantages and disadvantages of both processes can be seen. Vertical scaling is simpler, as it is centred around basically one analytical platform which makes management of the process quite straightforward. However, it is severely limited by the availability of upgrades and is not as safe in terms of redundancy. [19.]

Horizontal scaling is much more complex, demanding more attention to the planning process. Adding more machines requires an intricate level of organisation in order to ensure that the machines are seamlessly running in parallel to each other. Nonetheless, horizontal scaling has much more scalability potential, as one can add more machines as the demand on analysis increases. This is a major benefit to Smart Cities because the amount of data is always expanding and will require more computational power. [19.]

5.1.2 Security and Privacy

The location in which data is stored and analysed must be secure. An exceptionally large amount of data gathered in a Smart City operation will likely involve very confidential information. The access to this data by non-approved individuals could pose severe risks to both the government as well as individual inhabitants. [20, p. 11.]

When data is analysed from a cloud-based storage system, there are risks in transportation and the holding of such a large and varied amount of information in one central location [21]. When transporting the data, it is possible that it will pass through a number of networks where network security may not be verifiably protected. Governments may choose to hire privatized companies for data storage services. This means that the government is at the mercy of these companies' security practices, which may not be adequate. [20, p. 9.]

For these reasons, analysing data at the edge device may be more secure but come with additional challenges. For instance, protecting device-specific information, such as its address, that is necessary during the transport of data is very important. If malevolent entities are able to discover sensitive information about the device, they may gain access to confidential cached data. Additionally, it is important to be able to segregate user-specific device information in the case that a device's data is contributed to by several different users. [21.]

5.1.3 Interoperability

Hardware and data interoperability is essential to any Smart City project due to the vast differences in data types to be collected and in varying device setups. In edge computing, it is possible that devices with distinct architectures may need to interact with one another and additionally with the cloud as well. This may present communication issues when constructing a system that utilizes a variety of devices altogether. [18, pp. 7-8.]

It is essential for a linking layer to exist during the analytical process of a data bank comprised of widely sourced data. This part of the process would put priority on identifying connections between the data and mapping out a cohesive, meaningful data structure ready for more complex analysis. [7.] Ensuring that the algorithms used for analysis can efficiently handle such a diverse batch of data is essential to any Smart City looking to benefit from large-scale evaluation [20, p. 17].

5.2 Techniques

While there are many different techniques by which huge amounts of Smart City data can be analysed, this section will focus on just a few sets. Edge and cloud computing both have specific use cases in a Smart City setting. The same sentiment applies to batch and stream communication between edge devices and cloud services, in that they each have individual advantages to the analytical process. [15.] Additionally, artificial intelligence is utilized to perform the analysis in various Smart City applications [2, p. 162].

5.2.1 Edge and Cloud Computing

Data storage at the edge and on the cloud can be determined by which type of analysis is meant to be performed. There are significant differences between the reasoning to do computing at either end.

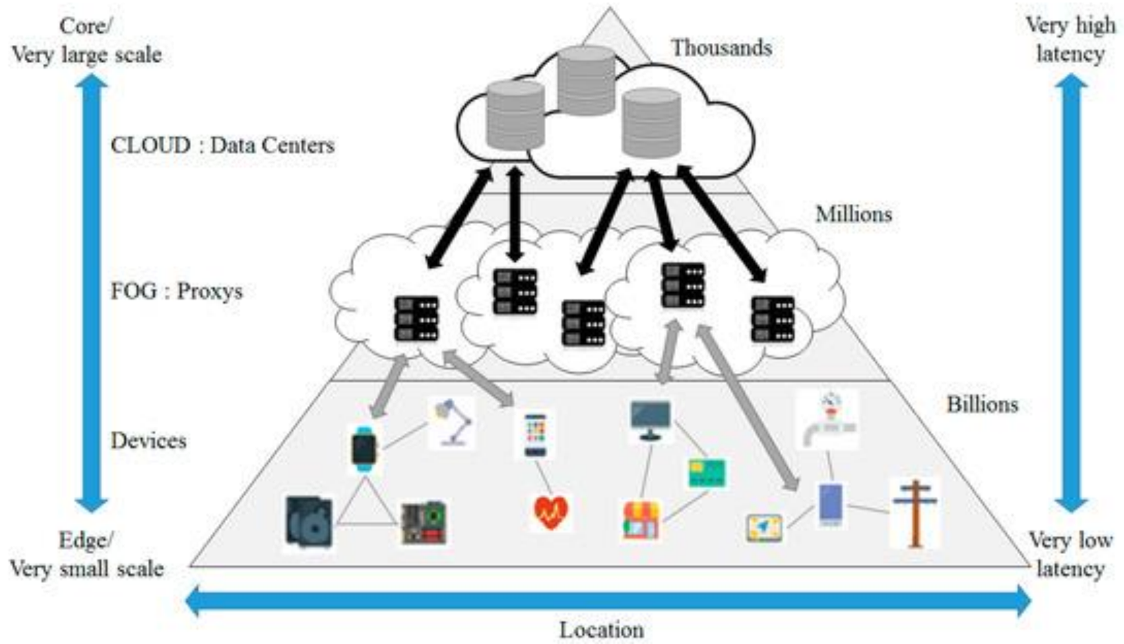

Figure 5. The Structure of Fog Computing. Copied from Cha et al. 2018 [22].

In Figure 5, the scale and overall structure of a Smart City's analysis sector can be seen. In a Smart City setting, edge devices make up the bulk of the system, with the fog layer acting as a communication platform between these devices and the data centres that will perform more intensive analysis. Each fog node would be able to support many edge device communications, which would then relay the data to even fewer data centres capable of handling much larger amounts of information. [22.] Edge devices generate massive amounts of data and when communicating such a large volume of information from the edge device, through the fog, and eventually on to the data centre, latency easily becomes a considerable issue. Edge computing can not only help to reduce the burden of increased traffic on a network, but it also offers the unique ability to do real-time data processing, thus reducing time wasted on latency issues when completing the full communication process. [15; 18, p. 9; 21].

The capability of running analysis at the edge allows for faster computations, which is important in cases where time is absolutely critical. For instance, consider a scenario in which a Smart City video surveillance operation captures a traffic accident. If the edge device is performing some type of targeted analysis constantly on the video feed, dispatchers would be alerted, and emergency services sent to the location. While en route, the dispatchers could have the ability to tap into the more elaborate analysis offered by the edge device. Using this evaluation to assess the environment surrounding the wreck, they would then be able to communicate this to the responders providing valuable information to prepare them about the site of which they are on their way to assist. [15; 18, p. 9.]

While edge computing is particularly good at real-time analysis, cloud computing can provide the long-term studying of trends. For instance, in the previous example of the car wreck, the larger, complete dataset could then be sent to the cloud for more in-depth assessment. Analysis of the traffic over a much longer amount of time could potentially provide insight on general large-scale traffic trends, possibly even those that led to the incident. These assessments could then lead to meaningful changes in the city to affect the flow of transportation. [15.] Furthermore, because of its usefulness in providing big-picture views, cloud computing is a possible source of collaboration between different Smart City sectors such as health, environment and safety. [20, p. 5.]

It is highly likely that the answer for big data analysis does not lie in either edge or cloud computing, but rather a combination of the two. Edge computing is exceptional at taking care of the time-sensitive analytics, while cloud computing will be used to identify much larger tendencies. [21.]

5.2.2 Batch and Stream Communication

There are several different ways in which data can be transmitted to the cloud. Two of these methods are batch and stream. Some analysis benefits from, or is even critically dependent on, fast response. Streaming communication is important here, as it offers the ability to quickly communicate smaller amounts of (likely already moderately analysed) vital data. [15; 19.] However, in the interest of investigating much larger trends from mass amounts of data, it is not necessary to transmit data as quickly and can thus be sent in sizable batches at more periodically timed intervals [15].

Each of these methods have their benefits and Smart Cities will likely rely on both types in collaboration with each other. The methods also each uniquely offer challenges to a city’s network infrastructure that cities must consider during implementation [15.]

5.2.3 Artificial Intelligence Analytics

Artificial Intelligence (AI), as defined by the Cambridge Dictionary, is the “the study of how to produce machines that have some of the qualities that the human mind has, such as the ability to understand language, recognize pictures, solve problems, and learn” [23]. AI methods are a major sector of a Smart City’s big data analysis due to its ability to predict human behaviour and the benefits that skill comes with [2, p. 162]. As the data in a Smart City grows, there is only so much that human monitoring and analysis can accomplish. Employing advanced AI analytics, the burden of requiring such a heavy human presence would be removed, thus allowing the machines to do the bulk of the diagnostic work. [24.]

AI offers advantages for active, real-time analytics, including the ability to monitor and evaluate video feeds in order to identify license plates of vehicles or to perform facial recognition in public spaces. It can also be incredibly useful in traffic situations in order to minimize traffic accidents and help to streamline public transportation routes. [24.]

Additionally, AI has advantages in predictive analysis as well. AI is exceptionally good at identifying patterns across large amounts of data, which is not feasible by purely human monitoring. For instance, identifying patterns in public transport ticket sales can reveal data not just on how effective a public transport system is (and the subsequent data-based restructuring of routes, stops, etc.), but also on its role in traffic congestion and how it affects the lives of citizens who use it. [24.]

6 Benefits of Mass Data Analysis

Alone, the idea of mass data collection can seem daunting. This is understandable as there are quite often stories in the news regarding colossal data breaches. These events can leak many different types of user records including financial, residential, and classified personal information. Regardless of these threats, mass citizen data collection and processing has the potential for incredible benefits regarding Smart Cities. These positive impacts can range from very broad ideas such as an overall improvement to transportation systems within a city, to very specific cases for example disaster readiness and response. [2, pp. 155-156.]

6.1 City Transportation

While inner-city transport is an extremely broad topic, it affects nearly every part of city life. Personal vehicle traffic congestion may be more common for some megacities, where others may have an efficient public transportation system already in place. In either case, improvements to the system are vital for rapidly growing epicentres. Some cities already have a start on combatting these issues. Emphasis on ridesharing, city-rented bicycles or scooters, and adding additional types of public transport are all ideas that work towards a smarter city. Mass data collection and analyzation of a city's entire transportation sector could potentially have substantial benefits to the environment as well as the safety and coordination of a system that moves millions of people every day.

6.1.1 Vehicle-to-Everything Communication

Vehicle-to-Everything Communication (V2X) is an important concept in the advancement of Smart Transportation.

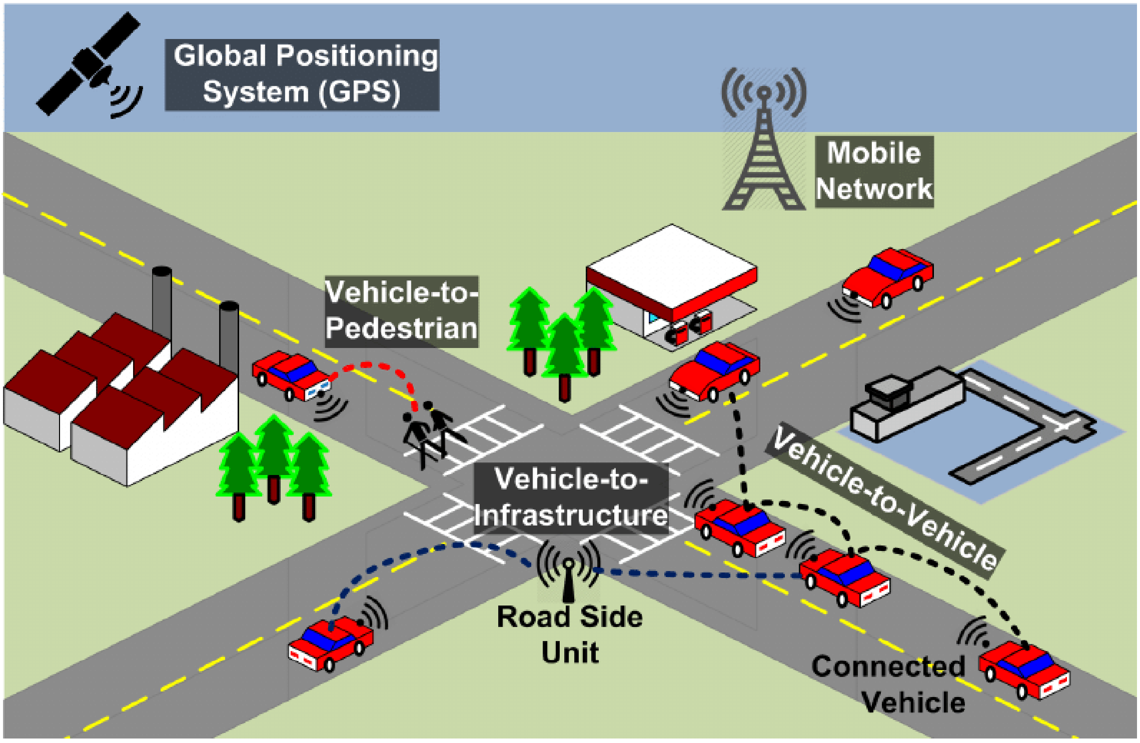

Figure 6. ITS V2X Communications. Copied from Hamida et al. 2015 [25, p. 384].

As seen in Figure 5, V2X covers several different types of linking communications between vehicles and infrastructure, pedestrians, and other vehicles in the same area. Though new advancements may update the technology, for now V2X originated from IEEE 802.11p standard, which belongs to the "IEEE Wireless Access for Vehicular Environments program (WAVE)" [26]. V2X works well for vehicles regardless of speed and outstanding weather conditions that cause issues for other types of transmission. In addition to this, it also has the ability to surpass technologies that are reliant on an established line of sight such as cameras or radar systems [26]. Because it is low latency, it is perfect for vehicle communications where real-time, fast decisions must be made at all times. When a vehicle is present on this type of network, it would have the ability to connect with objects like parking meters, traffic lights, crosswalks, and even citizens in the vicinity with access to the same network likely via smartphone applications. [25, p. 384.]

6.1.2 Personal Vehicle Traffic Congestion

In places where personal vehicles are still more widely used, such as in the United States, traffic congestion has had an effect on global warming via Greenhouse Gas

(GHG) emissions. According to reports done by the United States Environmental Protection Agency, transportation was responsible for 28% of GHG emissions in 2018. Though this sector includes all modes of transportation, the biggest contributors to GHG emissions were those of passenger cars and light-duty trucks. [27.] Not only does traffic congestion heighten GHG emission output, but it can also have an effect on the frequency of vehicular accidents.

V2X communication could be a potential answer to these issues. This type of connectivity would allow not only for advancements largely in traffic flow and congestion reduction, but also could lay the foundation for city-wide autonomous vehicular travel as well. The data provided by potentially millions of entities interacting with each other and many connection points throughout a city would provide a big picture view of what traffic looks like at all hours of every day. This data could shed light on trends that could be addressed via changes to infrastructure and automated traffic flow control in order to provide a more streamlined transportation sector. Additionally, real-time analysis of this data could be conveyed to drivers, via mobile applications for example, in order to give the most recent information on existing and predicted street conditions and advise them on the best course of action moving forward. [25, p. 389.]

The evaluation of this type of data could allow city planners and engineers to enhance objects such as traffic lights, toll roads, and pedestrian crossings through V2X communication. Optmised streets would lead to a better flow of personal vehicle traffic, in return reducing GHG emissions, improving fuel consumption and minimising traffic accidents. [2, p. 156.] These ideas focus on the benefits mass data collection and analysis can provide for personal vehicle travel. However, this is just one part of a city's transportation system.

6.1.3 Public Transportation Efficiency

Public transport plays a crucial role in many large cities worldwide. For many cities with one of these systems already in place, several different modes of public transportation exist. Amongst these are vehicles that have the ability to move multiple people at once such as buses, light rails, trains, and subways, and other options that are more individualised. Almost all of these methods must exist alongside each other as well as with personal vehicles and pedestrians alike. This level of coordination is incredibly complex. [2, pp. 155-156.]

V2X can once again be a helpful resource used to provide valuable, real-time information to passengers in order to help them plan their routes more efficiently and safely. Some of these benefits include, but are not limited to, improvements to a city's ride-sharing opportunities, more accurate transportation capacity projections, and traffic management capabilities by use of traffic flow statistics. One such application, Bus Capacity Prediction, used sensors and cameras on board buses in Singapore to keep track of the number of passengers present. People not yet on the bus could then use the app to see if the bus may be too full to ride and in turn gave them the opportunity to re-route their journey based on these statistics. [28.] Using technologies such as this, traffic could also be streamlined in a way that would grant certain types of vehicles (such as public transportation mediums) prioritization which would in turn result in less latency in their routes and more reliability on a city-wide scale.

Additionally, data analysis on city-wide public transportation might provide insights to subjects such as how often citizens use public transport and how efficient the current routes are. Studying these trends could lead to significant changes to the timetables and routes that would make the system as a whole much more effective. A more efficient public transportation sector could reduce waste in resources while simultaneously increasing citizen usage figures. [20, p. 5.]

6.2 Utility Usage and Preservation

Utility availability is a solid foundation of wellbeing and health. Urban population density is rapidly growing, and the aim of most Smart cities is to become sustainable in terms of resource management. It is important for governments to consider the impact concepts such as growing population density and climate change have on the city's ability to sustain itself. This will likely come in both enhancing existing resource infrastructure or building new communities with Smart resource planning implemented from the very beginning [6, p. 1; 29.] In either case, the movement towards Advanced Metering Infrastructure (AMI) by use of Smart utility meters is an integral part of many Smart City settings [30].

6.2.1 Smart Metering

The rollout of Smart meters in a city requires consideration both on the hardware and the ways in which the sensors will communicate data for further analysis. Both of these matters have several requirements.

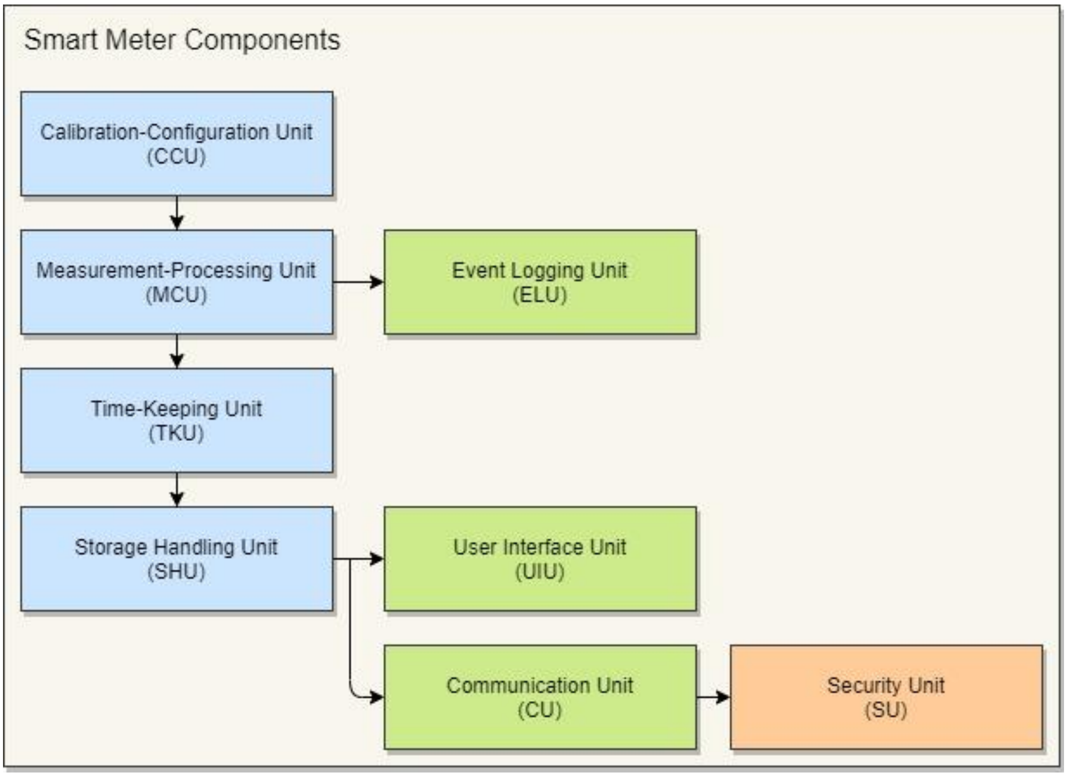

Figure 7. Fundamental Building Blocks with Hardware/Firmware Layers. Self, based on Smart meter data [30].

In Figure 7, the various components that make up a Smart meter can be seen. The CCU is where the user can set specific constraints and run calibration to make sure the meter is properly measuring values such as voltage, and watts. properly. The MCU is the section of the Smart meter that receives and handles signals through use of analog-to-digital converter(s) and digital signal processing. In most cases, software embedded in the device will then be applied to run desired tasks on the data. During all of these processes, the TKU is very important in order to keep track of timing and time-stamp critical events. [30.]

The SHU is vital to field devices. Inside the SHU, Smart meters often make use of either Electronically Erasable Programmable Read-only Memory (EEPROM) or Non-Volatile Random-Access Memory (NVRAM) [30]. These types of memory are important in distributed, embedded devices because they have the ability to retain the contents of their storage for long lengths of time without power [31]. In the case of Smart metering, configurations and recent readings are able to be stored in the SHU so that there is redundancy in case of power failure or disconnections [30].

The UIU is typically comprised of an LCD screen with some type of navigational buttons attached. It is used to provide the user with easy access to important panels such as calibration, various configurations, and diagnostic information. The CU and SU both take care of the device's communications, making sure that they are transmitted successfully and securely. [30.]

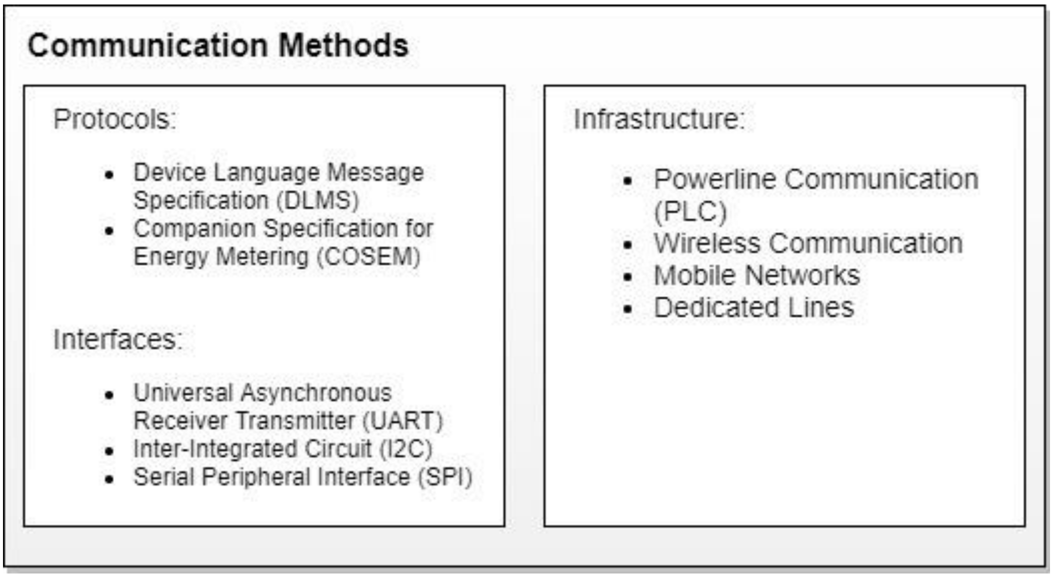

Figure 8. Smart Meter Communication Methods. Based on Smart meter data [30].

In Figure 8, a deeper look at the communications sector of a Smart meter can be seen. There are many ways in which data from the Smart meter devices can be communicated. PLC is economically viable, as it relies on the existing powerline infrastructure in a city to deliver fast transmission rates. However, utilising existing infrastructure that handles other types of transmissions can quickly become messy and put restrictions on performance. On the other hand, wireless communication offers a simple, low-cost solution. Dedicating specific lines to Smart meter communication is highly efficient in both transmission rates and security but is incredibly costly and difficult to implement in an already established city. [30.]

As seen in Figure 8, Smart meters may apply protocols such as DLMS and COSEM made specifically for Smart meter devices in order to regulate transmissions for more efficient communication. Additionally, they can typically communicate through various interfaces such as UART, I2C and SPI depending on the circumstances. [30.]

Furthermore, when considering these communications, there are several important goals to meet in a Smart City setting. The technology should have the following qualities:

- Economically viable
- High transmission range
- Reliable, built-in security

When dealing with a massive number of devices spread over the size of a modern city, costs, the range at which communication can occur, and physical/network security are all incredibly important. [30.]

6.2.2 Smart Water and Electricity

Water is a resource that is essential to human life and must be meticulously managed in large cities. On a larger scale, agriculture plays a significant part in a city's water usage. Wide-scale distribution of sensors on farms could provide data on overall crop moisture levels. When such a large amount of data is processed, the trends could give instruction to farmers, via mobile networks, on how much water should be in use at any given time. This process could save potentially massive amounts of water when considering the size and number of farms included in the network. [2, pp. 158-159.] It is not only important to a city's food supply, but it is absolutely necessary for individuals as well. While water leakage is nearly impossible to prevent when reflecting on the size of a megacity, the placement use of previously mentioned Smart meters throughout a city's water supply grid could allow for real time reporting of these leaks and consequently lead to quicker reaction times and an overall downward trend of loss [2, p. 158].

Smart metering is not limited to just water, but also to another valuable resource: electricity. There have already been many advancements in this area, such as in Georgia, the USA. In this example, Smart metering has allowed for utility companies to distantly check a customer's electricity meter without having to expend time and money to physically be on the property. When companies are able to remotely contact the meters, it is also much more accessible for them to quickly restore power to these properties after an outage. Modern Smart meters are also capable of two-way communication, allowing for the meters to send diagnostic material regarding the target sensor as well as other sensors of the same type back to the data centre for additional analysis. [29, 30.] Furthermore, if most of the city could be a part of the Smart meter network, it would be much easier to get mass amounts of electricity usage data at a more frequent rate than ever before. The trends provided by such a system could lead to a better understanding on the electricity usage of a city as a whole, allowing for targeted campaigns on reduction efforts. [20, p. 6.]

6.3 Emergency Services

In addition to the conservation and improved distribution of utilities, advanced Smart Cities can vastly enhance emergency services. This ranges from processes such as improvements to first-responder abilities, to preparing for critical disasters in at-risk areas. Smart Cities can further develop many of these systems through the process of city-wide data collection and evaluation. For example, the advancement of the transportation sector can play a part here by allowing fast-tracked route planning and prioritization to emergency vehicles on the streets. Expanding this process could lead to faster and more effective response times. [25, p. 387.] Additionally, with increasing populations in natural disaster areas such as coastlines or earthquake zones, disaster preparedness can be impacted by the improved ability to better monitor weather related developments through bulk data processing [2, p. 159]. Furthermore, a city's crime sector could benefit from the expanse of surveillance in public places and communication between crime units [16].

6.3.1 Crime Statistics and Reporting

Crime is a definite concern in cities with large amounts of residents. The impacts of crime affect not only the individuals that commit them and their victims, but society as a whole. Smart cities can offer advantages to both prevention and response. For a real-life example, one can look to a crime prevention piloting programme that was conducted in South Korea a few years ago. A system was put into place that utilised mediums such as sophisticated video surveillance, alarms, and sound detection to monitor public spaces in an attempt to prevent crime as well as respond to emergency situations as they arise. One such case was the vehicle recognition system, which used cameras specifically equipped to capture a vehicle and its license plate number, make, and model. This helped officials to follow questionable persons as they moved throughout a city, whether the car was owned by a person of interest or the person had been behaving in a suspicious manner provoking police attention. [16.]

Another example is the use of sound detection in public spaces. This system was used and trained to identify abnormalities in public noise patterns, such as screams or accident-related commotions. When a sound of this nature was detected, the system was then able to use the location of the sensor to contact closed-circuit cameras nearby. The program would then alert the authorities responsible for reviewing and responding to the crime-prevention system with both the audio and video of the incident. [16.]

Both of these instances have the potential over time to allow for revealing trends in crime locations and frequencies as more data is collected. While there is a focus on real-time crime handling, successfully knowing in what places and what times crimes are more likely to be committed gives authorities the ability to counter these issues. This type of data analysis could lead to a more robust crime-prevention infrastructure.

6.3.2 Crisis Preparedness and Response

Smart Cities, and the data they can provide, are capable of greatly impacting a city's efficiency in disaster preparedness while also supplying a more comprehensive response after it has occurred. There are many ways in which a Smart City's data can contribute to these goals.

One aspect of preparedness is the of use sensors throughout the city to identify problem areas before catastrophic damage can be done. For instance, being able to monitor the status of a city's buildings would allow officials to know specific zones in which structures tend not to be up to code. A data bank of this magnitude could lead to targeted campaigns with a goal of providing maintenance to these unsafe structures. [32.] For another example of proactivity, one could look to Japan. There, areas that could be exposed to radiation (such as agricultural sectors and public spaces) are watched closely by SIM-enabled (Subscriber Identity Module) mechanisms. These devices tell the authorities when the areas have reached unsafe levels of radiation, prompting evacuation and public warnings. [17.]

In addition to preparedness, a city's ability to respond effectively to catastrophe is incredibly important. Extending information provided by mass data collection during a disaster to citizens within the area would be considerably helpful for their safety and wellbeing. For example, governments could develop useful mobile applications for tracking things such as power outages, the availability of petrol in nearby stations, or the real-time movement of disasters. Citizens could heavily benefit from this information by being able to check if they should be relocating, where resources (such as electricity, food, and water) are located, and how to safely make it to this new destination should they be advised to reposition. [32.]

7 Mass Data Handling Challenges

When a Smart City is managing large amounts of data, there are many concerns that affect nearly every part of the data collection and analysis process. Potential threats to citizen privacy can come from numerous sources such as hackers hoping to exploit personally identifiable information (PII), companies mining data from their consumers surreptitiously, and government surveillance on its citizens [33; 34, p. 10; 35, p. 2]. The public is not the only sector at risk, but governments also need to be on constant lookout for these types of security risks as well. Government breaches are a massive threat to any modern nation expanding the influence of Smart concepts. [36.]

Complications extend to infrastructure as well, where a Smart City faces issues pertaining to how exactly data can be collected, communicated, stored, and analysed safely. Not only do the networks need to be robust, but the physical infrastructure needs to be both reliable and central to the operation in order to handle the massive amounts of data produced by a large Smart City. [18, p. 9; 20, pp. 2-9.] Smart implementations are expensive and require extensive knowledge and planning in order to implement them securely and intelligently so as to reduce costs in construction and further maintenance [20, pp. 2-9].

7.1 Privacy Concerns

As increasing amounts of data are being collected from citizens, it is absolutely important to protect this data from harm not only for the consumer but also for the institutions that collect and subsequently use the said data [34, p. 1]. There are many forms of risks to the public, and this section will talk more in depth about two possible sources of harm: government surveillance and the handling of personal data by private companies.

7.1.1 Government Surveillance

It is important in a Smart City setting that inhabitants feel as though their privacy is respected in regard to both personal and impersonal data types. Personal data is considered to be concepts such as medical records or income whereas impersonal data is more general such as the traffic patterns or environmental factors affecting a city. Often times, the distinction between these two categories can become vague, and it is hard to gauge exactly how the public feels about government access and monitoring of these information types. [33.]

While the goal may be to gather widespread data and thus analyse it in order to study and reveal trends, government surveillance remains a highly sensitive topic that evokes strong emotions of the public. There have been many modern examples of governments using technology to perform invasive observation and analysis on citizens by using various forms of their data. In 2008, the city of Nice, France rolled out the most extensive public surveillance system of CCTV, earning much backlash from the community. In 2012, the government in Dresden, Germany tracked the movements of people throughout the city by monitoring citizens' mobile phones. They then logged the monitored activity, storing the data of these happenings. This action was also heavily criticized. Europe's General Data Protection Regulation (GDPR) has been introduced in recent years in order to help combat invasive usages of residents' data. [33.]

For another example of intrusion by government surveillance, one can look to the previously mentioned crime prevention piloting program in South Korea. According to the documentation, the need for sound detection arose from the illegality of using CCTVs as voice-recording surveillance. This violates the Privacy Act of South Korean citizens. In the crime prevention program, the sound-detection devices were used as a workaround for this law. The devices were trained to listen for abnormalities, but by doing this, the public was still subjected to their audio being recorded and stored for analysis. [16.] While the sentiment behind gathering this data is for the benefit of the public, it can still be viewed as an infringement on the privacy of citizens.

7.1.2 Private Company Data Handling

Making use of advanced, privatised data storage frameworks and analytical software will be a major part of a Smart City's diagnostic process. Much of the data being collected and analysed in a Smart City setting is personal and considered to be highly sensitive. This is a liability, as it requires private companies to handle this data securely which is lacking in many large data companies at present. [20, p. 9.] Additionally, private companies, such as those who deliver the services or physically build the technologies, may use techniques to accrue data for their own purposes through the practice of data mining [35, p. 2]. These practices may not be made apparent to the user without further investigation on their part.

Another major and prominent threat in today's society is that of data breaches. In one study conducted in 2015, it was concluded that violations of personal privacy through identification were the most detrimental for individuals as well as the institutions in which the breach occurred. These breaches did not just affect those who had their private information leaked, but also caused significant financial loss for the organisations. The data from the study suggested that large companies are hit with breaches repeatedly and amongst these are several different types of institutions including financial services, mobile providers, and social media websites. [34, pp. 7-9.]

In regard to data mining and breaches, consumers must worry about many different threatening scenarios. Not only can private companies take advantage of customer data for their own gain, but consumers have the added risk of security breaches leaking this data to malicious entities. The previously mentioned study suggests that customers' concern must go beyond just a company exploiting their private data and consider the implications of their personal information being leaked to "underground markets" where their privacy is much more at risk of harm. [34, p. 10.]

7.2 Security

Similar to privacy risks, threats to government and individual security comes in many different forms and is constantly expanding. As mentioned in the previous section, security breaches often result in substantial privacy violations. Security in a Smart City setting requires heavy consideration in both the structure and effectiveness of its attack preparedness and response. Data theft can be performed by many methods, be it hacking into vehicles to obtain global positioning system (GPS) information or illegally tapping into government databases to acquire sensitive PII. These types of attacks can have incredibly damaging effects and must be closely monitored in a Smart City environment. [37; 38.]

7.2.1 Threats to Citizens

There are a number of different ways that a Smart City inhabitant may be at risk of harm. As vehicles get smarter and have more access to the internet, they become increasingly vulnerable. Unrelated to data collection, there have been several student presentations on the ability to hack into smart cars, showing their power to access the door controls while the car was driving. Additionally, they were able to tamper with the steering and brakes. This is obviously a huge risk area, as cars are highly dangerous when mishandled. [2, p. 163.] However, the ability to hack into the cars provides other potential risk areas, such the capacity to access GPS information saved by users. Those who utilise their cars' GPS systems often save their key addresses for easy access. If their vehicle is compromised, hackers would be able to not only track someone's movements, but also have vital information on whether the person is home or away. This leaves the person open to attacks such as burglaries or physical injury in the case of stalker situations. [39, pp. 494-495.] If enough of these data points are collected, a comprehensive map of vulnerable residences and individuals could offer a reasonable market for criminals.

Furthermore, the tactic of phishing has become more sophisticated and continues to be a threat in the modern day. Phishing is when a malicious entity imitates a critical service, for example a bank, in order to retrieve sensitive information from a consumer such as their login credentials or PII. This is often done through electronic mediums, such as email. Being able to run successful phishing campaigns allows hackers to gain massive amounts of data on consumers, frequently leading to damaging outcomes. [40, p. 1090.]

7.2.2 Government and Corporation Breaches

One of the biggest threats to data security today is that of data breaches. Data breaches are when an unapproved entity gain access to "sensitive, protected, or confidential data". In a modern Smart City setting, governments consistently gather and evaluate data on sensitive topics such as government and military personnel PII, the economy, and the country's defence and safety measures. In the wrong hands, this type of information could be detrimental to the security of a nation. [36.]

The United States Office of Personnel Management issued a statement that in 2015, 4.2 million federal employees had PII stolen in a massive government breach. Among this data were highly sensitive pieces of personal information, including federal security clearances of those individuals. This information came from a wealth of background checks. While this is obviously a substantial risk to the government, 21.5 million people overall had their personal intelligence compromised from having some type of relation to the person of which the background check was conducted on. [37; 38.]

Data acquired through breaches can have many different uses. In the previous example, the personnel information stolen could have roots in identity theft or further government infringements using the federal clearances stolen. Another potential risk area for data theft is blackmail. In 2014, Sony found itself in the middle of a tricky situation in which malicious entities attacked the company and retrieved massive amounts of data, deleting the original copies from Sony's systems. They then proceeded to blackmail the company with threats of releasing their copy of the data to the public if their requests were not met.

[38.] There is no reason this type of exploit cannot be used to harm a government which could lead to highly dangerous and costly outcomes.

7.3 Financial Impacts and Data Value

Smart Cities have the ability ultimately to positively impact costs in a city over time. For instance, in the optimisation of traffic, the drastic decline in fuel usage has the potential for substantial savings. Additionally, the monitoring and distribution of resources could reduce waste in consumption, saving money in a Smart City's water and electricity sectors. [2, p. 156; 20, p. 4.] However, Smart implementations are an incredibly expensive investment both initially and for upkeep. New technologies can prove to be important assets to all parts of the data process but are often costly and take time and expertise to employ appropriately. Additionally, if these technologies are not planned and used properly, the city could suffer a huge loss in time and finances from both the initial operation and subsequent reconstruction. [20, pp. 2-9.]

Financial costs are not just limited to the rollout of Smart City concepts and cities need to consider whether the cost of such heavy data analysis is worth the outcome overall. While this topic is most suitable for a full cost-benefit analysis of Smart City implementations, the next section will lay out a prospective scenario due to the current scarcity of available data sources on Smart City data generation and potential associated costs.

Figure 9. Smart City – Multiple Applications Create Big Data. Copied from Cisco 2015 [41].

As seen in Figure 9, Cisco estimated that by 2019, Smart Cities with a population of one million people would generate over 180 million gigabytes of data. This data can come from a number of sources, both from inhabitants as well as Smart edge devices used to monitor various statistics. For this dataset, Smart Grid, Smart Hospital, and Public Safety will be used. Despite being expansive, Smart Grid sensors are estimated to only generate about five gigabytes of data per day, for a total of roughly 150 gigabytes per month. Smart Hospitals produce much more data, as there could be vastly more sources of data involving patients, capacities, etc. creating around three thousand gigabytes per day and ninety thousand per month. Finally, Public Safety is likely to use lots of CCTV streaming which makes it very heavy in terms of data generation, with an expectation of around fifty million gigabytes per day and 150 million per month. [41.]

It is difficult to find statistics on the costs associated with this level of data storage, and the finances will likely be affected by things such as the volume, types of analysis, and government contracts. For this reason, this dataset will utilise the basic Amazon Web Services (AWS) financial calculator to estimate costs.

Smart City Estimated Data Storage/Analysis Financial Impact

Sector	Daily Data Generation (Gb)	Monthly Data Generation (Gb)	Transmission Rate	Adjusted Daily Data Generation (Gb)	Estimated Monthly Cost (USD)	Annual Cost (USD)	References
Smart Grid	5	150	0.007	1.05	$72.45	$872.98	[41],[42]
Smart Hospital	3,000	90,000	0.007	630	$652.62	$6,518.35	[41],[42]
Public Safety	50,000,000	1,500,000,000	0.007	10,500,000	$44,656,633.72	$92,602,006.63	[41],[42]
Total	50,003,005	1,500,090,150	NA	10,500,631	$4,657,358.79	$92,609,397.96	[41],[42]

City-Specific Estimated Data Storage/Analysis Financial Impact

City	Population	Estimated Monthly Cost (USD)	Estimated Annual Cost (USD)	References
London	9,300,000	$43,313,436.75	$861,267,401.03	[41],[42],[43]
Singapore	6,000,000	$27,944,152.74	$555,656,387.76	[41],[42],[43]
New York	8,200,000	$38,190,342.08	$759,397,063.27	[41],[42],[43]

Figure 10. Estimated Smart City Data Generation Rates and Associated Costs. Based on Cisco Global Cloud Index [41], Amazon Web Services financial calculator [42], and Population Statistics [43].

In Figure 10, these estimates can be seen. First, the dataset only includes the sectors and values mentioned in the previous Cisco figure for Smart Grid, Smart Hospital, and Public Safety. The monthly rate was calculated based on a 30-day multiplier. Cisco has further estimated that out of all data generated, rough 0.7% will be stored/analysed, so that value was used as the transmission rate for calculating the adjusted data usage. After that, these values were plugged into AWS's financial calculator with the following assumptions:

1. Both raw and processed data would be stored in the full amounts

2. The data would be retained for 12 months

3. Every day of the month, a query of the full dataset would be made

These factors are what led to the values in the sixth and seventh column. At this stage, staggering costs can already be seen for instance in the Public Safety sector, with an estimated ninety-two million dollar annual expense. In the second table, these values have been applied to the populations of the cities discussed earlier to provide potential data storage and analysis costs a Smart City may face. The expenses associated with collecting and processing such high amounts of data are staggering. [41; 42; 43.]

While it is too early to tell in most cases, this will be a huge hurdle for many governments when it comes to assessing the value attached to such a massive financial investment. This dataset only covers 3 potential sectors of a Smart City and does not begin to address the costs associated with implementing these technologies, maintaining them, and replacing them over time. However, many entities maintain that Smart Cities will save governments, businesses and inhabitants much more money in the long run. A fairly recent study conducted by ABI Research concluded that governments of Smart Cities could save the equivalent of nearly five billion U.S. dollars annually, with main proponents being smart streetlights and buildings. It is also estimated that investments in Smarter water and electricity infrastructure could make a huge impact for citizens, saving them potentially the equivalent of 26 billion dollars. [44.]

While it would seem that Smart Cities have the potential to save large amounts of money for governments, it still may be too early to tell if the overall value provided by mass data generation, storage, and analysis will prove to be rewarding enough to justify that kind of financial loss. However, future technologies may provide the possibility for cheaper avenues regarding data storage and analysis.

7.4 Non-threat Constraints

Implementing big data in Smart Cities goes beyond just concerns of dangers to both privacy and security. Other major topics need to be considered such as network ability and physical constraints. Physically, a significant distance between all entities communicating on a network provides latency issues for data transmission. The size of resources such as a data centre for storing and processing big data may prove difficult to find accessible in a large, established city. [15; 20, pp. 2-9.]

Not only must the size and location of the facilities be considered, but also the types of devices used in the storage and analysis of data are very important. Storage devices in use should be exceptional and follow best practices for data encryption due to the large amount of PII likely accumulated on the devices. Redundancy is also important, and these storage devices should be tough in their ability to run under continuous stress, with additional backup machines that can take the load in the case of failure. Additionally, there need to be enough quality machines to be able to store and keep data safe for extensive time cycles. For instance, being able to access archived items such as video surveillance is useful in law enforcement, where personnel often investigate crimes for many months or even years. [19.]

Networks in a Smart City need to be robust for the sheer amount of data that will be consistently communicated. In 2017, it was predicted that nearly 21 billion devices would be interconnected by the end of the decade. This number of devices and the push for faster transmissions can put a huge strain on modern networks, especially for Smart Cities looking to provide city-wide wireless network access to their inhabitants. Edge computing can offer a solution to some network traffic, but there will still need to be advancements made in networks that can take not only user traffic, but traffic from distributed Smart sensors and devices. [18, p. 9.]

8 The Future of Smart Cities

In the future, the world will see increasing numbers of Smart Cities as populations and the ensuing need for Smart implementations rise. With these changes, there are many challenges that both inhabitants and governments will face. Governments will encounter cybersecurity threats, and a lack of public trust with their data. [33; 45, p. 51.] However, many modern Smart City advancements have already been seen in recent years. The findings from these current accomplishments will provide a foundation for the future of Smart Cities going forward.

8.1 Challenges

Smart Cities in the future will be confronted with many obstacles to overcome. Among these are limitations both by hardware and analysis software in the processing of data, low levels of public trust in Smart technologies, interoperability, and a lack of organised leadership. [20, pp. 10-11; 33; 45, p. 51.] Because Smart Cities are a relatively new concept and advancements in technology are always on the rise, it takes some time to work through these challenges.

Simply put, technology today is just not fully up to the task of analysing city-wide data accurately. In South Korea after piloting Smart crime prevention, the findings indicated that inadequacies in technologies used to analyse the sound data promoted false claims of crime events. While it did successfully identify legitimate events, having to additionally filter false alarms can interrupt the system flow and cause delays in response. Additionally, the data collected in this setting had very specific uses and did not transfer so easily to other sectors where more comprehensive data collection could be useful. In the future, it will be much more valuable to review how to expand on the information gathered, so that multiple regions of a Smart City can make use of the accumulated data. [20, p. 10.]

Governments in Smart Cities also have much to consider when it comes to moving forward. Because cybersecurity is so new to most governments, many have a deficiency when it comes to financing and organising a comprehensive cyber security department with experienced, dedicated personnel. Smart City governments absolutely need to devote financial and staff resources in order to have a cybersecurity sector capable of protecting against the wide range of threats. [45, p. 51.] Additionally, they must be able to overcome public trust issues concerning citizen privacy. In 2013, a survey in the United

Kingdom indicated that only 11% of the population felt trust in their own national government. A government needs the ability not only to win over the trust of its citizens in regard to their data, but also to maintain that confidence by ensuring citizens' privacy is not violated and that the data is always ethically used to better society. [33.]

8.2 Plans and Ideas for a Smarter Future

As population density grows in cities all over the globe, there has never been a higher demand for Smart implementations. Forthcoming Smart Cities should look a certain way if they reach the previously mentioned goals of citizen quality of life improvement, sustainability, and connectivity. [2, p. 147; 3, pp. 5-6.] Based on modern examples, one can imagine some concepts that may appear in the future of Smart Cities.

Smart Cities should have comprehensive data analysis from sectors all over the city and the follow up ability to act effectively on the evaluations received. They should have in depth assessments on different parts of society, so as to provide a better picture on how Smart Cities operate and to expose areas of growth. These findings should push governments to make improvements based on the data with the goal of providing citizens a better quality of life. [46.]

With GHG emissions contributing to global warming, Smart Cities will need a comprehensive way to address energy needs. Green energy practices need to be implemented, and a city's Smart Grid will likely be the answer to this issue. By enhancing the power grid, cities could gain the ability to analyse usage statistics. Applying this analysis, the grid could intelligently route power based on necessity, increasing capacity in places that require it and cutting back in less resource-intensive areas. Additionally, when power requirements are low, consumers can provide energy back to the city from their own renewable sources. [2, p. 158; 46.] This would be a great way for Smart Cities to work towards goals of sustainability.

It is possible that Smart Cities in the future may convert to fully autonomous transportation. These vehicles, utilising V2X communication, would provide inhabitants safe and efficient transportation options for moving around a city whether they are using personal vehicles or public transport. [46.] Some scientists believe that robust V2X implementation (even without fully autonomous transportation) may completely rework the way people drive, removing the requirement for road signals completely and relying solely on advanced analytics to control the flow of traffic [2, p. 156]. These are all ideas that have been somewhat implemented in places around the world already and may very likely contribute to the foundation of future Smart Cities.

9 Conclusion

In conclusion, a Smart City is an incredibly complex concept that requires much planning and consideration to be implemented meaningfully. Even though Smart Cities are comprised of many different ideas, improving quality of life is really the ultimate goal. For a lot of countries, this includes working towards a more connected society, with the ability to sustain itself. [2, p. 147; 3, pp. 5-6.] Examples of these modern Smart Cities can be seen all over the globe, for instance in London, Singapore, and New York [3, p. 11; 14, p. 703; 38].

One of the biggest attributes to becoming a Smart City is that of data collection and processing. Mass data analysis is key to understanding trends in a city that will provide governments with ideas on how to improve infrastructure and ultimately quality of life. [7; 16.] Smart Cities will need to employ both cloud and edge storage tactics in cooperation with each other to handle massive amounts of data, considering the distance between devices and data centres to reduce latency and network overload. [15; 18, p. 9.] Furthermore, analysis of such a large mass of data is a challenging concept. Because of this, there are requirements that must be met for a Smart City's analysis sector to be successful both initially and over time. The system must be structured in a way that allows for scalability, data privacy and security, and interoperability. [18, pp. 7-8; 19; 20, pp. 9-11; 21.] Artificial intelligence will play a significant role in surpassing human capabilities in examination of the data as well as the time invested in performing the assessments [24].

However arduous the process for mass data evaluation will be, it undoubtably has countless benefits for individuals and governments alike. Traffic in a Smart City would be greatly improved. Utilising V2X communication to connect vehicles with pedestrians and other nearby entities could lead to fewer GHG emissions and more streamlined, safe traffic flows. [2, pp. 155-156; 25, p. 387.] Smart sensors located throughout Smart Cities will relay information on concepts such as water and energy usage, allowing for governments to address distribution and leakage issues. [2, pp. 158-159; 20, p. 5.] In recent studies on Smart crime prevention, it was found that varying real-time analysis technologies can provide automated crime detection services, and then alert these events to the proper departments for quick response. [16.] Mobile applications sourcing data during an active city-wide emergency could be deployed by the government to warn and guide citizens to safety and important resources. [17; 32.]

While Smart Cities can absolutely use data for good, there are many risks associated with mass collection and storage of sensitive data. There have been examples in recent years in Europe where governments invested significant resources into invasively monitoring citizens and storing data on their movements. [33.] Additionally, citizens' privacy is at constant risk in a Smart City setting, not just from hackers but from institutions that they choose to share their data with. Data breaches are unfortunately very common, leaving individuals susceptible to attacks when sensitive PII is stolen and shared with malicious entities. [34, pp. 7-10; 35, p. 2.] Governments also have to be aware of constant attacks. Recent successful government data breaches have had detrimental effects, when massive amounts of highly confidential PII on government personnel was stolen. [37; 38.]

Cities must also consider constraints that are not threatening in nature, but just as important as data protection. There must be consideration in the size, location, and number of resources needed for a substantial amount of data processing. Furthermore, networks need to be able to handle the constant, excessive communication of data amongst users in the city and edge devices. [18, p. 9; 20, pp. 2-9.]

Smart cities of the future will face many challenges in technological capabilities, public outlook and opinion, and the protection of Smart resources [20, pp. 10-11; 33; 45, p. 51]. They will need to have thorough analytical processes and proactive views on using Smart technologies to the fullest in order to improve society [46]. While it will be very challenging moving forward, investing in Smart Cities globally will be essential to advancing civilization as a whole. There is no doubt that regardless of the risks, the benefits of achieving a more connected world will be immensely impactful for generations to come.

References

1 Cisco Annual Internet Report - Cisco Annual Internet Report (2018–2023) White Paper [Internet]. Cisco; 2020. URL: https://www.cisco.com/c/en/us/solutions/collateral/executive-perspectives/annual-internet-report/white-paper-c11-741490.html. Accessed 5 August 2020.

2 Mouftah HT, Erol-Kantarci M, Rehmani MH. Transportation and Power Grid in Smart Cities: Communication Networks and Services [e-book]. Newark: John Wiley & Sons, Incorporated; 2018. URL: https://ebookcentral.proquest.com/lib/metropolia-ebooks/detail.action?docID=5630272. Accessed 27 October 2020.

3 Greater London Authority. Smarter London Together: The Mayor’s roadmap to transform London into the Smartest City in the World [PDF]. Greater London Authority; 2018. URL: https://www.london.gov.uk/sites/default/files/smarter_london_together_v1.66_-_published.pdf. Accessed 19 February 2021.

4 Brian Copeland. New Digital Map Gives Local Residents and Visitors a New Way to Experience Queen Elizabeth Olympic Park [Internet]. Living Map; 2017. URL: https://www.livingmap.com/queen-elizabeth-olympic-park. Accessed 22 February 2021.

5 Empire State Development. New York Smart Cities Innovation Partnership [Internet]. Empire State Development, New York City Government; 2020. URL: https://esd.ny.gov/new-york-smart-cities-innovation-partnership#process-&-timeline. Accessed 10 February 2021.

6 Harmon R, Castro-Leon E, Sandhiprakash B. Smart Cities and the Internet of Things. 2015 Portland International Conference on Management of Engineering and Technology; 2015. URL: https://www.researchgate.net/publication/305183838_Smart_cities_and_the_Internet_of_Things. Accessed 19 December 2020.

7 Khan Z, Anjum A, Kiani S. Cloud Based Big Data Analytics for Smart Future Cities. 2013 IEEE/ACM 6th International Conference on Utility and Cloud Computing; 2013. URL: https://www.researchgate.net/publication/262412207_Cloud_Based_Big_Data_Analytics_for_Smart_Future_Cities. Accessed 29 December 2020.

8 International Programs - Information Gateway - U.S. Census Bureau [Internet]. United States Census Bureau; 2019. URL: https://www.census.gov/data-tools/demo/idb/informationGateway.php. Accessed 14 September 2020.

9 Sang KL, Heeseo RK, HeeAh C, Jongbok K, Donju L. International Case Studies of Smart Cities: Singapore, Republic of Singapore [Internet]. Inter-American Development Bank; 2016. URL: https://publications.iadb.org/publications/english/document/International-Case-Studies-of-Smart-Cities-Singapore-Republic-of-Singapore.pdf. Accessed 4 January 2021.

10 Government of Singapore. New Features, New Look for LTA’s My Transport.SG Mobile Application [Internet]. Land Transport Authority; 2018. URL: https://www.lta.gov.sg/content/ltagov/en/newsroom/2018/9/2/new-features-new-

look-for-ltas-my-transportsg-mobile-application.html. Accessed 22 February 2021.

11 One Motoring [Internet]. Land Transport Authority; 2021. URL: https://onemotoring.lta.gov.sg/content/onemotoring/home.html. Accessed 22 February 2021.

12 Liming B, Gavino A, Lee P, Jungyoon K, Na L, Pi T, Tan H, Buay T, Xiaoping T, Valera A, Jia E, Wu A, Fox M. SHINESeniors: Personalized Services for Active Ageing-In-Place [Internet]. 2015 IEEE First International Smart Cities Conference; 2015. URL: https://www.researchgate.net/publication/308605676_SHINESeniors_Personalized_services_for_active_ageing-in-place. Accessed 22 February 2021.

13 NYC Mayor's Office of Tech and Innovation. Building a Smart and Equitable City [Internet]. New York City Government; 2015. URL: https://www1.nyc.gov/assets/forward/documents/NYC-Smart-Equitable-City-Final.pdf. Accessed 22 February 2021.

14 Shah J, Kothari J, Doshi N. A Survey of Smart City Infrastructure via Case Study on New York [Internet]. Procedia Computer Science; 2019. URL: https://www.sciencedirect.com/science/article/pii/S1877050919317247. Accessed 22 February 2021.

15 Burton J. The Facts About Big Data Storage in Smart City Applications [Internet]. Seagate Technology; 2019. URL: https://www.securityinfowatch.com/video-surveillance/video-surveillance-storage/article/21089531/seagate-the-facts-about-big-data-storage-in-smart-cities-applications. Accessed 29 December 2020.

16 Park M, Lee H. Smart City Crime Prevention Services: The Incheon Free Economic Zone Case [PDF]. MDPI; 2020. URL: https://www.mdpi.com/2071-1050/12/14/5658/htm. Accessed 16 November 2020.

17 Smart City Resilience – Learning from Emergency Response and Coordination in Japan [PDF]. Japan Meteorological Agency; 2013. URL: https://www.gsma.com/iot/wp-content/uploads/2013/02/cl_SmartCities_emer_01_131.pdf. Accessed 14 September 2020.

18 Ahmed E, Ahmed A, Yaqoob I, Shuja J, Gani A, Imran M, Shoaib M. Bringing Computation Closer Towards User Network: Is Edge Computing the Solution? [Internet]. IEEE Communications Magazine; 2017. URL: https://www.researchgate.net/publication/317617524_Bringing_computation_closer_towards_user_network_Is_edge_computing_the_solution#pf3. Accessed 4 January 2021.

19 Shahat A, Elragal A, Bergvall-Kåreborn B. Big Data Analytics and Smart Cities: A Loose or Tight Couple? [Internet]. International Conference on Connected Smart Cities 2017; 2017. URL: https://www.researchgate.net/publication/317491579_Big_Data_Analytics_and_Smart_Cities_A_Loose_or_Tight_Couple. Accessed 7 January 2021.

20 Al Nuaimi E, Al Neyadi H, Mohamed N, Al-Jaroodi J. Applications of Big Data to Smart Cities [Internet]. Journal of Internet Services and Applications; 2015. URL:https://www.researchgate.net/publication/284196317_Applications_of_big_data_to_smart_cities. Accessed 15 October 2020.

21 Carr D. Edge Computing Vs. Cloud Computing: What's the Difference? [Internet]. The Enterpriser's Project; 2020. URL: https://enterprisersproject.com/article/2020/4/edge-computing-vs-cloud-what-is-difference. Accessed 22 December 2020.

22 Cha H, Yang, H, Song Y. A Study on the Design of Fog Computing Architecture Using Sensor Networks [Internet]. Sensors; 2018. URL: https://www.mdpi.com/1424-8220/18/11/3633. Accessed 7 March 2021.

23 Artificial Intelligence Definition [Internet]. Cambridge English Dictionary; 2021. URL: https://dictionary.cambridge.org/dictionary/english/artificial-intelligence. Accessed 13 January 2021.

24 Navarathna PJ, Malagi VP. Artificial Intelligence in Smart City Analysis [Internet]. 2018 International Conference on Smart Systems and Inventive Technology (ICSSIT); 2018. URL: https://ieeexplore-ieee-org.ezproxy.metropolia.fi/document/8748476. Accessed 13 January 2021.

25 Hamida E, Noura H, Znaidi W. Security of Cooperative Intelligent Transport Systems: Standards, Threats Analysis and Cryptographic Countermeasures [PDF]. Electronics; 2015. URL: https://www.researchgate.net/publication/279765559_Security_of_Cooperative_Intelligent_Transport_Systems_Standards_Threats_Analysis_and_Cryptographic_Countermeasures. Accessed 16 November 2020.

26 McLellan C. What is V2X Communication? Creating Connectivity for the Autonomous Car Era [Internet]. ZDNet; 2019. URL: https://www.zdnet.com/article/what-is-v2x-communication-creating-connectivity-for-the-autonomous-car-era/. Accessed 27 October 2020.

27 US EPA O. Fast Facts on Transportation Greenhouse Gas Emissions [Internet]. US EPA; 2015. URL: https://www.epa.gov/greenvehicles/fast-facts-transportation-greenhouse-gas-emissions. Accessed 27 October 2020.

28 Gao JH, Peh L. Automotive V2X on phones: Enabling next-generation mobile ITS apps [Internet]. 2016 Design, Automation Test in Europe Conference & Exhibition; 2016. URL: https://ieeexplore-ieee-org.ezproxy.metropolia.fi/document/7459429. Accessed 12 October 2020.

29 Pucilowski, D. Three reasons smart cities need smart utilities [Internet]. POWER-Grid International; 2019 URL: https://www.power-grid.com/der-grid-edge/three-reasons-smart-cities-need-smart-utilities. Accessed 16 November 2020.

30 Chakraborty S, Das S, Sidhu T, Siva A. Smart Meters for Enhancing Protection and Monitoring Functions in Emerging Distribution Systems [Internet]. International Journal of Electrical Power and Energy Systems; 2021. URL: https://www-sciencedirect-com.ezproxy.metropolia.fi/science/article/pii/S0142061520341715#s0010. Accessed 24 February 2021.

31 Christensson, P. Non-Volatile Memory Definition [Internet]. TechTerms; 2019. URL: https://techterms.com/definition/non-volatile_memory. Accessed 24 February 2021.

32 Matthews K. Smart Cities May Be a Key Natural Disaster Resource [Internet]. Smart City Hub; 2019. URL: http://smartcityhub.com/security/smart-cities-natural-disaster. Accessed 26 November 2020.

33 Van Zoonen L. Privacy Concerns in Smart Cities [Internet]. Government Information Quarterly; 2016. URL: https://www.sciencedirect.com/science/article/pii/S0740624X16300818. Accessed 19 January 2021.

34 Wheatley S, Maillart T, Sornette D. The Extreme Risk of Personal Data Breaches and the Erosion of Privacy [Internet]. The European Physical Journal; 2015. URL: https://www.researchgate.net/publication/277333872_The_Extreme_Risk_of_Personal_Data_Breaches_The_Erosion_of_Privacy. Accessed 19 January 2021.

35 Cui L, Xie G, Qu Y, Gao L, Yang Y. Security and Privacy in Smart Cities: Challenges and Opportunities [Internet]. IEEE Access; 2018. URL: https://www.researchgate.net/publication/326337187_Security_and_Privacy_in_Smart_Cities_Challenges_and_Opportunities. Accessed 12 January 2021.

36 Joseph RC. Data Breaches: Public Sector Perspectives [Internet]. IT Professional; 2018. URL: https://ieeexplore-ieee-org.ezproxy.metropolia.fi/document/7950860. Accessed 25 January 2021.

37 Cybersecurity Incidents [Internet]. U.S. Office of Personnel Management. Available from: https://www.opm.gov/cybersecurity/cybersecurity-incidents/. Accessed 25 January 2021.

38 Wang P, Ali A, Kelly W. Data Security and Thread Modelling for Smart City Infrastructure [Internet]. 2015 International Conference on Cyber Security of Smart Cities, Industrial Control System and Communications; 2015. URL: https://www.researchgate.net/publication/308806952_Data_security_and_threat_modeling_for_smart_city_infrastructure. Accessed 25 January 2021.

39 Elmaghraby A, Losavio M. Cyber Security Challenges in Smart Cities: Safety, Security and Privacy [Internet]. Journal of Advanced Research; 2014. URL: https://www.researchgate.net/publication/260559293_Cyber_Security_Challenges_in_Smart_Cities_Safety_security_and_privacy. Accessed 12 January 2021.

40 Aldairi A, Tawalbeh L. Cyber Security Attacks on Smart Cities and Associated Mobile Technologies [Internet]. Procedia Computer Science; 2017. URL: --https://www.researchgate.net/publication/317548513_Cyber_Security_Attacks_on_Smart_Cities_and_Associated_Mobile_Technologies. Accessed 12 January 2021.

41 Cisco. Cisco Global Cloud Index: Forecast and Methodology, 2014-2019 [PDF]. Cisco; 2015. Accessed 7 March 2021.

42 Amazon Web Services. AWS IoT Analytics Calculator [Microsoft Excel spreadsheet]. AWS; 2021. URL: https://aws.amazon.com/iot-analytics/pricing/. Accessed 7 March 2021.

43 World City Populations 2021 [Internet]. World Population Review; 2021. URL: https://worldpopulationreview.com/world-cities. Accessed 7 March 2021.

44 Ismail N. Smart Cities Could Lead to Cost Savings of $15 Trillion-Report Suggests [Internet]. Information Age; 2017. URL: https://www.information-age.com/smart-cities-lead-cost-savings-5-trillion-123469863. Accessed 9 March 2021.

45 Kitchin R, Dodge B. The (In)Security of Smart Cities: Vulnerabilities, Risks, Mitigation and Prevention [Internet]. Journal of Urban Technology; 2019. URL: https://kitchin.org/wp-content/uploads/2020/02/JUT-2019-Security-of-Smart-Cities.pdf. Accessed 12 January 2021.

46 Alexander D. 5 Things You Will See in the Future of “Smart City” [Internet]. Interesting Engineering; 2018. URL: https://interestingengineering.com/5-things-you-will-see-in-the-future-of-smart-city. Accessed 26 January 2021.