MPhil Southern Urbanism Dissertation

Data Driven Urbanism: Challenges in implementing open data policy and digital transparency in the City of Cape Town

By

Majaha Dlamini

Supervised by Professor Sophie Oldfield & Professor Nancy Odendaal

August 2019

Declaration

1. I know that plagiarism is wrong. Plagiarism is to use another's work and pretend that it is one's own.

2. I have used the Harvard convention for citation and referencing. Each contribution to, and quotation in this essay/report/project from the work(s) of other people has been attributed, and has been cited and referenced.

3. This essay/report/project is my own work.

4. I have not allowed, and will not allow, anyone to copy my work with the intention of passing it off as his or her own work

Signed by candidate

Date 30/08/2019

Acknowledgements

I would firstly like to acknowledge the opportunity that the Almighty God granted me to complete my studies. I have been provided and sustained by His grace.

I extend my gratitude to the African Centre for Cities for granting me the opportunity to be part of this inaugural MPhil in Southern Urbanism program. Completing this program would have not have been possible without the scholarship awarded by the centre through its Director Professor Edgar Pieterse whom I'm grateful to and admire unreservedly.

I am also grateful and appreciating the support of my family for understanding my decision to leave fulltime employment to focus on the program. I am also grateful to friends that I also regard as family for all the support they have provided me in various ways.

Last but not least, I would like to sincerely thank my supervisors, Professor Sophie Oldfield and Professor Nancy Odendaal for their unwavering support through providing supervision and moral support when it would all seem daunting.

List of Abbreviations

CoCT – City of Cape Town

HCI - Human-Computer Interaction

ICT – Information Communication Technology

GIS – Geographic Information Systems

NGO – Non-Governmental Organisations

OGD – Open Governmental Data

PSI – Public Sector Information

Abstract

As part of its quest to become the first digital African city, in 2014 the City of Cape Town adopted an open data policy, which was later coupled with an open data portal to make government data available for public access. This was touted as a novelty initiative as the City of Cape Town was the first African city to implement a policy of this nature. This open data initiative aimed at enhancing transparency and accountability as well as promoting inclusive economic participation for its citizens.

Open data project managers from the city and external industry experts working on open data initiatives were interviewed to understand the current the state of open data within the city and how it worked with other stakeholders. The study draws on these interviews to present the current challenges experienced by the city from the city's official point of view as well as from open data experts working closely with the city. To understand the practical experiences of how the city publishes data in its platforms, the study also extensively explored the city's open data portal, as well as examining and commenting on the documented open data policy guidelines contrasted and compared to current practical experiences.

To guide the objectives and analysis of the study, four key themes were adopted from literature; *context*, *use*, *data* and *impact*. *Context* focused on the overall context or environment at which open data in the city is provided as a public service, while *use* focused challenges on the uses of open data as well as it is users, *data* focused on the types of datasets published on the portal as well as the technical challenges in publishing them. Lastly *impact* looked at the expected benefits and goals of the city's open data policy.

The study through the themes highlighted the ongoing challenges at various levels that the city experience as they implement and develop the open data policy. Overall it was noted that open data was not a goal but continuous challenges were arising daily while implementing and developing the policy- while it was noted that various stakeholders within and outside government had to collaborate to effectively meet the required open data standards.

Keywords: Open data, Open Government Data, Digital Transparency, Open data portal, Data release

Contents

List of Tables

List of Figures

Chapter 1- Open Data: Novelty with high expectations

1.1 Introduction

In the quest of efficient and effective governance for city governments, open data has been seen as one of the tools to help governments to operate in transparent and open customs. Over the past decade there has been a dramatic increase in the adoption of Open Government Data (OGD) initiatives by governments from all levels; nationally and locally across the world. Halonen (2012:6) asserts that "*open data has been hailed as one of the most important public policies of our time, and the potential impacts of sharing such data cooperatively are enormous".* Governments, however are faced with several challenges in implementing these policies effectively. Penning policy into practice has been the core of these challenges facing many governments worldwide. Magalhaes and Roseira (2017) note that generally the release of open data by governments created expectations of increased public accountability, enhanced public sector efficiency, and encourage innovation and economic value. Conradie and Choenni (2014) interject and suggest that currently the release of government datasets is relatively new and there is less experience and knowledge so far about its barriers, costs and benefits. The objective of this study is to understand the current barriers – challenges limiting the city in implementing its open data policy and thereby its digital transparency.

In 2014 the then Mayor of the City of Cape Town Patricia De Lille similarly approved an open data policy for the city to increase its operational efficiency and improved public accountability. This policy addresses how datasets can be released online for public consumption openly at no cost and without any legal restrictions.

Post the implementation of the policy, to make the data accessible to public, an online open data portal was developed. On the portal the city makes accessible various datasets that can be downloaded for free, by anyone and for any use with no legal restrictions - concerning city services and opportunities such as tender information, census data, crime statistics, transport data, land administration to basic service and infrastructure – healthcare facilities, road networks, government buildings and water networks, etc.

Before the adoption of this policy, government datasets would be hidden from the various city departments and the public would not have access to it. This came with many challenges for the city and its civilians. The city generated significant volumes of data, however this information would be hidden in departments and was difficult to access by civilians. Similarly, there were various data access policies and procedures within the city that hindered public access to data (CoCT Open data policy, 2015).

To oversee the critical operations and practical implementation of the policy, the mayor appointed an Open Data Steering Committee comprised of internal city officials and two external members with open data and related expertise. One of the key functions of the steering committee is to ensure that requested data for release into portal is made accessible by the various contributing departments. The external committee members play a critical role in guiding the departments on issues such as release of data, data security – what data to release or not among other important factors. The open data committee has been met with resistance from departments as a result of incomprehensive internal data sharing policies. This meant that even internally city officials from the various departments had to be educated on the guidelines and importance of data release.

Kitchin (2015:1), asserts that data-driven, networked urbanism is the "*key mode of production for what have widely been termed smart cities".* While city governments are adopting data-driven urbanism initiatives and policies; there has been equally challenges and criticism for such. Kitchin (2015) further argues that while these initiatives provide solutions for urban problems, they do so within restrictions and in the service of particular interests.

Attard et al., (2015) supports that there has been a rapid growth recently of open data initiatives for governments with transparency and data reuse as the two main objectives. The release of government data has been greatly seen to bring many solutions for governments and its civilians. For Dvir (2018) suggests that the release of government datasets to the public can potentially strengthen the relationship between the government and its residents. However, studies have shown that there are several challenges facing open data's efficacy.

Conradie and Choenni (2014) assert that studies that relate to open data lack the focus on the challenges experienced at local level by public sector information professionals. Equally another important observation identified in literature are the impacts that open data brings.

Kereru and Chan (2015) argue that there is a compelling lack of evidence on the social and political impacts of open data initiatives, mainly due to the novelty nature it has. The two key issues raised above signifies the importance of this study; the study will provide an opportunity to understand the issues faced by the City of Cape Town Open Data Steering committee and other professionals in the public sector information. Secondly it will also provide the opportunity to assess the political and social impact it has brought at a local level since its inception. These discoveries will conclusively assist the study in understanding the entire ecosystem of the open data and how it operates as a continuum, and also understand what the possible benefits and impacts on a local context.

As a first African city to implement an open data policy, the study of the City of Cape Town will help understand and relate the challenges and issues of open data experienced at a local level.

1.2 Research Questions

To achieve this, the study will look at the following research questions;

1. What are the main barriers and limitations in implementing Open Government Data policy efficiently in the City of Cape Town?
2. Who is using Open Government Data? What motivates their engagement with ODG?
3. How is Open Government Data from the city's portal being used in practise? What technical processes are being adopted?

The research questions would be part of the empirical work of the study. These questions therefore seek to understand and assess how the open data policy is implemented in practise and the challenges experienced in the process.

To assess the open data policy systematically, four key assessments themes have been adopted from literature to guide the issues the study seeks to focus on. These are the themes; Context, Data, Use and Impact (Caplan et al., 2014).

Context – addresses the context or environment within which government data is released. Context can be national, local or a particular sector. In this study it will focus on the local context which is the City of Cape Town. Context covers important features such as

organisational context- the makeup of the city organisational structures and stakeholders, political will & leadership, legal and regulatory framework etc. (Caplan et al., 2014)

Data – involves the category and qualities of datasets made available on the city's open data portal. It focuses on the social, technical, and practical limitations of data openness.

Use – This component deals with the use of open datasets. It involves current and potential users of the open datasets, through accessing or providing. It also focuses on the purposes for which datasets are being used. Caplan et al., (2014) asserts that this part of the component focuses on the - *who*, *what* and *why* of open data initiatives.

Impact – the impact component addresses the potential benefits expect from using open data. Impact is assessed through social, political/governance, and economic dimensions. It focuses on what benefits does it bring to these dimensions (Caplan et al., 2014).

These four key themes will narrowly guide the broader research questions, to pinpoint the limitation and challenges in the context, the data, it uses.

To achieve the study's objectives, the research will have two approaches; an empirical desktop assessment of the open data portal as well as interviews. The former will assess the state of the portal as a repository for data accessibility and the qualities of the datasets made available. Various portal assessment methods adopted from literature will be used. This practical exploration and assessment of the portal will attempt to bridge the contrast of what is in policy documents and what is being done in practice – policy documents stipulate how the portal should work and how data must be stored and accessed. This assessment will afford the study an opportunity to assess the city's portal quality by standards of open data principles.

For interviews, various stakeholders will be interviewed, this includes the City of Cape Town open data officials, a member from the Open Data Steering Committee, industry experts and officials from open data Non-Governmental Organisations. Interviews with city officials will help the study understand the current organisational context, operations; technically, legally and otherwise, while the steering committee member will be expected to provide neutral and balanced inputs on the city's open data operations and limitations. Industry experts and NGOs

are also vital as the users of open data as they would detail how they use open government data and what impact and benefits they expect from it.

Findings from these two approaches will make up the complete research study to present its argument and analysis.

Algameli (2016) strongly suggests that open data officials should see open data as an ongoing process rather than a product. This study therefore seeks to understand the gap between the desired outcomes and current state of the open data policy of the City of Cape Town since it was incepted.

The following chapter explores in detail the open data literature. Open data themes and principles relating to objectives and context of this study is used from the literature.

The literature used in this paper examined open data initiatives, principles, arguments and deliberating on how evidence is produced, represented, and legitimized as reliable knowledge that reinforce urban governance decisions (Kitchin,2017).

Chapter 2: The anticipated benefits, promises and realities of open data

2.1 Introduction

This chapter focuses on the literature of open data; it provides, definitions of open data, background, as well as the debates and current state of open data for governments. The chapter introduces the critical terms, themes and principles used in open data literature- these terms, themes and principles will be used ground and guide the study's questions and objects.

Open data has various definitions, described and interpreted by various organisations and authors. Based on the Open Data Policy-Managing Information as Asset define open data refers to *"publicly available data available data structured in a way that enables the data to be fully discoverable and usable by end users"* (Dvir,2018). For the purpose of this study, open data will refer to data produced or published by government or government agencies, data can be freely used, reused and redistributed by any one, as well data that is machine readable by default (Doyle, 2015). This research paper will interchangeable use open data and Open Government Data (OGD). Open Government Data is a subset of open data, which is government-related data that is made accessible to the public (Kucera et al., 2013; Attard et al., 2015). The meaning of freely used means that the data is available with no cost for the public. Redistribution means that this data can be use and reused by anyone without any legal or license restrictions. Broadly, OGD is expected to improve the efficiency of public sector, with increased public accountability and innovation (Magalhaes & Roseira, 2017; Halonen, 2012). Lnenicka (2015) also support that the simple main reason for OGD to exist is for government data to be a sharable resource.

To make this data accessible online for public consumption, governments are required to provide online web portals where the public can view and download the government datasets. Lnenicka (2015) notes that the growing open data initiatives across the world have seen many governments launching open data portals to enable the release of datasets in an open and reusable format.

According to Kucera et al., (2013) the opening of opening of datasets is highly anticipated as a crucial enabler of open government.

This opening of data by governments has however been met with challenges in making it a success as anticipated, Coletta et al (2017:5) asserts that "if 'big data', 'smart cities' and 'data-driven cities' are merely useful buzzwords, they nevertheless evidence an expanding chatter of heterogeneous voices who are merging with and reshaping the urban environment."

To better understand what open data is in literature and what it intends to do, a brief understanding of smart cities and data-driven cities literature would be pivotal as it embodies open data and other similar initiatives. The section below gives more detail on the definition of smart cities, what it seeks to address and how it embodies and relates to open data.

2.2 Smart Cities: An antecedent to open data

"Cities which have embraced ICTs as a development strategy, being pioneers in embedding digital infrastructure and systems into their urban fabric and utilizing them for entrepreneurial and regulatory effect, have been variously labelled as 'smart cities', 'wired cities', 'digital cities', cyber cities or 'intelligent cities'" (Kitchin, 2013:1). The implementation of digital technological strategies has clearly helped cities to be touted as better, improved and more intelligent than before without all these initiatives. The idea of smart cities came as a way of transforming existing governance systems in cities; Osborne and Gaebler (1992) in their book Reinventing Government: *How the Entrepreneurial Spirit is Transforming the Public Sector* suggested that city governments change their hierarchal, in house service model of doing business and proposed new methodologies on how governments can achieve it (Goodsell, 1993). According to Courmont (2017:18), "open data and smart cities policies underscore the implementation of the ideology of information liberalism into the urban government". Courmont further highlights that this ideology was theorized by Benjamin Loveluck, a French scholar through a political genealogy of internet. He asserts that information is at the centre of contemporary liberalism and that it must flow freely in order to explain various challenges in the context of cybernetic theories (Courmont, 2017). "*The information liberalism is based on the assumption that data exist, are autonomous and can easily circulate"* (Courmont, 2017:183). This argument informs the foundation that governments produces significant amounts of data that needs to be shared and reused by anyone.

The term ‘smart city’ has been defined in various ways within literature but can largely be categorized into two distinct but related understandings as to what makes a city ‘smart’ (Kitchin, 2013). Kitchin further explores the categorizations, firstly describing that ‘smart city’ refers to the growing degree to which urban places are composed of ‘everyware’ (Greenfield, 2006); that is prevalent and ubiquitous computing and digitally instrumented devices fabricated into the very fabric of urban environments. On the other end, ‘smart city’, is understood to refer more broadly to the development of a knowledge economy within a city-region (Kitchin, 2013; Kourtit et al, 2012). The two juxtapositions are a critical evidence demonstrating the integrated strong relation between computing and knowledge (data); it demonstrates how the two-work hand in hand to achieve the ‘smart city’ narrative. In the understanding of smart city, it would be imperative to understand that it is not a concept existing in silo if it were to succeed. Batty et al (2012:482) also assert that “*smart cities are often pictured as constellations of instruments across many scales that are connected through multiple networks which provide continuous data regarding the movements of people and materials in the terms of the flow of decisions about the physical and social form of the city*”. Taken from this statement, it is therefore that data is one of the cornerstones that could be weave together the requirements and desires of smart cities. Batty et al (2012:482) further argues that cities however can only *be smarter if there are intelligence mechanisms that are able to synthesize and integrate this data to beneficial purpose, methods of improving operational efficiency, equity and sustainability of life in cities*.

It is clearly that the concept of the ‘smart city’ is eclectic and relate to many urban phenomena. Holland (2008) for instance shows that the concept includes many urban life aspects such as sustainable development, urban planning, technologies, and economic development and so on. Thus, also the term ‘smart’ have a wide range of meanings, connected with its various fields of application. This diversity of meanings is why Chourabi (2012) concludes that with all the credentials and potentials that smart city has, research has sparingly discussed the phenomena. This study focuses on the city’s open data policy as well as its digital transparency and the challenges experienced since its implementation.

2.3 The eclectic meanings and applications of open data

In the eclectic fraternity of the smart city concept, data is one of the fundamental components to realise the potential and success of a smart city. The introduction of open data has been well received by many city governments globally while it has been equally faced with challenges in its infancy stages. *"The release of Public Sector Information (PSI) by governmental organizations is getting increasing attention from local, national and international government levels*" (Conradie and Choenni, 2014:11). While many global cities have adopted the open data policy, the City of Cape Town is the first city to adopt and implement the policy in Africa. This motivates the researcher to do a further study to learn more about the local context into which the city has implemented the policy.

Open data unlike other forms of data comes with set principles, practises and processes for it to effectively serve the purpose that the city governments intend. The City of Cape Town Open Data policy (2015), defines open data as "data that can be freely used, shared and built-on by anyone, anywhere, for any purpose". According to Jimmy Oh (2013), the Open Knowledge Foundation (2012), defines *Open Knowledge* to be "content that people are free to use, re-use and distributed without legal, technological or social restrictions." It is clear that 'open data' or open knowledge is not like any other data, drawing from the two definitions; it comes with conditions and regulations to qualify it as 'open'. Such guidance would be imperative in guiding and tracking the open data phenomena for this study. These conditions are universal and standard for any government implementing their open data initiative as guidelines that they need to follow in order to regard their data as open. The study will work with these universal definitions and principles to assess the city's policies, i.e. are they met, and if not – what are the challenges leading to that.

In literature open data is termed in various ways, some being more specific; Gonzalez-Zapata and Heeks (2015) refer to it as Open Government Data (OGD) and to define the term, the authors dissected it into three key entry points to OGD; Government data, Open data and Open government: Government data - approaches emanate from government being the sole largest collector, user, holder and producer of information about citizens, organizations or public service delivery (Gonzalez-Zapata & Heeks, 2013; Heeks, 2006). Challenges with this kind of data is the way in which it can be managed as a resource within the public service.

(Otjacques, Hitzelberger, & Feltz, 2007). In contrast Open Data include approaches that are not sector-bound but are established from the information and communication technology – (ICT) related innovation and diffusion which potentially enables much greater accessibility of data than previously (Braunschweig, Eberius, Thiele, Lehner, 2012). The main challenge is incapacitating barriers to this accessibility through innovations in the manner data is collected, stored, processed and output (Janssen & Zuiderwijk, 2014).

The dissection of the term open government data by Gonzalez-Zapata and Heeks (2015) is well explained and conceptualizes it perceptively, stating requirements and challenges of each component. In these various data components, they highlight challenges such as data management as resource with the public government sector, barriers to accessibility to the data, and empowering of citizens through open government.

Here I adopted this approach and focused on the challenges highlighted in the City of Cape Town's own implementation of the open data policy. Their terms assisted the study in breaking down the various meanings of open data and its challenges relating to the city. In this paper, open data is used interchangeably with Open Government Data (OGD) and in the context of this study open data refers to the latter – data made available by government agencies.

2.4 The anticipated benefits and challenges in open data

Open data initiatives have been identified to come with benefits to change and transform how cities are governed and controlled. Literature has extensively covered these benefits and some of the challenges that face open data initiatives. Conradie and Choenni (2014:11) importantly mentions that "data release by governments is still novel and there is little experience and knowledge thus far about its benefits, costs and barriers". The novelty of open data initiatives signifies the significant gap in literature and in practice locally; it must also be noted that the City of Cape Town open data portal is the first of its kind in Africa. This study will prove significant in laying the foundation for other upcoming similar studies in a local context.

Open data allows government at all levels, local, provincial and national to create, promote and execute information-based policies (Tran and Scholtes, 2016). As OGD aims to promote innovation, efficiency and effectiveness in government services, it helps to create collaboration across and within public agencies and departments. OGD also acts as an innovation and improved service delivery platform, e.g.; law enforcement maps crimes using census demographic and housing data (Tran and Scholtes, 2016).

Keseru and Chan (2015), assert that government openness provides remarkable benefits for our societies through increased state responsiveness, reduced corruption levels, empowering local and disadvantaged citizens or service delivery and service utilization. Gonzalez-Zapata and Heeks (2015) supports; opening of government data is also perceived to be helpful in reducing cost and improving quality of internal process, such as the formation of policy. The study will award an opportunity to assess the claims coming from literature.

The main challenges that have come with the implementation of open data initiatives has been mainly caused by its novelty or rather it being a new concept which makes it difficult to compare to pre-existing similar approaches. Keseru and Chan (2015), assert that recently more academic literature also suggests that evidence on the political and social impact of open data initiatives is extremely rare. The authors also argue that the second edition of the Open Data Barometer (2015), a global platform assessing open data initiatives pointed out that most evidence remains anecdotal and explains output rather outcomes and impact. It would be also be vital for this study to look deep into the outcomes and impacts of the City of Cape Town Open Data initiative.

Another challenge that would be critical to explore in this study is the data personnel, data professionals at the local level. Conradie and Choenni (2014), argues that so far open data studies lack the focus on the challenges experienced at the local level by public sector information. *"There is an absence of understanding on the local government levels on the impact, barriers and opportunities of open data release"* Conradie and Choenni (2014:10). This motivates the research explored by this study; to understand these barriers at a local level. While the study will look on impact and opportunities, its main focus would be on understanding barriers affecting the release of data.

2.5 Key factors in implementation of Open data policy

According to Ubaldi (2013), establishing legal and regulatory framework is critical to enable Open Government Data; however, even when such frameworks exist, ODG initiatives are often undermined by problems of implementation, technical challenges and administrative delays. Technical; insufficient information about the data? How reliable? Is it useable? Access; how do you know data is available? Data published but cannot be found and not in user friendly formats (Ruijer et al., 2017). Ubaldi (2013) further asserts that legal, technical and financial limitations, among others may limit data accessibility and reusability. Conradie and Choenni (2014) concur that while possible barriers and potentials of open data described by other researchers, a profound understanding of the primary processes for these barriers is missing. Addressing these various barriers related to policy, technology, organisation, finances, and legal framework is crucial to establish an ecosystem, and build a sustainable and working model for OGD initiatives that can produce the desired benefits (Ubaldi, 2013). Below the different challenges faced in open data by governments – described by Ubaldi (2013). These challenges would help guide the study in identifying the challenges faced by the City of Cape Town

Various policy challenges may represent critical barriers to the implementation of OGD initiatives. Disclosure policies may, for instance, restrict data transparency and copyrights may result in lack of clarity over ownership of government data (Ubaldi, 2013). Lnenicka (2015) strongly argues that the effective implementation of OGD takes dedicated and sustained policy attention. OGD brings information that government policies are based on to the fingertips of any citizens with access to the internet (Trans and Scholtes, 2016). Policies should enable accessibility and address the lack of procedures and standards on dealing with OGD i.e. dealing with the limitations of tools to make data open as well as validation structures and guidelines.

In its open data policy document, the City of Cape Town highlights some of its key challenges such as data storage and distribution; city's information is often hidden in department archives or difficult to access as well as certain data access policies and processes with government impede public access to information. Dvir (2013) asserts that there are numerous challenges to have an effective OGD program in place. One type of challenges is related with

the storage retrieval of the data; retrieval barriers are chiefly a challenge of the whole open data ecosystem. These barriers are stemming from the lack of proper policy implementations and strategies. As OGD policies aim to raise transparency, they have the potential to allow the public to assess government policies and hold it accountable (Trans and Scholtes, 2016).

Government data are often scattered as every public agency has its own set of data, formats and standards (Ubaldi, 2013). For civilians to use and interact with OGD, it must be free, over the internet in open, structure, and machine-readable formats that are downloadable in bulk (Trans and Scholtes, 2016: 10). An important component of accessibility is knowing the source of what the user is searching for, and in many cases, where to start searching can pose a challenge (Ubaldi, 2013).

Another technical barrier in accessing OGD is the technical capability and skills of users; Dvir (2018) highlights what he terms Human-Computer Interaction (HCI) factors of open data applications that are developed for the public using open data. Enhancements in the design and accessibility of OGD datasets has enabled the use of open data to new and less skilled users; however, people might require assistance in making sense of published OGD (Dvir, 2018). Algameli (2016:97), strongly concurs that 'most people may lose the benefits of OGD due to lack of technical skills'. Conradie and Choenni (2014) concede that, the challenges e.g., technical issues, privacy or law- related with data release require expertise from various branches of knowledge. For the success of open data initiatives by governments, various personnel with various data knowledge need to be involved to ensure open data standards are met; be it technically, organisationally or legally.

Governments recognized open data as a strategy to increase transparency, improved civil engagement, economic welfare as well improved policy-making and public decision-making (Algemili, 2016). As OGD initiatives increase steadily, the growth has emerged with both opportunities and challenges. According to Ubaldi (2013), several economic and financial challenges are hindering the rapid development of OGD initiatives in several countries. The main challenge in this is on how to develop a business case and financing model for collecting, converting and distributing public data in the spirit of the accessibility principle, that OGD should be free (Tran & Scholtes, 2016; Ubaldi, 2013).

Ubaldi (2013) outlines the financial challenges in detail as; collection and provision of data, converting large volumes of data into reusable formats, posting full datasets in open source formats on government websites and designing a new data framework to assess costs and benefits of OGD.

Governments have issues about the cost of opening up OGD, although such cost together with the cost of data production- has not been appropriately assessed so far. Making data open is not always free, especially when government strictness, as there are undoubtedly costs associated with the creation and presentation of open data that needs to be considered and accounted for.

Ubaldi (2013) further iterates the costs in detail; there is a huge commitment and investment in government agencies to acquire new skills, train employees, purchase of technologies, and upgrade of network infrastructure, that need to be accounted for. To add on that there also costs related to human-resource; costs for organising and preparing data to be published online, especially when data is published on a special portal that may require an IT and design team.

Conradie and Choenni (2014) discovered that the approach in which data is collected, stored and used are all important for open data success. However, highlights that majority of the local governments lack structures for leveraging their data release initiatives. Converting open data can have cost implications, specifically when there is a high-level use of proprietary software (Ubaldi, 2013). Converting public data to usable web and linked data formats is often time consuming and thus costly. These additional costs can result to some reservations from government agencies to convert public data.

To implement a successful OGD initiative requires careful development of data framework to minimize cost effectively. Ubaldi (2013:33) *"insists that it is difficult to develop an OGD business case when there is a lack of both cost clarity and a clear measure of the gains from opening up government data. When government provides reusable data, the practical costs of re-use, adaptation and innovation by third parties are significantly reduced".*

Governments are still faced with several key organisational challenges (Ubaldi, 2013); Organisational challenges such as the structure of organisations involved the different roles they fulfil, teams or personnel in the open data policy or program (Caplan et al., 2014).

Due to the complexity and cross-cutting nature of OGD, governments are required to put in place the suitable organisational structures. On the emphasis of the point aforementioned by Caplan et al (2014) about teams and personnel involved in OGD, Ubaldi (2013) suggests that having a ministry or other agencies dedicated in collecting the various datasets from various government departments that will then be made public has been considered as a way to i) maintain joint work to strengthen data integration across various parts of the public sector, ii) help better capacities across government to deal with emerging issues (e.g. transparency, privacy) and iii) ensure that decision makers of data release do so in a rigorous and consistent approach.

Open data's potential for privacy security and civil rights violations has been at the forefront with more focus in the recent past (Tran & Scholtes, 2016). Empowering independent external bodies to demand and publish information procurement, budgets, and expenditures is considered important to ensure data transparency (Ubaldi, 2013). In response to these issues scholars, policy makers and activists have proposed measures to moderate the dangers of open data (Tran & Scholtes, 2016). Ubaldi (2013) concurs that several governments have considered the formation of independent ethics and governance groups to oversee policies and procedures for improving the use of OGD.

According to Ubaldi (2013), the legal arena surrounding data sharing and opening is without a doubt complex. Support for open data promotes transparency as a model of promoting good government, however taken to extremes, transparency can weaken government or upset other values (Tran & Scholtes, 2016). For governments to have consistent and robust legal framework in place is paramount to facilitate government data accessibility and re-use and to secure and to improve secure data sharing between civil users and public authorities for better results, impact an informed better policy making. Ubaldi (2013) concludes that updating policies and rules is crucial to properly address problems related putting government data online.

2.6 The potential of open government data portals

This section discusses the emergence and importance government open data portals.

According to Lnenicka (2015:589), *"the open data portal is a web-based system used to collect existing data from multiple sources that may be different in formats and publish these data*

on user-friendly dashboards that users may view, download and access via Application Programming Interface (API)". ODG initiatives and in particular the development of OGD portals have flourished since the mid-2000s both at central and local governments. Governments across the world make their data available through these platforms, however, disappointingly the use of this data is lagging (Ruijer et al., 2017; Ubaldi, 2013). Nayek (2018) concur that the number of portals is growing quickly; and that the reason behind the growth being that government data are becoming more easily accessible for various uses. Another critical reason that the author presents is that OGD is expected to improve decision making for both government and citizens. By making government data accessible on open data portals, the governments are giving it back to civil society, which indirectly paid for its creation with their taxes in the first place Lnenicka (2015).

According to Nascimento (2018), a various number of models of quality evaluation for open data portals have been proposed over the years. Mainly these models frequently involve manual evaluation processes, which is time consuming and costly to maintain an ongoing evaluation of open data portals. Nascimento (2018) on his paper suggests the use of automated evaluations of open government data portals. Reiche et al (2014), argue that manual evaluation by human is the most important because it is the human who will make use of the repository. The mere publishing of the data on the web is not enough, to truly advance OGD, the data portals need to fulfil certain legal, administrative as well technical requirements (Braunschweig et al., 2012).

Government portals are playing an important role in the accessibility of open data to public, however they are not enough alone to address the challenge of accessibility. Conradie and Choenni (2014) assert that from a user perspective, major barriers are the access to proper datasets and the proper use of these datasets; all these would be well facilitated with functioning government data portals. The authors further argue that datasets are often published in a fragmented manner, and placed on several websites, which in some case are hard to find. Attard et al (2015) critically retort that after engaging with OGD portals, users still did not consider that transparency and access to information have been achieved.

Lnenicka (2015) concludes that the first problem to be identified and solved when working with any data is where to find it. Kučera (2015) claims that open data portals are a solution that should be used to effectively improve discoverability of free available datasets.

2.7 Open Data Barometer Global Report 4th Edition findings

This part of literature review will focus on the recent findings of the Open Data Barometer Global Report 4th Edition (ODB 4th Global Report). The Open Data Barometer is a global measure of how governments are publishing are publishing and using open data for accountability, innovation and social impact, according to their website. According to the recent report, the fourth edition, the ODB covered 115 countries and jurisdictions. This is section is there to give an indication of the current status quo of open data globally by a trusted open data institution. Below are the findings of the report that were found from the various nations studied:

2.7.1 Government data not open

On this 4th edition, the Barometer assessed over 1725 datasets from 15 different sectors from 115 countries. Furthermore, it was discovered that only 7% of the data is completely open, and only one of every two datasets are machine readable with only one in four datasets having an open license.

2.7.2 Incomplete and poor-quality datasets

According to the report, open government data is generally incomplete, out of date, low quality and fragmented. In many instances, open data portals or catalogues are manually updated as the result of informal data management approaches. Procedures, responsibilities and timelines are often unclear among government institutions responsible for this type of work. This results to the overall data management and publication method weak and disposed to multiple errors. The report further highlights that of the 115 governments surveyed, 79 have an open government data portal, though data is frequently published on sources other than the official open data portal.

2.7.3 Political will: The making or breaking of open data initiatives

Political momentum is vital to introducing and scaling up open data. Political will needs to be translated into legal and policy foundations. Open data initiatives can also struggle in

situations where leaders fail to develop greater reforms that encourage a culture of openness. Countries such as Canada, Japan, Korea and Mexico have significantly shown their political will as they steadily progressed in their Barometer rankings, contrary to countries such as Costa Rica, Ecuador and Rwanda who have now lack further government engagement compared to their initial efforts on open data initiatives.

2.7.4 Publishing of data that people need

The 4th edition report founds that open data portals often do not contain the data people really want and need. The report further suggests that governments must invest in data that people need (e.g., data on budget, spending, contracting, and company registers).

2.7.5 Less promotion of inclusivity and equity

Previously research evidence has shown that open data is contributing to economic growth and creation of new businesses, however little or no evidence is there for its impact on social inclusion. The report found that populations with lower income or less political power are frequently excluded from consultation and decision-making processes concerning open data and generally lack internet connectivity and skills to access open data.

From this literature, the study adopts various meanings and principles of open data to understand how they are implemented on the City of Cape Town's open data policy. Various authors have identified the desires and challenges of open data. The city's early implementation of the policy makes it a great case to study to understand their policy desires and gaps in meeting them. This would therefore allow the study to understand the current challenges experienced by the city on a local context as identified in other studies from the literature.

Chapter 3: The City of Cape Town's open data initiative – A study of policy implementations and digital transparency

3.1 Introduction

The City of Cape Town adopted and enacted an open data policy which has been running for the past five years. The policy was mainly adopted to increase transparency and improved accountability within the city through public data release as data was hidden in departments and difficult for the general public to access. The policy had several desired outcomes or potential benefits upon implementation such as making data accessible for free, enhanced transparency, and entrepreneurial innovation.

Upon the implementation of the policy, an open data portal was developed as a platform where all the various datasets would be shared. Since the initial development of the portal, there has been upgrades made on it in terms of the user experience and the increase of datasets shared.

While the policy has been seen as a step in the right direction with the number of users increasing on the portal. Fewer or no studies have been done to review the state of the policy and its implementation as well as the use of the open data portal.

This study therefore seeks to engage the city officials and other various stakeholders within the open data sector to find their perception on their experiences and challenges in fully implementing this policy as anticipated.

To achieve this, the study would do oral interviews to get profound experiences and challenges within and outside the city – the inclusion of external stakeholders will give a complete understanding of open data challenges in Cape Town as the city does not exist in isolation.

This would allow the study to understand the gaps between the desired results and the current open data policy challenges facing the city.

3.2 Research Methodology

According to Kumar (2011:42), "*anything that becomes a means of collecting information for your study is called a 'research tool' or a 'research instrument', for example observations forms, interview schedules, questionnaires and interview guides".* It would therefore that this study adopted two various 'research tools' for collecting information for the study. The research's main aim was to discover and understand the meanings articulated by the participants. The research aims focused on the current state of open data policy and digital transparency and its implementation in the city of Cape Town. The study was steered by the following research questions:

1. What are the main barriers and limitations in implementing Open Government Data policy efficiently in the City of Cape Town?
2. Who is using Open Government Data? What motivates their engagement with ODG?
3. How is Open Government Data from the city's portal being used in practise? What technical processes are being adopted?

These questions were guided by the following open data assessment tools proposed by Caplan et al. (2014); Context/environment, Data, Use and Impact. Interview questions were guided by these four components; relevant questions were asked under each component to get understanding. Caplan et al (2014) suggests that with the growth of open data, multiple various means have emerged to measure various features of open data implementation, readiness, outcomes and impacts.

A qualitative research tool was used as the methodology to carry out the research. This method was chosen because it unearths situations through profound human interactions. The empirical perusal of the open data portal was chosen as it would afford the researcher a personal experience of how the portal works.

The next section paragraphs provide details of the justification of using qualitative methods – interviews in particular as well as the justification for desktop assessment of the open data portal.

According to Kumar (2011), the core focus of qualitative research is to explain, explore, understand, discover and clarify situations, feelings, perceptions, attitudes, values, beliefs

and experiences of a group of people. The multi-layered character of qualitative research allows researchers to develop a complete picture of the topic in question (Denzin and Lincoln 2005). Based on Hancock (2009) qualitative research is focused on constructing explanations of social phenomena. It helps to understand the social aspects of the world and pursues to answer questions about:

- Why people behave the way they do
- How opinions and attributes are formed
- How people are affected by events that revolve around them
- How and why cultures and practises have developed the way they have

According to Kumar (20110:137), "interviewing is commonly used method of collecting information from people". There are several definitions of interviews. Based on Monette et al. (1986), an interview consists of an interviewer reading questions to respondents and recording their answers. Hancock (2009) describes the interviewing process, interviewing can, on one end, be structured, with questions organised and presented to each respondent on the same way using a strict fixed order. On the other hand, interviews can be entirely unstructured, like a free-flowing conversation.

In this stuy a semi-structured approach was used. Kumar (2011) asserts that during interviews, a researcher has the freedom to choose the format and content of questions to be asked as well as choosing the wording of questions, the manner to ask the questions as well as the order in which they are to be asked. This study used a semi-structured approach. Hancock (2009), asserts that preparing for semi-structured interviews involves setting up a "topic guide" which is a list of themes the study wishes to discuss. This is not a fixed set of questions and should not limit the interview, be able to give room to be flexible while allowing follow up points of interests for both interviewer and interviewee. The guide also allows the interview to have written prompts to ensure that necessary grounds or themes are covered. While Kumar (2011) argue that on the other extreme, structured interviews offers the researchers predetermined set of questions, with the same words and order of questions as defined in the interview schedule. The one main advantage that Kumar (2011) highlights about structured interviews is that it requires fewer interviewing skills compared to unstructured interviewing.

In all interviews for this study, a "topic guide" was prepared with four key themes that guided the research; context, data, use, and impact. Each interview followed these four themes with varying questions. Questions depended on the interviewee organisation and seniority or experience.

The following paragraph provides a detailed justification for the desktop empirical assessment of the open data.

According to Lnenicka (2015), an open data portal is one of the resolutions that need to be used to meaningfully advance the discoverability of free available datasets. However, emerging number of issues prohibits public use of open data such as poor explanation of meaning of data, and the lack of knowledge to make sense of the data. Based on Braunschweig et al (2012), the abstraction of valuable datasets coming from government open data sources require quality evaluation as data quality plays an important role in the use of open data portals.

This study for these reasons did an empirical desktop assessment of the city's open data portal, following guidelines of previous open data portal assessments from the literature. Doing the assessments by the researcher was an added value to practically assess the open data portal as part of the City of Cape Town's digital transparency strategy. This afforded the research study an opportunity to measure the city's portal based on used and tested methods.

3.3 Sampling and data collection

This section give details on how the research respondents and data was collected. According to Kumar (2011:182), "sampling is the process of selecting a few (a sample) from a bigger group (the sampling population to become the basis for estimating or predicting the prevalence of an unknown piece of information, situation or outcome recording the bigger group". Findings obtained from the respondents (sample) are called sample statistics (Kumar, 2011). In this research, the sample was chosen based on industry relevancy and the information needed. The study also sought factored the balance of opinions when sampling. Respondents' selection was diversified to have a broader information base. Respondents from the study was composed of the following individuals:

- The City of Cape Town open data officials; two respondents,
- The City of Cape Town Open Data Steering Committee member,
- A local open data NGO, two respondents
- An open data expert

Sampling for the research was saturated to this number of respondents. Various reasons resulted to this; the open data subject is relatively new – meaning that there would be fewer experts on the subject, the city of Cape Town open data policy is only five years and not much information has been gathered to research it intensely, and lastly, there are also fewer organisations working on open data around Cape Town. Kumar (2011), supports this succinctly, in quantitative research a sample is guided by the desire of the researcher to choose a random sample, while in qualitative research, the researcher is guided by their judgement as to who is likely to provide

Data was collected using interviews. To collect the data from respondents, responses were recorded using a voice recorder as well taking down important notes. According to Mohajan (2018), the researcher is the principal tool for data collection and analysis. Kumar (2011) suggests that during data collection and analysis phases, the key issues identified should become the foundation of formulating chapters. Notes taken during interviews were issues concerned important, or recurring from other respondents, this guided the data analysis and follow up those issues in subsequent interviews.

Other data collection methods such as questionnaires were not explored. One of the reasons were because the sample size was not big enough, questionnaires often have identical questions for every respondent in the sample – no flexibility to change questions or to get more views and feelings on the questions asked.

3.4 Data analysis

According to (Hancock, 2009 & Kumar, 2011) data collection and data analysis progress simultaneously. Hancock (2009), asserts that in theory data analysis should be done at the same time as data collection to enable the researcher to polish the research question and data collection methods in the event of new findings, however in reality, this is difficult to achieve. This is because transcribing recorded interviews consumes time, with analysis even

taking more time. Based on Kumar (2011), the researcher needs to define how they plan to analyse interviews to draw meanings from what interviews have said about certain issues.

In this research data analysis was derived from the common or recurring themes derived from the interviews, while guided the main the four themes of the research. Interviews were recorded and transcribed. From the transcriptions, key themes were drawn to be included as part of analysis. Strong themes recurring from various respondents were asked in detail in subsequent interviews to get more details on the issue. Following this method insured that the researcher obtains enough information for analysis for a certain issue. The emerging issues were then slotted under the main four research themes; context, data, use and impact. The issues were analysed and discussed describing the challenges under each theme.

The study's objectives were to understand both open data policy and how it is implemented in practice – digital transparency and thereby identifying and assessing the challenges and limitations to this. Information retrieved from city officials and other stakeholders informed the study of the current challenges and limitations of open data in the city. The reasoning for coupling these various methods was seen befitting for the objectives of this study to understand comprehensively the city's operations at policy phase – by city open data officials as well at the delivery phase – whereby data is delivered to the public.

In the following chapter, results of the research are presented using the abovementioned methodologies.

Chapter 4: Open data – When policy meets practice

4.1 Introduction

The success of any open data initiative implemented by a government would have to stem from a good open data policy, coupled with an open data portal to make the data accessible to the general public. Exploring the City of Cape Town's open data policy and its digital transparency strategy i.e. practice. This study explores the current gaps and challenges between the desired outcomes and current policy implementation strategies and the challenges experienced since the policy was adopted.

This chapter explores and outlines how the City of Cape Town's open data policy is currently translated into actionable practices, the study takes two approaches; Firstly, the study explores the policy's implementation context – operations and the challenges experienced by the city open data project managers to implement the policy effectively as desired. To achieve this, interviews were done to understand the current challenges and gaps; four key open data assessment themes were adopted from literature were used in this study to guide the interview questions asked to respondents, the four themes are; context, data, use, and impact. Having these helped to understand the challenges at various levels and contexts. Findings of these are discussed in detail and analysed together with the challenges associated with each of the four themes.

Secondly, it explores the city's open data digital transparency strategy through the explorative desktop assessment of the city's online open data portal. This is important to understand how the city makes data accessible, what kind of datasets are made available as well understanding the qualities of the datasets among other factors important when publishing open government data. Results of these factors are presented and analysed together with the challenges experienced. This assessment will bridge the understanding of what is being done practically as oppose to only focusing on policy.

Applying these two approaches will help the study to understand the current operations and its desired outcomes i.e. understanding policy desired outcomes and the city's current open data operations. This will therefore afford the study to completely understand the challenges

both in policy and in practice. The sections below discusses the interview findings and challenges.

4.2 The technical operations and policy tensions

The City of Cape Town's open data policy document have several specifications about technical requirements needed for data to be release efficiently and effectively. Technical specifications such as data being open by default i.e. data released in line departments without any bureaucracy to the portal. This section discusses the challenges whereby policy guidelines are not met by current technical operations.

In this section, the study presents and analyses its interview findings.

Four key themes were identified from literature; context/environment, data, use, and impact which guided the objectives and questions of the study.

Focus groups were interviewed based on these themes. Common and opposing opinions were identified from the various focus groups and synthesized. The process was as follow:

Each of these four themes have subcomponents that guided the questions in more details. These themes perfectly fitted the objectives of the study- to research and understand the current policy challenges. *Context* helped in understanding the broader environment of open data internally and externally of the City of Cape Town. While *data* examines the technical and legal requirements of how data is released and current challenges and gaps around that. *Use* clarified issues such as who are the users, and what uses are they putting open data to. While *impact* assessed how open data is benefiting the intended public and to question if it really does. Each of these themes extracting from the interviews from respondents are discussed in detail as well as the associated challenges in each.

4.2.1 Context – The internal and external organisational relations

The theme *context* sought to understand the current status of Open Government Data within the City of Cape Town, this included understanding other stakeholders interacting with the city within the open data environment. The theme *context* is further discussed with a focus on four subthemes; organisational, social, legal, and regulatory as well as political or

leadership This will help the study fully understand the open data *context* completely in various levels.

i) Organisational structures and operations

The City of Cape Town's open data initiative located in the city's Corporate Services Directorate and administered under the department of Information and Knowledge Management. The Director of Information and Technology oversees all data related matters, including engaging with internal and external stakeholders, and leading project administration among other functions. There is an Open Data manager, who oversees the operational and technical aspects of open data at the city, including managing systems and processes. In support of these functions, there city also has an Open Data Steering Committee, the committee is responsible for considering and assessing appropriate content categories to be published on the open data portal. The committee is composed of two representatives from the public who have an interest or are involved in open data initiatives.

The city also engages with various stakeholders within the open data environment. To engage the public and other relevant stakeholders, the city arranges conferences, workshops, and hackathons. According to the city Progress Report, dated 24 August 2018, the city arranged a Water Hackathon in April 2018 to engage entrepreneurs to develop water consumption solutions, the city also welcomed students from the University of Manchester where the progress of Open Data Policy was presented during their visit. The city also takes part in open data summits and conferences, example the IEEE Open Data Summit held on 5th February 2018 which was attended by various stakeholders from the open data ecosystem.

In the inception of the open data project, various stakeholders co-existed in the form NGOs as well as key individuals, who engaged with the city to see the project implemented. These organisations are a sounding board to the city's implementation of open data. A key organisation that is prominent within the open data ecosystem and working closely with the city is an NGO called Openup- are advocacies of open data, working on various government datasets published for public consumption freely.

Generally, the open system ecosystem is still growing in Cape Town and South Africa in generally. Apart from the engagements that the city has been involved in, the city officials did allude that the uptake of open data is gradually developing, both internally and externally,

which has been a challenge in taking the initiative to the next level. Internally, the main challenge for open data officials is persuading other departments to release their data, making them understand the value and impact of releasing data; sighting risk and legal consequences. City officials, especially from infrastructure departments are afraid to release data, as it may expose city's infrastructure to vandalism and theft. Equally there are often laws that prohibit data release such as the Protection of Personal Information Act (POPI), the city must always ensure that datasets released cannot lead to identification of person.

Similarly, externally, the city faces external organisational challenges, the officials highlighted that, at the moment there is no existing open data community with a single voice that would make data requests known in a collective, pointing out that currently data requests are single case based. To solve this, the city officials alluded that more needs to be done to strengthen the ecosystem by both parties i.e. the city and external stakeholders (public).

What seem to transpire is that the few NGOs working on open data tend to focus on their own project deliverables with a less focus on strengthening what the city is already doing. Officials from Openup did suggest that, while they do work with government, it is often difficult to integrate into what they have already started, stating that "the challenge is that when government implements policy, there is a clear start and end, which makes it difficult to introduce something difficult" (Interviewee 2, 3 May 2019). While open data NGOs consume the city data, the working relationship between the two sectors is yet to blossom to a productive one, whereby the sharing of data, skills and resources are interoperable.

Post the implementation of the open data policy, the then city mayor appointed an Open Data Steering Committee. The open data steering committee is composed of two external members coming from relevant sector bodies, appointed to serve for three years (CoCT Open data policy, 2015). The committee is chaired by a Mayoral Committee member responsible for the corporate services including representatives from other directorates. These two committee members are not remunerated by the city for the activities they do for the steering committee, stipulates the policy documents.

The Open Data Steering Committee meets at least once in two months, confirmed one member. The committee member on the interview confirmed the duties that they do for the committee; one of the main duties is ensuring that data is open and released. In the meetings, the committee assesses all the datasets that are requested on the portal, and the reasons

given by departments to release or not to release them; i.e. users request datasets and those requests are subjected to approval. The committee would then asses the legitimacy of the reasons given not to release datasets; the committee member made an example, that departments cannot withhold data because it was not clean, asserting that this is not a legitimate reason not to release data, and thereby the committee makes a ruling for the data to be released. However, the committee highlighted that one of the suggestions that they made to the city officials was to develop criteria for departments to decide when to and when not to release data, this was to make it easy and efficient for every department across board to act on same practices and principles. Prior to this, the committee had to review many datasets refused for release due to intangible reasons.

The other main challenge is that while the city produces a lot of data, departments often do not have human capacity and time to clean and prepare data. Often data preparation and dissemination from departments must be done by officers that already have their main daily duties while data preparation is an ad hoc task. The committee member strongly suggested that the city should invest more on data personnel to ensure the continuous release and publication of data. This would mean that the city has dedicated data specialists focusing strictly on open data, and possibly every department involved in data sharing should have personnel responsible for data preparation and dissemination, this would ensure that data is timely prepared and released upon request, this would ensure that there is accountability and transparency within the city departments.

ii) Social – Interaction with user groups

The social subcomponent seeks to cover questions relating to the social aspects of the city's open data initiative such as; to what extent is open data being introduced and implemented? With the focus being civil society groups, small business and individual citizens – to question their capacity and freedom to the city's open data.

Through the city's quest to make a social impact and to reach to as many people as possible, city officials have highlighted some challenges in terms collaborating with civilians as a unit. From the perspective of civilians currently the city deals with isolated cases of data requests from civilians. The officials suggested that this makes it difficult to know what people want collectively, and for them to engage with the city as a unit. Similarly, civil society groups are playing their role for their own organisations to open data. This somehow thus leaves the city

as mere data custodian with no strong working relationships with the broader community. The city has also conceded that forming meaningful partnerships has been a challenge; example being that once the city releases data on the portal not much is done as a follow up to engage with civilians and civil society groups, other than the few hackathons and workshops they have hosted so far. A similar challenge has been experienced in reaching out to small enterprises in terms of them developing solutions using the city's open data. In response to this challenge the city officials alluded that the city needs a "counterpart" that has keen interest in open data to work together in partnership.

It is certainly apparent that the city does need to identify their user base and seek to unify it with a single voice, without compromising diversity.

iii) Legal and regulatory systems

Open data universally is governed by laws and regulations and the City of Cape Town is not unique. It should be determined what legal frameworks are in place affecting access to data and its usage.

For data to be released within legal constraints, certain laws and regulations must be considered, which often not an easy process as different legislations may vary for various datasets. The city has a legal department responsible for assessing and ensuring relevant laws are followed and implemented when realising datasets. However, this is a delicate process as the city can be found releasing information they should not. One overriding law is the upcoming implementation of Protection of Personal Information (POPI) Act which stringently protects personal information. With any data that the city release, it should ensure that there are traces that lead to the identification of a person, an example to this would be release of residential addresses and thereby exposing the property owner's details. Certain laws and regulations are hindrance to data release, emanating from department policies and some from the general national constitution, the city officials affirmed, however mentioned that various ways should be discovered to counteract these laws and regulations to ensure a sustainable supply of open data.

A major limitation to release data within the city departments highlighted by city officials; was that data is not open by default, the sharing and movement of data from internal departments is filled with bottlenecks that make it difficult to share data internally and

externally. One interviewee suggested that for the city to have a well-functioning and sustainable data sharing system, there should be an equal policy focus on sharing data externally (portal) and internally (within departments) asserting that a well operation internal data sharing system will make it easier to also share data externally via the open data portal. The participant further asserted that the open data portal is not a benefit for civilians but that also city officials in their departments can greatly benefit from it. The city officials concurred to this as something they are already working on; the city is setting a system that will enable departments to share internal data; be it infrastructure or policies, the city officials confirmed.

iv) Political will/Leadership

Political will or leadership support is key in the implementation of open data policy, especially if it is to succeed. This section explores the political and leadership role in the implementation of open data policy in the City of Cape Town. The open data was presented and proposed by the then city Executive Mayor; political representation was involved from the beginning when the mayor proposed it to the council. A Mayoral Committee (MAYCO) on behalf of the political representation worked closely with the Mayor, the Corporate Services, and the MAYCO.

Up to this far, the political and leadership has been well received even though there are other challenges. A non-governmental respondent, alluded that for there has been a political will for the open data to be where it is now, asserting that it would not be where it is if there was no political will. City officials, similarly confirmed the support that they had received from the political leadership of the city since the inception of the initiative. However, city officials concede that there was a minor challenge at the beginning, their main challenge was keeping the balance between realistic or practical deliverables while appeasing the political constituency into what open data can deliver. This somehow led to exaggerated promises and the potential of open data, to which resulted in disappointments to certain quarters.

The section below addresses findings from interviews about the various components of data from the city's portal and policy.

4.2.2 Data – The accessibility and possibilities of data in practice

This section covers the components of data and how it is made available- this covers aspects such as technical, practical, technical, legal and social – how is data practically accessed and the challenges associated with that.

The city has an open data portal where data is published for public access. Data is aggregated from the various department or sectors of the city, such as transport or social development. The open data portal is administered by the Geographical Information Systems and Department of Information (GIS&DI).

The datasets are made available in different data formats such as; KML, Shapefile, and CSV files- with structures tables and pivot tables. APIs are also made available as part of the portals sharing methods.

Data provided on the portal is reusable without any license restrictions according to the city policy documents. However, city officials alluded that data was not yet open by default – this means that currently data is still requested through applications from the departments owning it.

Below is a graph displaying the distribution of datasets by themes or sectors available in the city's open data portal;

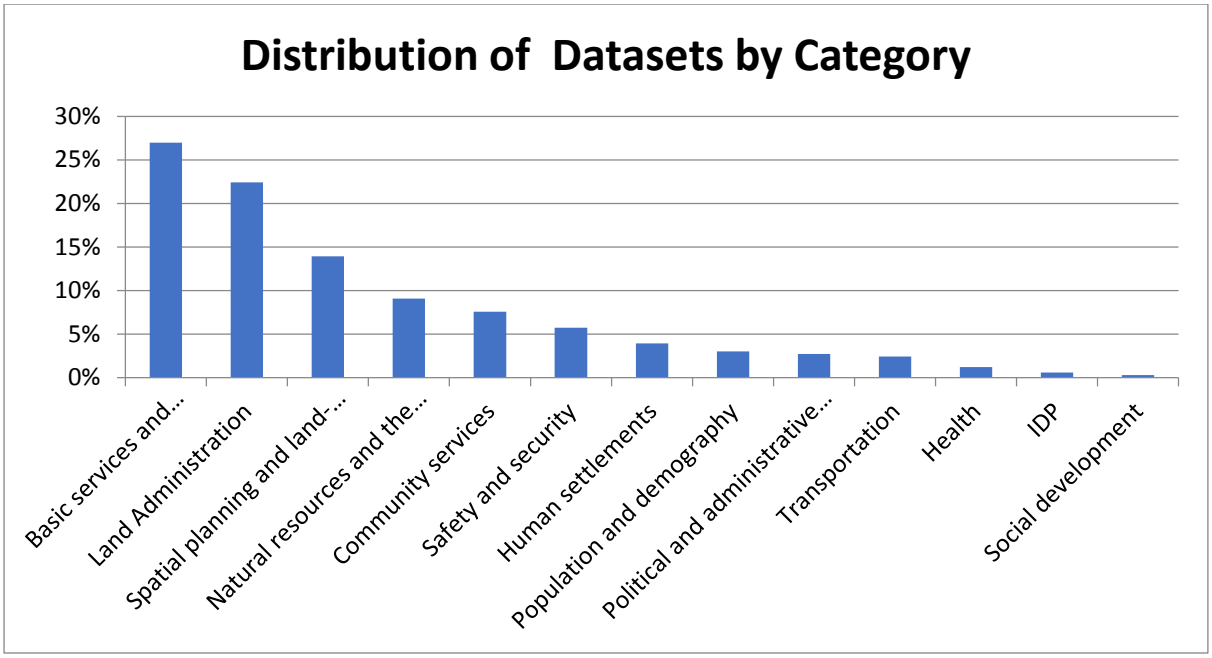

Figure 1: Open data portal distribution of datasets by theme (City of Cape Town, 2018)

According to the city, there are over 119 datasets housed in the open data portal with still more datasets to be integrated into the portal.

No.	Dataset Name	Download Frequency
1	Tender awards	2067
2	Resorts	1051
3	Street addresses	1023
4	Budgets	904
5	Air quality	887
6	Town survey marks	879
7	Clinics	819
8	Land parcels	806
9	Cape Town Industrial Survey	779
10	Cemeteries	619
11	Parking zoning	559
12	Digital elevation model	538
13	Integrated rapid transit (IRT) system bus routes	506
14	Official planning suburbs	469
15	Fire stations	467
16	Building footprints	455
17	Aerial photography	445
18	Community parks	444
19	Electricity service regions	440
20	Subcouncils	450

Table 1: Top 20 popular datasets as at 03 March 2017 (City of Cape Town, 2017)

The table above showing datasets are frequently downloaded. According to city officials, data download frequencies are also influenced by stimuli happening at the time, for example; when there was drought in Cape Town, users downloaded frequently the “City of Cape Town

Water Consumption Data". Similarly, electricity data was downloaded frequently during load shedding.

According to the City of Cape Town Progress Report released on the 24th August 2018, the city's Open Data Steering Committee approve the release of additional 20 datasets on their meetings of 27 March 2018 and 13 June 2018 to be published on the portal. Further 53 more datasets were requested via the open data portal's "suggest a dataset" function.

According to the city officials, the city has an ongoing challenge in terms of releasing data within departments, as also stated in the policy documents; data is often hidden from view in departments' archives, this is often due to siloed systems devised by departments for their own internal operations without any considerations of data sharing. The city open data officials highlighted that they have to continually teach departments about the importance of sharing the data publicly. The apparent reason that departments suggest hinders their ability to share data, is the concern of risk; there is a fear that releasing certain datasets publicly such as infrastructure data will expose it to vandalism and theft. Because of this, data is not open by default, one respondent asserted this is the case currently because data within departments is released with permission bloated with red tape. City officials also acceded that data should be open by default and as it stands it is not as far as their departments are concerned. The city officials however suggested that more engagements and workshops with departments will help in finding mitigating some of the concerns. Once the various city departments are well informed about the risks and benefits of releasing data, publishing of data to the portal will be easy and thereby opening data by default.

The city officials have also highlighted that one of their challenges is the balancing of supply and demand, especially in the early stages of the initiatives. Knowing which data users would need was a challenge, however the city officials mentioned that datasets had to be released incrementally to understand the public demand. The use of the data suggestions function has helped in informing city officials which datasets users are interested in. An interviewee from an open data NGO reaffirmed that there was a demand of open data from the public, however suggested that such demand needed to be coupled with the skills to support it. Often the challenge is that data users are not equipped to interpret open data. This is also a concern that city officials highlighted, which somehow contributes to the city not publishing some information that may be misinterpreted and used against the city. A big factor to this could

also be the newness of the concept of open data, especially from the user perspective. The NGO interviewed mentioned that they have an Open Data Academy whereby they teach open data users such as journalists how to analyse and interpret open data into their reporting.

4.5.3 Use – Who uses open data and for what purpose

This section covers the details on who uses the city's open data and the various use cases that are applied using the data.

i) The end users

Open data should be used by anyone for any purpose for free. On asking city officials who uses open data, the officials indicated that, while they have a clue who uses their data through the requests they receive, it was difficult to know the groups using it and their purposes. However, they indicated that some of the people requesting data are from these industries; architecture and planning, academics, researchers, students, and other NGOs. City officials highlighted in most cases requests and downloads are coming from spatial data professionals; this could be attributed to the richness of spatial data the portal has. To track users, a manager from an organisation called DataFirst alluded that there are software systems that are used to track portal users that they use on their own portal. While this could be useful in terms tracking open data users, this would be somehow added task and resources for the city, this would however another respondent strongly suggested that this would correctly inform the city on which data to release.

Respondents from an open data NGO mentioned that open data is used by various people and when releasing data, they are not targeting anyone, their data is requested by general civilians, researchers, students, and journalists amongst other users. In as far as users are concerned, skills to use and interpret the data was one of the major challenges, alluded the respondent. Further affirmed that data usage and provision depend on the "persona" (user), and the data type provided; raw data versus synthesized; data versus information; a CSV versus graph (derived from the CSV) – "it is about knowing your audience" (Interview 2, 3 May 2019).

Another respondent strongly suggested that "access creates demands" (Interview 4, 4 June 2019). The responded was reacting to the question of, to whom should data be made

available; that government should make all kinds of datasets available for anyone for any use. As an NGO, their purpose was to collect various datasets including OGD and make it open and accessible to anyone in demystified and easy way than it would be published by a government agency. One example they made was making national budget data easily understood by people not versed with finance and budget knowledge.

Below is a table displaying the distribution of users from other countries other than South Africa;

No.	Country	Portal Views
1	United States	1306
2	Spain	999
3	*Europe	509
4	France	382
5	United Kingdom	335
6	Switzerland	296
7	Netherlands	222
8	Canada	210
9	Germany	132
10	Ireland	112

Table 2: Country origin and count of portal views from outside South Africa between 1 July and 31 December 2017 (City of Cape Town, 2017)

While majority of the portal views, 81% are from South Africa, the city of Cape Town portal has also shown a growing interest from outside South Africa, with other countries also interested in its developments, with majority views coming from the United States. City officials mentioned that there is a positive feedback and demand coming outside the country. It would have been useful to have the distribution of South African users.

The city officials highlighted that the growing interest from foreign users was during the Cape Town drought in 2017 with international countries requesting data and searching information on the drought conditions at the time. To respond to the demand the city developed a web applications display water consumption data of the City of Cape Town.

ii) Purpose – Reasons data is requested from the city

Open data published by government is used in various ways to achieve various goals. The city's open data is used in multitude ways ranging from business to academic research data as seen on table above. Based on the table above, users on the portal were mostly downloading tender awards followed by resorts and street addresses. Tender awards are generally downloaded by local businesses doing or seeking doing business with government. For resorts were downloaded due to the high demand of affordable holiday accommodation in Cape Town or by developers looking at expanding their resort portfolio around the city. Many businesses depend on street addresses for various business outputs, such as delivery services or spatial planning – for this the city of Cape Town mapped a comprehensive GIS data downloadable on a shapefile and CSV that users can consume. According to the city reporting documents, for the period 1 July 2017 to 31 December 2017, the portal had 106 datasets published for download. From this, the most downloaded datasets on the portal are "Tenders awarded with a value of more than R200 000" followed by "City of Cape Town Water Consumption Data".

From this period, 1 July 2017 to 31 December 2017, a further 32 data requests were received from the portal's "Suggest a dataset" function. Below are the reasons that users provided for requesting datasets;

- Property profiles
- Writing articles
- Property development analysis
- Data visualization
- Research and analysis purposes
- Planning
- Water consumption awareness (City of Cape Town, 2017)

Once the suggestions have been listed on the portal, they are then processed, within the current capacity, and the relevant departments. The suggestions are assessed together with the steering committee which involves several steps including, availability, privacy or other issues related to public data release. When data is received, the status of all the suggestions and processing times can be viewed on the open data portal under the 'suggested dataset' list (City of Cape Town, 2017). Further, the period 1 January 2018 and 30 June 2018, the portal

had 119 datasets available for downloading. In this period, a further 53 data requests received on portal. For this reporting period, below are the reasons users gave for requesting datasets;

- Impact of commuter travel
- 3D modelling
- Groundwater monitoring
- Public engagement/participation
- Verification of land ownership
- For planning research and analysis purposes
- Surveying

In this reporting period, of the 119 datasets available for downloads, the most downloaded datasets in the portal were the "City of Cape Town Water Consumption Data" followed by "Tenders awarded with a value of more than R200 000". It should be noted that between the two reporting periods the 2 most downloaded datasets remained the same, with only changing rankings; "Tenders awarded with a value of more than R200 000" ranking high in the period between 1 July 2017 and 31 December 2017 while during the period between 1 January 2018 and 30 June 2018, the dataset of "City of Cape Town Water Consumption Data" was downloaded the most.

iii) Activities – Practical applications of open data

Individuals and organisations putting open data to use for various reasons in different sectors. Evidence based usage i.e. products and solutions built on open data have been limited, confirmed a city official. Currently the city depends on the "Suggest a dataset" function where users provide reasons for requesting datasets. The city officials did however mention that from time to time, they do search what individuals have been doing with the open data through internet searches. Among other activities that city officials mentioned that individuals created was, web applications for water consumption monitoring – a map application that a user developed to display various locations and their water consumption around the city. Another user used the city's fire incident data to develop a tool that predicts where the next fire outbreak may occur, especially in the dense low-income areas of the city.

Reason for requesting data by users	
1 July 2017 - 31 December 2017	**1 January 2018 - 30 June 2018**
Property profiles	Impact of commuter travel
Writing articles	3D Modelling
Data visualization	Groundwater monitoring
Property development analysis	Public Engagement/Participation
For research and analysis purposes	Verification of land ownership
Water Awareness	For planning research and analysis purposes
Planning	Surveying

Table 3: Reasons provided for requesting datasets, for both reporting periods – 1 July 2017– 31 December 2017 and 1 January 2018 – 30 June 2018, City of Cape Town (2017)

The table above shows the comparison of reasons for requesting datasets for the two reporting periods. From these reasons provided by users it would be inferred that these are the actual activities that individuals do with the city's open data.

It would be of great interest to have a count of each reason given to track how many individuals are performing a certain activity. This would be of importance for the city in knowing what data are in demand and for what activities. This would also further help the city in understanding the sectors in which they need to provide data literacy, thereby improving data quality and timely provision. From the two reporting periods, the activities or reasons given are different and are assumable the most given reasons on the portal. There is a recurring occurrence on the interest of property information, provided as "Property profiles" and "Verification of land ownership"- this data is on demand confirmed city officials, however it is highly private as it contains information that can lead to a person's identification, and for this reason, the city has had challenges on how they can provide this on demand data while not compromising the personal information of individuals. Another interview also confirmed that there is a great demand of data for land ownership and land value, especially at national level, asserting that there is no existing functional nation land registry and that national government desperately needs especially for rural areas. The interviewee further affirmed that while the city knows who owns what in terms of land, there still need to be an efficient way of providing this information to those needing it as it their legislative right. Property information sourced from the city for free ends being sold by third parties in the form of property reports. Another frequent reason for data requests on both periods was

“Research and analysis purposes” – in this case research is done by various organisations, ranging from NGOs, private sector or educational institutions etc.

The apparent challenge that came across respondents on the “use” of open data was the skills of the users. Users needed to have skills to support the supply of open data. The consensus amongst respondents is that open data is a steadily growing phenomena in Cape Town or South Africa in general. Supply almost have to get to a point where it understands the demand, i.e. the city understanding what individuals need and for what purpose. This would assist the city in fully understanding its user base and its needs, such as the skills needed to interpret the data. The other one main challenge that city officials was the existence of a “user community”, alluding that if one existed, this would make the city’s job much easier as concerns would be raised from a unified voice. With the gradual uptake of open data, with growing access, there would be a growing demand with a growing user base.

One of the other reasons leading to this alluded the committee member was the availability of metadata, whereby users would discover what data is available. The committee member further mentioned that initially it was difficult to suggest data that users would request data without fully knowing what the city have. The respondent further highlighted that it would be ideal whereby data is just made available and people download it for whatever reason, data open by default. This is the issue of supply and demand, how does the two meet each other; in response to that the interviewee, mentioned that constant supply of data will encourage the supply as oppose to the converse; that when the supply is abundantly accessible there would be demand and therefore data must be released timely, freely and openly.

In terms of adopting and implementing the open data policy, the committee member commended the City of Cape of Town and its leadership. She suggested that often governments develop and setup data infrastructure and systems before policy is drafted and enacted. The city after enacting the policy, it was then an open data portal infrastructure was put in place with an incremental publishing of datasets.

The section below addresses the potential impact of the city’s open data initiative and its possible benefits.

4.2.4 Impact – The benefits and success measures

After all is done; initiatives and policies implemented, measurable progress must be in place and studied to understand the efficiencies and effectiveness of those policies and initiatives. OGD literature has strongly condemned the lack of evidence of impact on OGD initiatives.

Similarly, the four-year old Open Data Policy of the City of Cape Town have not yet provided any proven evidence of impact. City officials admitted that currently there has not been any concrete evidence of impact brought by the open data initiative, they affirmed that this could also be because of the recentness of it and that many people are still learning about open data. The city officials strongly suggested that, research would soon be needed to effectively evaluate in the implementation of the policy. Two other interviewees concurred that little focus or research has been done to assess OGD impact both from government and NGOs perspective. Both respondents from open data NGOs, alluded that they were internally initiating strategies to assess the impact their organisations have on open data.

i) Social impact

Government Open Data is often touted as one of the strategies that would bring and inculcate transparency in governance, equal social benefits, improve public services etc. thereby improving the social life of civilians. For now, to assess impact or potential impact be it social, economic or political – it can only be looked at through the lens of the activities that individuals partake on using open data. Activities such as developing solutions on water consumption awareness; such solutions can have a great social impact, but still after development they need to reach many users who would be able to use and understand them.

ii) Political/Governance impact

There are also expectations that OGD would help improve government efficiency and accountability. However, in the case of the City of Cape Town open data, this is the one main challenge they have; efficiency must be achieved internally – departments need to be accountable for their sharing of data together with efficient systems put in place. Open data would have such impact if it completely open i.e. open by default. One respondent importantly asserted that open data would not produce efficient governance on its own, mentioning that civilians and other relevant stakeholders would still need to know how governance works and thereby understanding how can open data help achieve it even better.

iii) Economic/Commercial impact

The city's open data policy emphasises the role played by data in the economy and meaning that data is a crucial feature in competitive economies. It further highlights how entrepreneurs can use large datasets to develop new products that solve civic challenges. Currently this is contrary to what is happening. City officials confirmed that currently there has been a disappointing interest from entrepreneurs using their open data to develop products. However, the officials alluded that often entrepreneurs have great ideas to build products but are limited by financial resources to develop them further, and that the city was not at a position to assist them in that regard. Similarly, respondents from NGOs could not attest to any existing economic or commercial benefits through entrepreneurial exploits. The city officials have however affirmed that this one of the aspects of open data they would like to see thriving in the coming near future.

Generally, there is a lack of evidence of impact for open data initiatives globally, especially at local government level as established from literature. This study through its participants confirmed that currently there have not been documented evidence on the impact that OGD brings or is at least expected to bring. City admitted that there is currently there has not been any impact studies of the open data policy. The officials attributed this to the fact that the policy is still new and suggested that with time they would be able to start accessing the impact of the policy.

For every implemented policy, there should be measurable success guidelines to gauge its progress and effectiveness, thereby enabling measurable impact assessments. However, the political and leadership will of the City of Cape Town should be acknowledged in implementing the policy; both locally and nationally, other upcoming open data organisation; governmental and otherwise are taking great cue from the City of Cape Town, one example is the Open Data Durban whom the city officials confirmed their great working relationship.

The following section covers in detail the desktop assessment of the city's open data portal.

4.3 Desktop assessment of the open data portal: Datasets, formats and quality

As part of the research methods, the examination of the city's open data portal was done. This section seeks to answer the third research covering, the technical data processes being implemented on the portal when making data accessible.

The portal is currently on its beta version – it is one last version before the main version is made live for complete public consumption.

Here datasets available on the portal were assessed from the different categories as set out on the portal.

According to the City of Cape Town Open Data Portal website the beta version has the following functionalities;

- Ability to view spatial data on a map an in tabular format before downloading
- Filter data based on specific attributes before downloading'
- User can mark certain layers as favourites for easy access and related datasets will be listed with all search results.
- Users can also use the API to extract data programmatically
- Automated updating of datasets allowing access to the most recent data.

The assessment of the portal focused on the data representation, availability of metadata and the availability of different formats etc. (Nayek, 2018). To carry that out these are the components to be looked in detail from the portal; Categories, Formats, Metadata for Datasets, and Purpose.

These terms are explained below as described by Nayek (2018);

Categories – an area, topics or sectors of datasets provided by the city

Formats – a file format is a standard approach that information is encoded for storage.

Metadata for dataset – metadata is defined as about data. Metadata is beneficial for understanding and interpreting the contents of datasets.

Purpose – What is the main objective of opening data for the public?

These above terms are adopted and explained from Nayek (2018) as they perfectly fits what the study is examining on the open data portal. The table below displays the categories of the various datasets available on the portal. It also represents the number of datasets available in each category and the type of formats the data is available.

4.3.1 Categories

Categories	Content Type		Total Number of Datasets
	Document Link	Feature Layer	
Basic Services and Infrastructure	11	10	21
Community Services	2	13	15
Demography and Statistics	0	0	0
Economic Development	3	7	10
Finance	9	0	9
Health	3	2	5
Human Settlements	0	0	0
Imagery	0	1	1
Land Administration	3	8	11
Natural Resources & the Environment	0	11	11
Political and Administrative Boundaries	5	6	11
Safety and Security	2	1	3
Social Development	4	0	4
Spatial Planning	3	15	18
Transportation	0	4	4

Table 4: Categories count of datasets available in City portal

On the landing page of the city open data portal beta version datasets are aggregated in categories. Currently there are 15 data categories on the city's portal. Each category encompasses various common datasets e.g. transport related datasets would be under the category "Transportation". Each category was assessed individually to determine the data formats it contained. This is denoted by its content type – which simply describes whether data is provided in document or layer format described as "Document Link" or "Feature Layer" on the portal. This was important to understand the openness of the portal's datasets as described by open data standards and principles – "data must be machine readable"- meaning that data should be easily downloadable and usable without the use of special hardware or software. The total number of datasets available in each category was also recorded. This was important to understand which categories had more datasets available on

the portal to understand the datasets that are in demand and those with less supply. Two categories were not yet populated with any datasets, the “Human Settlements” and “Demographic and Statistics”. The reasons for this are beyond the scope of this study- the possible explanation could be that the portal is still under construction and is not fully live, it is in its beta version, meaning it is one version before the main version to be released for complete public consumption.

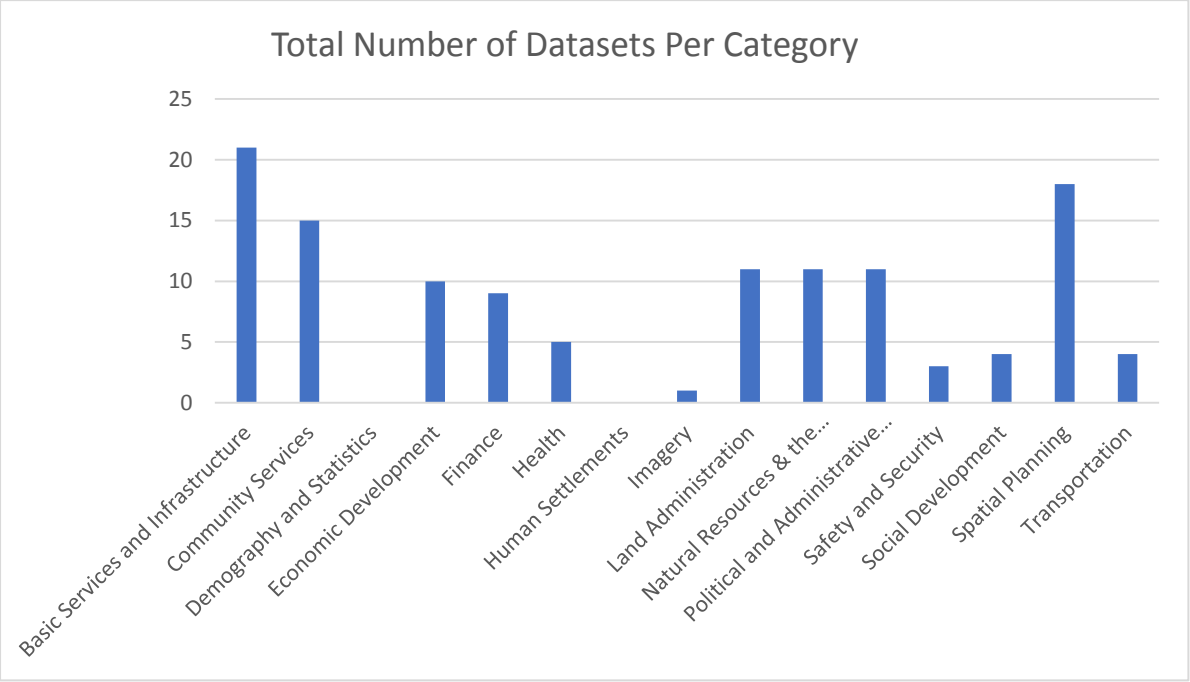

Figure 2: The current number of datasets per category

The graph above depicts the total number of datasets available per category, this allows us to see which categories or sectors contributed less or more data in the beta portal. Basic Services and Infrastructure contributed more data Demography and Statistics, and Human Settlement had no data at all in the portal.

Remarks: The beta open data portal has 15 topics/sectors/categories, these topics had a total of 123 datasets, composed of either feature layers or document links. Feature layer datasets had a map, where you can visualize the data on a map and on a table before downloading it. With this content type it also allowed downloading in my multiple data formats such as spreadsheet, KML, API or in Shapefile, majority datasets were in feature layers 78 and 45 document link. Document link generally are downloadable in Spreadsheet.

Majority of the topics were populated with information except for two categories; Demographics and Statistics and Human Settlement. It was discovered that the categories did not only kept spatial data, but also other non-spatial datasets.

4.3.2 Data formats

Majority formats of datasets found on the portal are in these 3 main formats; excel spreadsheet, KML, Shapefile and APIs. Spatial (point vector) datasets can be downloaded in all four formats, while all non-spatial datasets were only available in spreadsheet.

4.3.3 Metadata for datasets

Metadata	Document Link	Feature Layer
Title	√	√
Description	√	√
Type	√	√
Date Created	√	X
Last Updated	X	√
Abstract	X	√
Purpose	X	√
Credits	X	√
Constraints	X	√

Table 5: Metadata for datasets available in City portal

Metadata for datasets available in City portal

Remarks: As clearly depicted from the table, metadata is not equally provided or presented for the two content types; document link and feature layer. Datasets only provided in document links i.e. spreadsheets only had limited metadata presented on the portal. Generally, information given for this was title, description, type and date created. Critically it did not cover the purpose, credits (which department contributed the data) and the constraints the data had. While on the other hand feature layer content type generally provided enough information about metadata, this came from a separate provided next to the dataset which takes you a separate page with metadata information.

4.3.4 Purpose

As stated on the city's open data portal website, the overall purpose for providing open government data; *"By sharing this data, the city aims to increase transparency in its processes*

and actively involve residents and other stake holders in local government, as well as promote economic opportunity". Generally, each of the datasets in all content type were given their purpose for use.

Other web applications

Other than the open data portal, the City of Cape Town, also publishes web map applications to make some information available for consumption. The city has the following applications displayed on their website; an application to find properties, streets, suburbs, subcouncils and wards in Cape Town, an application for Economic Areas Management Programme (ECAMP), an application for recycling and waste-wise services, and application which provides information on household water use to name just a few. The applications are published through map visualizations for consumption. It must be noted the intention of the applications is there but not enough to convey the information it intends. Maps do not come with dynamic interactions; generally, the map visualizations are not intuitive enough for someone not skilled in using maps.

4.4 Data Portal Assessment Framework

To understand the practical experiences of how the city publishes data in its open data portal, the study extensively analysed the city's portal through a desktop assessment; examining and exploring data quality issues such as data formats, data availability, data metadata etc.

The assessment framework used here was guided by some of the open data principles as adopted from Attard et al., (2015). Below is a table showing analysis of the city's beta version open data portal. Importantly it should be noted that these analyses were based on the current status of the beta version portal.

Each of the portal's dataset categories were analysed based on the questions as shown on the table below and a quality indicator was given under each data category as explained below the table.

This exercise was done to assess how the city policy on publishing data is put into practice, i.e. how policy meets practice.

Category/Sector	Does the data exist?	Is the data in digital form?	Is it available online from in any form?	Is the data machine-readable?	Are data identifiers provided?	Is data timely & up to date basis?	Is the data publicly available?	Is data for free?	Is the data openly licensed?
Basic Services and Infrastructure	🟢	🟢	🟢	🟢	🟢	🟡	🟢	🟢	🟢
Community Services	🟢	🟢	🟢	🟢	🟢	🟢	🟢	🟢	🟢
Demography and Statistics	🔴	🔴	🔴	🔴	🔴	🔴	🔴	🔴	🔴
Economic Development	🟢	🟢	🟢	🟢	🟢	🟢	🟢	🟢	🟢
Finance	🟢	🟢	🟡	🟢	🟢	🟢	🟢	🟢	🟢
Health	🟢	🟢	🟢	🟢	🟢	🟡	🟢	🟢	🟢
Human Settlements	🔴	🔴	🔴	🔴	🔴	🔴	🔴	🔴	🔴
Imagery	🟡	🔴	🟡	🔴	🟢	🔴	🔴	🔴	🔴
Land Administration	🟢	🟢	🟢	🟢	🟢	🟢	🟢	🟢	🟢
Natural Resources & the Environment	🟢	🟢	🟢	🟢	🟢	🟢	🟢	🟢	🟢
Political and Administrative Boundaries	🟢	🟢	🟢	🟢	🟢	🟢	🟢	🟢	🟢
Safety and Security	🟢	🟢	🟢	🟢	🟢	🟡	🟢	🟢	🟢
Social Development	🟢	🟢	🟡	🟢	🟢	🟡	🟢	🟢	🟢
Spatial Planning	🟢	🟢	🟢	🟢	🟢	🟢	🟢	🟢	🟢
Transportation	🟢	🟢	🟢	🟢	🟢	🟢	🟢	🟢	🟢

Table 6: Open Data beta portal quality assessment

The table above displays a quality assessment of the city's open data portal. Questions meeting open data principles were used for assessment using colour coding. Green colour indicated good quality, orange – medium quality, while red indicates poor quality or principle not met. This was done based on the overall assessment of datasets in each category.

Below are more open data standards explained and analyses how the portal faired against each of them.

Data accuracy - By accuracy, it is meant the degree to which data or metadata records describes the respective information (Kucera et al., 2013). Generally, data did not have enough metadata provided e.g. which department contributed or produced the data and with enough details of its intensions.

Data completeness – According to Attard et al., (2015) a dataset is regarded complete only when it contains all the information required to have the needed representation of the described data. Datasets from the city portal had instances where it is not complete, i.e. some records not populated or sparsely populated, however datasets were complete, particularly feature layer datasets.

Data consistency – In a dataset, a field is considered consistent if the respective values are selected from a constant set of options (Attard et, 2015). An example of this the use of "Jan" and "January" interchangeably, this is regarded as inconsistency. From the datasets analysed, a nature of this was not encountered. However, the inconsistencies on dates was found on different datasets, some were written in number formats and some written in word format.

Timeliness – From the datasets analysed most datasets were up to date i.e. not more than two months since they were last updated. Some categories (Basic Service & Infrastructure, Health, Safety & Security and Social Development) datasets were last update a year or two ago. This could be attributed to the fact that these are immovable government infrastructure that do not change often. Timeliness pattern was not immediately recognised, i.e. different datasets had different dates of release and updating.

Data accessibility – As described by the authors (Ochoa & Duval, 2006) accessibility is has two measures; the *cognitive accessibility* – how easy it is for a data user to understand the published datasets. The second measure is the logical accessibility – defined as the ease with which the relevant dataset is discovered through a data portal or catalogue. The datasets analysed were fairly easy to understand, however with the limited metadata, some were not intuitive as to what purpose do they serve. Datasets were easily discoverable on the portal, as it they were aggregated in sectors or categories.

Data openness - According to Kucera et al., 2013) data must be open both technically and legally. The city's open data portal follows most of the eight open data principles, e.g. easily usable data formats, use and re-used with no legal obligations and is accessible through the

portal. While there could still be improvement in meeting completely the principles, the intentions and representation of opening up government data can be seen.

Having review the required principles and frameworks from the open data portal following section discusses the current potential challenges in publishing open data on the city's portal. The challenges discussed are from the researcher's observations while doing the portal desktop assessment.

Chapter 5: Cape Town's open data experiences – Potential value balanced with challenges in practice

5.1 Introduction

The release of datasets by governments can be seen as a conduit to releasing data-driven urbanism, promoting efficient governance, a process with the potential to strengthen relationships between government and its constituencies. However, studies have shown that there are several challenges affecting the release of datasets and open data initiatives by governments (Dvir, 2018). The case of the city of Cape Town's open data policy has shed a light of some of the challenges coming with the implementation of open data policy. Ubaldi (2013) supports that there are several challenges that may be linked with the implementation of OGD initiatives. Disclosure policies for example, may restrict data transparency and copyrights which can in turn lead to lack of clarity over government data ownership, thereby limiting public's right to use government data (Ubaldi, 2013).

Understanding open data at a local level was pertinent for this study as the City of Cape Town is the first African city to adopt an open data policy. Conradie and Choenni (2014) affirm that open data studies lack the focus on the challenges experienced at local level by data professionals. Importantly the authors justify the focus on a municipal level instead of a national level, asserting that data is largely collected at a local level and thus emphasising the need to support and study local governments for the success of open data. This study has afforded current and future research a foundation to understand some of the challenges experienced at a municipal level.

To categorically identify these challenges in detail, four themes that emerged from literature review are used to assess the city's open data implementation; *Context, data, use* and *Impact*. To do this, I focused on the city's policy challenges as well as the open data portal. This was done to understand the gaps between policy and practice i.e. how is data published in the portal as stipulated in policy documents and universal open data standards.

To study what happens in practice, the study embarked on a desktop assessment of the city's open data portal where government datasets are made available. The assessment was a practical review of the city's open data portal, looking at issues such as data quality, availability of metadata, machine-readability and other open data requirements as set out by open data standards as displayed in chapter 4 table 6. This practical assessment was done as an approach to corroborate policy with practise – the city's digital transparency. Open data has to be provided online and certain requirements and principles need to be met to ensure efficient access to the data government releases.

Government open data portals have several technical requirements necessary to ensure satisfactory user experiences when searching and downloading datasets. The study perused the portal using various open data requirements and principles adopted from the literature to assess the portal, issues such as data quality and data formats were assessed and analysed. Digital transparency is concerned with how governments make their services easily accessible on the web; Lnenicka (2015) affirm that enabling the discoverability of data is a prerequisite to reaching the full potential of open data.

This section below further explores the analysis of the four themes based on the study's main findings. Each of the themes are used to analyse the findings.

5.2 Context – A struggle to collaborate across like-minded organisations, institutions and individuals

The theme *context* focused on the overall environment of the open data initiative within and outside the city. It sought to understand the relationships within the city as well as with its constituencies. This theme also had four subthemes which were explored in detail in the findings; organisational, social, legal, and regulatory as well as political or leadership will. *Context* sought to understand the existing open data operations without any assumptions or anticipated challenges. When analysing context, it was critical for the study to look at it from various components to ensure their entire context or environment of the city's open data initiative is well understood; it was critical for the study to understand what is the current organisational context i.e. what are organisation challenges is and how they contribute to the

barriers of opening data. Similarly a legal context had to be understood, i.e. what are legal barriers in data release, answering questions such as data ownership, data privacy etc.

When considering organisational input, Algemili (2016) strongly asserts that open data initiatives need a collective and continuous effort and an open data policy is certainly the foundation, however other problems emerge at both organisational and technical levels. When assessing the organisation context of the city's open data initiative several challenges were identified. The research discovered that, presently the city does not have personnel dedicated with data management within the various departments contributing data. As a result, the study discovered that data was not open by default, meaning that data has to be requested with justifications from the various departments to be released. It was thus discovered that open data advocacy was still lacking interdepartmentally. Data open by default meant that data requested internally or externally had to be released without applications or justifications unless there threatening legal or security issues. At this moment, the study discovered that this was still a challenge with some of the departments, but data officials alluded it was something gradually changing and department were releasing data much better than initially.

Organisational challenges also focused on external relationships that the city has within the open data environment. This looked into the city's partnerships and how they work together in advancing open data in the city. Through the study of the City of Cape Town, it was established that the city had existing relationships with some open data NGOs and other relevant stakeholders. Several studies have emphasized collaboration as a critical component in the success of open data projects. Ubaldi (2013:34) strongly asserts that "*the creation of the right ecosystem – i.e. a community of key actors – is essential not only to reap the economic benefits, but also to generate the value of OGD initiatives in social and political terms".* City officials admitted to the importance of fostering strong relationship with external stakeholders and alluded that they were a key part in the development of the policy. Having strong working relationships between the city and other relevant stakeholders would certainly be a benefit to the growth and development of the city's open data initiative. However, what also emerged from the study was that NGOs sometimes found it difficult to work closely with the government, citing rigid long-term policies and practices – NGOs seek to work with issues on the ground with a room for flexibility. With the rigid working policies,

NGOs suggested that this often led to siloed strategies and development of open data projects. Attard et al., (2015:408) strongly supports that "*governments and public entities are sharing data on the internet at an astonishing pace. Yet, there is a lack of agreed upon standards for data publishing"*. This is evident from the findings of this study, an NGO interviewed as part of this study is collecting government data and other datasets but had their own portal where they publish and present data differently from the city. While this can be good for expanded knowledge and variation, there is also a risk of compromising the principles and objectives of open data. For instance, users do not have to be confused of the expectations and what open data is i.e. freely accessible data that is machine readable and can be used by anyone for anything. Any public data holder would have to uphold and keep to these principles. Having agreed upon standards between governments and public entities will ensure that open data beneficiaries receive data that is of high quality with great value for reuse and distribution.

Lastly, refocusing on the subtheme *social,* under the main theme *context. Social* component focused on the participation and collaborations of the city's open data project with the civilians. This component focused on how the city interacts with the community of users through feedbacks, where the user would have engagements with the city in issues such as requesting and suggesting datasets as well contributing to the governance of the city's open data. Attard et al., (2015) suggest that the challenge to achieving the full potential of open data published in portals is their use, or lack thereof. This means that there has to be an existing broader open data community of users or interested parties interacting with the city in issues relating to open data.

City officials mentioned that, currently there is no existing known structured community of open data users with a unified voice and vision. Attard et al., (2015) affirm that through participation, various stakeholders can participate and contribute in the governance of the open data portal for instance, suggesting what datasets to release, or rating datasets and functioning of the portal itself. Having a formidable community, would ensure and enable that there is participatory governance around the development of the open data policy and that the city is accountable and transparent through the datasets it publishes for public consumers.

5.3 Data, a continuous process not a product – Figuring out technical specifications and data expertise

The publishing of datasets is the cornerstone of a successful open data initiative. Through this theme; *data,* the study focused on the issues of data, such as quality, and relevance. As data issues can be eclectic, the theme *data* focused on the following subthemes; legal, practical, technical and social openness of data. Ubaldi (2013) confirm that legal, technical and financial difficulties among others, may limit data accessibility and reusability.

The primary key for the city's open data policy to be successfully implemented is the publishing of datasets to the portal. To ensure steady growth and process implementations, the city incrementally increased datasets from its portal. This effort aimed to gradually understand the demand of open datasets as the public gets to know more about the service. Today the city releases hundreds of various datasets ranging from transport, health, statistics and demography etc. In order to minimize risk be it legal or security, the city had to ensure that data undergoes various data verification processes; this included ensuring that the city does not publish data that it does own or data that is at risk of revealing private personal information, other issues included the identifying security and risk potential for data on the city's infrastructure.

Before data is even published on data portals, certain legal aspects need to be ironed out, potential legal barriers to publication of datasets need to be identified (Kucera & Chlapek,2014); such as issues around licensing, copyrights and protection of private information. Ubaldi (2013) supports this statement confirming that the legal arena surrounding data sharing and opening is unarguably complex.

Verhulst and Young (2017), comment on data quality; data quality is a prevalent challenge to positive impact, both in developed and developing countries, the latter being where this challenge is the greatest. The authors assert that data quality issues are expressed in a number of ways, such as inaccurate information, outdated data, lack of completeness, or even corrupted datasets. The quality of datasets released by the City of Cape Town had some quality issues, not so malicious but need improvement. Some datasets provided were often not in structured standardized manner; information on tables was not structured in easily

understandable and synthesized manner. These datasets were not elaborative or intuitive on the information they provided; one of the study's responded also commented on this that data provided as raw information in tables is often not immediately understood by users. This meant that some information had to be synthesized and be presented in various simple forms such as charts or graphs.

Another issue identified was that it was not possible to view information on tables before downloading it i.e. users needed to download each file they wanted to view- a practise against open data principles. Users needed to download files of information they are sure they wanted. While the portal is at beta version, these are suggestions that can be taken going forward for future updates.

Lnenicka (2015) suggests that for user to be able to assess the quality of data, they needed to understand the nature of the data as data holders cannot anticipate all users and their activities, and thus emphasizes that the provision of better quality metadata is equally important as the quality of datasets themselves. From this study's findings, the city was far off the mark from achieving these requirements as their portal and initiative is relatively new.

Concerning legal aspects of opening data, Ubaldi (2013) reveal that there are several circumstances that in which government datasets may – rightfully from a legal perspective not be completely open to public. Firstly it is whereby information has legal exceptions on premises such as national security or privacy, and is therefore not publicly released, even when requested.

Finding a balance between publishing datasets and protecting private information would prove to remain complex for the city. For the scope of this study, the focus was on basic legal requirements such as; is the data free for use, reuse, redistribution and if it government owned without any license or copyright restrictions.

Another major issue that emerged from the study was the preparation and release of data from the various contributing departments. Open data project managers revealed that often they were faced with departments that are reluctant to release data, citing reasons such as risk and privacy of government infrastructure asserts. City officials highlighted that releasing data about infrastructure such as power plants ran the risk of vandalism and theft, these datasets would then be halted from public access. Kusera and Chlapek (2014:35) also support

that "*detailed data about infrastructure (power plants, dams and transmitters etc.) might be used to cause damage to the infrastructure"*. The study also discovered none of datasets of this nature.

The most striking difficulty for departments however was the lack of dedicated data specialists working on data requests. Departments often did not have enough expertise and time to prepare and release data. This resulted in backlogs and many datasets not being published. Data management and preparation for publication seem almost as an ad hoc exercise within the city's departments. Some departments released data as and when they can, this shows a clear indication of how much capacity and culture that is currently lacking as far as open data is concerned. This is what Verhulst and Young (2017) refer as "cultural and institutional roadblocks"- these roadblocks are manifested in the institutional culture that continues to be sceptical of openness or even the absence of well-trained personnel capable of identifying and acting on the potential of open data (Verhulst and Young, 2017). The case of the City of Cape Town is indeed testament of the need to eradicate these roadblocks in organisations seeking to implement open data initiatives.

A lack of intentional investment into open data emerged from the study, the city has not yet invested heavily on expertise and infrastructure to take the initiative to the next level. While the study did not establish the reasons to this, financial constraints could be an attribution to this. One responded on the interviews, working closely with the city's open data policy, retorted that the city has not intentionally invested financially into the open data initiative – this included mainly investment in well-trained data expertise. The respondent further mentioned that, hiring and training open data personnel is an expensive exercise but would be highly critical and recommended if the city wanted to succeed in implementing the policy effectively.

Ubaldi (2013) insists that several financial challenges are hampering the rapid development of OGD initiatives. There is a huge commitment and investment by governments as they *need to acquire new skills, train employees, purchase new technologies, and upgrade network infrastructure, which need to be accounted for"* (2013:32). This confirms the efforts that go behind the implementation and development of an open data initiative and that it takes substantial commitment and strong will to invest. Lack of financial investment may undermine the development and implementation of open data by the city, and risking it to

remain an “ad-hoc” service as it is currently. To make data provision a mainstream operation within the city as oppose to be an ad-hoc service, departments need to be provided new skills, be trained and provided with new technologies to be able to carry out their data requirements.

It also emerged that, departments also had technical difficulties to collect and prepare some datasets due to infrastructure limitations. City officials revealed that some datasets were big or too complex to prepare. Some datasets were too complex in terms of technicalities to clean and make them accessible and if were to made available would need heavy human, capital and infrastructure investment which the city did not have. City officials also revealed that sometimes datasets were too big to aggregate and publish on the portal for public use.

The study considered in detail how data is made available with the focus on the type of the data published, its quality and formats. This was done through a desktop assessment of the city’s open data portal. To turn the potentials of open data into benefit, it is important for potential users to be able to easily discover the published datasets. Government and other open data holders provide portals to make this available and certain technical and legal requirements need to be met to ensure efficient use when searching for datasets in these portals. The study embarked on an assessment on assessing the city’s portal. The main purpose for the desktop assessment was to examine the alignment of policy specifications with technical practice. Lnenicka (2015) reaffirm that the effective development of progressive open data portals necessitates them to be evaluated systematically, in order to under them better as well assessing the different types of value they generate and recognize the needed improvements for increasing this value. The study assessed the city portal which is at a beta version, this means that four years down the line since the implementation of the policy, the portal is still not yet at its complete state. There is a lot that need to be taken into consideration when designing an efficient open data portal and an open data initiative in general. The study therefore looked at issues such as data quality, data formats, metadata, machine-readability, visualization capabilities, data search-ability function etc. – this section generally focused on the usability i.e. the user experience the portal has currently. Open data portal quality standards and principles adopted from literature were used to gauge the city’s portal. While it might have been prematurely to assess the portal while it is in beta version, city officials alluded that it was not too different from the live version that would be released.

Quality of data and the types of data published is a key component in the success of an open data initiative. The study discovered that the city provided datasets from various government sectors ranging from transport to health. Some datasets were downloaded more than the others. It emerged that many downloads were from small business interests, downloading tender opportunities advertised by the city followed by water consumption reports. The latter, it emerged that it was mostly downloaded during the drought that hit the city.

While the city has ensured to incrementally increase the supply of datasets on the portal, more could still be done; providing quality datasets with value to the public and that provided timely with an improving infrastructure.

5.3 Use: The struggle to build a participative open data audience

Once data is published into the portal, it is expected to be put into use by the targeted audience users, which is the public. Post the adoption of the open data policy, the city developed an open data policy

This theme concentrated context of use the city's open data, it also included the various category of users, the purposes for which datasets are used as well as the activities undertaken; this theme addressed the *who, what* and *why* of the open data initiative (Caplan, et al., 2014).

From the study it emerged that there is a gradual growing user base of the city's open data through its portal. The city takes records of downloads and users logging-in on the portal; it also emerged that the portal also attracts users from other countries across the globe. City officials alluded that international interest on the city's data increased dramatically during the 2017 drought in Cape Town whereby users were interested in downloading information regarding the drought conditions. Verhulst and Young (2017) suggest that as the open data ecosystem expands globally, there is now ever growing focus on understanding the demand side of open data i.e. the stakeholders who will make use of government datasets. The city reiterated this statement, that to ensure effective and sustainable data release to grow their user base certain systems had to be in place at all the time, this included having a technically functioning open data portal where the users will search for government datasets.

The portal had a userbase from various fields such as small businesses, researchers, surveyors and NGOs etc. City officials however alluded that it was difficult to know what exactly users do with the data except from taking hues on the reasons given by users when requesting for datasets not yet published on the portal. Some of the reasons that users gave ranged from research, planning and mapping, and data visualization. This made it difficult for the city to manage the supply of datasets meeting the demands needs. However, in response to that, one responded suggested that "supply creates demand", alluding that the city needs to increase its supply and that would increase the demand. This also meant that the city did not know what datasets to release for the users, Verhulst and Young (2017:32), strongly suggest that a clear problem definition can help unearth which data sources may add value and inform strategies for collecting or accessing such data.

Zuiderwijk and Janssen (2013) confirm that results of data reuse are not discussed and minimal feedback and comments is given back to governments (data providers) and that adversely impacted supporting policy and decision-making. This raises questions on the existence of an open data community of open data users that would interact and share feedback of their experiences using the data and portal. Therefore, to tackle this challenge, the city would have to find an effective way to form an interactive and proactive open data community.

While the user base has been steadily growing, city officials were not completely satisfied on uptake of the portal use. Lnenicka (2015) confirms that despite governments actively promoting the use of their portals and data by organising events such as challenge competitions, the response to take advantage of this has been underwhelming, especially for innovative activities. This challenge was also discovered through this study, city officials confirmed that there was still a huge lack of open data uptake especially for entrepreneurial ventures. Lnenicka (2015) concludes to say that this brings the questions about the reasons barring the interests to innovate using open data.

The main concern that emerged from the study for the theme *use* was the uptake of open data through the use of the portal and creating innovative products for economic exploits. Verhulst and Young (2017:20) strongly argue that "*the activities enabled by access to open data in many ways are only limited by the imagination and skills of the actors on the side of the equation*" (users). This can come with user capacitating efforts whereby users are

equipped to use data to interpret and building innovative solutions - something which the city currently has no capacity to deal with. Ubaldi (2013) supports that, uneven dissemination of skills, time and resources have a significant impact on who uses government released data as well the reach they can have using it. There almost has to be a balance of thought about the various users targeted for open data as well as the usability of its digital infrastructures. Davies (2010) asserts that the current Open Government Data community tend to be very technology focused and that this may risk losing sight of the real end user. Algemili (2016) also support that many people are at the risk of losing the benefits of open data due to their lack of technical skills. The design of datasets and portal infrastructure had to be considerate of the skills of the prospective users of data, and this mean that the city would have to understand and know its user base.

To avert the lack of skills to interpret datasets and using portals- locally the study discovered that some NGOs in the open data sector space have developed open data literacy programs to help users understand and interpret open data for their respective endeavours. A growth of these type of open data literacy will undoubtedly bridge the gap of data supply and development of impactful innovations. This further goes back to the point of lack of financial investments on open data; sufficient investments would ensure that the city is able to run workshops, hackathons and industry collaborations on regular bases to promote and make open data awareness.

With these challenges in the midst, there are some successes since the policy was implemented. Today, for instance, the portal currently attracts not only local users but also international users from various countries have shown interest on the city's data on the portal. More uptake would also be expected as the portal is launched to the live version with increased numbers of and varied types of datasets. Algemili (2016) concludes that open data issues are ongoing, that they need continuous improvements at all stages, from data collection up to visualization.

5.4 Too early to assess impact - The challenge of defining intended goals and clear targets

As part of the assessment of the city's open data challenges, the study looked at the possible impact that the city's open data has since its inception. *Impact* focused on the following subthemes; social, political, and economic impacts. Of the four themes, *impact* was the most underwhelming in terms of the research findings and literature in other studies has covered a lot on the reasons for this.

Studying the current or potential impacts of the city's open data initiative for this study was in a quest to discover the goals and benefits of the initiative as intended by the city's policy. Finding this would then enable the study to draw the question who are the intended beneficiaries of open data and for what ends is it provided to them.

City officials alluded that it was difficult to track and assess impact and they have currently not embarked on the process of tracking it. Verhulst and Young (2017) note that open data is a double-edge sword in several ways; the way open data is made publicly accessible ensures that citizens have access to datasets for any reason, while it also means that identifying those activities and capturing their impacts is very challenging, particularly for resource-restricted governments. The lack is also accounted by Kesuru and Chan (2015) revealing that despite all the benefits that open data can come with, there is however an incredible scarcity of evidence on the social and political impact of open data. In the context of Cape Town through the study, other than the city officials more than one respondent confirmed that tracking impact was indeed a challenge in their open data organisations and that it is something they are yet to start assessing internally. Verhulst and Young (2017:21) suggest that to "*gain some meaningful sense of the impact of open data releases, data holders can seek to develop indicators tied to the problems open datasets stand to address"*. Surely there must be a certain value attached to any datasets released and thereby expected benefits and impacts attached to it. Assessing impact would be simple for data holders when quality data with usable value is released. The authors, further state that a meaningful capture of open data impacts through indicators of success would possibly prove critical for maintaining the political will required for open data initiatives to be sustainable. Chan and Keseru (2015) also add that most open data initiatives have tendency to function in environments where the influence of other

stakeholders and initiatives is needed to realizing sustainable change, which makes it more difficult to display the connection between the initiatives activities and the impacts its seeks to obtain. The two statements made by these authors emphasize the fact that open data initiatives are almost dependent on other operations to be successful and thereby achieving their desired impacts.

According to Verhulst and Young (2017), there are four types of open data impact; *improving governance, empowering citizens, creating economic value,* and *solving public problems*. However findings of this study were not arranged by these types. In none of the four types, the city has confidence they have seen its impact. City officials were mainly concerned about lack of economic value created through open data; there was a lack of entrepreneurial innovations undertaken through their data. *Solving public problems*- there were a few open datasets from that had obvious impact on solving immediate public problems, one that come to mind is the release of water consumption and restrictions to individual households, with this, citizens were able to track their household water consumption through updated maps and report. This was effective and beneficial as it alerted citizens to reduce their water consumption and thereby discontinuing the then imminent drought effects in the city.

The study has afforded the researcher an opportunity to uncover the various issues affecting effective implementation of the City of Cape Town's open data policy. According to Verhulst and Young (2017), open data initiatives frequently fail to build an audience and develop successfully over time if they not intentionally define the intended benefits of the use of open data as well as clear target goals.

5.5 Conclusion

In conclusion, Algemili (2016) asserts that for open data holders such as governments, to confront of the issues of open data, project manager would have to perceive open data as an ongoing process than a product. Open data initiatives require constant improvements at all levels as it matures. Learnings from study's findings will contribute significantly to the current Open Government Data practise for other local cities that are yet to implement an open data policy.

The case of the City of Cape Town open data initiative shows us that there is indeed a lot that goes into implementing an effective open data portal. Findings confirmed that up to this stage, open data operations at the city were at an implementation stage. In this context it showed that it was more important to understand the *context* of the city's open data i.e. understanding the technical, legal and organisational environment of open data. This was important for the study to completely understand how processes and relationships came together, and it was discovered that improvements were still needed.

At the stage of the policy it was still early to focus on *impact,* this was confirmed by underwhelming responses from the study on *impact.* The newness of the city's open data initiative limited the assessment of *impact* as not much has been done to consider impactful. Literature has however argued that expected open data goals and benefits must be stipulated from the beginning. Not having long sight to the future of what impact the city want to see, could be detrimental as they would not fully understand what datasets to provide and for what reason. This calls for the city to begin to think about assessing the impact of their open data policy as this would also inform them of their progress and challenges.

While there has been lack *impact* focus, the publishing of *data* and its *use* were more important at this stage for the city. While there has been efforts made to publish datasets and developing and open data portal, a lot still had to be done to insure that data is opened; with investing more on well-trained personnel and improving infrastructure. For the city's open data initiative to succeed, a strong leadership-will would be needed to intentionally build capacity of open data within the organisation. Once this is achieved, the use of open data by the public would be much easier; more datasets would be released by the city, openness to community engagement would thrive – currently the city is lacking interactions with the open data portal users. Several open data studies have confirmed that the success of government open data initiatives depended largely on public participation, whereby users suggests datasets and engage on issues of data governance. The case of Cape Town has demonstrated to us that this is still lacking and solutions had to be found. Proposed solutions to expand open data users and engagement are arranging open data competitions, hackathons and workshops.

Findings and analysis from this study have demonstrated that open data challenges, opportunities and potentials can be easily understood when analysed on different themes as

done with *context, data, use and impact.* These themes have afforded the study an opportunity to understand the city's current challenges and gaps in policy and practise. Understanding the city's open data *context* gave strong indications of where the city is at in terms of organisational operations.

While it emerged that there were currently no *impact* assessments on the city's open data initiative, building and expanding an audience would still be a challenge if there are no defined benefits and clear targets. Setting up this would ensure that there is understanding of what service is being delivered to public and for what purpose. To increase and maintain user base, among other things, user research has to be instituted, it has been suggested that upfront mapping and understanding of relevant stakeholders as well as an assessments of their needs can facilitate more targeted open data-driven interventions (Verhulst and Young, 2017).

The City of Cape Town being the first African City to implement an open data policy, findings of study revealed that while the city has done well to introduce the policy, a lot still had to be done for the initiative to reach its potential. These findings will undoubtedly contribute to this under-studied subject especially on a local level as suggested in literature. Through this case study of the City of Cape Town, other local cities and municipalities that are yet to implement an open data policy would be able to learn the potentials, challenges, and opportunities that an open data initiative come with. These findings will also contribute immensely on future literature on future studies of open data for local and other African cities.

A comprehensive study on open data users and impact for future studies would be recommended as this study was limited by the fact that the policy was relatively new to fully understand these components.

Conclusion

The City of Cape Town generates large volumes of valuable data that represent a resource that can be potentially used by various sector bodies, ranging from individuals to enterprises as well as other government agencies. Government Open Data has therefore been seen as a means to make this data available and easily reusable by anyone for any purpose. To achieve this, the city enacted the open data policy to enable data sharing with these various entities. An open data portal was developed as part of their digital transparency to make the data accessible digitally. While at the process, the city has had challenges in effectively implementing the policy.

In this study, we presented the potentials, challenges and barriers of open data. The paper mainly focused on the challenges that the city has experienced since the adoption of the policy. A focused framework was adopted from literature with these four key themes; context, data, use and impact. The four themes formed the method of identifying and analysing the city's open data challenges.

The main finding of the study found that the actual challenge was the fact that open data was still a fairly new concept locally. In this infant stages the city had challenges in organisational culture such as scepticism and reluctance towards data openness.

The study suggests that, for the city to take their open data initiative to the next level, significant investment with a strong will would have to be implemented on human capital, infrastructure resources, and also on open data public awareness. Ensuring the investment in these resources will ensure that open data reaches its potential growth and thereby making the intended impacts.

As the City of Cape Town is the first city in Africa to adopt an open data policy, researching the Cape Town case was very crucial for understanding this novelty initiative, especially in an African context. The study provided us an opportunity to understand the challenges of open data from both policy perspective as well as in practice. Findings from this study will not only contribute towards the City of Cape Town, but will also contribute in understanding other African cities that are yet to implement an open data policy.

References

Algemili, U.A. 2016. Outstanding Challenges in Recent Open Government Data Initiatives. *International Journal of E-Education, E-Business, E-Management and E-Learning.* 6(2):91-102. DOI:10.17706/ijeeee.2016.6.2.91-102.

Attard. J., Orlandi. F., Scerri. S. & Auer.S. 2015. A Systematic review of open government data initiatives Available: https://www.researchgate.net/publication/281349915_A_Systematic_Review_of_Open_Government_Data_Initiatives

Braunschweig K. Eberius J. Thiele M. Lehner W. 2012. The State of Open Data Limits of Current Open Data Platforms. In: 21st World Wide Web Conference 2012 (pp. 1-6). New York: ACM.

Batty, M., Axhausen, K., Giannotti, F., Pozdnoukhov, A., Bazzani, A., Wachowicz, M., Ouzounis, G. & Portugali, Y. 2012. Smart cities of the future. *The European Physical Journal Special Topics.* 214(1):481-518. DOI:10.1140/epjst/e2012-01703-3 Available: https://www.openaire.eu/search/publication?articleId=dedup_wf_001::af61f2f09fac7f34007a6d9886a7d085.

Blok, A., Marquet, C., Courmont, A., Minor, K., Young, M., Hoyng, R. & Nold, C. 2017a. Data Platforms and Cities. *Tecnoscienza. Italian Journal of Science & Technology Studies.* Available: https://hal.archives-ouvertes.fr/hal-01703100.

Blok, A., Marquet, C., Courmont, A., Minor, K., Young, M., Hoyng, R. & Nold, C. 2017b. Data Platforms and Cities. *Tecnoscienza. Italian Journal of Science & Technology Studies.* Available: https://hal.archives-ouvertes.fr/hal-01703100.

Buranarach, M., Krataithong, P., Hinsheranan, S., Ruengittinun, S. & Supnithi, T. Nov 7, 2017. A Scalable Framework for Creating Open Government Data Services from Open Government Data Catalog. ACM. 1-5.10.1145/3167020.3167021. Available: http://dl.acm.org/citation.cfm?id=3167021 .

Buss, T. 2015. *Open Data Policy and Practice.* Available: https://www.igiglobal.com/chapter/open-data-policy-and-practice/112968

Carrasco, C. & Sobrepere, X. 2015. Open Government Data. *Social Science Computer Review.* 33(5):631-644. DOI:10.1177/0894439314560678 Available: https://journals.sagepub.com/doi/full/10.1177/0894439314560678.

Charalabidis, Y., Alexopoulos, C., Alexopoulos, C. & Loukis, E. 2016. A taxonomy of open government data research areas and topics. *Journal of Organizational Computing and Electronic Commerce.* 26(1-2):41-63. DOI:10.1080/10919392.2015.1124720 Available: http://www.tandfonline.com/doi/abs/10.1080/10919392.2015.1124720.

Chourabi, H., Taewoo Nam, Walker, S., Gil-Garcia, J.R., Mellouli, S., Nahon, K., Pardo, T.A. & Scholl, H.J. Jan 2012. Understanding Smart Cities: An Integrative Framework. IEEE. 2289-2297.10.1109/HICSS.2012.615. Available: https://ieeexplore.ieee.org/document/6149291 .

Corrêa, A.S., Paula, E.C.d., Corrêa, P.L.P. & Silva, Flávio Soares Corrêa da. 2017. Transparency and open government data. *Transforming Government: People, Process and Policy.* 11(1):58-78. DOI:10.1108/TG-12-2015-0052 Available: http://www.emeraldinsight.com/doi/abs/10.1108/TG-12-2015-0052.

Tran, E. & Scholtes., *OPEN DATA LITERATURE REVIEW.* Available: .

Gabriel, R.P. 1998. Jimmy, Jimmy, Oh Jimmy Mac. *Ploughshares.* 24(4):85. Available: https://search.proquest.com/docview/1297371256.

Goëta, S. & Davies, T. 2016. The Daily Shaping of State Transparency: Standards, Machine-Readability and the Configuration of Open Government Data Policies. *Science, Technology and Innovation Studies.* Available: https://hal.archives-ouvertes.fr/hal-01829314.

Gonzalez-Zapata, F. & Heeks, R. 2015. The multiple meanings of open government data: Understanding different stakeholders and their perspectives. *Government Information Quarterly.* 32(4):441-452. DOI:10.1016/j.giq.2015.09.001 Available: https://www.sciencedirect.com/science/article/pii/S0740624X1530006X.

Government information quarterly. 2001a. *Government Information Quarterly.* 18(4):375-378. Available: https://www.sciencedirect.com/science/article/pii/S0740624X01000934.

Government information quarterly. 2001b. *Government Information Quarterly.* 18(4):375-378. Available: https://www.sciencedirect.com/science/article/pii/S0740624X01000934.

Government information quarterly. 2001c. *Government Information Quarterly.* 18(4):375-378. Available: https://www.sciencedirect.com/science/article/pii/S0740624X01000934

Halonen, A. (2012). Being open about data: Analysis of the UK open data policies and applicability of open data. Finnish Institute in London. Retrieved from http://finnishinstitute.cdn.coucouapp.com/media/

Kalampokis, E., Tambouris, E. & Tarabanis, K. 2011. A classification scheme for open government data: towards linking decentralised data. *International Journal of Web Engineering and Technology.* 6(3):266. DOI:10.1504/IJWET.2011.040725.

Kitchin. R. 2014. The real-time city? Big data and smart urbanism. *GeoJournal.* 79(1):1-14. DOI:10.1007/s10708-013-9516-8 Available: https://www.jstor.org/stable/24432611.

Kitchin, R. 2015. *Data-driven, networked urbanism. Programmable City Working Paper 14.* Available: https://www.openaire.eu/search/publication?articleId=od_____223::5821c645ff5fb2be6cfa7ae5e3ee70b5 .

Kitchin, R., Coletta, C., Evans, L., Heaphy, L. & MacDonncha, D. 2017. Smart cities, epistemic communities, advocacy coalitions and the `last mile' problem. *It - Information Technology.* 59(6):275-284. DOI:10.1515/itit-2017-0004 Available: http://www.degruyter.com/doi/10.1515/itit-2017-0004.

Kucera, J & Chlapek. D. 2014. Benefits and Risks of Open Government Data. *Journal of Systems Integration.* 5(1):30-41. DOI:10.20470/jsi.v5i1.185 Available: https://search.proquest.com/docview/1690665526.

Leszczynski, A. 2018. Digital methods I. *Progress in Human Geography.* 42(3):473-481. DOI: 10.1177/0309132517711779 Available: https://journals.sagepub.com/doi/full/10.1177/0309132517711779.

Lourenço, R.P., Piotrowski, S. & Ingrams, A. 2017. Open data driven public accountability. *Transforming Government: People, Process and Policy.* 11(1):42-57. DOI:10.1108/TG-12-2015-0050 Available: http://www.emeraldinsight.com/doi/abs/10.1108/TG-12-2015-0050.

Magalhaes, G. & Roseira, C. 2017. Open government data and the private sector: An empirical view on business models and value creation. *Government Information Quarterly.* DOI:10.1016/j.giq.2017.08.004 Available: https://www.sciencedirect.com/science/article/pii/S0740624X17302629.

Marrone, M. 2018. Smart cities. *Business & Information Systems Engineering.* 60(3):197-213. Available: http://www.econis.eu/PPNSET?PPN=1027050530.

Marvin, S., Luque-Ayala, A. & McFarlane, C. 2016. Smart urbanism. London and New York: Routledge. Available: http://bvbr.bib-bvb.de:8991/F?func=service&doc_library=BVB01&local_base=BVB01&doc_number=028445663&sequence=000001&line_number=0001&func_code=DB_RECORDS&service_type=MEDIA .

Nascimento, F.R.A., Cesar da Rocha, J. & Garcia, A.C.B. 2018. Automated Evaluation of Open Government Data Portals. *International Journal of Electronic Government Research.* 14(3):57-72. DOI:10.4018/IJEGR.2018070105.

Nayek, J.K. 2018. Evaluation of Open Data Government Sites: A Comparative Study. *Library Philosophy and Practice.* : 1. Available: https://search.proquest.com/docview/2163342298.

Ruijer, E., Grimmelikhuijsen, S., Hogan, M., Enzerink, S., Ojo, A. & Meijer, A. 2017. Connecting societal issues, users and data. Scenario-based design of open data platforms. *Government Information Quarterly.* 34(3):470-480. DOI:10.1016/j.giq.2017.06.003 Available: https://www.sciencedirect.com/science/article/pii/S0740624X17302009.

Reiche K. J. Höfig E. Schieferdecker I. 2014. Assessment and Visualization of Metadata Quality for Open Government Data. In Proceedings of the 4th Conference for E-democracy and Open Government (Cedem 2014) (Vol. 1, pp. 335–346).

Seo, J. 2016. *An evaluation of Esri's tapestry segmentation product in three Southern California communities: Manhattan Beach, Santa Monica, and Venice Beach.* Available: https://search.datacite.org/works/10.25549/USCTHESES-C40-220747

Vanky, A. 2015. *The Elusiveness of Data driven urbanism. 103rd ACSA Annual Meeting Proceedings, The Expanding Periphery and the Migrating Center.* Available:http://vanky.co/hello/wp-content/uploads/2016/04/Vanky-ACSA-103.22.pdf

Verhulst, G. 2017. Open Data in Developing Economies: Toward Building an Evidence Based on What works and How. Available https://muse.jhu.edu/book/57263

United Nations, New York. 2013. *Guidelines on Open Government Data for Citizen Engagement.* Available: http://workspace.unpan.org/sites/Internet/Documents/Guidenlines%20on%20OGDCE%20May17%202013.pdf

Zuiderwijk, A. & Jansen, M. 2013. "A Coordination Theory Perspective to Improve the Use of Open Data in Policy-Making", *in the proceedings of the 12th IFIP WG 8.5 International Conference: EGOV 2013, Sprinter, Heidelberg, 2013 pp.38-49*