

South African Cities Network (SACN)

Open Data Almanac for Cities

Concept Paper

14 September 2015

DRAFT

Glossary

COGTA	Department of Cooperative Governance and Traditional Affairs
ISO	International Standards Organisation
NT	National Treasury
PAIA	Promotion of Access to Information Act 2 of 2000
SACN	South African Cities Network
SALGA	South African Local Government Association
SoCR	State of the Cities Report
SOCs	State Owned Companies
UCLGA	United Cities and Local Governments of Africa
WCCD	World Council on City Data

Contents

1	Introduction	3
1.1	The Concept outlined	3
1.2	Purpose of a Central Database	4
2	Points of Departure	5
2.1	Key assumptions	5
3	Defining the SACN Open Data Almanac	6
3.1	Premises for making City and other Data Open	8
3.2	Principles of an SACN Open Data Almanac – Outcome/Output	9
3.3	Principles of Open City Data for SACN – Input	12
4	Structure of Future System	13
4.1	Operational Mandate to SACN	14
4.2	Overarching Data Collection Process	16
4.3	System Requirements	16
4.4	Key Stakeholders & Systems in City Data	17
5	Funding & Financing the Open Data Almanac	20
5.1	Seed Funding is currently available in the Data Acquisition by Cities	20
5.2	A Five Year Value Proposition	20
5.3	A Base for Leveraging Direct Funding	20
6	Towards achieving Open Data Almanac of the State of the Cities Report (SoCR) 2021 – Approach to Implementation	22
7	Conclusion	23
8	Annexures	24

Figures

Figure 1: The Components of the SACN Open Data Almanac.....	7
Figure 2: Concept Structure of the SACN Open Data Almanac	13
Figure 3: New streamlined data collection process	16
Figure 4: SACN Data Repository - Conceptual Design	17

Tables

Table 1: Proposed Categories of 'Open' Data	7
Table 2: Proposed Categorisation of Data and Metadata on the SACN Open Data Almanac.....	7
Table 3: Outcome/Output Principles of an SACN Open Data Almanac	9
Table 4: Input Principles of an SACN Open Data Almanac.....	12
Table 5: SACN member and non-member municipalities who should participate in the Open Data Almanac	17
Table 6: South African Government role players in City Data	18
Table 7: State Owned Companies (SOCs) and other Entities playing a role in City Data	18
Table 8: Private Data Service Provider Companies	18
Table 9: Private Consulting Companies working with City Data	18

1 Introduction

The South African Cities Network (SACN) has for a number of years, administered the capturing and distribution of data on various statistics and development indicators of different Cities and Municipalities in South Africa. This has mainly aligned with the generation of the State of Cities Report which is compiled and released every 5 years.

The SACN Urban Indicators Status Quo report revealed a number of key findings relating to challenges in the current data collection and reporting environment.

The report found that in general the existing approach to data collection, extraction and reporting was complicated and uncoordinated amongst the different spheres of government. Public, state-owned and private entities that are currently involved in data collection or reporting have or are in the process of developing their own stand-alone databases for their own purposes. These databases are fulfilling certain mandate and/or legal reporting and oversight requirements and are, at times, for public consumption. However, of utmost concern is that:

1. Nearly all these initiatives are based on a range of publicly available data;
2. Are built on the presumption that municipalities must/should and can lead in populating these datasets; and
3. Nearly all initiatives – separately – seek to (re)present some common, comparable view of an element or sector of local government and cities in South Africa.

This approach means that government departments and entities and private sector agents alike are collecting similar data which creates apparent duplications in effort, time and even in expenditure. Often, incongruent and inconsistent findings for the same cities and indicators arise because of differing focuses, purposes and methodologies employed.

In line with the report's initial findings and immediate responses, this report aims to outline the design and possible implementation of a future system for the SACN Centralised Urban Indicators' Database for Metropolitan Cities which may be call the data almanac.

1.1 The Concept outlined

The SACN Open Data Almanac is conceived as a system into which various participating entities and stakeholders will contribute. The purpose of the system is to create an open, shared, maintained and organised database of information about South African cities and large towns.

This paper proposes a plan to design and implement a central open database in which all data collected by, for and about municipalities will be stored, managed and updated. The envisioned SACN Open Data Almanac does not become the custodian of the said data, but it does become the central repository for finding and aligning related datasets. The goal is to grow the Almanac's role into a reporting portal for SACN member organisations.

Public data would be free of charge on the envisioned Almanac, but the intergovernmental reporting and PAIA compliance reporting modules would offer rights managed access only to those who need it, i.e. municipalities and related stakeholders.

The data and metadata stored on this central open database would be coordinated and integrated, but would also be easily governed and monitored by professionals, stakeholders and members of the public using it.

The buy in from SACN member cities and stakeholders for this new system is essential to the successful implementation and operation of the Open Data Almanac, but it is not a pre-condition.

1.2 Purpose of a Central Database

The broad goals of the **SACN Centralised Urban Indicators' Database** is to improve service delivery operations and ease the burden of current reporting requirements. It is hoped that the availability of a public managed and centralised database may facilitate more informed decision-making and contribute to better planned and run cities.

Specifically, the main objectives of SACN establishing the Centralised Urban Indicators' Database for Cities is to:

1. Digitise the data and metadata used to compile the State of Cities Report 2016 and hereafter into an Online Data and Metadata Almanac;
2. Improve ease of access of current and comparable information about South African Cities Network (SACN) member cities; and
3. Ease the burden of intergovernmental reporting and compliance to maintaining public data for South African Cities Network (SACN) member cities by:
 - a. pre-populating as far as possible known national, provincial and other reporting requirements; and
 - b. Supporting the role of the information officer in terms of the Promotion of Access to Information Act (PAIA) (2000).

In fulfilling the above objectives the SACN serves the mandate of its member cities, first, and will seek to harmonise and align its efforts as far as possible with initiatives and directives of key city stakeholders such as COGTA, SALGA and National Treasury, among others.

2 Points of Departure

2.1 Key assumptions

This concept paper assumes that certain foundations actions are or have been resolved in order to establish the rightful place of the SACN Centralised Urban Indicators' Database for Metropolitan Cities.

Firstly, it is assumed that COGTA, SALGA and National Treasury (a Steering Committee of City Stakeholders), at least, have confirmed their envisaged purpose and mandate for the SACN. This would be resolved to identify that the Centralised Urban Indicators' Database is one of the initiatives that would add value to member cities and would be the basis for harmonising and aligning urban data initiatives across South Africa and Southern African government. This mandate would affirm SACN's leading role as the think tank analysing the various elements of cities, especially:

- among metropolitan municipalities (members and non-members);
- with the external environment of cities (WCCD, UCLGA, UN Habitat, etc.); and,
- with inter and intra departmental structures across all levels of government

Secondly, it is assumed that a robust discussion is being facilitated to consistently identify those indicators that metropolitan municipalities and stakeholders deem necessary and relevant to monitor for the express collective purpose of benchmarking and comparisons both locally and internationally. This includes known and open identification, evaluation, review and acceptance procedures.

Lastly, this initiative is assumed to be couched in the vision of an **Open Data Almanac of the State of the Cities Report 2021**. The initiative around a public database of city indicators and data requires a longer term view, preferably tied to the terms of office of SACN board members. It is assumed that the prevalence of a cooperative attitude among the various public organisations dealing with data and indicators has given rise to practical efforts to coordinate the energy and resources already dedicated to this effort among the various cities, if not the various parts of government.

A detailed plan for the Public Almanac of the State of the Cities Report 2021 is in place with key milestones for its development. This plan directs the SACN with its public and private sector partners in developing a comprehensive data almanac; providing information on topics that the cities and stakeholders feel are currently relevant whilst developing the monitoring and reporting regime required to provide real and relevant data. An initial contribution and/or funding plan has been compiled to allow the establishment of the initial system.

3 Defining the SACN Open Data Almanac

The **SACN Open Data Almanac of the State of the Cities Report 2021** is envisioned as a ground-breaking initiative collecting data and metadata for more than 210 measures across fifteen (15) metropolitan and other significant municipal service areas¹.

The purpose of SACN Open Data Almanac is to foster and support a culture of service excellence in municipal government by creating new ways to measure, share and compare performance data, operational practices and improve compliance to open data practices. SACN Open Data Almanac aims to be a source of credible information to assist elected councils, senior management, officials and citizens to understand how their municipality is performing over time and in relation to others.

Defining Open Data

The definition of openness of data and content means, “*Open data and content can be freely used, modified, and shared by anyone for any purpose.*”² Further, it can be defined as structured data that is machine-readable, freely shared, used and built on without restrictions.

The most important points of open data are:

- **Availability and Access:** the data must be available as a whole and at no more than a reasonable reproduction cost, preferably by downloading over the internet. The data must also be available in a convenient and modifiable form.\
- **Re-use and Redistribution:** the data must be provided under terms that permit re-use and redistribution including the intermixing with other datasets.
- **Universal Participation:** everyone must be able to use, re-use and redistribute. There should be no discrimination against fields of endeavour or against persons or groups. For example, 'non-commercial' restrictions that would prevent 'commercial' use, or restrictions of use for certain purposes (e.g. only in education), are not allowed.

The Component Parts

The SACN Open Data Almanac would comprise of four key parts in line with the objectives that the SACN would seek to achieve (Section 1.2 above refers). The parts are a digitised and online data and metadata almanac for the 2016 SoCR; any current and comparable information about SACN member cities; and, two modules to facilitate intergovernmental reporting and public data compliance as per the Promotion of Access to Information Act, 2000 and its applicable regulations.

¹ The envisioned Almanac – when fully developed – will include the full list of SACN Urban Indicators and will represent the data collected by the 8 metropolitan municipalities, namely City of Cape Town; Ethekwini; Ekurhuleni; City of Johannesburg; Nelson Mandela Bay; City of Tshwane; Mangaung; Buffalo City; and 7 large towns of Sol Plaatje; Msunduzi; Mbombela; Polokwane; Rustenburg; Mafikeng and Emfuleni.

² <http://opendefinition.org/>

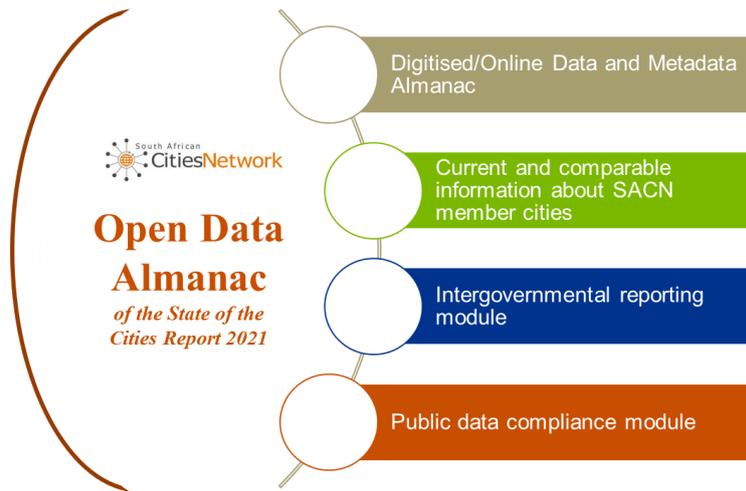


Figure 1: The Components of the SACN Open Data Almanac

These components respond directly to the broad needs expressed by respondents and participants in the SACN Urban Indicators project.

Free and Open Data

The concept of an Open Data Almanac for municipal data does not imply that all data should be free and open to all, but rather that it needs to be regulated and managed through proper governance structures that are put in place. In short, this means that stakeholders that need access to the data need to be given proper access rights that are managed and documented, including the signing of a confidentiality agreements regarding sensitive information, if needed. An added benefit from this regulated process should be that the users of open data can be tracked and managed resulting in informed data management decisions.

It is suggested that the SACN lead the development of the Open Data Almanac by defining “open” as connoting free, unlimited access to those who need it. This implies restricting access to data and metadata to two main categories, namely ‘Free’ and ‘Rights Managed’.

Table 1: Proposed Categories of 'Open' Data

Category	Description
Free	Public data will be shared free of charge without application or registration, and without requiring details of the user. Public data includes information on which public services are run and assessed; policy decisions are based; and other data that is collected or generated in the course of a municipality undertaking service delivery. Where applicable, this data will be anonymised.
Rights Managed	Refers to various access control technologies that may be used to control the use of digital content on the Open Data Almanac in line with confidentiality and user rights policies that may apply to municipal reporting. This means the use, modification, and distribution of works in the Intergovernmental Reporting and PAIA Compliance modules will be controlled.

More categories of information may be added as needed. The information in the various components of the SACN Open Data Almanac may be categorised as illustrated below.

Table 2: Proposed Categorisation of Data and Metadata on the SACN Open Data Almanac

	Free	Rights Managed
--	------	----------------

- | | | | |
|----|--|---|---|
| 1. | Digitised/online data and metadata almanac for the 2016 SoCR | ✓ | |
| 2. | Current and comparable information about SACN member cities | ✓ | |
| 3. | Intergovernmental reporting module | | ✓ |
| a. | <i>Synopsis of Intergovernmental reporting</i> | ✓ | |
| 4. | Public data compliance module | | ✓ |
| a. | <i>Synopsis of Public data compliance</i> | ✓ | |

A Comparative Platform

The purpose of SACN Open Data Almanac is to serve as a comparative platform that does not attempt to provide an evaluation of, or explanation for, each municipality's results. This is proposed in order to support the analytical purpose of the State of Cities Reports. Further, questions about specific results should be directed to the respective municipalities through their appointed Chief Information Officers and/or the SACN Knowledge Manager Reference Group (KMRG) representatives.

3.1 Premises for making City and other Data Open

South Africa has established and elaborate premises for making public and other government information and data open.

The Constitution, 1996

The South African premises for open data are rooted in section 32(1)(a) of the Constitution of South Africa (1996) that stipulates that, "...everyone has the right of access to any information held by the State." This right of access to any information held by a public or private body may however be limited to the extent that the limitations are reasonable and justifiable in an open and democratic society based on human dignity, equality and freedom as contemplated in section 36 of the Constitution.

Making information in the hands of the organs of state publicly available through the SACN Open Data Almanac may be interpreted as alleviating the administrative and financial burden on the parts of the state in giving effect to its obligation to promote and fulfil the right of access to information as contemplated in section 32 (2) of the Constitution.

Promotion of Access to Information Act, 2000

In the same vein, the purpose of the Promotion of Access to Information Act (PAIA) is to give effect to the constitutional right of access to any information held by the state, as well as information held by another person that is required for the exercise or protection of any right.

The motivation for giving effect of the right to access to information is to foster a culture of transparency and accountability both in public and private bodies and to promote a society in which the people of South Africa have effective access to information to enable them to more fully exercise and protect all their rights.

PAIA (2000) outlines some key obligations of organs of state and sets out a framework that the SACN Open Data Almanac should support. This framework relates to obligations on all public bodies to comply to Chapter 2 Publication and Availability of Certain Records. These obligations are surmised as follows:

- 1) The obligation to appoint an information officer as per Column 1 of Schedule 1 or 3 to the Public Service Act, 1994 (Proclamation 103 of 1994);
- 2) The obligations to appoint a deputy information officer that can be delegated (Sec 17);

- 3) The obligation to voluntarily disclose and make automatically available – at least annually to the Minister – certain categories of records (Sec 15) of the public body that are automatically available without a person having to request access in terms of PAIA;
- 4) Duty to assist requesters of public information (Sec 19), free of charge, especially in completing the Forms of Requests (Sec 18)

For the purposes of PAIA, each public body must, subject to legislation governing the employment of personnel of the public body concerned, designate such number of persons as deputy information officers as are necessary to render the public body as accessible as reasonably possible for requesters of its records.

The Urban Indicators project has demonstrated a lack of capacity and resources to fulfil this requirement by municipalities and other organs of state alike. This makes it an ideal opportunity for the SACN Open Data Almanac to fulfil this requirement through cities delegating this function to dedicated resources on a funded, shared services basis among member cities.

3.2 Principles of an SACN Open Data Almanac – Outcome/Output

The principles below would guide the outcomes/outputs of an SACN Open Data Almanac. These are surmised from a range of sources referenced below and serve to identify those conditions that must be achieved for the data and use that comes out of the system.

Table 3: Outcome/Output Principles of an SACN Open Data Almanac

	Principle	Principle Described	Application for SACN Open Data Almanac	Possible Implications or Considerations
1	Responsiveness	Public data policy and practice will be clearly driven by the public and municipal businesses that want and use the data, including what data is released when and in what form	The growth, development and refinement of the SACN ODA will be driven by the public and municipalities who use the data published	Much of the SACN ODA will be evolving to meet the needs of ODA users and funders
2	Machine readability	Public data will be published in re-usable, machine-readable form.	Datasets released by the SACN Open Data Almanac should be stored in widely-used file formats that easily lend themselves to machine processing (e.g. CSV, XML).	-
3	Licensing	Public data will be released under the same open license which enables free re-use, including commercial re-use to increase openness and minimize restrictions on the use of the data.	Maximal openness includes clearly labeling public information as a work of the government and available without restrictions on use as part of the public domain.	There may be a need to review data subscriptions by Cities in order to ensure that maximum value is being obtained from such resources.

	Principle	Principle Described	Application for SACN Open Data Almanac	Possible Implications or Considerations
4	Accessible including mobile accessibility	Public data will be available and easy to find through a single, easy-to use, online access point that is also accessible via mobile devices	The single online access point must be secured and will be the portal for publishing all successive iterations of the SACN Open Data Almanac. Due consideration to be paid to presentation, access, sharing and downloading via mobile devices as well.	-
5	Sense-check & Evaluation	Public data will be published using open standards, and following relevant recommendations of the World Wide Web Consortium (W3C)		-
6	Alignment and Harmonization	Public data from different departments about the same subject will be published in the same, standard formats and with the same definitions	This is in line with the PAIA Sections 15 requirement for cities to publish Form D: AUTOMATICALLY AVAILABLE RECORDS AND ACCESS TO SUCH RECORDS as City of Cape Town and Msunduzi have done.	-
7	Re-usability	Public data underlying the Government's own websites will be published in re-usable form		-
8	Low or No Usage Costs	The Government of South Africa and SACN member cities should release the raw data on the Open Data Almanac site free of charge. This may not apply to specific reports and analyses that may require manipulation of data.	Funding of the SACN Open Data Almanac cannot be driven by user payments, however can be funded from contributions of member cities and other stakeholders in the interests of promoting democracy, transparency and accountability.	Costs of managing the system will be funded by the rights managed components of the Almanac.
9	Use of Commonly Owned Standards	Release data quickly, and then work to make sure that it is available in open standard formats, including linked data forms	This may require that the all data may initially be released as MS Excel format downloads first and foremost, even whilst it is being prepared for certain (re)presentations via tools and/or analysis.	-

	Principle	Principle Described	Application for SACN Open Data Almanac	Possible Implications or Considerations
10	8.License-free	Public data will be freely available to use in any lawful way	Because government information is a mix of public records, personal information, copyrighted work, and other non-open data, it is important to be clear about what data is available and what licensing, terms of service, and legal restrictions apply. Data for which no restrictions apply should be marked clearly as being in the public domain	-
11	Non-discrimination	Public data will be available without application or registration, and without requiring details of the user	Non-discriminatory access to data should enable any person to access the data at any time without having to identify him/herself or provide any justification for doing so.	This means that data categorized as 'FREE' will not require any registration or management of access rights.
12	Encouragement of Data Re-use	Public bodies should actively encourage the re-use of their public data	This may require promotional campaigns to popularize and acknowledge re-use of public data.	-
13	Publish Inventories of Data Holdings	Public bodies should maintain and publish inventories of their data holdings	SACN may follow the publication of data in terms of PAIA Form D.	-
14	Publish Metadata, Meanings, Definitions and Methodologies Applied	Public bodies should publish relevant metadata about their datasets and this should be available through a single online access point; and they should publish supporting descriptions of the format provenance and meaning of the data	This will be done in the Almanac, although attention may need to be paid to how this may be presented for ease of use and accessibility.	-

The above is compiled using the following Open Data Principle Resources:

Government of the United Kingdom – Government Service Design Manual

<https://www.gov.uk/service-manual/technology/open-data.html>

<http://data.gov.uk/library/public-data-principles>

<http://data.gov.uk/blog/public-data-statement-of-principles>

Government of Canada - Open Data 101Open Data 10bsOpen10

<http://open.canada.ca/en/open-data-principles>

Ten Principles for Opening up Government Information

<https://sunlightfoundation.com/policy/documents/ten-open-data-principles/>

The Annotated 8 Principles of Open Government Data

<http://opengovdata.org/>

Open Definition

<http://opendefinition.org/od/>

3.3 Principles of Open City Data for SACN – Input

As there are outcomes/output principle, there are also principles for managing the inputs into the SACN Open Data Almanac. These are surmised below from research and findings from the Urban Indicators' Status Quo report.

Table 4: Input Principles of an SACN Open Data Almanac

	Principle	Principle Described	Application for SACN Open Data Almanac	Possible Implications or Considerations
1	Regular Reviews for Relevance	Develop, test and continually refine data collection mechanisms for impact indicators	Data collection mechanisms for impact indicators such as the QoL/CSS need to be standardized to ensure that it can indeed be utilized and is comparable.	Where relevant, the data may need to be certified through the NSS with the SACN facilitating the process of standardizing the methods of collecting this data.
2	Scaling the Almanac for Manageability	Select a smaller, more practicable and manageable number of indicators that we think are necessary and have valid data attached to them.	The approach to building the online Almanac must be incremental and grow from a strong base.	-
3	Permanence	The capability of finding information over time is referred to as permanence. Information released by the government online should be available online in archives in perpetuity. For best use by the public, information made available online should remain online, with appropriate version-tracking and archiving over time	Metadata is as important as the data itself and the Almanac should focus on ensuring its completion. The Almanac should avoid updating, changing or removing information without any indication that an alteration has been made. An archive of previous data needs to be made available in some way.	-
4	Sense-check & Evaluation	Commitment to growing relevance and extent of data sets available	The SACN Open Data Almanac must be supported by known and rigorous processed for evaluation, prioritizing and focusing the data being released. There are many datasets available, those to be included must shed light on the state of development and the organizations managing large settlements	-

4 Structure of Future System

The urgent need for a central database for collecting municipal data as well as the sharing of data with the municipalities was a concept received well by various national government departments, parastatals and institutions involved with City data.

This report assumes that agreement has been obtained from all cities, stakeholders and other data role players on establishing a central database for urban indicators, as well as how it should be funded, governed and managed.

Concept Structure

Therefore, the structure of the SACN Open Data Almanac has data and metadata at its core. The focus at the core is to establish data governance and monitoring protocols alongside reporting standards, beyond simply collating, organising and presenting the data. This first layer is considered critical for preserving data integrity.

The internal data governance activities will further be supported by regular independent and external verification of the data, metadata and application of data governance guidelines. These will take the form of reliability reports to interrogate the structure and utility of the data and metadata in the Open Data Almanac.

It is expected that training will be required and provide to ensure that data governance, monitoring and reporting standards are met and consistently improved.

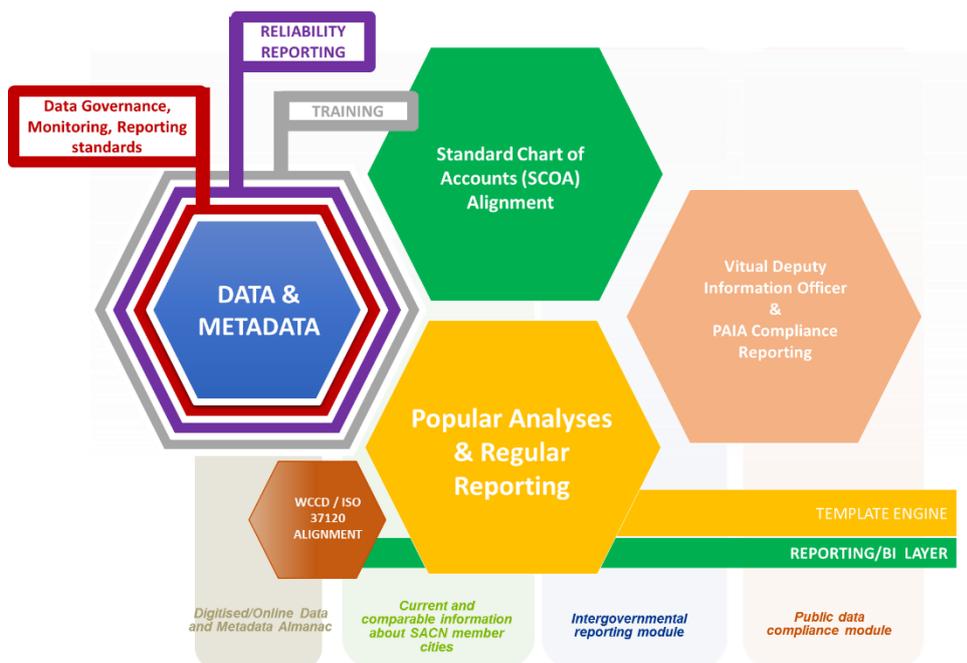


Figure 2: Concept Structure of the SACN Open Data Almanac

Reliability Reporting

Reporting was one of the key findings that came out of the KPMG Urban Indicators Project Status Quo. It was suggested that the sporadic and inaccurate data was a result of irregular reporting, lack of clear processes and inconsistent methods used to arrive at data points.

A data and metadata reliability report is intended to be a quick reference guide to how useful certain sets of data may be for typical purposes.

The core Open Data Almanac is envisioned as a simple digitisation of available data and metadata currently informing the chapter compilers of the SoCR 2016. The core Almanac will also provide some alignment with WCCD/ISO 37120 initiatives.

Whilst the core of the Almanac aims to present current and comparable information about SACN member cities, this will be built on alignment with Standard Chart of Accounts (SCOA) initiatives; popular analyses and regular/regulated reporting

requirements of cities and local government. A reporting/business intelligence layer will need to be developed and deployed in the system to achieve this.

Lastly, the development and deployment of a template engine in the Almanac would unlock the utility of the Almanac for municipal users as it should drive the development of the Intergovernmental Reporting and PAIA compliance reporting modules,

The system developed and managed as above will place the SACN as the ideal virtual Deputy Information Officer as contemplated in PAIA, 2000. PAIA compliance is a known challenge in cities. Therefore, fulfilling this requirement virtually can be done by SACN for contribution-driven participation in this part of the Open Data system.

4.1 Operational Mandate to SACN

This concept paper takes a clear position on the operational mandate of the South African Cities Network given by cities. That mandate points to providing increasingly transparent and robust support to member cities on sourcing, managing and navigating the data and reporting burden they currently face.

Reduce the cost of buying data

The first burden the SACN must address is the cost of buying data. The Cost Benefit Assessment (CBA) report identified the responsibility of the SACN to undertake collective bargaining on behalf of cities with respect to data service providers. The CBA report postulated that conservative collective bargaining by the SACN could save a typical SACN municipality an average of R 89 471.20 (20%) per annum on data acquisition costs. However, a more drastic collective bargaining approach by SACN, could save the same municipality up to R 262 453.59 (60%) per annum on data acquisition alone. The above collective bargaining would release existing resources for the development and maintenance of the Open Data Almanac.

Provide a consistent benchmark of overall municipal performance

With the above achieved, the Open Data Almanac is pitched to assist cities who sign onto the SACN with more than just a platform for sharing comparable public data. This is done through simply placing similar data side by side on a spreadsheet. And, later, this can be improved by deploying a reporting/business intelligence layer that can include a data visualisation tool(s) to provide live and varied analysis of data and metadata on the Open Data Almanac.

Participating in this part of the Open Data Almanac provides value to a municipality because basic planning, development and indicator data points are subjected to recognised data governance procedures. This increases the reliability of planning and reported information over time hopefully providing a consistent benchmark of its overall performance rather than having several disparate sectoral views as is currently the case.

Reduce the burden of intergovernmental reporting

The SACN is also well positioned to be the insightful intermediary between local government and all other spheres of government. Urban Indicator Project Status Quo findings confirm that the voice of local government is weak when it comes to determining the scope, scale, frequency and format of required regulated reporting to national and provincial organs of state fulfilling their oversight functions.

Taking the above as a given, the SACN's hosting of an Open Data Almanac places the bulk of a member city's data repository at its disposal in electronic format. Whilst the Almanac cannot assume to have the most complete record of city statistics, it can go a long way in easing the reporting burden

by pre-populating required reporting templates of other spheres of government. To do this, the Open Data Almanac has to develop a template engine that:

- 1) Digitises the most commonly required reporting templates, on request by member cities;
- 2) Pre-populates said templates with updated city data from the Almanac for finalisation by paid up SACN member city representative; and
- 3) Offers a report publishing facility that creates an archive of the same information and report for future use and analysis.

The above functionality builds a reporting repository for member cities that also provides the continuity and insight into regulatory reporting that is often missing in municipalities with high staff turn overs and persisting low levels of competence in some areas.

Contribute to Increased Transparency, Compliance and Accountability

Lastly, the Promotion of Access to Information Act (PAIA), 2000 represents an opportunity for SACN to formalise its role as a virtual Deputy Information Officer for member cities on a shared services basis.

It is noted that only two municipalities – City of Cape Town and Msunduzi – seem to be complying with Promotion of Access to Information Act (PAIA), 2000 regulations to publish voluntary disclosures as per legal requirement. Compliance to PAIA regulations is challenging as it requires the appointment of an information officer, his/her deputy and maintaining a secretarial function to ensure that various requests for information are fulfilled as per PAIA regulations. Fulfilling this obligation is complex, requires regular expenditure on human resources and the repeated competent execution of time-bound legal processes. Currently, every municipality seeks to fulfil this requirement through their own means.

However, the SACN through hosting and developing the Open Data Almanac equipped with its intergovernmental template reporting capability, can thus further facilitate municipality's compliance to PAIA and its regulations in the following three ways:

- 1) The SACN Open Data Almanac can be listed as one of the manners of accessing public member municipality records. The Almanac will be in a position to assist member municipalities to provide evidence of fulfilling citizen's constitutional right of access to a range of information held by the municipality and that may be required for the exercise or protection of any rights. Implementing the SACN Open Data Almanac assists municipalities to comply with this requirement, but does not remove the responsibility and accountability for doing so from municipalities.
- 2) The SACN should offer its member cities a dedicated Open Data Almanac staff complement to serve in the delegated role of their respective Deputy Information Officers. The legislation does not preclude member cities from jointly appointing the SACN in its Deputy Information Officer for the express purpose of achieving and maintaining compliance in terms of PAIA.

For the SACN, this is an opportunity to provide a paid service to member cities whilst also alleviating the burden of directly employing information officer personnel. This arrangement allows a municipality to ensure that access to the bulk of its information is rendered within the prescribed periods and that the required annual reporting may be compiled for submission to the Justice Minister. Specifically, the annual compilation of:

- Voluntary disclosure of Automatically Available Records and Access to Such Records in terms of Section 15 of PAIA (Act no. 2 of 2000), and Regulation 5A (Annexure A refers);
- Information Sheet Reports of Requests for Information in terms of section 32 of PAIA (as per Annexure B)

The SACN Open Data Almanac will include a voluntary subscription to the Virtual Deputy Information Officer service alongside that of the public data compliance module. SACN can also provide training of information officers and deputy information officers as a service to municipalities.

- 3) The SACN Open Data Almanac can also assist member municipalities through digitising the PAIA forms of request for access to a record, contemplated in section 53(1) of the Act. These requests for information must be made on stipulates forms and using defined Forms to describe and detail a request.

The digitisation of these forms would be a once of exercise, whilst its promulgation in electronic format on the Open Data almanac and elsewhere contributes to increasing the transparency and accountability of member cities by improving these few elements of compliance to PAIA regulations.

The above are the key elements of the operational mandate of the SACN.

4.2 Overarching Data Collection Process

The aim of the basic design of the SACN Open Data Almanac is to centralize, simplify, decrease costs, and eliminate the duplication of effort as far as possible.

It is presumed that if all city data and metadata is collected by a team of individuals within a specific time frame and collated and verified within one place the costs of repeating this process over time, will decrease.

There will be an initial outlay to implement the process and build the infrastructure. As mentioned the above the design below is basic outline of what the overarching data collection process could look like.

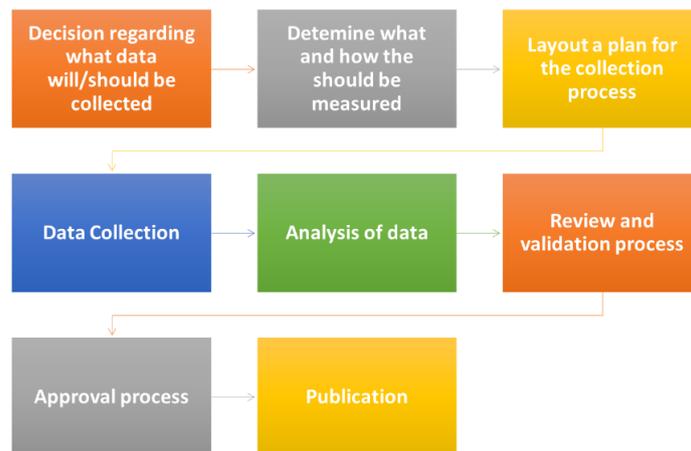


Figure 3: New streamlined data collection process

The new streamlined data collection process is explained in Annexure A below.

4.3 System Requirements

The SACN Open Data Almanac can be developed in a variety of ways given the range of technologies available. These options have not been explored in this report, but would be typified by the data repository conceptual design schematic below.

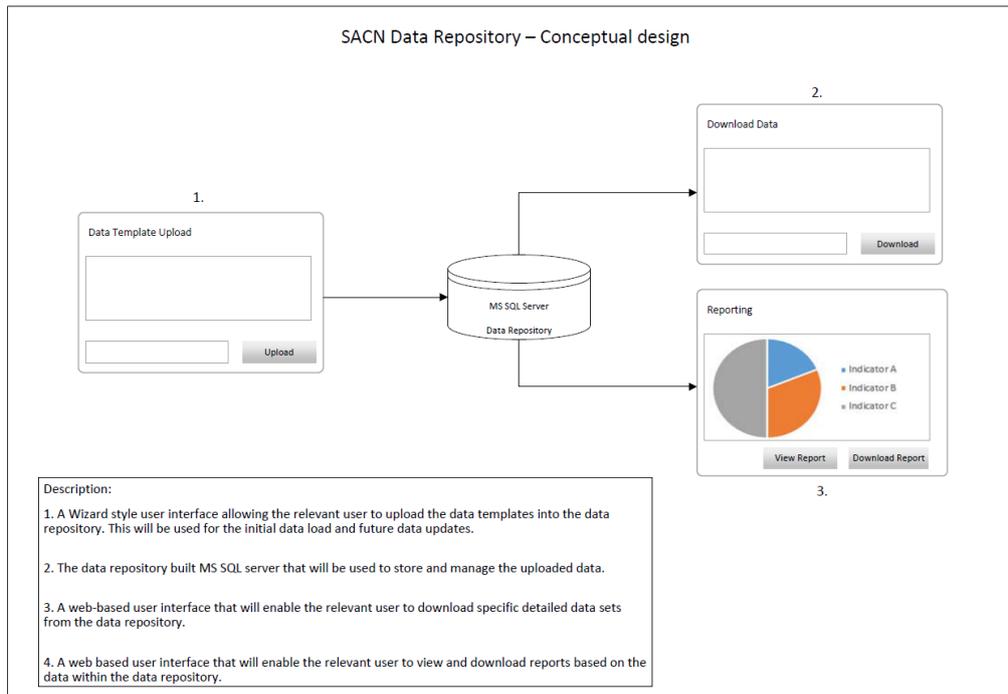


Figure 4: SACN Data Repository - Conceptual Design

Detailed system requirements will be developed when concept has been accepted and specific information about system operating, hosting and use environments is known.

4.4 Key Stakeholders & Systems in City Data

Metropolitan and large town municipalities are the focus of the SACN mandate. Rather than a focus in growing ‘the club’, there is merit in pursuing some of the larger and fastest growing secondary cities to become SACN members, namely economies of scale and information sharing.

For the purposes of this report, the following fifteen (15) municipalities are identified as SACN members and non-members, but are all viewed as targeted members to be gained, retained and regained.

Table 5: SACN member and non-member municipalities who should participate in the Open Data Almanac

Member Municipalities	Non-Member Municipalities
Ethekwini Metropolitan Municipality	City of Cape Town Metropolitan Municipality
Ekurhuleni Metropolitan Municipality	Sol Plaatje Local Municipality
City of Johannesburg Metropolitan Municipality	Mbombela Local Municipality
Nelson Mandela Bay Metropolitan Municipality	Polokwane Local Municipality
City of Tshwane Metropolitan Municipality	Rustenburg Local Municipality
Mangaung Metropolitan Municipality	Mafikeng Local Municipality
Buffalo City Metropolitan Municipality	Emfuleni Local Municipality
Msunduzi Local Municipality	

Regardless of the above, there remains several levels of stakeholders with regards to City Data in South Africa. Some stakeholders are also custodian of city-level data and operate various systems in fulfilling their mandates. These are tabulated in turn below.

Table 6: South African Government role players in City Data

South African Government Departments	Systems and Databases Administered
National Treasury (NT)	Vulindlela
Cooperative Governance and Traditional Authorities (COGTA)	Back2Basics Information System (B2BIS)
Statistics South Africa (StatsSA)	SuperCross
Department of Environmental Affairs (DEA)	
Department of Water & Sanitation (DWS)	Green Drop / Blue Drop
Department of Rural Development and Land Reform	SPISys
South African Local Government Association (SALGA)	Municipal Barometer
Provincial Treasuries	
Provincial Departments of Public Works	
Provincial Departments of Housing/Human Settlements	

Table 7: State Owned Companies (SOCs) and other Entities playing a role in City Data

State Owned Companies (SOCs) and other Entities	Systems and Databases Administered
National Energy Regulator (NERSA)	D & G Forms
Global City Region Observatory (GCRO) – Gauteng Only	GCRO Barometer
CSIR	SPISYS
South African Weather Service (SAWS)	South African Air Quality Information Systems (SAAQIS)

Table 8: Private Data Service Provider Companies

Private Companies - Data Service Providers	Systems and Databases Administered
IHS Global Insight	REX data
Quantec	EasyData
Municipal IQ	A web-based data and intelligence service specialising in the monitoring and assessment; benchmarking; and comparison of municipal performance
BMI Techknowledge	Commercial and industrial research
UNISA Bureau of Market Research (BMR)	Socioeconomic and market research

Table 9: Private Consulting Companies working with City Data

Private Consulting Companies	Systems and Databases Administered
-------------------------------------	---

Private Consulting Companies	Systems and Databases Administered
Africa Scope	Proprietary software, systems & datasets.
Palmer Development Group (PDG)	Proprietary software, systems & datasets.
Urban-Econ	Proprietary software, systems & datasets.

5 Funding & Financing the Open Data Almanac

Funding and financing the SACN Open Data Almanac is based on initial contributions by cities that can be motivated for by savings already in the way cities acquire their data. Whilst it is believed these should anchor the basic Open Data Almanac that will offer free city data to the public, opportunities to provide paid for services are:

- Paid Capture and Pre-population of Regulatory Reporting Templates for Data Custodians
- City Specific & Non-regulatory Analyses and Reporting
- Virtual Deputy Information Officer & PAIA Compliance Reporting

5.1 Seed Funding is currently available in the Data Acquisition by Cities

This concept paper assumes the SACN will have succeeded in leveraging approximately R 4 000 000 to R 11 000 000³ over the next five year period (2016 -2021) from savings the existing member municipalities stand to make by participating in collective bargaining led by the SACN. This was outlined in the elaborated Scenario 4 in the Cost Benefits Assessment report.

Therefore, the resources require are understood to already be in the current system of data acquisition and that these are available for progressively establishing the SACN Open Data Almanac. This includes system development for the first few iteration and may include appointing dedicated resources and personnel to manage growing responsibilities and roles in relation to data and services to cities that may be required in the fully fledged Open Data Almanac.

5.2 A Five Year Value Proposition

The value proposition of the Open Data Almanac is one that should be tied to the next political term, thus resonating with the value and support that could be provided to municipalities in the 2016 – 2021 term of office.

The conceived system is aimed at addressing the following cost/financial drivers that exist within the data acquisition and management environment for city data:

- **Procurement** – these costs refer to the costs associated with subscriptions and ad hoc payments from sources other than Statistics South Africa
- **Legislated reporting** – these are the costs associated with the primary research and dedicated timeframes applied during the reporting and strategy development process of all levels of government
- **Time** – these are the costs associated with the time used to gather, compile, interpret and/or manipulate the data by both Government and other organisations
- **Use** – these are the costs associated with time applied to data used for both other organisations and for Government outside of the reporting framework

5.3 A Base for Leveraging Direct Funding

The Open Data Almanac will require initial funding for development and for sustaining itself through the initial phases. However, the proposed implementation plan seeks to ensure that the rights managed component could grow to fund itself by providing valuable service to member cities.

³ Possible savings were identified amongst the nine SACN members over the next five years (i.e. IDP term of 2016-2021) ranging from R 805 240.76 to R 2 362 082.34 that could be leveraged from each municipality immediately for data acquisition and management. This amount was multiplied by the 9 existing municipalities and assumed available to be dedicated to establishing the SACN Open Data Almanac.

Paid Capture and Pre-population of Regulatory Reporting Templates for Data Custodians

There remains real options to have paid for participation by data custodians whose templates can be digitised in the intergovernmental reporting module that is yet to be explored. This approach does not preclude also charging the end user for reporting and other specialised services being provided by the Almanac.

This option will have to be explored carefully as it may raise that cost of data acquisition for public sector and other users of data which would may jeopardise the openness and relatively free access to city data that the seeks Almanac to encourage.

City Specific & Non-regulatory Analyses and Reporting

It is believed that all the urban indicators used to compile that SoCR 2016 report should be made available free to the public in MS Excel format and in as disaggregated detail as possible. This will achieve the vision of the SACN of having a truly electronic and open online Data Almanac.

It is conceived that the Open Data Almanac will also host link – at least – to the large variety of public data sources available free of charge from other organs of state. In time, it is conceived that the regulatory reporting of member cities may possible be made available free of charge through the same platform⁴.

However, there are various reports that are city specific and not part of the regulatory reporting regime, but that may be automatically generated via the Open Data Almanac. It is proposed that a common, scheduled of fees may be charged for such reports much in the same fashion that other report generation services charge, i.e. ranging from R10 to approximately R40 per unique report. The registration of user details may be required for these services as this allows the capturing of user payment details and histories in order to ensure users do not pay multiple time for the same reports.

Virtual Deputy Information Officer & PAIA Compliance Reporting

The compliance reporting required of all organs of state in terms of the Promotion of Access to Information Act (2000) is an important opportunity for providing a required service by all metropolitan municipalities. This is detailed above in the operational mandate of the Open Data Almanac. It is evidenced that during August 2015 not all SACN member cities were currently complying with PAIA and its applicable regulations, i.e. SACN member cities had not gazetted “Compulsory Voluntary Disclosure of Automatically Available Records and Access to Such Records” nor “Compulsory Information Sheet in terms of Section 32 of PAIA”.

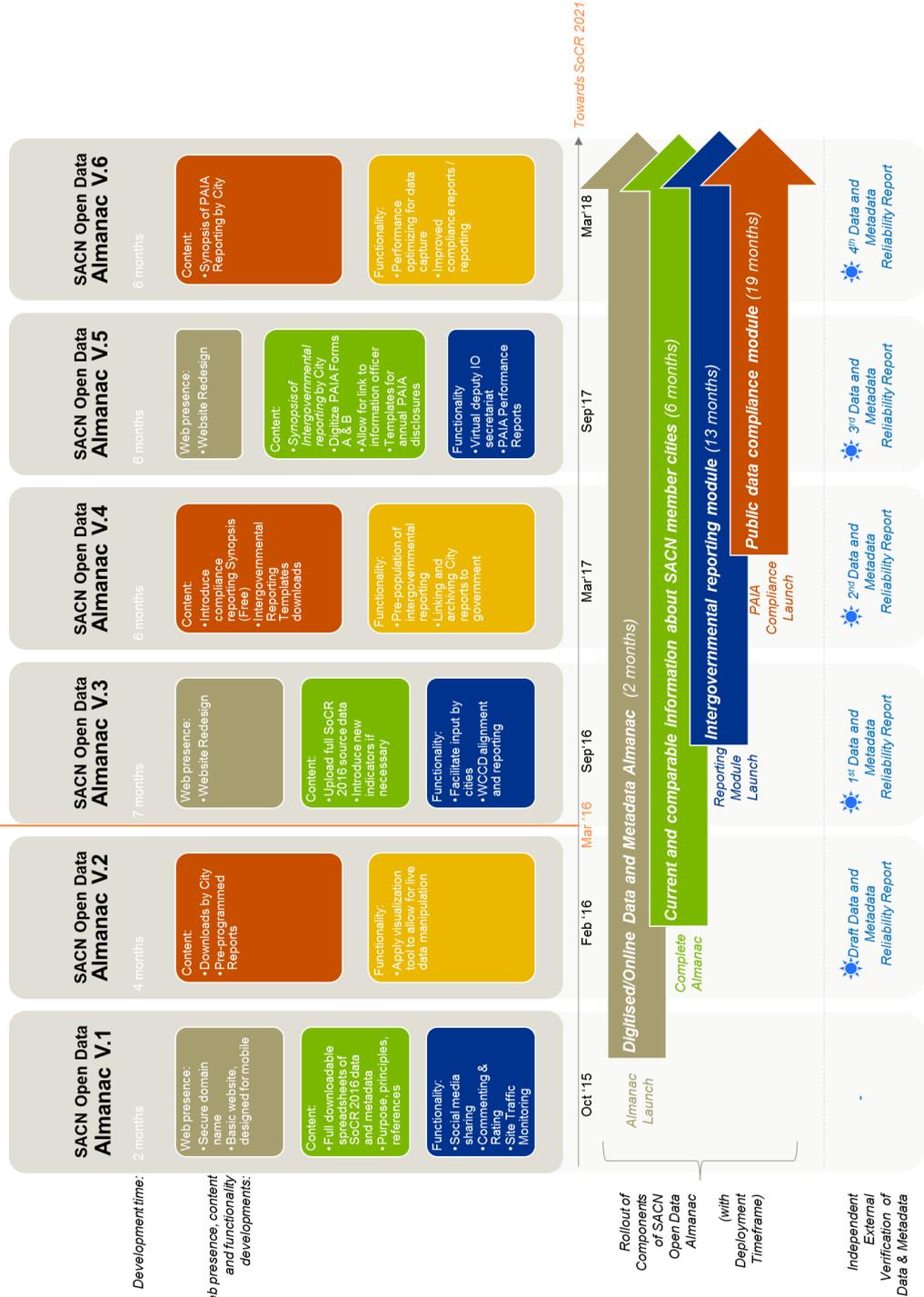
The envisaged role and mandate of the Open Data Almanac is to drive transparency and contribute to improve democratic and free access to information about member cities. The SACN is well positioned – once the Open Data Almanac is operational – to fulfil the legally required responsibility of the Deputy Information Officer as envisaged in the PAIA and its applicable regulations on a shared services basis and virtually, i.e. from a centralised basis serving subscribing cities. It is envisaged that this service should be provided on commercial terms as the SACN can execute this service consistently, more cost efficiently and possibly more effectively as it can build its institutional capacity to ensure that this competency is met at all times.

⁴ This will require a board decision and official sanction from member cities, possibly on a report by report basis. It is recommended that all regulatory reporting remain confidential and specifically for member city purposes until express approvals for its dissemination are confirmed by the SACN Board and city leadership.

6 Towards achieving Open Data Almanac of the State of the Cities Report (SoCR) 2021 – Approach to Implementation

The approach to implementing the SACN Open Data Almanac is a long term programme of improving the structure, management and, separately, the quality of city data and metadata.

Implementation Plan: Open Data Almanac *Development & rollout towards of the State of the Cities Report 2021*



7 Conclusion

The above report presents a concept of how the South African Cities Network (SACN) Open Data Almanac for Cities can be achieved. The focus of this report has necessarily been on how the structure and management of the Almanac can be instituted and improved upon over the next 30 to 60 months.

This paper recognises that although mandates and strategies of different city data stakeholders may differ, their ultimate objectives are similar. An open, well governed and largely free database is needed to serve all stakeholders with verified and consistent data as required and in the preferred formats.

Further, there are also parallel processes of improving, aligning and harmonising the indicators, measures and datasets that are required of local government. These are not reflected on in this paper.

The formation of the SACN Open Data Almanac requires a commitment from, at least, the existing SACN member cities. Its implementation will go a long way to creating cohesion between the different actors in the city data field.

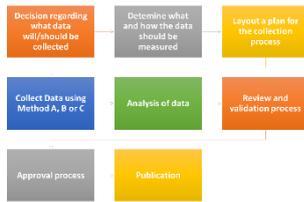
The wider application and relevance of the intergovernmental reporting and PAIA compliance components requires the cooperation and involvement of other entities and stakeholders in the data/metadata value chain. This represents a new level of maturity in the commitment to the Open Data Almanac.

Whilst the coordination of the abovementioned entities is vital to the robust functioning of the SACN Open Data Almanac, it is not a pre-condition. Current SACN member cities as led by the SACN board are in a position to establish a robust framework to support the collection; monitoring and dissemination of city data and information.

The SACN is well positioned to create value that all of South African government and the private sector actors can leverage off of.

8 Annexures

A New Streamlined Data Collection Process



The new streamlined data collection process is envisioned as one of the key protocols to be followed to get data into the SACN Open Data Almanac. The importance of this process is that it would be a repeatable and confirmed method of verifying how, where and why certain datasets are collected in the way they are. This process will be applicable for new indicators and for in-depth review of existing indicators and data points.

Decision regarding what data will/should be collected

- At this stage the various stakeholders and interested parties should determine what should be measured.
- This would ideally occur once every five years at the Urban Conference for the State of the Cities Report, however can happen periodically as the KMRG provides guidance.

Determine what and how the data should be measured

- At this stage the type of indicators and how their impact will be measured should be decided.
- This allows for consistency across the various spheres of government and the inter and intra departments within the environment
- It will also assist with comparison and benchmarking exercises
- This would also occur once every five years at the Urban Conference for the State of the Cities Report

Layout a plan for the collection process

- The entity in charge of collecting the data (essentially the data collectors & managers within the SACN Open Data Almanac) would determine when and how this data will be collected and collated for analysis
- This collection process would occur once a year to ensure consistency and allow for lags in gathering process

Collect Data using Method A, B or C

- The act of collecting data can take three forms:
- **Method A:** Data collection using data collectors and field marshalls would be trained individuals that both understand the data and its eventual purpose. This likely requires the roll-out of meeting/workshops within the stakeholders to provide context and understanding
- **Method B:** Data collection through information sharing sessions with data owners run by a team separate to the data collectors & field marshalls.
- **Method C:** Data Collection via email, data extract and/or obtaining a data dump

Analysis of data

- A dedicated team within the SACN Open Data Almanac would attend to the analysis
- The team will receive the data from the field, collate it, provide high level analysis and maintain the database.

Review and validation process

- A separate team within associated with the Open Data Almanac would be able to review and provide validation. Else, an appointed panel with the applicable knowledge to review and provide validation will be engaged.
- This verification would be additions or improvement to the datasets and would be presented as evidence in published data and Metadata Reliability Reports

Approval process

- This function could possibly sit with the review and validation panel of the SACN Open Data Almanac as they are best placed to provide approval.

Publication

- The release of the data and analysis into the web-based Almanac would be freely accessible to those who require such data.
- A record of the data that has been reviewed and approved should be returned to the analysis team for record-keeping and publication on the web-based data portal.

Source: KPMG

B Example of Compulsory Voluntary Disclosure of Automatically Available Records and Access to Such Records – Msunduzi Local Municipality

C Example of Compulsory Information Sheet in terms of Section 32 of PAIA – South African Human Rights Commission

Contact us

Andile Skosana

Associate Director – KPMG Cities Centre of Excellence

T +27 (0)82 719 2677

E Andile.Skosana@kpmg.co.za

Hermann Schutte

Senior Manager - KPMG Cities Centre of Excellence

T +27 (0)82 718 8550

E Hermann.Schutte@kpmg.co.za

Zunaid Khan

Senior Manager - KPMG Cities Centre of Excellence

T +27 (0)76 511 3375

E Zunaid.Khan@kpmg.co.za

Taryn Ramputh

Junior Analyst - KPMG Cities Centre of Excellence

T +27 (0)76 153 8439

E Taryn.Ramputh@kpmg.co.za

www.kpmg.co.za

© 2015 (member firm legal name), the (member firm jurisdiction) member firm of KPMG International, a Swiss cooperative. All rights reserved.

The information contained herein is of a general nature and is not intended to address the circumstances of any particular individual or entity. Although we endeavour to provide accurate and timely information, there can be no guarantee that such information is accurate as of the date it is received or that it will continue to be accurate in the future. No one should act on such information without appropriate professional advice after a thorough examination of the particular situation.

The KPMG name, logo and “cutting through complexity” are registered trademarks or trademarks of KPMG International Cooperative (“KPMG International”).

