CITY OF CAPE TOWN
ISIXEKO SASEKAPA
STAD KAAPSTAD

CITY OF CAPE TOWN
Data Strategy

August 2023

Making progress possible. Together.

EXECUTIVE SUMMARY

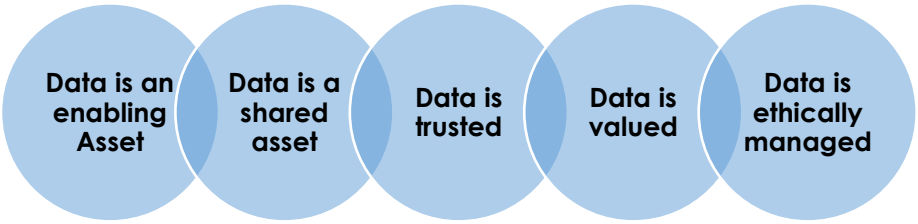
The purpose of this strategy is to ensure that data is leveraged to its full potential, contributing to evidence-based decision making to ensure the City of Cape Town is fully effective and responsive to residents and business's needs. It contains the vision, principles desired outcomes, objectives, and enablers that are needed to achieve this.

Data Strategy Vision:

Data is used to deliver better public services to all people in Cape Town

The principles that inform this Data Strategy are:

The desired outcomes of this Data Strategy are:



The following strategic and foundational enablers will guide the work of the CCT:

Contents

List of text boxes

List of figures

List of tables

Definitions

Application Programming Interfaces	An interface or mediator between systems.
Custodianship	Custodianship involves assigning responsibility for the collection and management of individual information sets on behalf of users.
Core Application Review	The CCTs core application review programmes is the plan for replacing and transforming its core applications to enable the municipality and residents to operate more effectively in a digital medium.
Dashboards	Are collections of individual measures, reports, and graphical displays that summarise key metrics.
Data analytics	Data analytics refers to the end-to-end process of extracting data from various storage technologies and transforming it into insights that support decision-making. The analysis component of data analytics is critical in generating insights that inform evidence-based decision-making.
Data Steward	Person designated to be responsible for managing certain datasets.
Data Custodian	Person designated to be responsible for managing the technical environment in which the datasets are stored.
Enterprise Content Management	Enterprise Content Management is the systematic collection and organisation of information that is to be used by a designated audience – business executives, customers, etc. which is neither a single technology methodology or process, it is a dynamic combination of strategies, methods, and tools used to capture, manage, store, preserve, and deliver information supporting key organisational processes through its entire lifecycle.

Enterprise Information and Knowledge Management (EIM)	The systematic, City-wide management of information and knowledge as critical assets of the organisation.
Enterprise Resource Planning	Refers to the type of software that organisations use to manage day-to-day business activities such as accounting, procurement, project management, risk management and compliance, and supply chain operations. At the time of drafting this strategy, the CCT used SAP is the ERP.

Abbreviations

APIs	Application Programming Interfaces
CAR	Core Application Review
CCT	City of Cape Town
CDO	Chief Data Officer
CDigO	Chief Digital Officer
CKAN	Comprehensive Knowledge Archive Network
CM	City Manager
CSIR	Council for Industrial and Scientific Research
DCC	Data Coordinating Committee
DGC	Data Governance Committee
DGWG	Data Governance Working Group
ED	Executive Director
ERP	Enterprise Resource Planning
GIS	Geographic information systems
IDP	Integrated Development Plan
ODTP	Organisational Development and Transformation Plan
PAIA	Promotion of Access to Information Act, 2000 (Act No. 2 of 2000)
POPIA	Protection of Personal Information Act, 2013 (Act No. 4 of 2013)
SARB	South African Reserve Bank
SARS	South African Revenue Service
Stats SA	Statistics South Africa

1. Introduction

Data is an important part of everyday life that is essential to providing Capetonians with the services and information that they need. Data have the power to enable to City of Cape Town (CCT) to make better decisions to ensure the public resources entrusted to the CCT are used in the right way, at the right time, to produce the most meaningful outcome for residents and business.

Evidence-based decision making for strategy and operational planning, and the delivery, and monitoring of service delivery, is a key component of this. Furthermore, good data practices and management allows the CCT to produce stronger and better-informed policies, as well as improved delivery of effective, equitable and inclusive programs and services.

The CCT developed an internal administrative Data Strategy in 2018. Much progress has been made since this important step in transforming CCT administrative data into meaningful and relevant business information that effectively supports decision-making, increases efficiency, and enhances transparency. Through its implementation, the CCT has learnt the value of having accessible data, for rapid response to both shocks and stresses, impacting on Cape Town. Importantly, the CCT has built a sound foundation for the next phase of its maturity journey. The CCT is ready to institutionalise and bring to scale the advancements made in the data environment, particularly during a time when it is undergoing significant changes in its core applications through the Core Application Review (CAR) Programme.

Through the implementation of this Council approved Data Strategy (2023), its first public data strategy, the CCT will be empowered to extract the full value of Data. It sets out how the CCT will acquire, manage, govern, and use data to be more operationally effective and responsive to resident and business needs. Implementation will offer value to residents and all stakeholders by making quality data more accessible and usable, helping stakeholders gain insights into their business and operations, make decisions, and provide fit-for-purpose solutions.

Data Strategy Vision: Data is used to deliver better public services to all people in Cape Town

In the future, people and businesses have access to services when and how they need them. Data makes this future possible because it is stewarded and used by the CCT for the public good, such that it contributes to current Integrated Development Plan[1] (IDP) vision (2022-2027) of a City of Hope for all - a prosperous, inclusive and healthy city where people can see their hopes of a better future for themselves, their children and their community become a reality.

[1] *City of Cape Town, 2022. Five Year Integrated Development Plan – July 2022 – June 2027*

*When Cape Town experiences a shock, like the COVID-Pandemic, the CCT response is underpinned with **enhanced evidence-informed decision-making**. Data is readily available and integrated into analysis of the issue and how it impacts on resources, the economy, infrastructure and ultimately residents and their communities. As a result, people and organisations receive the support they need quickly because the government can respond in a coordinated and timely way. Impacts on women and other vulnerable groups are better understood and programs and services are constantly improved because data is embedded throughout the life cycle of a project from planning and design to implementation and measuring the impact.*

*All residents are able to access the data they need to make informed choices around how they access and consume public services and are comfortable in the knowledge that their data is safe and kept private. There is transparency and openness in how the government operates and uses data and this helps people feel that they have a **trusted and accountable government** that can deliver for them while protecting their privacy.*

*Good data spurs innovation and helps people make informed choices about things that matter. The private sector and civil society extract **greater public value from the CCT open data portal** by leveraging data on things like housing, transportation, and energy to develop tools that help the public confidently make decisions in their everyday life.*

Text box 1: What do we mean when we say evidence based decision making?

What do we mean by evidenced based decision making?
Evidence -Based Decision-Making is a process for making decisions about a program, practice, or policy that is grounded in the best available data and research evidence and informed by experiential evidence from the field and relevant contextual evidence.[2]
What do we mean by 'extract the full value of data'?
Data is available, accessed, analysed, contextualised and visually represented to create actionable information and insights.
What do we mean by data quality?
Data Quality refers to both the characteristics of data and to the processes used to measure or improve the quality of data. Data is considered high quality to the degree it is fit for the purposes data consumers want to apply it[3]

[2] *While the CCT is committed to being an evidenced-based organisation, data and analysis are not needed for every single decision. It is accepted that some decisions require basic insights and can draw on experience while others require much more in depth analysis. There is a need to maintain flexibility to meet urgent demands, and some courses of action make sense no matter what the data say.*

[3] *DAMA UK Working Group. 2013. The Six Primary Dimensions for Data Quality Assessment: Defining Data Quality Dimensions. United Kingdom: DAMA UK. Available at: https://www.whitepapers.em360tech.com/wpcontent/files_mf/1407250286DAMAUKDQDimensionsWhitePaperR37.pdf.*

1.1 Strategic alignment to the existing policy and legislative ecosystem

The IDP clearly explains how the CCT plans to build a capable, modernised and administratively efficient government. Rich and reliable information is key to ensuring that the decisions the CCT takes in implementing the IDP, are evidenced based. A foundational objective of the CCT is to lead through evidence-based decision making in support of IDP Objective 16: Capable and Collaborative government.

The City has committed to the allocation public resources in a way that achieves maximum public benefit through better use of data and data analytics for long-term planning and decision making. The Data Strategy is strongly aligned to this commitment and enables this commitment by further institutionalising and scaling a culture of evidence based decision making, underpinned by data throughout the organisation.

This Data Strategy expands on IDP Objective 16: to build a modernised and administratively efficient government and the Evidence Based Decision making programme (IDP project: 16.3), which prioritises investment in robust decision-making systems. Through the implementation of this Data Strategy, the CCT will make progress towards its commitment to establishing an open, and transparent city government. This programme includes broader access to CCT data via the municipality's open data portal, and the development of appropriate data governance standards in line with international best practice and the Protection of Personal Information Act 4 of 2013. Figure 1 illustrates the role of data in, and the strategic alignment to the IDP objectives. Table 1 lists a selection of data related programmes extracted from the IDP. Table 2 lists the supporting Council approved policy ecosystem.

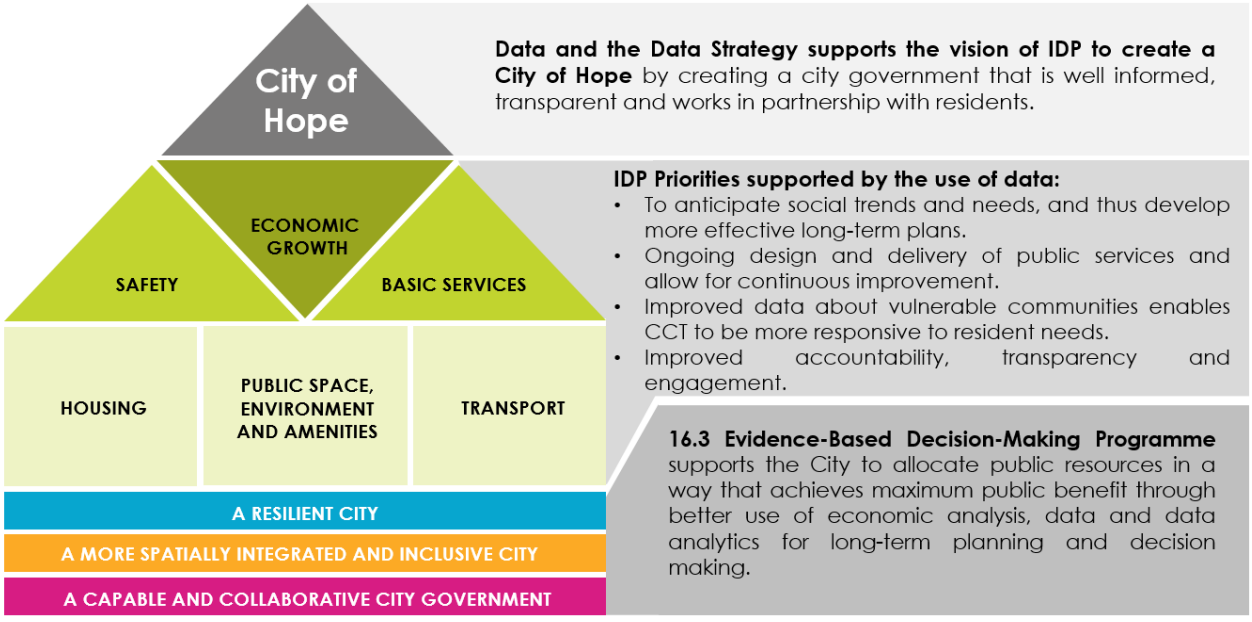


Figure 1: Strategic Alignment to IDP

Table 1: Data Strategy Alignment to IDP Priorities

IDP Priority	Data Related Programme
	*1.5 Consolidated land pipeline and release programme:* Land planning and release to support economic growth and more affordable housing. *1.5.A. Data-driven land management initiative*: The City will establish and maintain a consolidated database of its landholdings, land planning acquisitions and reservations. This database will include vacant and underutilised City-owned land and will inform decisions on the repurposing or release of this land for development.
	*4.3/5/8 Excellence in Water, Waste and Energy service delivery programmes:* Data-driven asset maintenance project *4.1 Utility Business Model Reform programme* Adapt the City's utilities to be sustainable and future-fit.
	Safety *5.2 Safety technology programme:* Adopt new technology to target policing resources effectively. *5.2.A Technology safety partnerships project* *5.2.B Incident, crime and emergency detection project* *5.2.C Digital evidence management project*
	8.2 Informal settlements upgrading programme: 8.2.B. Informal settlements data improvement initiative The City has good data to inform planning and customer service in formal areas. Yet data to inform the provision of basic services and security of tenure in informal areas is relatively limited. The City will develop robust data systems for informal settlements, recognising that this is crucial for responding to service delivery disruptions and adapting services to changes in population density and the physical characteristics of the area.

Table 2: Policy Instrument Ecosystem

Policy Considerations
• Integrated Knowledge Management Policy (2015) • Open Data Policy (2020)
Legislative Considerations
• Promotion of Access to Information Act, 2000 (Act No. 2 of 2000) • Protection of Personal Information Act, 2013 (Act No. 4 of 2013) • Electronic Communications Act, 2005 (Act No. 36 of 2005) • National Archives and Record Service of South Africa Act, 1996 (Act No. 43 of 1996) • Spatial Data Infrastructure Act, 2003 (Act No. 54 of 2003)

2. The City's Data Ecosystem

2.1 Data in context of Cape Town

In coming decades Cape Town will need to navigate a set of profound and complex spatial, social, economic and resource-based transitions, if it is going to realise sustainability, resilience, and its aspiration to be a *City of Hope* for all. Establishing how best to achieve these transitions requires the collective efforts of a wide variety of public and private stakeholders harnessing collective skills, resources and knowledge. Complex problems such as poverty, unemployment, crime, climate change, aging public infrastructure, and operational sustainability of the CCT functions all contribute to the urban development challenges and policy- context in Cape Town. Never before has there been such a need for actionable insights.

Data plays a crucial role in striving to solve complex urban problems by providing insights into the underlying issues, helping to identify patterns, and enabling evidence-based decision-making to optimise resources, and in a crisis, to make quick but informed decisions. For example, data analysis can help the CCT:

- 'measure' the city to understand current needs and demands, such as measure flows of traffic to optimise transport and minimise congestion;
- 'predict' needs to target and allocate resources to meet future demands such as the spatial distribution of development to support infrastructure planning,
- 'weighing' of trade-offs and detailed understanding of costs and benefits, such as prioritising one infrastructure project over another, and
- 'evaluate' policy or projects such as data on crime patterns to determine if interventions in high-risk areas have had the intended impact overtime.

Moreover, data can facilitate public engagement and collaboration. By sharing data with the public, the CCT can foster greater transparency and accountability, and encourage collaboration with citizens and other stakeholders to solve urban problems.

2.2 What is data and where the CCT gets it

The CCT generates and holds vast amounts and diverse array of data, including spatial data, operational data, transactional service data, and data collected from or about residents and businesses. Text box 2 provides an overview of data definitions and some examples of CCT Data and sources.

Text box 2: Different types of data support evidence-based decision making in the CCT

What we mean by data?

Data is raw, unorganised facts that need to be processed. Data can be something simple and seemingly random and useless until it is organised. Generally when referring to data in the CCT context, this refers to information about people, things and systems. While the broader definition of data covers records and documents, the strategy will not focus on the management of records and documents.

Administrative Data

Administrative data is derived as a by-product of existing processes and systems. This is mostly generated through the CCT Enterprise Resource Planning (ERP) system SAP, for example, financial, supply chain management and HR data.

Transactional Data

Transactional data is information that is captured from transactions, and is stored and processed in real time. This data supports the ongoing functioning of the CCT and provides an insight into the state of the organisation at a given time.

The CCT gathers data about various transactions that ensure Cape Town functions, such as:

- how much water and electricity households consume in order to bill accurately and fairly;
- about the number of wheelie bins needing collection in order to plan the routes and timing of waste collection;
- the maintenance status of infrastructure; and
- water quality data at key points in the water system in order to ensure the safety of water for drinking.

Other kinds of data that exist in the CCT, but also data beyond the CCT

- **Sensor data** is the output of a device that detects and responds to some type of input from the physical environment, such as air quality or water pressure.
- **Demographic data** is information about groups of people according to certain attributes such as age, sex, and place of residence.
- **Economic data** are data that describe an economy to give insight into productivity or financial health of markets and the implications thereof.
- **Socio-ecological dynamics data** is data about how individuals interact with and respond to the environment around them, and how these interactions affect society and the environment as a whole.
- **Urban mobility data** is data about all of the movements of both people and goods that occur in Cape Town via public or private transportation.
- **Note on spatial data:** spatial data is a significant data source, derived through the ESRI Geographical information Systems (GIS) software - where the map layers are provided by ESRI, or created by departments, and the spatial data is self- generated by City staff in the

field. Spatial, quantitative, and qualitative data is routinely collected by line departments and fed into SAP and ESRI for monitoring and evaluation, and reporting. Importantly, spatial data can be understood as a particular attribute of data collected such as the spatial attribute of asset for asset maintenance, drinking water quality, or customer satisfaction.
Thus, the Data Strategy speaks to the value of the various forms of data in serving both: • Operational mandate – whereby the purpose behind data generation and consumption is tied to a specific business mandate or requirement (internal or external to the CCT) and informs these decisions. • Exploratory function – whereby data is consumed to tackle more complex and exploratory business questions and requirements (internal or external to the CCT) to inform tactical and strategic decisions

2.2.1 CCT data sources and data gaps

As a major South African metropolitan municipality, the CCT is dependent on many sources of data which is used to inform decision making. These data sources include:

2.2.1.1 CCT departments

The CCT line departments are the main source of data. This is because the CCT staff generates most of the data needed for execution of planning and operational activities. The majority of this data is available in the CCT ERP and GIS systems, although many departments still rely on data capturing in spreadsheets.

2.2.1.2 CCT Assets

Many of the modern assets the CCT uses to deliver services are capable of generating data. This data is often used to monitor and detect certain events or to determine the status of an asset's life. Although these sources of data require extensive analytics and interpretation it often offers the most direct value to the organisation in informing planning activities.

2.2.1.3 Contractors and Suppliers

The organisations the CCT does business with also possess a lot of data that the CCT uses and depends on. Examples of these include technical data about products and services, financial data and sensor data. In order to ensure that this data is shared with the CCT, contractual agreements are required to govern the generation and sharing of data.

2.2.1.4 Partners

In helping to answer business questions partners' use, clean and reformat CCT data and in some instances combine it with unique datasets they have access to. These are valuable datasets that improve the CCTs data quality. Having a set of minimum data standards and

visibility of the steps involved in cleaning and reformatting the datasets is key. This includes public organisations for example, Stats SA, SARS, SARB, CSIR, etc.

2.2.1.5 Members of the public

Much of the data the CCT utilises comes from members of the public. This might be through service notifications, correspondence or collaborative agreements with organisations working with the CCT. Since this data is not always subject to the CCT data governance it usually requires reformatting, cleaning-up and interpretation before it can be fully utilised.

2.2.1.6 Data gaps

That being said, there are several data gaps in the CCT environment. A large proportion of people and activities in Cape Town are very difficult to gather data on due to the lack of direct touch-points with CCT government, particularly within neighbourhoods which are more informal. In addition to this lack of *representative* data, the CCT does not have *current* data on certain aspects, for example on employment, both formal and informal.

A dearth of data shared from other governmental entities and a lack of data at the right level and type, because of infrequent data collection, data lags or small sample sizes collectively impacts on the CCTs ability to build a sound evidence-base. Survey data largely comes from Statistics South Africa or other spheres of government, but, apart from the census (held every 10 years), the data is often not at granular enough scales to be useful at the city level. When the City needs to procure data from an external source, this is done through contracts or tenders.

In some cases, the City is able to obtain survey data from research institutes, or administrative data, for example mobile phone data from private sector companies. That said there is a need for the City to build strategies to fill these gaps and make use of alternative, higher frequency data sources to improve planning and delivery of services. For example, establishing mechanisms for gaining machine readable access to the Deeds Office, or Home Affairs to enable the CCT to provide better services to residents.

Text box 3: CCT Bi-annual Survey

The CCT's Policy and Strategy Department, Research Branch, is initiating a bi-annual city-wide socio-economic survey in Cape Town to obtain accurate, reliable, frequent, up-to-date and sub-metro level data on households (in various dwelling unit types) to inform metropolitan CCT planning, policies and decision-making. These surveys are one of the most important sources of socio-economic and urban development data. Although national surveys from Statistics South Africa are valuable and of high quality, the CCT as a metropolitan municipality, has limited influence on the types, forms and scale of these surveys, alongside the questions, spatial representation, frequency and availability of the data.

2.2.2 Increasing amount of data and its security

The CCT also understands that as new technologies reshape the lives and work of residents and business owners, the volume of data continues to grow exponentially helping drive innovation and shape decisions. It is important that the CCT take decisive steps to ensure that it is adapting to, and making use of, a fast changing data and technology landscape for the benefit off the public. In the same vein, the CCT needs to ensure appropriate governance of increasing amounts of data while accelerating digital transformation. The CCT needs to build on the areas of competence and experience within the organisation. Part of this endeavour will be to do better at the basics of gathering, managing and using data for decision making and operational performance management, internally and with external partners.

The other aspect for the CCT is to lead by example in the management, security and use of the data that Capetonians entrust to them. This data must be used and managed as a strategic asset in an ethical and secure manner that respects privacy and generates trust. As such, it is essential to build and maintain public trust through transparent data practices and the protection of privacy in accordance with policies and legislation.

2.3 CCT as a data ecosystem

This Data Strategy focuses on enhancing how the CCT creates, protects, uses, manages and shares data to improve how it makes decisions on policy and programmes. The CCT data practices have the potential to yield benefits for the lives of Capetonians and support businesses, researchers and the non-profit sector. This means the City needs to share data internally, and make useful data available externally. The strategy builds on current data initiatives to ensure complementarity, coherence and transparency, so that emerging opportunities are understood and acted upon. The CCT is also ensuring it is ready for future application of data from new technologies.

Figure 2 illustrates the CCT's aim for its data ecosystem. By viewing data as a shared asset, the CCT commits to:

- making data accessible and useable for all directorates,
- all directorates having sufficient capacity and capability for producing and utilising data, and
- having core specialist skills to enable data engineering and data analytics with appropriate quality controls.

The base of Figure 2 is the foundation, and illustrates a City-wide data platform which is naturally built upon the digital foundation which includes the City's network, data centre and cloud solutions. The foundation is within the Information Services and Technology (IS&T)

domain, which is leading the CAR programme of the CCT, and will determine the digital architecture.

There is a strong link between the foundation, as the systems that provide data, and the layers above the foundation where data can be used. This foundation enables the Data Strategy. Importantly, there are functional mechanisms to access data through an Application Programme Interface (API), which means that the API gateway can translate a request from a user in such a manner that it will *call* the system correctly. (In some cases, that might be ill-advised or very difficult which may result in the introduction of a caching layer). Data is made available to and between Directorates and departments for sharing, following the departure point that data is a CCT shared asset.

Furthermore, data is made available to and shared with entities external to the City via the Open Data Portal depending on the type of data agreements in place and levels of access and permissions. Through this mechanism the CCT will both make data available, and access data from external entities. The CCT intends to position itself in such a way that it benefits all stakeholders, and in addition to utilising its own data, the CCT will also strive to actively utilise data that is produced elsewhere. Conceptually, there is a single point of entry (API Gateway) for these systems.

The arrows demonstrate the different ways of sharing and utilising data within the City and with external stakeholders. Each arrow is bidirectional: the CCT both shares its own data and utilises data produced by others. The governance arrangements, explicated in Chapter 5 of this strategy will unpack the parameters for data sharing across multiple entities.

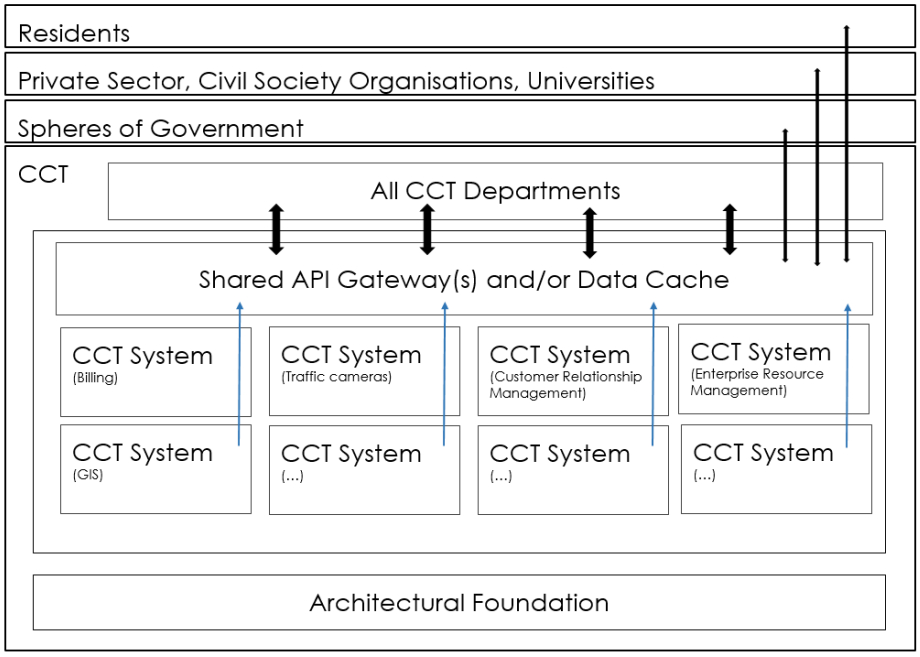


Figure 2: CCT as a Data Ecosystem

2.4 Data Users and Consumers

As a CCT line department, resident, business-owner, researcher and active citizen, this Data Strategy offers something for everyone. All stakeholders can have confidence that the CCT decisions have the potential to be based on sound evidence and rigorous analysis while being able to access the data they need. Table 3 provides an overview of the users of data in the CCT data ecosystem.

Table 3: Users of data

The strategy will make more data available to **CCT line departments**, helping them to enhance operational efficiencies, and solve problems they face using evidence based decision making. There are typically two types of decisions, which include: • Operational decisions which include repeatable work flow process related decisions, tools that support the logistical decisions relating to the day to day allocation of human and capital resources, analytics and predictions relating to large volumes of customer data or business continuity functions. • Strategic planning decisions which include policy and strategy decisions, medium term and long term planning and budgeting decisions, using macro trends descriptive tools, cross-sector predictive models (such as land use and mobility models) and financial impact models.
The **General public** will have progressively greater access to data about CCT service delivery practices and public policy issues, to support greater, and more informed participation in the decisions that affect them.
Researchers will be enabled with access to datasets to support research on various topics, for example social sciences, economics, and public health
Intergovernmental partners will achieve enhanced coordination of planning efforts, operations and outcomes due to better data sharing.
Non-profit organisations will be able to access CCT data and data products to identify community needs and create programs to address them.
Businesses and **private** sector will be able to: • Access to data for market research, such as identifying consumer trends, business opportunities, and product development. • Access to data for application development. • Access to data and information about the plans and investment pipelines of the CCT.

3 The City's data journey

The CCT is committed to evidence based decision-making. An approach that seeks to ensure that decisions are informed by the best available evidence and that evidence is targeted to the business questions which matter most. Data management and analytics forms an important foundational and knowledge-generation linked component of the CCT's efforts in driving evidence into decision-making processes. Figure 3 depicts the suite of mechanisms through which the CCT is driving evidence based decision-making.

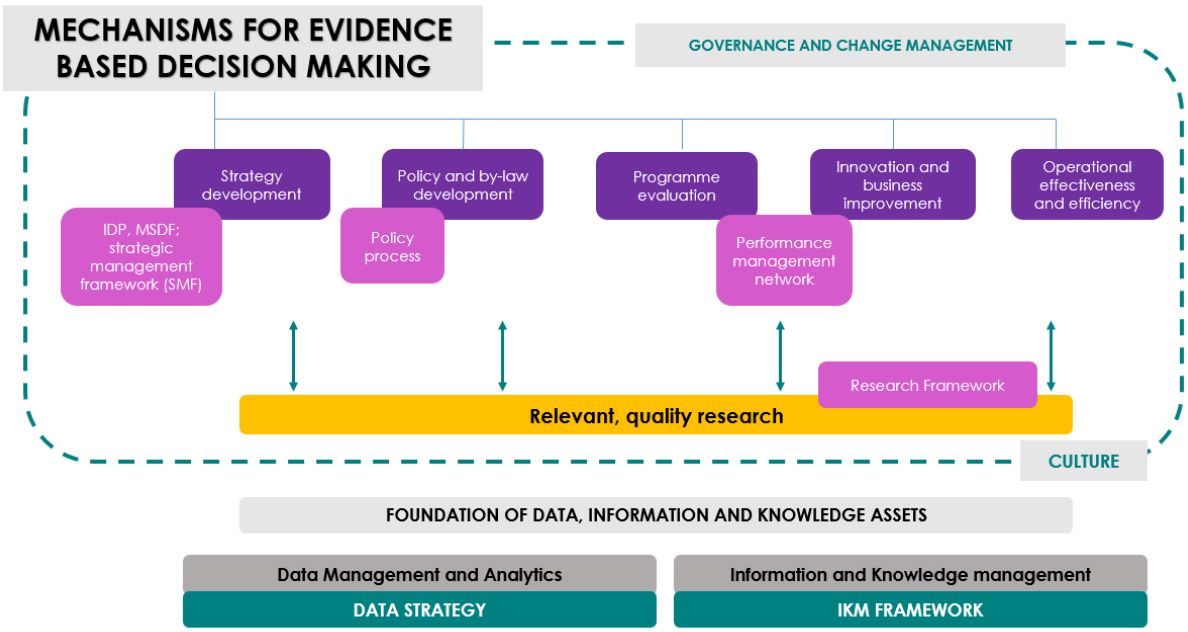


Figure 3: Mechanisms for Evidenced Based Decision Making in the CCT

There are multiple mechanisms for evidence based decision making resting on a foundation of data, information and knowledge assets. Importantly this draws on data and management analytics (guided by the Data Strategy[4]) as well as information and knowledge management (guided by the Integrated Knowledge Management policy) to enable strategy and policy development, programme evaluation, business improvement, and operational effectiveness and efficiency. Taken together and within City's governance and culture context, this leads to more effective City decisions.

This Data Strategy is a key lever for success going forward, but it is important to note that the use of data insights as a key enabler of CCT outcomes, has a long history. The story begins at the start of the CCT official journey as a Unicity, in 2000. Over time, significant technology, data and research reforms have enabled the improved use of data and research to achieve CCT

[4]*City of Cape Town (2018), Data Strategy. (The internal Data Strategy is discussed in greater detail below).*

outcomes. For simplicity, this journey is divided into two phases with key milestones identified, phase 1: Foundational, and phase 2: Building Blocks, are summarised at a high level in this section[5].

3.1 Phase 1 (2000-2015): Foundations of this Data Strategy

As early as 2000, the CCT embarked on its journey of implementing a range of significant and pragmatic reforms towards to the goal of a functioning Information and Technology (IT) infrastructure for improved public administration and service delivery. The CCT was the first South African city to embark on a City-wide Enterprise Resource Planning (ERP) system rollout. Overtime, the investments enabled reforms to centralise, share and better manage information - including data and research. One automated system would now manage the CCT's business processes, enabling a transition from the previous disjointed, paper-based system to a comprehensive, more integrated system with enhanced transparency within and between departments. With digitised records and systems, data and information would now be available in real time for faster operational decisions.

By consolidating and centralising information internally, the CCT has been able to embrace a more holistic and integrated approach, which supported the sharing of information more broadly with the public, and also to better manage, leverage, and apply the research and data work being done external to the CCT. The approach adopted was not without challenges, and the lessons, captured elsewhere, inform the CAR process going forward as the CCT considers using more agile and intuitive software solutions tailored for specific functions where appropriate, as well as systems that allow for better data integration.

A key takeaway is that this foundational phase revealed the value of digitally produced data and its utility for multiple functions, rather than just an output of the business process. The early prioritisation of internal administrative efficiency enabled the CCT to derive more value from the availability of data and overcome internal data silos. There are two key advancements (discussed below) in the policy ecosystem which had material impacts on the data environment.

3.1.1 Integrated Knowledge management

The CCT's Integrated Knowledge Management (IKM) work was led by an IKM department, which has a focus on development information, and also leading on the data cleansing and data centralisation on ESRI and SAP.

[5] *Further detail can be found here: Wright, C; Primo, N; Delbridge, V and Fortuin, K. 2023. Case Data and research as key enablers of city outcomes: A case study of the City of Cape Town (2000-2022). Available: https://www.globalfuturecities.org/resources/use-case-approach-city-cape-towns-data-strategy*

In 2011, the IKM department developed the CCT first policy targeting better research management: the Corporate Research Management Policy Framework[6]. This laid the groundwork for a key advancement in the data management eco-system, the adoption of the IKM Policy[7].

3.1.2 Open data policy and portal

A key activity of the CCT, as it has progressed through its data maturity journey, has been to make the CCT data more accessible externally. The second key advancement in the data management eco-system was the Open Data Policy[8] approved in 2015, outlining the legal provisions, policy parameters, stakeholder involvement, and regulatory context for a public use Open Data Portal[9]. The goal is to make certain datasets available to all, free of charge, in a usable and easily accessible format, without legislative boundaries and in a manner that mitigates risk for the CCT. The custom design of the portal was strongly enabled by the existing IKM and IS&T infrastructure and skills base in the CCT.

In many regards, the Open Data Portal is a major success, and demonstrates the CCT early commitment to transparency and openness. However, there are limitations and challenges, which provide insight for future iterations of the portal, and provide a guidance for and this Data Strategy. The following issues remain:

- Limited frequency and automation of data uploads to the Open Data Portal.
- The portal is not setup to receive data from outside the City, for example data that is cleaned by external researcher.
- Not all data is in a machine readable format.
- Data is not readily findable and accessible by default within the organisation.
- Data quality is perceived to be low, and thus departments are reluctant to share it.
- Fear of how data will be interpreted by people not working with it daily.
- Fear regarding the release of certain data might expose the City to a security or business risk.

The CCT is however committed to openness and transparency as a core value, and objective and guiding imperative for the Data Strategy. The CCT welcomes active engagement from citizens and commits to enabling progressive access useful data sets. Thus the ethos of the strategy in an effort to continually improve and drive enhancements across the CCT.

---

[6] *City of Cape Town Strategic Development Information & GIS Department (2011), Corporate Research Management Policy Framework and Guidelines.*

[7] *City of Cape Town Strategic Development Information & GIS Department (2013), Information and Knowledge Management (IKM) Policy Pertaining to Development Information.*

[8] *City of Cape Town. 2015. Open Data Policy.*

[9] *See: https://odp-cctegis.opendata.arcgis.com/*

3.2 Phase 2 (2016-2022): Building Blocks of this Data Strategy

Over the years, the automation of CCT systems and improved knowledge management have resulted in the generation and consolidation of large amounts of data, research information and reports. While the focus was initially on systems and operations, the CCT matured into viewing this data and research as assets in their own right - generating value and contributing to evidence-based decision making, and more effective governance and policymaking. However, the CCT still struggled with siloed management and use of data limiting potential impact and public benefit.

3.2.1 Institutionalising Data Science and Data Analysis

In response, the City's Organisational Development Transformation Plan (ODTP), approved in 2016/ 2017, and the IDP (2017-2022) recognised the need a transversal approach and rapid decision making in response to an increasingly dynamic operating environment. The Organisational Performance Management department was created and brought in data science expertise, tasked with applying advanced data tools and techniques to help the organisation modernise its data analytic capabilities and improve its operational performance and strategic outcomes. The purpose was to ensure policy was explicitly linked to some form of evidence or data, and the evaluation of policy effectiveness was to have measurable criteria.

At the same time the CCT firmed up process for policy development and operational decision making through the transversal Policy Process and Strategic Management Framework, which both drive evidence based, strategy led budgeting. Other reforms include the corporate programme and project scoping/framing and monitoring/tracking via the Corporate Portfolio, Programme and Project Management Department and its stage-gate analyses. This is complemented by the introduction of a SAP module that enables the CCT to capture and manage all capital projects, making analysis of project information possible. These have brought about organisation-wide improvements in the use of data for decision making and their respective evolution and institutionalisation offer valuable lessons for the Data Strategy, especially as it relates to governance arrangements for implementation.

In the wake of the 2017 Drought, the value of insights generated through data analytics became even more visible [see Text box 4 for an overview of the City's Drought Response and the Role of Data]. This was a substantive catalyst for building data capacity and capability in the CCT. Data expertise evolved into a Data Science unit within the Organisational Performance Management department, a first for South Africa local government. Their purpose was to apply advanced data tools and techniques to help the organisation

modernise its data analytic capabilities and improve its operational performance and strategic insights.

In the absence of a data analytic platform to enable data science at the time, the unit developed an open-source data analytics platform that facilitates agile operations and leverages cutting-edge technologies for the benefit of the organisation and it residents. By opting for open source, the team were able to take advantage of the collaborative approach within the open source community. Open source also allowed the CCT to easily collaborate with the technical community in Cape Town, as well as globally across industry, academia, and government.

The Data Science environment is now used by subject matter experts across the organisation. Data pipelines are created for line departments use which allows specialists to do data analytics. Following this lead, many line departments have matured their data capacity and capabilities. Scaling and institutionalising this environment is a key focus of this Data Strategy.

At the same time, the CCT Economics Analysis Branch, reoriented to use a wider range of economic analysis tools to inform decision making (in addition to monitoring the Cape Town economy). A key reform was the introduction of systematic use of cost benefit analysis and economic modelling tools to support infrastructure project preparation, as well as major policy and programme decisions. This advance was made easier by the improved availability and quality of administrative data.

Text box 4: Data Use for Planning and Policy during Cape Town's Drought

During Cape Town's 1-in-300-year drought (2015–18), the CCT needed to engage with its own data in new ways to support decision-making in a time of crisis. The CCT used real-time data on water usage, dam storage, and pipe condition to internally monitor water levels, target maintenance, and drive down consumption. In addition, a critical component involved sharing and visualizing data on water supply and consumption to keep Cape Town's population informed and spur collective action to conserve water. Two tools were developed to enable this:

The CCTs Water Usage Map ('Green Dot Map') *reflected household water consumption levels and neighbourhood water usage, using the CCTs water billings data. Households would be awarded "green dot status" for exceptional water saving, enabling peer-to-peer monitoring and competition to motivate behaviour change (reducing overall household water consumption).*

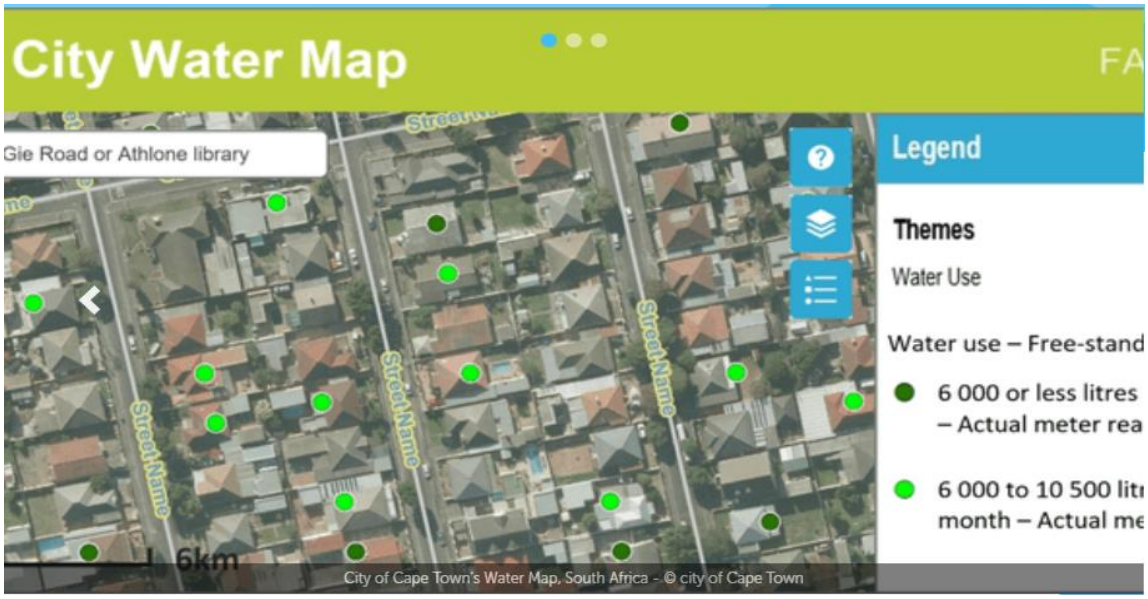


Figure 4: Snapshot of the Green Dot Map

The Water Dashboard *reflected the CCT's weekly dam storage percentages, weekly dam level changes, and average daily water production. The data was sourced from existing administrative data from the Bulk Water department. This innovation continues to be used for weekly updates in 2023.*

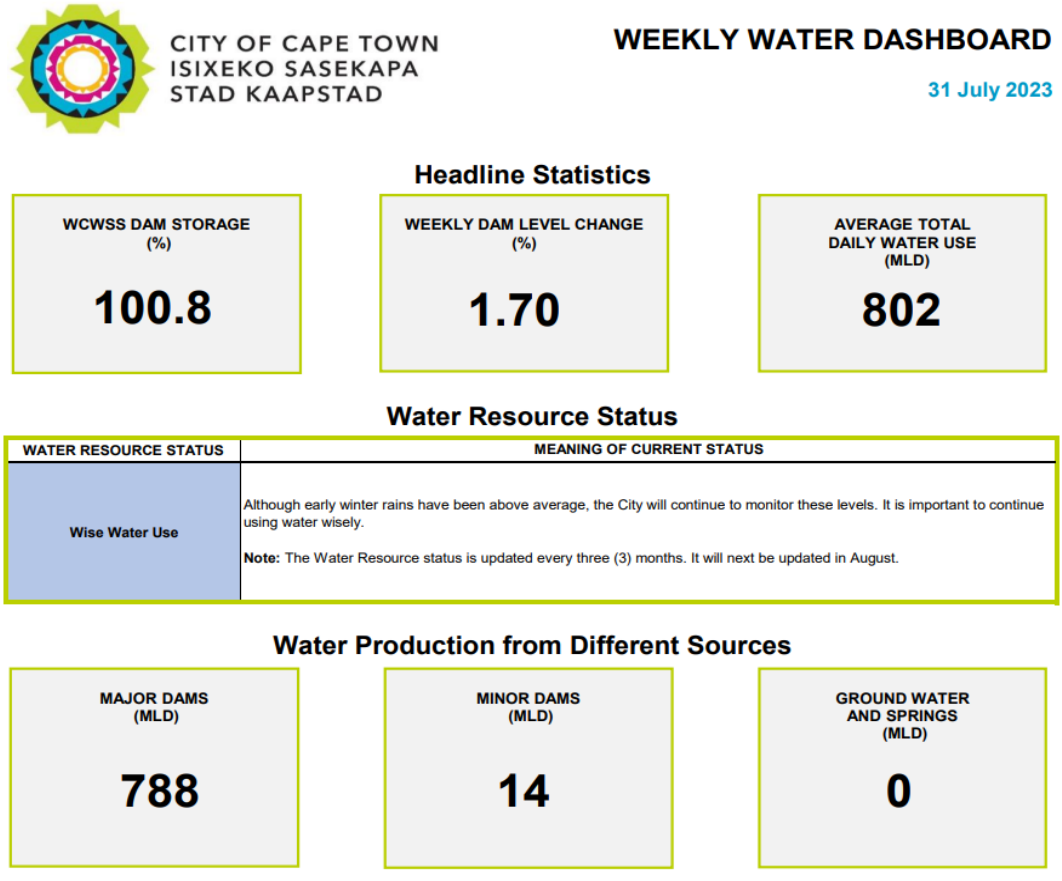


Figure 5: Snapshot of the Water Dashboard

Building these tools required cross-departmental collaboration between Water, Communications, Legal, Finance, IS&T, and IKM's corporate GIS unit. CCT's Research unit, in consultation with other CCT colleagues in various departments including Policy, Spatial Planning, and Data Science helped to review approaches, source data, and adapt indices to inform planning alongside external research partners, and also assisted the Water and Sanitation department with their large volume of research requests at the time.

By 2018, as a result of these initiatives, the city's consumption had reduced from 1.2 billion MLD to 516 MLD, making Cape Town the first city globally to reduce consumption by more than 50% in 3 years. Subsequent research has also shown that the DayZero communication campaign (including its data sharing and visualisation efforts) was more effective at encouraging water saving than changes in water tariffs. Furthermore, by harnessing and monitoring detailed, high-frequency data, the City was able to increase the efficiency of existing water infrastructure. For example, pressure reducing valves (PRVs) were used to provide data on pressure and flow in order to inform and optimize the operation of the water infrastructure. Through the use of PRVs, demand and leakage within a distribution system was managed more efficiently by reducing pressure in a discrete zone. The installation of PRVs provided valuable data that helped to improve reliability of service delivery, decrease infrastructure damage and water losses, and forecast and budget for repairs.

The drought highlighted the complexity and importance of spatially explicit, real-time data access, the value of visualization, and the issues with combining different data sources. Data sharing was crucial to CCT's success in reducing water consumption to a level that allowed CCT to make it through the drought. The drought also illuminated data challenges, such as data sharing between departments, data ownership (some key data sets were owned by contractors), and the translation of technical data nuances for public communications.

These challenges and lessons are key informants to this Data Strategy.

3.2.2 The internal Data Strategy

The City's internal Data Strategy was drafted and approved in 2018. Having built on over a decade of data readiness work using IKM and IS&T, the strategy offered direction and grounds for embedding data into City practices, attempting to push the organisation beyond reporting and towards critically engaging with data for prediction or impact evaluation.

The City's internal data strategy, was developed in response to several challenges in terms of data availability, collection, preparation, validation and utilisation, which impacted on the decision-making processes. It was aimed at addressing several barriers to realizing the full potential value of data including data isolation, lack of visibility across the organization, poor

data quality, uncertain accountability, lack of consistency in data security, and a lack of capabilities in data analytics.

The internal Data Strategy identified six pillars associated with existing challenges the CCT faces in the data environment. These are listed in Table 5. The CCT Data Strategy implementation plan created project teams to prioritise its six pillars. The Data Coordinating Committee (DCC) was established to govern the implementation of the internal data strategy, and the first Chief Data Officer (CDO) was appointed to chair the committee.

Table 4: Internal Data Strategy Enablers and Objectives

ENABLER	Objectives
Valuing Data:	An organisation that **recognises the value of data** for strategic and operational decision making.
Data Collaboration:	An environment where data is treated as a **shared public good** – to be used in the ways that best enable the CCT to meet the ever-evolving needs of its customers.
Data Partnerships:	Harness external **expertise and support** to enhance CCT's technical data capabilities in a cost effect way, as well as to gain insights through the utilisation of CCT data.
Data Capabilities:	Drive data-driven decision making by enhancing the data generation, management and analytics **capabilities of CCT staff.**
Data Architecture:	A sustainable and enabling **data architecture** that is Forward Looking and able to support the growing need for new environments (both internal and external) for data sharing and analytics.
Data Governance:	**Improved data governance** to ensure one source of the truth, taking into account the need for **data quality, integrity and security.**

Thus, the strategy was designed to enhance collaboration around data, improve accessibility, establish data quality assurance capabilities, define data roles, and build data analysis capabilities to enhance evidence-based decision making.

Strategy implementation gained momentum when the CCT secured three-year Technical Advisory[10] services with the United Kingdom Foreign, Commonwealth and Development Office (UKFCDO) and executed by a consultant consortium. Data Strategy provided an opportunity to improve the overall maturity of the CCT data system. Specifically, among other things,

[10] More information can be found here: *https://www.globalfuturecities.org/resources/use-case-approach-city-cape-towns-data-strategy*

preliminary data governance arrangements were established, maturity assessment conducted, and a capacity matrix was created while capacity and skills were developed.

'Use cases' were used to demonstrate value whilst learning rapid, valuable lessons about what needs to change. These were demand-driven test use cases that exemplify how the principles within the internal data strategy result in improvements in planning and service delivery in the lives of the residents of Cape Town. The tools and processes developed during this three year collaboration showed the potential impact of embracing data science, advanced analytics and economic analysis to solve CCT challenges and bring multiple stakeholder groups into the fold of evidence-based decision making. What follows is an overview of the CKAN Data Flows and Storage Platform, and the Informal Settlements Use cases.

3.2.2.1 Open Data Flows and the use of CKAN

At the time of developing the use cases, the CCT did not have standardised ways of managing data, resulting in users of data (internal and external to CCT) not being able to find data they need for their analyses. As described above, the CCT Open Data Portal fulfilled the function of providing data to public data users but together with the challenges listed, there was no means to enable receiving data for external parties. The prospect of a data catalogue that would overcome these challenge, that can be used internally and externally, serves as data discovery tool, and provide data on a granular level, thus posed an exciting opportunity for the CCT.

The CCT explored various data catalogues that can harvest data, pulling data from various data storage technologies, federating the data into one searchable catalogue, with metadata and data management functionalities.

The value of this approach is:

- Discovery – reduce the time spent on discovering datasets.
- Evaluation – ability to evaluate the datasets suitability for an analysis use case.
- Access – provide a seamless user experience while securing privacy, and compliance sensitive data.
- Collaboration – encourages collaboration internally between CCT staff and also externally with academia and industry.

The CCT decided to pilot the CKAN data porta[11]. CKAN is an open-source data management system for powering data hubs and data portals. It can be used to share private data with

[11] *Prior to the collaboration with the UKFCDO, a collaboration between CCT and researchers at the Abdul Latif Jameel Poverty Action Lab (J-PAL) and University of California, Santa Barbara (UCSB), enabled a technical support to the CCT with purpose of exploring software options. The CKAN solution was a product of this collaboration.*

both internal and external stakeholders, with an ability to set different levels of data access to different users.

CKAN implementation was accelerated and advanced, and used in the CCT's COVID-19 pandemic response and has become a key enabler for storing and sharing critical data internally and externally in a secure and governed manner. (The COVID-19 pandemic necessitated both immediate operational decisions as well as adjustments to longer term strategies and plans. Due to the significant amount of data collation and two-way transfer required between the various response stakeholders, the CKAN data portal was deployed to act as a central transfer node for crisis decision making that needed to happen at a pace where much is potentially unknown and high risk, and so support provided optimisation and efficiency to mitigate risk and improve the speed and accuracy of decision making. Text box 4 provides insight into the CCT COVID-19 Pandemic Response and the Role of Data).

Further work is underway to determine how best to scale the data catalogue by ensuring data management and governance in the CCT is standardised. Importantly, this platform meets the needs of internal users, lowering the cost of in-house data discovery and exchange. At the same time, internal efforts are being leveraged to streamline data sharing with external researchers whose research proposals have been vetted by CCT staff. Using the platform, researchers are able to search for available data, securely download data to which they have been granted access, and contribute metadata, analysis code, and other files.

The benefits of streamlining the data sharing process include an increased volume of research on policy relevant questions, greater accountability of researchers to report findings and share cleaned data or new data sets with CCT, and a more secure and standardized approach to transferring data to researchers.

It is important to note that CKAN is also just a tool and without the necessary structures, technical people and processes to standardise and automate data management the CCT will be no better off than what we currently are. This is key consideration for the Data Strategy and informs our strategic enablers and commitments over the next several years.

3.2.2.2 Data equity for informal settlements

People need to be counted and included in strategic planning, operational responses, and planning considerations. Formally housed people and people who consume formal services are often linked to a specific data set, whereas informally housed people are not. The informal settlements use case focused on addressing this gap.

The business problem to be solved in this use case was the absence of a mechanism to connect household level survey data to the Informal Settlements spatial data set. This problem was related to the way in which survey data is captured and the general inefficiencies in the

approach. Challenges included the paper-based survey methods, and manual, labour intensive process resulting in poor data quality with errors and repetitions while the external datasets were not integrated in CCT systems. This traditional approach is both costly and cumbersome to the City.

A key factor of this use case was identifying a more cost effective way to collect quality data, outside of the traditional system. The objective of understanding how such systems can be developed and deployed within the City's existing data architecture.

The approach sought to develop a system to capture and store social survey data electronically and link it directly to the spatial data set (settlement polygons and individual dwelling points). This way, the City is able to seamlessly produce Certificates of Residential Address based on data obtained from social surveys to link people to dwellings within settlements. This provides reliable and accessible social data that informs planning, delivery and social accountability.

The use case highlighted a tension between an historic method of procuring 'long term' data solutions for 'permanent' use within the City versus experimenting with alternative products and modifying these to suite a particular purpose, at a particular time, at low cost and with relative speed. A key realisation was the friction between the security parameters and approvals process, and line departments need to obtain the most accurate validated information for transversal use across the City.

The City learnt the value of alternative approaches to collection quality data which can be used to better allocate scarce resources. A key consideration going forward, especially when data is collected in collaboration with external partners, is ensuring data collection methods used, are aligned with internal systems.

Importantly, this use case showed the possibility of increased integration of planning, delivery and management for inclusive development while considering voices and experiences of women and other vulnerable groups. In addition, it showed the possibilities of improved decision-making and governance that uses data and evidence to understand drivers and barriers to development, including that of inclusive and climate resilient development, and the triggers and impacts of shocks.

Text box 5: The role of data in the COVID-19 Crisis Response[12]

The Covid-19 pandemic presented an unanticipated crisis for the CCT. Building on its experience of the drought, staff and working groups who were involved in the overall data

[12] *Cole et al. (2021). Managing city-scale slow-onset disasters: Learning from Cape Town's 2015–2018 drought disaster planning, International Journal of Disaster Risk Reduction.*

and research effort were drawn on for the CCT Covid-19 response. The COVID-19 Coordination Committee (CCC), chaired by the City's Chief Data Officer, was established, and meant that the CCT had a single executive leading the response, while working across the City's traditional silos. This allowed data to be used across the silos and a consistent data-driven decision approach to prevail.

Tools were developed to enable understanding of information flows, risks, and trade-offs across the broader planning efforts. A dedicated data workstream was established to identify and access relevant data, and the data science team together with other research units were drawn on in modelling various scenarios. Data sharing agreements were also urgently established between the City and the provincial Western Cape Government's Department of Health (WCGov DoH) to access needed data.

A number of City data tools were designed for organisational readiness, responsiveness, and continued service delivery:

- *A clinics model to depict clinic capacity, overflow, and impact on staff and materials.*
- *An HR capacity data tool to reflect daily staff status (i.e., sick leave, working from home, in quarantine, etc), which informed line manager decisions.*
- *An operational dashboard to depict and compare international and national pandemic statuses, behaviour insight data from resident social media responses, and an overview of business continuity from a multi-department co-ordinating committee, which informed senior leadership decisions.*

Data projections also informed related policy and planning responses for example:

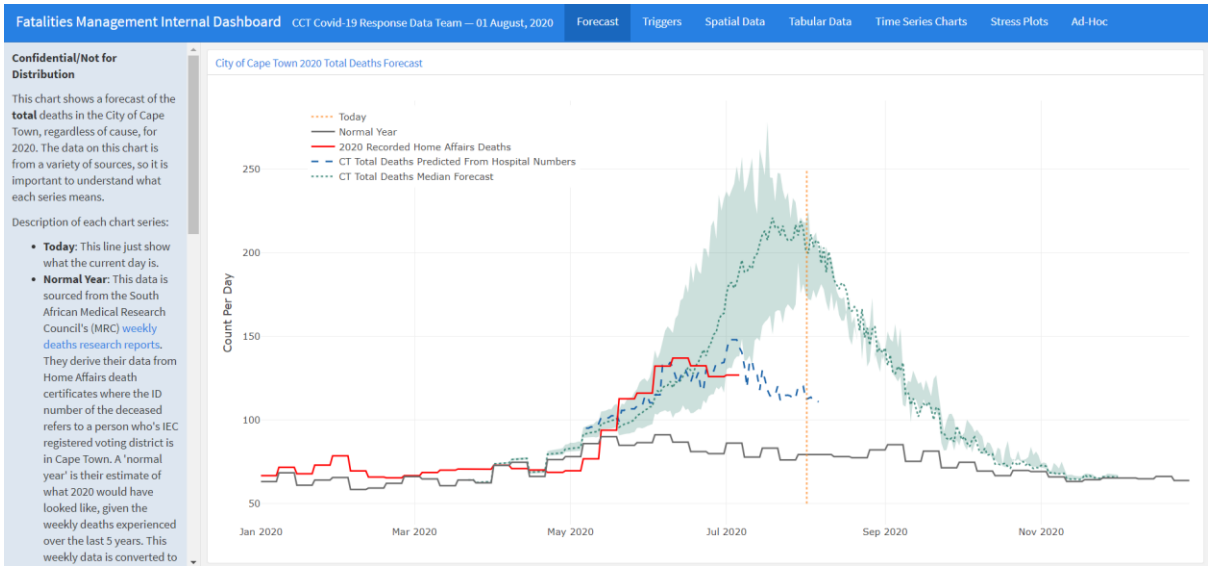


Figure 6: Snapshot of the Fatalities Management Dashboard

*A **fatalities management dashboard** was developed in May 2020 for internal use representing overall deaths - weekly, daily, per neighbourhood, and overall for Cape Town. Total deaths in a normal year were also recorded and compared with Home Affairs recorded deaths and with deaths predicted from hospital numbers, in order to get a more holistic idea of the pandemic fatality impact.*

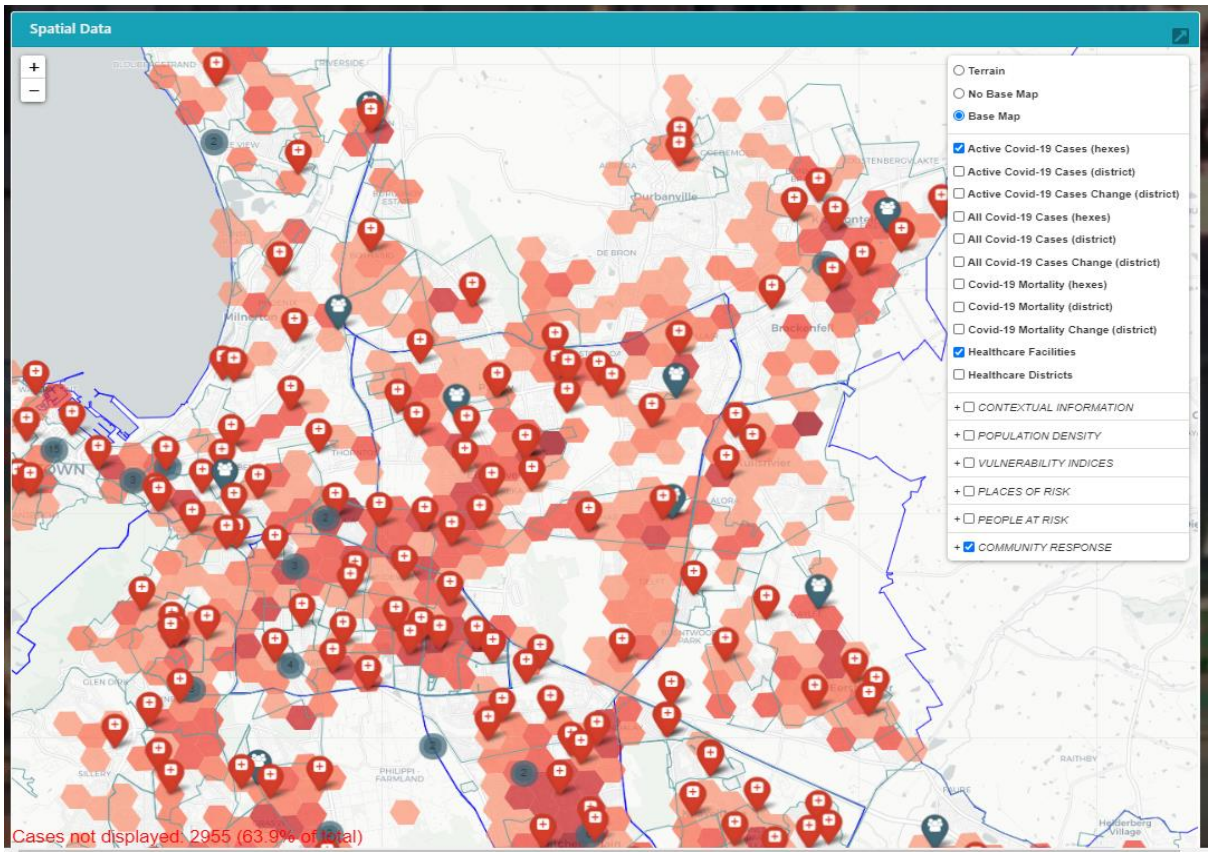


Figure 7: Snapshot of the Vulnerability Viewer Dashboard

*A **vulnerability viewer dashboard** and map were built in July 2020, layering Covid-19 data with spatial data on vulnerable people (>55yrs) and places (e.g., old age homes, high density informal settlements, high risk economic nodes) at the lowest level of disaggregation, as well as the SVI. This allowed for targeted, strategic interventions in those areas, building on the vulnerability work done during the drought.*

3.2.3 Lessons from data led crisis responses

From the CCT's experience during the COVID-19 and 2018 drought events, data played an integral role in guiding operational decision-makers. These processes stemmed from establishing foundational elements such as prior investment in flexible data systems, capabilities, data maturity and, crucially, partnerships across related internal and external government departments. A key lesson was the important role in developing data agreements

across multiple stakeholders to minimise barriers to sharing data for analysis and tool development. The value-add to rapid response in times of crisis together with efficiency gains for resource allocation has been clearly demonstrated through the City experience and inform this Data Strategy. Another critical is having sufficient specialist data expertise inside the CCT and the value of flexible open source tools that enable agile response. Building a data ecosystem that is agile and enables the sharing of data as described in chapter 1, is a priority for this strategy.

3.3 Review of progress against the internal Data Strategy

The internal Data Strategy identified six pillars associated with existing challenges the CCT faces in the data environment. These are listed in Table 5. The CCT Data Strategy implementation plan created project teams to prioritise its six pillars.

The CCT has made significant progress against the associated objectives and the whole organisation has advanced in several directions with data. The lessons learnt during the operationalisation of the internal strategy reveal and highlight where more work remains to be done. This is discussed at a high-level in table 6.

Table 5: Internal Strategy Progress and Lessons for the Data Strategy

Data Governance and Data Management To date, the CCT has invested a lot of thought and attention into data governance, and roles and responsibilities across the data lifecycle. On reflection, there are relatively strong data governance protocols in place, for example, in the form of user permissions for certain data sets. Furthermore, the CCT has a strongly structured custodianship organogram, with defined roles and responsibilities. In recognition of the need for compliance to National legislation, specifically the Protection of Personal Information Act (POPIA), through the Data Strategy implementation process, a POPIA Workstream was established and has successfully managed City's POPIA processes, and put in place structures, processes and procedures to prevent and respond to breaches of personal information. The City has developed a compliance framework for the Protection of Personal Information Act (POPIA) which includes conducting a risk assessment, providing awareness and training, developing procedures for accessing personal information, and appointing information officers.

As the CCT develops in maturity, it is expected that stakeholders will become more familiar with data governance but it was found that this can be accelerated and enhanced through training and awareness programmes.

Lessons that inform this Data Strategy

- The need for clear roles and responsibility between IS&T and those responsible for leading the use of data within the CCT.
- The need for an escalation and resolution mechanism where there are competing priorities between IS&T and data users.
- The need to focus and work towards aligning the various department roles and processes through data governance training, and standardisation across the CCT.
- Strengthening internal data governance and enhancing understanding of data privacy, security, and POPIA to protect against the misuse of data and to ensure compliance.
- Data quality is still a challenge, and setting up effective, quality data pipelines should be a major focus moving forward.
- In reviewing the datasets against the detailed business procedural documents, data management processes are not fully standardised.
- Establishing clearly defined data management and governance roles, which are fully aligned best practices to improve decision making at business and corporate level.
- Strengthen capabilities of data custodians to ensure the capacity to describe the data in detail and the ability to identify data-driven partnerships and cross-departmental collaborations.

Capabilities

Although, capacity building and capabilities enhancements was a central component of the data strategy to ensure sustainability, the overall data maturity and capacity of the CCT is inconsistent. There are pockets of excellence, while other parts of the organisation have lower levels of data maturity.

Findings suggest that individual departments lack broad capacity for team members to fully engage with and explore their own data. Capabilities often rest in a set of individuals that would need to enable others to access and engage with the data. Consequently, many individuals are unable to explore departmental data easily and thereby bring insights and opportunities. This is juxtaposed to the reality that in the absence of a resourced, central data analytics platform, staff cannot fully engage and explore of data.

Training interventions however have increased the capacity to create, access, manage and use data, and departments across the CCT are making better decisions and, ultimately, better serving residents and business.

Lessons that inform this Data Strategy

- The organisation as a whole does not have a uniform starting level of maturity and will not shift consistently along a maturity journey.
- A purposive set of interventions, defined in a programmatic approach, are needed to drive maturity, at varying levels and to varying degrees for enhancement across the CCT.
- Development of data capability should aim to be user-focussed and practical.
- The need for a centralised data team and experts to lead the capability journey, as a compliment to the role of IS&T in building enabling systems.
- Value may be added by approaching this in an integrated manner rather than along silos or traditional business lines.
- Identifying gaps and consideration for the inconsistencies in access to resources and data capabilities of the organisation, and address these gaps via various options, including upskilling, training and recruitment.

Data Culture and Analytics

Experience has demonstrated that a CCT-wide value of data, which enables and supports the strategy, will be critical to realising the desired outcome. Foundations for valuing data have been laid in parts of the organisation but work remains. A key takeaway is the need to acknowledge that every level of the CCT, from leadership to operations, technical and strategy departments as well as line departments, play a role in enabling evidence based-driven decision making.

Progress has been made with the support of the Data Science, a unit established with the aim of assisting the CCT in harnessing its administrative data resources, gaining valuable insights, and mitigating uncertainty associated with shocks and stressors. A data analytics environmental was created for line departments use, which allows specialists to do data analytics. Scaling and institutionalising this environment is a key focus of this Data Strategy. At the same time, many line departments have matured their data capacity and capabilities.

Lessons that inform this Data Strategy

- The need for a clear leadership in data engineering and data analytics for the organisation located within Future Planning and Resilience Directorate.
- Although the CCT has a well-defined stewardship organogram, with clear roles and responsibilities, custodians require support to ensure consistent describing of the data in detail.
- There is value in continually reinforcing the roles and responsibilities for data across its lifecycle to remove any uncertainties about accountability, over time.

- The critical role of analytics in the evidence-based decision-making can be further institutionalised. While it exists across the organisation, the CCT has varying levels skills in the analytics domain.
- Supporting and enabling pockets of excellence in line departments.
- There are many strategic and operational decisions which require that the CCT be able to access, manage and analyse data that is generated both internally and externally to CCT, as well as to use analytical skills and tools that are external to CCT.
- The importance and opportunities in new and emerging data collection and sharing techniques, in terms of automating data pipelines and improving data quality.
- Automating data flows and enabling central, accessible sharing of data greatly supported decision-making, especially in times of crisis.

Data Architecture

The CCT has learnt a great deal regarding Data Architecture in implementing aspects of its internal strategy. That said, existing barriers to technological innovation, remain which has had an impact on the pace at which the CCT is able to make data accessible for multiple users and uses. There are several standalone systems lacking integration within the CCT, which means that it is not possible to have sight of and access data from a single entry point.

Within the CCT the collection and extraction of data (which is not standardised) requires institutional knowledge spread across departments. Currently, there is no single-entry point to data for either humans or machines which can hamper the ability of the CCT to rapidly load and assimilate data from a variety of sources to conduct analysis. Within the CCT the bulk of mission-critical data is well-centralised however it is not necessarily accessible for easy data sharing and use.

Lessons that inform this Data Strategy

- The need for a digital foundation that enables the access and use of data, per the spirit of this strategy.
- Potential systems and processes that allow the selection and introduction of appropriate and safe new technology in the city environment. (Currently, the CCT as an entity is, by necessity, cautious around introducing new technology into the environment.)
- Automating data flows and enabling central, accessible sharing of data to support decision-making.
- Ensure wider understanding of data privacy, security, and POPIA across the CCT.
- Potential for a single entry point for accessing data from disparate data sources.
- Integration with the CCT's authentication and authorisation management system to assist with the governance, management and security of data.

Internal Data Strategy Implementation

On reflection of the progress made, two central challenges relating the implementation structure have emerged. The first is that all of the members occupying roles in the DCC are carried out by the CCT staff in addition to their formal roles. The second relates to the need for adequate authority, decision-making power, transversal collaboration and, especially resourcing, which the DCC does not have.

The CCT has learnt valuable lessons about what it might take to operationalise this Data Strategy over the past 5 years, and is now ready for a governance structure that is able to drive a process as the CCT grows in maturity so the CCT is able to scale the use of administrative data and extract its full value. This Data Strategy sets out a set of arrangements for a fit-for-purpose governance structure to support the vision for data to be used to deliver better public services to all people in Cape Town.

The progress on the objectives provides a solid foundation for the Data Strategy contained in this document. It also reveals the key challenges that need to be addressed. These can be summarised as:

- Having sufficient data specialists across the CCT;
- Inadequacy of integrated data systems and processes across departments;
- Data capabilities and governance maturity is uneven among departments;
- Moving beyond reporting to critically engaging with data for prediction and measuring impact to a culture of understanding and improvement;
- Data gaps on informality and others, such as economic data on firms; and
- Achieving the right balance between open data and data security and protection.

The need to systemically manage data in the same way as any other critical asset, such as infrastructure, and collectively embrace the notion of data as a shared asset remains. This is essential for strategically managing data to contribute to the CCT's priorities. These realities have informed the Data Strategy.

3.4 Considering Different World Views

In pursuing evidence-based decision making, differing views on how to achieve the strategic imperatives of the organisation have become apparent. In building the CCT approach to data, we have undertaken a review of these imperatives together with necessary safeguards to mitigate unintended consequences. This is demonstrated in Table 6 below.

Table 6: Strategic Imperative and Requisite Guardrails

Strategic Imperative	Safeguard for consideration	City Approach
The CCT is committed to progressively increase the sharing of data with external partners while incorporating diverse sources of external data.	Data security is imperative to protect against unauthorised use and disclosure.	• CCT will develop partnerships agreements that build accountability around data quality, structure and data security. • The City will endeavour to make data available in a machine readable format following open standards.
The CCT needs to be enabled to extract the full value of data through data analytics as a key component of new evidence-based approaches.	The CCT needs to protect the privacy and interests of residents as well as the security of the CCT's critical infrastructures.	• Continue expanding role and capability of analytics across all departments in the organisation, and encourage experimentation and innovation. • Adopt an ethical approach to data, stringent security measures and robust data governance. • Build a central leadership and technical capability on data engineering and analytics
The CCT needs to establish good data practices throughout the organisation through the establishment of standards and rules.	The CCT needs to not place undue burden on those tasked with data generating, maintenance and analysis	• Well-maintained datasets become valuable when they are used to inform decisions and achieve operational efficiencies. • A fit-for-purpose governance structure and set of governance arrangements will ensure the operationalisation thereof. .
The CCT wants to be on the cutting edge of analytics, enabled through openness, agility and speed. A key components is the opportunity to remain flexible with continual iterations, rendering the best outcome through a learning and evolutionary process.	The CCT acknowledges that maintaining structure and reference to generally accepted approaches and methodologies is important for continuity and integrity of the IS&T system.	• The CCT accepts that for the organisation to become data driven, a data culture that ensures the enhancements of capabilities and upskilling of staff, is critical. • The CCT will structure itself to deliver on the growing needs of the organisation. • The CCT will explore processes and technologies that enable it to be agile and responsive to changing needs and circumstances, fostering innovation and efficiency in its operations. • The CCT will where possible rely on internal resources and capacity to realise the objectives of this strategy and carry out analytics.

4. The City of Cape Town's Aspiration for Data

The CCT aspires to extract the full value of data to be a truly evidence-based organisation, and a data-driven local government, which uses data to:

a. Gain better operational insights so that the IDP priorities, and CCT strategies and plans can be delivered and challenges can be overcome, and enable quick and sufficient responses to shocks and stresses;
b. Anticipate what the future of Cape Town may be like so that future challenges and uncertainties can be reduced and managed to avoid future pitfalls and that opportunities are anticipated and leveraged thus improving long term planning;
c. Enable better decisions in the face of increased complexity and time pressures;
d. Make use of advanced analytics to serve the needs of the CCT, and residents and businesses of Cape Town;
e. Leverage more external data sources to overcome existing data gaps; and
f. Enhance and broaden its current artificial intelligence initiatives and capabilities.

4.1 Guiding Principles

The Data Strategy will directly respond to the CCT's strategic and operational needs by ensuring that data is accessible, staff are enabled and empowered to use data confidently and safely, and all stakeholders can access data in a manner that is useful and productive. At the same time the CCT commits to ensuring the private information of Cape Town residents is protected. In an effort to realise this, the following guiding principles have been identified:

a. **Data Transparency** means that City officials should be transparent about what data they collect and how it is used. This includes being clear about the purpose of data collection, who has access to the data, and how it is stored and secured.

b. **Data Explainability** is an important related value and considers the extent to which analytical methods and algorithms can be explained in an accessible manner. To meet this standard, the City should be able to explain what the data does and how it was designed in plain language to a non-expert audience

c. **Data Accountability** means that there are effective governance and oversight mechanisms for any data initiative. Accountability therefore relies on mechanisms for scrutiny, data governance in relation to long term oversight and public scrutiny are built into data project lifecycles.

d. **Data Fairness** relies on data being obtained through fair means and used fairly. This relies on the purpose of data use being clear, and consent obtained prior to any

analysis occurring. It is crucial to eliminate any data initiative's potential to have unintended discriminatory effects on individuals and social groups

4.2 Data vision and desired outcomes

As per the IDP, the CCT will continue to enhance policy and strategy formulation and strategic planning to ensure sound and effective policy decisions and strategy-led budgeting. The intention is for all CCT planning and budgeting to be based on sound evidence to ensure effective and responsive service delivery to residents. The CCT will allocate public resources in a way that achieves maximum public benefit through better use of economic analysis, data and data analytics for long-term planning and decision-making with a particular focus on infrastructure management, operations efficiencies and crisis response.

Data Strategy Vision

Data is used to deliver better public services to all people in Cape Town

Data Strategy Desired Outcomes

Data is leveraged to its full potential contributing to evidence-based decision making to ensure the CCT is fully effective and responsive to residents and businesses needs

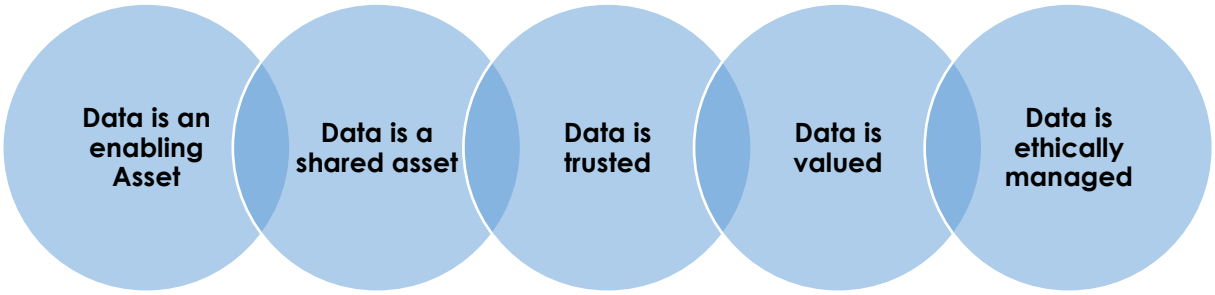


Figure 8: Desired outcomes of the Data Strategy

a. CCT data is a collection of public assets that are managed in a way to **enable maximised public benefit** and interest, and the organisation's vision and mission.
b. The CCT data **belongs to the organisation as a whole** and should be available for use across departments. By **sharing its data internally**, the CCT aims to promote transparency and accountability in government, and generate useful insights by using

data sets together. CCT data is also **shared externally** with residents, businesses, organisations and all who want to use CCT data to better themselves and Cape Town.

c. CCT data is **managed ethically** throughout their lifecycle to eliminate bias, ensure fitness for use, data management and protect the privacy of our residents.
d. CCT data is **trusted** through governance arrangements that improve data quality and overall **integrity** and **reliability** of data and the analytic practice.
e. **Data is valued** for evidence based decision making. The CCT leadership will seek the evidence with which to base decisions on and demonstrate an appetite for evidence driven decisions by encouraging and supporting the development of competencies that harness data and new technologies.

In realising these desired outcomes, the CCT believes in the value of leadership within this domain. It is imperative that stakeholders have clear roles, responsibilities and guidelines on data, including a clear lead on driving the use of data in decision making. This requires leadership, data expertise, and the empowerment of data stewards and line departments across different parts of the organisation that actively use data for operations and analytics. Executives and senior managers must champion analytics work in order to drive a culture of evidence based decision making to achieve operational efficiencies.

4.3 Theory of Change

Over time, the CCT has effectively executed diverse data initiatives to showcase the principles and concepts of its internal data strategy. There is still work to be done to improve data use and management, but at the same time the CCT is ready to institutionalise and scale advancements.

Extracting the full value of data for the CCT, will depend on responding to foundational challenges which include culture, data maturity and the need for embedded capabilities within departments. As a change process, harnessing the value of data will require leadership, fostering a culture of understanding, action, improving and innovation, attracting and empowering talent and staff across the CCT functions, and an operating model to deliver on the desired data outcomes.

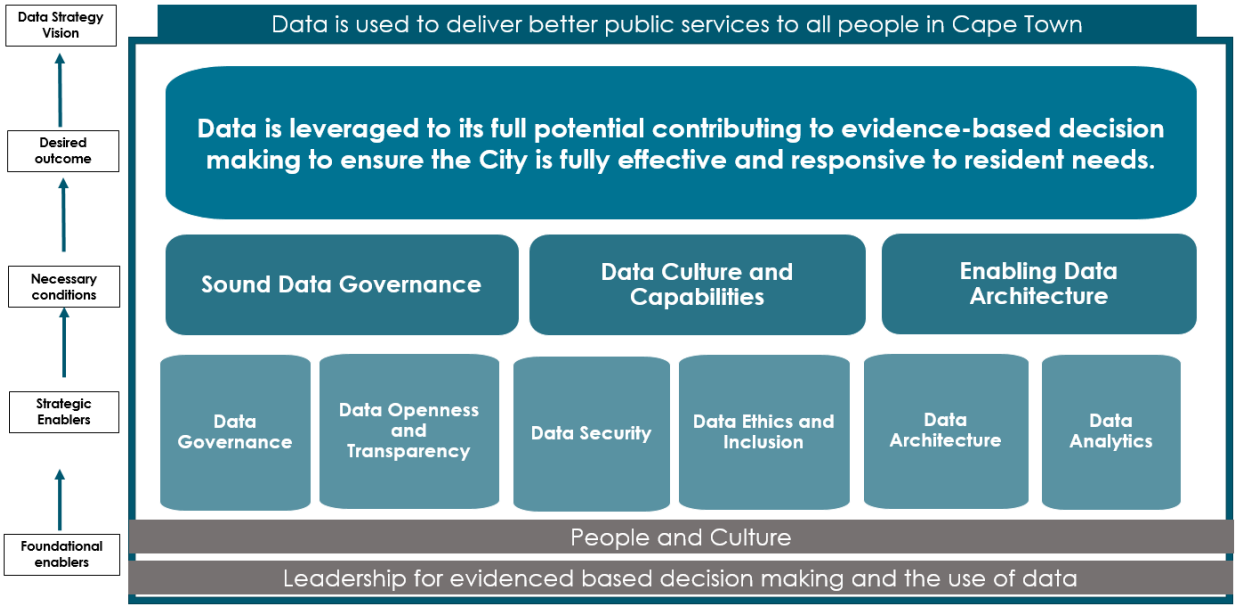


Figure 9: Theory of Change for the Data Strategy

4.3.1 Necessary Conditions

The theory of change highlights three essential conditions that need to be established to achieve the vision and desired outcomes.

a. **Sound data governance:** an established set of CCT-wide, implementable guidelines, rules and standards ensures that the organisation has a transparent and shared understanding of what good practice is with regards to data generation, management and analysis as well as a governing structure with requisite authority. This includes clear roles, responsibility and assigned authority for leading data analytics in the CCT.

b. **A data culture that prioritises capabilities**: improved understanding and capacity to gather, structure, analyse, use and share data to drive an improved culture of evidence based decision making and operational decisions on behalf of our residents. This culture is underpinned by an established organisational value for data, as a shared asset, which is supported by well-resourced, in-house data analytics capability.

c. **Data Architecture**: a robust data architecture is necessary to unlock the full potential value of data and enable the analysis of a wide range and large volume of data, in increasingly diverse formats and closer to real-time.

4.3.2 Strategic Enablers

The necessary conditions listed above will position the CCT to realise better outcomes for all stakeholders, residents and businesses of Cape Town. However in order to establish these three conditions, a number of changes need to be affected. This strategy has identified enabling

objectives to focus efforts on for the next 3-5 years in a phased way, which have the greatest potential return on effort.

These enabling objectives are outlined in the theory of change (figure 4) and elaborated on in Text box 3.

Table 7: Summary of Enabling Objectives for the Data Strategy

Data Governance: Data governance exists at the right levels to ensure that data is managed holistically as a strategic asset
Data Openness and Transparency: Progressively make relevant data sets available to the public and streamline sharing of datasets and analysis between CCT departments.
Data Security: Secure personal information and sensitive data, including critical infrastructure data, and ensure that its use is appropriate and justified
Data Ethics and Inclusion: Ethical and responsible data environment which provides insight into the realities affecting vulnerable groups and that supports greater inclusion in data driven decision making and service delivery
Data Architecture: a data architecture that enables analysis of a wide variety and volume of data, in different formats, and increasingly closer to real-time
Data Analytics: Establish analytical practices and tools, through leadership and technical capacity, to encourage analytics.

4.3.3 Foundational Enablers

The foundational enablers identified are essential factors that will provide the necessary support for the implementation of the Data Strategy. These enablers are the foundation upon which the activities of this strategy rest and are crucial to creating the conditions for success.

Leadership	People and Culture
Leadership for evidence based decision making and the use of data is critically important and needed, with clearly defined roles, responsibilities, accountabilities, and decision-making authority to govern data. Importantly, the CCT will rely on existing decision making processes, to inculcate improved data use for evidence based decision making, to build data maturity and drive implementation.	Having the right **people and culture** plays a central role in ensuring that the CCT is able to leverage the full value of data for decision making. . It is recognised a critical mass of skilled professionals and developed embedded technical capabilities across the organisation is foundational to the success of this Data Strategy. A functional model that the CCT has experience in, is utilising a 'hub and spokes' model wherein expertise and capability is centralised while providing

	support, leadership and quality assurance, together with a community of practice, to distributed nodes within the organisation.

4.4 Structure of the remainder of the Strategy

The 6 strategic enablers form the heart of the Data strategy and are linked to commitments from the CCT to realise the desired outcomes and objectives together with list of the immediate priorities. Some actions build on existing work of departments across the CCT that have shown leadership and made significant progress. This Data Strategy can increase the visibility of, and scale, these efforts for broader use. Other actions propose new activities that complement work currently underway or target specific challenges that need to be addressed to make way for future improvements.

Many of the actions focus on the first steps needed to drive coordinated change across the CCT. Consultation, governance and development processes will be undertaken, as appropriate, within each initiative as work advances. This work will require collaboration and input from across the CCT and will be successful only if all departments align and integrate the direction and intended outcomes of the areas of action into their own practices. More detail on departmental responsibilities will be included in the implementation plan that will follow the Strategy.

The Data Strategy enablers are detailed in the following chapters. Each enabler is linked to a commitment from the CCT and lists several immediate priorities for action. Each chapter describes the strategic enabler, offers an indicative assessment of the maturity within the CCT, (see table 7 for a guide guide), articulates the CCT commitment and lists a series of immediate priorities.

Table 8: Key for the indicative maturity of the CCT per enabling objective

Maturity of the City	
	**Expansion of existing programme:** The programme is currently being implemented but will be expanded or scaled up, with budget and resources for implementation allocated.
	**Implementation in progress:** The programme is currently being put into effect, with budget and resources for implementation allocated.

5. Data Governance

Outcome: Data governance exists at the right levels to ensure that data is managed holistically as a strategic asset	**CCT Maturity Status**
	Expansion of existing programme

CCT commitment: *to manage data as a strategic asset using well defined plans and processes throughout the data lifecycle, driven through a governance structure.*

Data governance systems will allow the CCT to define business processes, roles, and responsibilities needed to institutionalise its commitment to data quality and use, while protecting the data assets of the CCT.

To achieve this commitment, the CCT will establish and enhance the following data practices:

a. Formalisation of data governance arrangements, with the various data roles being well defined and an established network of data stewards.
b. Clear roles and responsibility between IS&T and those responsible for leading the use of data within the CCT together with escalation and resolution mechanism.
c. Data is regularly assessed against an established data quality framework, and formal data quality initiatives will be planned and implemented.
d. Master and reference data frameworks are implemented to improve data quality.
e. A documented organisation data model and data flows, and a data inventory maintained.

The CCT is committed to building a shared understanding of what good practice is with regards to data generation, management and analysis across the organisation.

Immediate Priorities

5.1 Develop a Corporate Data Management System

Develop a common language around data management practices across departments and establish a consistent set of governance roles and responsibilities.

5.2 Enable the enhanced sharing of data through clear governance standards and processes

Establish a common classification system for data to ensure that the sharing of data is enabled through appropriate levels of access and permissions, whilst still safeguarding personal information as per the requirements of POPIA.

5.3 Govern analytics products to build excellence and capabilities and promote use

Encouraging high value analytics experimentation through processes that promote excellence and the ability to share analytical products.

6. Data Openness and Transparency

Outcome: streamline sharing of datasets and analysis between CCT departments, and progressively make relevant data sets available to the public	**CCT Maturity Status** **Implementation in progress**

CCT commitment: *to create an environment where data is treated as a shared public good to be used in the ways that best enable the CCT to meet the ever-evolving needs of residents and businesses.*

There are two key aspects to the Data Openness and Transparency Strategic Enabler:

- **Accessible data within the CCT:** Datasets must be open, available and in a usable condition for sharing internally in order to combine them into a product for decision-making, and better equip CCT departments to analyse and respond to uncertainty and complexity.
- **Open data:** The CCT will progressively make appropriate data sets open and accessible for businesses, academics, and private citizens to encourage the development of tools and services that can complement public infrastructure.

To achieve this commitment, the CCT will establish and enhance the following data practices:

a. CCT data is openly shared between CCT departments as a default, subject to restrictions in terms of agreed security, privacy, and confidentiality guidelines.
b. The CCT enables broad access to relevant dataset via the open data portal -
    i. According to governance standards in line with international best practice and the POPIA.
    ii. According to prescriptions and in the spirit of PAIA which enables residents to access information held by the City.

Immediate Priorities

6.1 Enable public access to datasets via enhanced open data platform

Enhance the Open Data platform for publication of select datasets that are easily accessible to the public. The platform will feature those datasets frequently requested by the public, proactively identified and loaded. The platform will feature intelligent, scalable, and robust search methods for data access and extraction, keeping pace with the growing library of datasets. The CCT will continue to work with the users of the open data portal to prioritise data sets which should be made available on the open data portal over time.

6.2 Strengthen data partnerships and opportunities for collaboration

Enhance mechanisms for the integration of external datasets with CCT datasets for analytics purposes, including the development of a platform to enable access to CCT data that is not classified as 'open'[13] and enable contributions to CCT data.

6.3 Collaborate with the analytics community

The CCT will make use of open source[14] technologies and techniques in order to foster collaboration with the global tech community on data initiatives.

6.4 Ensure data-related tenders are implementing open standards

All CCT contracts with service providers must outline the provider's roles and responsibilities regarding data sharing and maintenance associated with the service.

7. Data Security

Outcome: Secure personal information and sensitive data, including critical infrastructure data, and ensure that its use is appropriate and justified.	**CCT Maturity Status**
	Implementation in progress

CCT commitment: *to secure the personal information of residents who entrust their sensitive information to the CCT, and take actions toward full compliance with POPIA.*

CCT employees both directly and indirectly handle and have access to sensitive data. Robust data security practices is critical for privacy, which goes hand in hand with building trust, both within the CCT, and with residents that the CCT serves. It is important that residents can be assured that their privacy is respected and protected, by both the CCT and any third parties who may provide services to CCT or access the CCT data.

To achieve this commitment, the CCT will establish and enhance the following data practices:

a. The training of employees on the importance of protecting personal information and guidelines for working with personal information is underway.
b. Maturation of data governance processes and structures to oversee the securing of the CCT data.
c. Protection of privacy and enabling people to exercise their rights in terms of POPIA.

Immediate Priorities

[13] *Some data may only be available at an aggregate level due to confidentiality requirements and compliance with POPIA.*
[14] *The CCT will develop an Open Source Framework and operating model to ensure sustainability and appropriate governance.*

7.1 Manage requirements of and move toward full compliance with the POPIA processes and procedures

The CCT to build on existing work undertaken to ensure compliance with POPIA and align that with work done in data governance broadly, and the data roles in particular. The CCT will educate employees on the importance and regulations of the POPI Act and ensure that the CCT can live up to the trust of its own employees and residents.

7.2 Robust data security practices

Data security is managed through a risk-based approach that secures data based on:

a. data sensitivity awareness and the level or risk associated with various datasets,
b. user roles (including need to know), and responsibilities, and
c. data classification with associated levels of access and permissions for data users and systems.

7.3 Limiting data breaches and data loss

The CCT will seek to systemise data loss prevention processes and be responsive to data breaches, with learning and enhancements made based on lessons learnt. The CCT has established a data breach incident response team to respond to the loss of personal information.

7.4 Levels of access and permissions

Data deemed to be sensitive will be securely managed in line with different levels of access and permissions attributed to data users and systems. Residents need to be assured that their privacy is respected and protected, by both the CCT and any third parties who may provide services to CCT or access the CCT's data.

8. Ethical Data Management and Inclusion

Outcome: An ethical and responsible data environment which provides insight into the realities affecting vulnerable groups and that supports greater inclusion in data driven decision making and service delivery.	**CCT Maturity Status**
	Expansion of existing programme

CCT commitment: *mainstreaming responsible and ethical use of data across its lifecycle and ethically acquire, process, disseminate, use, store and dispose of data and maximising the potential of data to bridge divides.*

As we become increasingly connected and data-driven, the CCT recognises it is important to include ethical considerations and issues pertaining to inclusion in the Data Strategy.

Data ethics is about the value judgements and approaches we make when generating, analysing and disseminating data. It is useful in this context because it provides the norms of behaviour that promote appropriate judgements and accountability when acquiring, managing, or using data, with the goal of protecting residents rights, minimising risks to individuals and society, and maximising the public good [15].

At the same time, the CCT is committed to creating a data environment which provides insight into the realities affecting vulnerable groups and that supports greater **inclusion** in data driven decision making and service delivery.

In considering data ethics and inclusion, the following definitions are useful:

- **Bias**: bias can lead to disparate access to opportunities on the grounds of several human characteristics such as race, age, gender or other grouping characteristic and should be discouraged.
- **Unfairness**: Unfairness in data analysis is to use data in a way that creates or reinforces bias. Analysis makes conclusions or predictions in ways that deny due process, deprive people of property or liberty (even temporarily) without transparency or human review, or make decisions that appear self-indulgent or unpredictable or aid undemocratic governance is perceived as unfair.

The Data Strategy has the potential to be a significant enabler of inclusion by allowing a better understanding of the different needs, capacities and opportunities that different social groups have. It will also support the design of appropriate interventions that address these different needs and enable opportunities. However, in order for the implementation of the Data Strategy to enable inclusive policy, programmes and projects – the way in which data is generated, governed and analysed by the CCT needs to have a strong citizen-focus and specific attention must be paid to women and other vulnerable groups.

To achieve this commitment, the CCT will establish and enhance the following data practices:

a. Ethical data lifecycle management where the concepts of bias, fairness, transparency are considered.
b. Transparency about what data is collected and how it is used, including being clear about the purpose of data collection, who has access to the data, and how it is stored and secured.
c. Accountability for the use of data and the ability to demonstrate that data is being used in a responsible and ethical way.

[15] The Open Data Institute (n.d.): *Data Ethics Canvas* – Retrieved from: https://theodi.org/article/data-ethics-canvas/ (Accessed online 23.)

d. Greater understanding of women and other vulnerable groups through data disaggregation and other practices that offer insight on vulnerability.
a. Enhancing capabilities to use disaggregated data in programme design, implementation and monitoring.

Data ethics has a direct relationship with sound data governance and it is through data governance that ethics and inclusion will be practised.

Immediate Priorities

8.1 Confirm and adopt ethical standards, principles and process

Ethical data management forms a key component of data governance and will be enacted through the governance structures established to ensure that supportive ethical guidance is available and ethical standards are known and practised across the CCT. These will be driven through the CCT's policies and standard operating procedures, as well as through the oversight of the governance structures.

8.2 Ensuring data is representative and reliable, and all stakeholders are visible

Reliable data is important for analysis to be representative and inform decision making optimally. An important aspect is establishing means to gather granular data to understand the needs and experiences of women and other vulnerable groups.

8.3 Empower staff with the skills and knowledge in the ethics of data

The CCT has invested a lot of thought and attention to data governance and roles and responsibilities such as data custodians, data owners, data users, etc. These role players are to be routinely empowered with skills and knowledge in data ethics and will formally hold the ethical data management responsibilities in accordance with their relative responsibilities for embedding and mainstreaming data ethics and data inclusion.

9. Data Architecture

Outcome: a data architecture that enables analysis of a wide variety and volume of data, in different formats, and increasingly closer to real-time.	**CCT Maturity Status**
	Expansion of existing programme

CCT *commitment*: *create a sustainable and enabling modern data architecture that supports the sharing of data, and the ability to use administrative data for decision making.*

An architecture which supports the principle of *Data as a Shared Asset*, is central to driving a culture of data informed evidence-based decision-making. This Data Strategy forms the

foundation of and informant for any existing system's data architecture, as well as any future systems' data architecture design and development. There are two guiding principles, which are as follows:

Data is growing exponentially. Innovative and advanced infrastructures are required to combine the potential of data and power of large-scale computing into a coherent capability for data analysis.	The linking of the various systems into a single entry point is necessary to address and enable various data analysis needs of the City.

To achieve its commitment, the CCT will establish and enhance the following data practices:

a. Ensuring a digital foundation that enables the access and use of data, per the spirit of this strategy.
b. Providing the solutions, processes and capabilities which allow for data to be made available to address the analytical and reporting needs of the CCT.
c. Housing data in such a way so as to support easy, shared access and processing to enhance day-to-day efficiencies, as well as advanced analytics.
d. Design a data architecture that considers data and systems resilience, and is able to meet the demands of increasing data volumes and workloads.
e. Flexibility to ensure that the architecture can serve diverse needs.
f. Adoption of open data standards.

By establishing a collaborative data architecture that facilitates cooperation and efficient access to data between internal and external parties the potential exists for the CCT to extract the full value from the data and practice the principle of Data as a Shared Asset.

Immediate Priorities

9.1 Design, build or procure data solutions to ensure operational efficiency and sustainability

Cater for varied data sources and business needs as well as ensure adequate staffing capacity to working with new capability.

9.2 Mature the CCT Data Infrastructure and enable the data practitioners to collaborate

Ensure that data architecture supports easy access to data as well as user-friendly interfaces and enable data practitioners (external to the CCT) to collaborate with the CCT. Through its commitment to a modern data stack with strong user bases and active data communities of practice, data users are able to actively participate in product development by using the product early, suggesting new features, and contributing to improvement in code.

9.3 Practice least privilege

The CCT will exercise the principle of least privilege. This means that a person or system should be given only the privileges and access they need to complete the task at hand and nothing more. Importantly, the CCT will not use this principle to make whole data sets off limits for legitimate analysis to improve outcomes for residents.

9.4 Ensure new systems or applications enable the Data Strategy

The CCT has committed to, through the CAR Data and Analytics tender, to ensure that software or hardware provided through this tender meet two functional priorities, which include: Authentication and Authorisation Management (Identity Life Cycle) and Network Accessible Programmatic Interface – API.

10. Data Analytics

Outcome: analytical practices and tools, through leadership and technical capacity, to encourage analytics.	**CCT Maturity Status**
	Expansion of existing programme

CCT commitment: *to enable evidence-based decision-making through the implementation of analytic practices.*

Principles for data analytics practice in the CCT:

To achieve this commitment, the CCT will establish and enhance the following data practices:

a. Capacitate and enhance the skills of subject matter experts in line departments.
b. Develop guidelines on basic data analytics and analytics standards.
c. Making datasets available to the organisation according to prescribed legislative parameters, together with metadata and ownership information.

d. Attracting and developing data analytics skills in order to unlock the next level of data value for the organisation.

Immediate Priorities

10.1 Approving the various administrative instruments for data analysis and analytics

Enabling the implementation of analytic practices through approved administrative instruments is a priority and the CCT will develop the appropriate guidelines and standard operating procedures that enable this.

10.2 Enhancing CCT data analytics skillset and computing capacity

Enabling the delivery of a range of analytic insights required by all levels of decision-makers, including specific skills and capacities relating to spatial and economic analysis.

10.3 Encouraging the reproducibility of analytics and the use of development environments

By building analytic processes that encourage an analytics culture and allow technical successes to be repeated over time the CCT can enable the sharing of the methods and products of data analysis. Encouraging experimentation is a crucial part of data analysis, where data users, specialists, and domain-specific professionals can experiment with approaches to data preparation, reporting, and analysis.

10.4 Improve and systematising the way the CCT develops, tests and embeds analytical tools

The CCT will establish ways of working with that aim to operationalise analytics using an agile and collaborative approach between the various analytics teams across the organisation, and ensuring expertise is available to enable, support and upskill where needed.

10.5 Leveraging spatial data for meaningful analytics

Effective spatial analysis and capturing geo-tags for all CCT data whenever possible, at the lowest possible geographic level, to allow for the analysis of data at different geographic scales, such as suburbs, wards, and other spatial units.

11 Foundational Enabler: People and Culture

Outcome: A critical mass of skilled professionals and developed embedded technical capabilities existing across the organisation underpinned by a data culture that recognises the value of data	**CCT Maturity Status**
	Expansion of existing programme

CCT commitment: *The CCT is committed to valuing data as a shared asset, and seeks consistent improvement of service delivery through continual data enhancement, experimentation and inquiry.*

To achieve this commitment, the CCT will establish and enhance the following data practices:

a. Addressing inconsistent data maturity across the organisation through tailored interventions.
b. Building capabilities in a variety of data roles to effectively deliver the CCT Data Strategy objectives.
c. Practice of data management as asset management that requires effort to manage and maintain.
d. Utilisation of data for strategic and operational decision making and risk management across the organisation.
e. Cultivating an appreciation for the value of data for all stakeholders who participate in the data lifecycle.
f. Transparently presenting data analysis results and explaining analytical choices.

Through this, the CCT is seeking to enable a culture that is open and shares by default, to create an environment that recognises all CCT staff as data agents, and is able to hire, retain, cultivate and empower the right talent and capacity.

Immediate Priorities

11.1 Building foundations of skills, tools and data flows amongst specialist users

CCT aims to enable domain specialists to work more efficiently and effectively with data, and help the organisation to extract value from data. Importantly, the CCT will aim to fill the skills-gap by acquiring scarce skills, and ensure the capabilities exist to support an agile data organisation.

11.2 Building culture and skill through actual projects and decisions

Develop data systems and processes across departments and building evidence to assess the impact of its policies and programs. A data culture will be development through skills transferal between technical data specialists to data staff in departments, working closely on their data needs and projects.

11.3 Improving access to training and recruiting more of the right skills

The CCT will prioritise the training and nurturing of staff to ensure that they value the power of data as an enabler for progress. The CCT will provide staff with the necessary knowledge, access, and tools to inform better decision-making. CCT will develop data competency frameworks that describe the competencies required in terms of technology and tools, as well as the particular skills needed to excel in a data role or pursue a career in data.

11.4 Resourcing core data analytics functions to provide support across the organisation

The CCT is committed to resourcing core data analytics functions to provide support across the organisation. This is underpinned by the principle that where possible and more cost effective, the CCT will do analytics in-house, subject to budget considerations.

12 Foundational Enabler: Leadership and Governance Structure

Outcome: CCT leadership seeks the evidence with which to base decisions on and demonstrate an appetite for evidence driven decisions, supported by a fit-for purpose governance structure.	**CCT Maturity Status** **Expansion of existing programme**

CCT commitment: *the CCT is committed to providing clear leadership to coordinate the effective use of data to drive improvements to service delivery.*

In realising these desired outcomes, the CCT believes in the value of leadership within this domain. It is imperative that stakeholders have clear roles, responsibilities and guidelines on data, including a clear lead on driving the use of data in decision making. This requires leadership, data expertise, and the empowerment of data stewards and line departments across different parts of the organisation that actively use data for operations and analytics. Executives and senior managers must champion analytics work in order to drive a culture of evidence based decision making to achieve operational efficiencies

Implementation of good data management requires an agreed governance structure, which provides the clear guidance and leadership to make strategic and tactical decisions related to the Data Strategy. This is the governance body for the Data Strategy and Implementation plan, responsible for implementing, monitoring, and reviewing the progress made against the strategy. Importantly, achieving the objectives of the data strategy requires organisational change, responsiveness and efficiency in decision making. The governance structure considers the following as key:

- Respecting the authority and accountability of the City Manager and relevant Executive Directors (ED) for decisions within their areas of responsibility.
- Respecting the authority and accountability of the ED responsible for Department IS&T and knowledge management for decisions relating to IS&T systems and governance and IKM.
- Respecting the role of Future Planning and Resilience Directorate for corporate leadership on strategy, advancing evidence-based decision making and long-term planning.

12.1 Data Governance Committee

The Data Governance Committee (DGC) provides the strategic guidance for data in CCT by driving improvements across all business units and serves as the ultimate advisory authority on data for CCT. The DGC prioritises and makes proposals, for approval through the relevant CCT processes, on data policies, standards and initiatives against budget and

resource constraints. The DGC is also expected to advise on strategic business investments to support business growth pertaining to data governance and related initiatives as well as to monitor effectiveness of governance on a continual basis.

The DGC is the forum where differences or disputes are escalated for resolution. Finally DGC will guide the work of the Data Working Group (discussed below) to ensure priorities and projects are strategically determined and have leadership support.

The DGC is chaired by the delegated Chief Data Officer (CDO), with director and management level representation from various business units and corporate functions. Representation may include IS&T Directorate executive participation, and data management will be represented as well as representation from other CCT governance bodies that are stakeholders in data decision making. Importantly, it will include a lead specialist advisor on data architecture and data engineering who can advise the CDO, Chief Digital Officer and DGC.

12.2 Role of a Chief Data Officer

The City's CDO will provide a central point for City-wide leadership on the use of data in evidence based decision making and the implementation of the Data Strategy. This includes data management, governance and integration. Thus the CDO will provide technical and strategic leadership and advise the City Manager and executive team on data strategy related matters. The role of CDO will be appropriately resourced to be able to influence data related organisational strategic decision making and planning activities. This is will be separated from the role of Chief Digital Officer. The CDO will serve as the interface on data strategy related matters with the Mayor, City Manager and Executive team. The CDO makes decisions after consultation with the DGC.

The CDO will be assigned authority, by the City Manager, to:

12.2.1 Access to all data (according to prescribed security standards)

This includes prescribing the requirements of the system to ensure data access and data quality to enable analytics.

12.2.2 Lead on data use and maturity, and support data analytics in the CCT

The role of the CDO is to lead on data and drive enhancements throughout the CCT through development of appropriate administrative frameworks and quality assurance, as well as through the establishment of a community of practice that builds excellence.

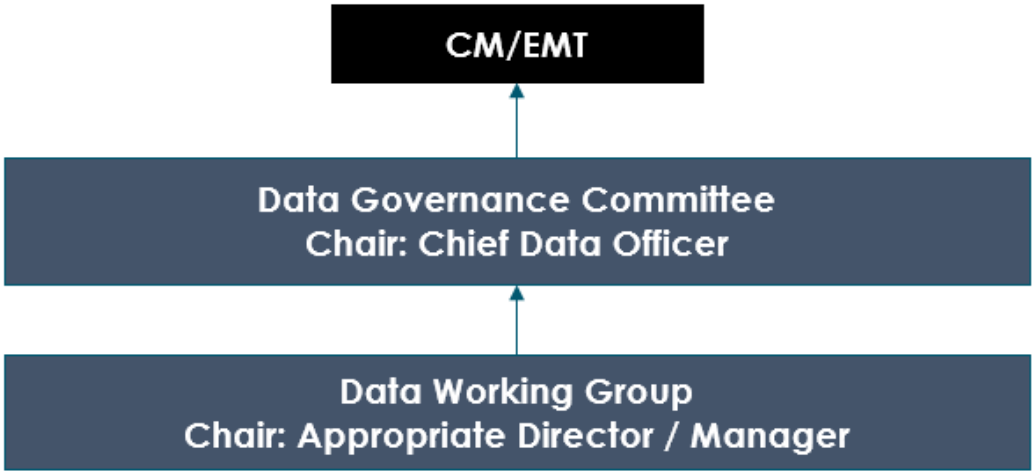


Figure 10: Governance Structure

12.3 Data Working Group

The Data Working Group (DWG) is chaired an appropriate Manager or Director and consists of individuals who are knowledgeable about CCT's data. These individuals can be business leaders and managers, data custodians, curators, stewards and owners. Membership of the working group also includes a Data Quality Lead.

With the guidance and support of the DGC, an implementation plan, with phases will be developed by the DWG which attributes timelines and resources to the principles and activities outlined in this strategy. It will seek to address the need for any additional resources or enrichment of existing posts that might emerge. The implementation of these needs is subject to budget being available.

12.4 A Supporting Operating Model

The DGC and the DWG together form the basis of a delivery structure. Importantly, the realisation of the CCT priorities, data strategy objectives and value proposition, requires a delivery and operating model that can execute and operationalise the data strategy. It is expected that the CCT needs a delivery and operating model that can transform the way the CCT works with data across multiple data sources and organisational structures. To this end, the CCT will establish an operating model that allows the CCT to institutionalise and bring to scale the advancements made to date. The specifics of this operating model will be determined and defined through administrative processes, by the City Manager.

Immediate Priorities

12.5 Establish the Data Governance Committee

The DGC Terms of Reference will be drafted, and the DGC will be established as part of the Data Strategy Implementation Plan to ensure strategic guidance for data, as approved by the CCT.

12.6 Establish the Data Working Group

The DWG Terms of Reference will be drafted, and the DWG will be established as part of the Data Strategy Implementation Plan, as approved by the CCT.

12.7 Appoint the Chief Data Officer

A CDO will be appointed by the City Manager, with requisite delegated authority, to provide a central point for leadership on the use of data in evidence based decision making and the implementation of the Data Strategy.

13. Risks to the Data Strategy

13.1 Diminishing Prioritisation of Data

Business functions tend to assign a low priority to the creation and maintenance of data in relation to their other activities (keeping the water flowing and the lights on). This is especially during stressful and time constrained periods or situations.

Risk Impact: Data quality deteriorates as it receives less attention, resulting in data quality and completeness backlogs. As these backlogs grow the amount of effort to catch up while maintaining day to day activities overburdens staff, eventually resulting in the abandonment in attempts to close the data quality gaps.

13.2 Responsibilities associated with Data Custodianship and Management

The role of data stewards is a big responsibility**.** Experience from other organisations that attempted to implement the role of a data stewards into their line departments have shown that very few individuals are willing to accept the responsibilities and additional work that go with this role (especially on top of their existing responsibilities).

Data stewardship is troublesome to establish and maintain leading to inconsistent levels of data quality between departments.

Risk Impact: If the data custodian roles are forced upon certain individuals, they tend to form a negative association with the role and the initiative as a whole that could have a negative impact on the implementation of the data strategy due to insufficient resource allocation, with requisite skill or capacity.

13.3 Reducing the administrative burden of governance compliance

Adherence to the governance principles embedded in CCT data strategy should not become a deterrent to the sharing of data. CCT should seek innovative ways to reduce the administrative burden involved in the creation, submission, validation, description and access of data, both internally and externally.

13.4 Data Management – lack of specialisation across CCT directorates

With current placement of critical business information function, in many cases as subcomponent of other departmental division, decisions regarding information and systems are being taken by managers not equipped to make those decisions. The City will continue to build embedded capacity and capabilities while enhancing centralised advisory capacity through appropriate structures.

13.5 Overburdening of staff responsible generating data

Collecting data takes effort and preparing the data analysis takes even more effort. The more detailed the data required, the more effort is required to collect and prepare the data. Relying on office and field staff to generate and prepare data to a specific level of detail while maintaining the status quo with their current job assignments may easily overburden them with additional administrative work. Some staff might be required to generate and manage data as part of their after-hours work at difficult times in difficult places (maintenance/ emergency staff).

Risk Impact: Overburdened staff may neglect parts their data management duties or their line operational duties in an effort to meet certain commitments. This will jeopardise overall data quality and may create change resistance.

13.6 Over-engineering the problem

Investing in data expertise, tools and systems that are underutilised and do not lead to operational efficiencies for CCT is a risk. Already there exists a number of tools and platforms activated but which remain underutilised because their use was not encouraged through change management activities, such as training.

Risk Impact: Without due consideration given to change management, the implementation of this Data Strategy will result in fruitless and wasteful expenditure which does not aid in realising data as an asset for the City

13.7 Going it alone

CCT needs to seek out constructive partnerships in implementing this Strategy, in order to increase the data analytics capacities available without the City needing to acquire these skills in-house at great expense.

Risk Impact: The extent to which CCT is able to realise data as an asset for decision making and driving operational efficiencies is tied to its ability to acquire all of the data analytics tools and capacities itself. This will lead to delays in realising the benefits of data.

Annexure A: Strategic Alignment

The City's Data Strategy is a critical enabler of the following strategy outcomes and this revised strategy will outline the data governance, architecture and capabilities required to realise these outcomes:

IDP (2022-2027)

- 1.5 Consolidated land pipeline and release programme:
    - 1.5.A. Data-driven land management initiative
- Excellence in Water, Waste and Energy service delivery programmes: Data-driven asset maintenance project
- 4.1 Utility Business Model Reform programme
- 5.2 Safety technology programme:
    - 5.2.A Technology safety partnerships project
    - 5.2.B Incident, crime and emergency detection project
    - 5.2.C Digital evidence management project
- 8.2 Informal settlements upgrading programme:
    - 8.2.B. Informal settlements data improvement initiative
- 16.3 Evidence-based decision-making programme
    - 16.3.B. Data-driven performance management initiative
    - 16.3.C. Economic analysis project

Inclusive Economic Growth Strategy

- 5.9 Leveraging data as an economic asset
    - C) Investigating the long term transition of the Open Data Portal into an inter-operable and extendable platform via an interface, allowing individuals, private organisation sand other spheres of government to interact with it, process applications through it, and even contribute data to it.
- 7.5 Developing a user friendly job-seeker database
    - A) Consolidating the various databases into an integrated and user friendly electronic system, accessible by job-seekers and employers alike.

Human Settlements Strategy

- **F. Data systems mapping in informal settlements:**
    - 2.2.6 (b) The City must establish a governance model between the role players of spatial mapping, informal settlements upgrading, data capturing on the

Housing Needs Register, and informal settlements record keeping that ensures the coordination of data relative to informal settlements.

Climate Change Strategy and Action Plan:

Research, data and knowledge management are a foundational element of the City's framework for action within the climate change strategy – two specific actions with implications for the City's data environment include:

- Action 14.2 Unlock energy transition with data – in support of a more distributed energy system
- Action 21.1 Compile a baseline carbon footprint measurement for the operations of
- the CCT Urban Mobility, SPE and Human settlements directorates