South Africa

land portal

State of land information in South Africa

Uncovering South Africa's land information ecosystem

Our mission

To build an information ecosystem for land governance that supports better informed decision and policy making at national and international levels.

Our objectives

To improve documentation, mapping and monitoring of land governance issues through the provision of a widely used platform which includes structured information, tools and services.

Promote, inform and enrich global debate and practice on key land issues while providing further awareness on selected thematic areas of central importance to land governance.

Raise awareness on Open Data principles, support the creation of a solid data infrastructure and build the capacities of information providers, in order to strengthen the flow of land governance information at all levels.

Acknowledgements

Land Portal would like to acknowledge the work of the Council for Scientific and Industrial Research in scoping and documenting the South African information landscape. Substantive content from the full CSIR report, along with the findings from the Land Portal's open data compliance assessment, are presented in this final report. The CSIR report authors were M. Napier, A. Rosenfeldt, and A. Cooper. CSIR contributors to the scoping exercise included B. Crankshaw, A. le Roux, and A. Breytenbach.

State of land information in South Africa

Uncovering South Africa's land information ecosystem

L. Mey
C. Tejo-Alonso

March 2020

Contents

Executive Summary

With this State of Land Information Report we seek to provide an overview of existing data and information on key land issues. Our aim is to uncover the many different sources of land data and information in South Africa and thus provide a basis to substantiate, refute or nuance the often-repeated rhetoric that there is a lack of land data. To this end, we developed an original scoping and assessment methodology building on existing internationally recognized and well-known frameworks. For the very first time, we systematically reviewed and categorized the entire landscape of data and information related to key land topics in South Africa, assessing over 104 land resources from 59 different sources. This robust scoping exercise not only allowed us to see trends and gaps when it came to land data collection, but also prompted us to provide very practical recommendations to improve visibility and usability of data and information, thus seeking to improve the land information ecosystem in the country.

The statement that there is a lack of data can partially be refuted: **our scoping exercise shows that 67% of key land resources are available as statistical or geospatial data, not documents**, as was the case in our previous SoLI studies in Kenya, Tanzania, Uganda and South Sudan. However, we did not assess how complete or accurate this data is, which remains an area of further work. Another important caveat to saying that there is "no lack of data" is that 60% of the key resources were either from before 2019 or the publication date was unidentifiable. This is a significant constraint for these resources to be useful or used.

Another important aspect that defines the usability of a resource for a user is knowing the source of the data or information. **In this particular scoping study, the government of South Africa was identified as the main provider of data (over 60%)**. This is of course limited to the results from this scoping study, and might not be wholly reflective of the ecosystem. It is clear, however, that the government of South Africa publishes a lot of land data and there are many activities around this topic. Clear divisions and responsibilities when it comes to data custodianship between different government departments, in particular, was found to be an area where improvements could be made. Our findings in South Africa differ from the four East African countries in many ways, but one where the results are the same is the underrepresentation of National Civil Society perspectives in our scoping exercise. They account for a mere 8% of the total resources identified in this study. This suggests that their data and information– which they undoubtedly have–is not published in a way that makes it visible for a wider audience. However, this suggestion warrants further and separate research specifically into the data and information provision by Civil Society Organisations (CSOs) in South Africa.

Availability of Data and Information							
Key Category	Data available?	Representation of Sources					Data up-to-date?
		Government	Research Institutions	National CSOs	International Organisations	Other	
Land Tenure Data	✗	✓	✗	✗	✗	!	✓
Land Cover, Use & Management	✓	✓	✗	✗	✗	!	✗
Land Disputes	✓	✓	✗	✓	✗	!	✓
Human Settlements	✓	✓	✗	!	✗	✗	✗
Land Markets & Financing	✓	✓	✗	!	!	!	!
Land, Climate Change & Environment	✓	✓	!	✗	✗	✗	!

✓ = good practice ! = room for improvement ✗ = poor practice

Our research further shows that the knowledge is published online (97%) and it is mostly available for free (75%)). The rudimentary access to data and information there seems to be in a good state in the South Africa Data and Information Ecosystem, with a notable exception of the fact that for many of the key resources identified (40%) there was still a login barrier or some kind of requirement to identify oneself prior to accessing the data. With regards to more sophisticated accessibility criteria (such as provision of metadata, use of standards and open licenses, etc.), South Africa scores considerably higher than the four pilot countries of the SoLI research, but there is still room for significant improvement.

Accessibility of Data and Information								
Online	No (log in) barriers	Free (unpaid)	Metadata	Standards	Downloadable	Open License	Machine-readability	(Linked) Data URIs
✓	!	✓	!	!	✓	✗	!	✗

✓ = good practice ! = room for improvement ✗ = poor practice

Overall, the health of the South Africa Land Data and Information Ecosystem is scored with 60/105 points.

Recommendations

The score of 60/105 is considerably better than what we've seen in Kenya, Tanzania, Uganda or South Sudan. This does not mean there is not considerable room for improvement. An important caveat we want to reiterate is that this is the result of the assessment of the resources identified during this study; there is a limited view there on what is not present. Another caveat is that this study only covers availability and accessibility of data; there is no assessment or judgment on the completeness or accuracy of the data. The following recommendations are made to improve the state of the Land Data and Information Ecosystem in South Africa:

- **Continue with the process of clarifying custodianship of governmental datasets** within government to ensure proper maintenance of the respective databases;
- **Consider facilitation of more equitable access to data by removing login requirements or payment barriers wherever legally possible**, at least for certain groups, and/or certain elements of the data;
- Support & enforce data publishing practices to **include a minimum set of metadata** with each publication, dataset or other type of information published by any type of information providers;
- **Support & enforce the use of standards** when publishing data and metadata to promote the usability as well as interoperability of data and information in the South African data & information ecosystem;
- **Institute a system of publishing downloadable 'raw data' alongside tableaus and visualization methods** published online to enable more re-use of data and information;
- **Apply open licenses to published data and information** to allow for more meaningful and in depth use, re-use and modification of data and information to increase its impact, and most importantly, consider licensing and publish it along with the data and information;
- **Apply unique identifiers** to key elements of the data to ensure consistent and reference to the data and information, and allows for more efficient exchange within the data ecosystem;[1]
- **Commission specific research and action into availability of data and information from civil society organizations** or NGOs, to gain a further understanding in their data and information supplies as well as sharing practices.

1 SANS 1876:2016 provides rules for unique feature identifiers in South African geospatial datasets, but the standard has not yet been implemented.

Introduction

With this **State of Land Information Report**, we seek to provide an overview of publicly available data and information on key land issues, from not only the government, but also other sources. The aim of the research is to uncover the many different sources of land data and information at the country-level and help to identify actual data and information gaps, with a view to establishing a baseline for targeted 'information-based' interventions to improve the information ecosystem.

What sets this research apart from other monitoring initiatives, is that the focus is on the database or dataset and its sources; the value or content of the information is not our main focus. Our belief is that data quality, accuracy and reliability lies in the judgement of the user. For the very first time, we look at the entire landscape of a country to see trends and gaps when it comes to land data collection, as well as how accessible it is on the World Wide Web. The State of Land Information report concludes with -where necessary- concrete recommendations to data and information providers to improve their data sharing practices, to help establish a functioning, inclusive and democratized ecosystem of data.

The centuries-long influence of colonial and apartheid planning in South Africa that included the practices of zoning land on the basis of race leading, in many cases, to land dispossession and extensive forced removals (Platzky & Walker, 1985), has meant that land and the control of land have been, and remain, central national issues with important economic, social, environmental and political dimensions. Given this context, it is perhaps fairly obvious why opening up access to information about all aspects of land would be one of a number of crucial elements needed to support and take forward the national, land transformation project.

Land reform has again, in the last couple of years, become a high-level, national priority and was a central issue in the recent national elections leading to the establishment of the sixth democratic parliament in June 2019 (Parliament of the Republic of South Africa, 2019). The sometimes charged debates surrounding complex policy-making processes can certainly benefit from access to more reliable data and information as an evidence-base to allow issues to be more responsibly addressed.

When discussing regulations designed, for example, to control the ownership of property by non-South Africans (Samasuwo, 2004), or progress in changes in patterns of land ownership since 1994, or land expropriation without compensation, for all these types of issues, access to relevant and reliable data is essential if regulatory or other interventions are to be designed well.

The recent release of the findings report of the Presidential Advisory Panel on Land (Mahlati et al., 2019) which was the result of many months of engagement and consultation, has again focused attention on the importance of overcoming, in practical ways, the considerable challenges experienced so far by the state-led land reform programme.

Why is this report useful?

The report is useful as a tool for any land governance work that requires access to data and information. For example, a researcher may use this report to identify gaps in information and identify research priorities accordingly. A land practitioner working at the global level may use the information sources as a basis to monitor land governance performance against international indicators. Private companies may find the report a useful starting point in due diligence processes prior to working or investing in a country. Local information providers may identify weak links in their data sharing practices and implement concrete recommendations. Governments can use the outcome to establish or strengthen their policies that aim to increase access to data and information by citizens. Ultimately, we hope the report will make data and information more visible and usable by any potential user and thus improve the local information ecosystem from the bottom-up.

Importance of transparency in land data and information

Why are open access to land data, and information transparency, crucial for South Africa today? We can make a general, in-principle argument for open access to land data, and that it would be 'good' for democracy, accountability and improved government. In this instance, there is more than an in-principle case for open access to data because South Africa's Constitution, adopted in 1996 (SA Government, 1996), enshrines the right to access to information. According to Clause 32:

1. Everyone has the right of access to
   - any information held by the state; and
   - any information that is held by another person and that is required for the exercise or protection of any rights.
2. National legislation must be enacted to give effect to this right, and may provide for reasonable measures to alleviate the administrative and financial burden on the state. (Republic of South Africa, 1996)

This Constitutional obligation is further affected by legislation such as the Promotion of Access to Information Act 2 of 2000 (PAIA) that obliges the state to make unclassified information available to the public.

The next section will expand on later legislation designed to improve management of, and access to, land and spatial data.

According to Manona, "At least five constitutions – of Kenya, Panama, Poland, Serbia and South Africa – expressly extend the right to information to state-owned enterprises and/or private entities that exercise public functions as well as to public authorities." Further to this, Manona points out that "South Africa is also among the initial eight founding members of the Open Government Partnership (OGP) in an initiative that was formed in 2011 with a view to providing an international platform for domestic reformers "committed to making their governments more open, accountable, and responsive to citizens."" (Manona, 2019, p. 8)[2]

For any country, but particularly for South Africa given the country's land history briefly alluded to above, some of the advantages and benefits of transparency in land data and information would include:

- decision-makers being able to measure progress in spatial transformation away from the past patterns of racial segregation and unequal access to land and secure tenure, and towards a more egalitarian future;
- the open sharing of data substantially assisting in what is referred to in government circles as improved inter-governmental relations (or "IGR"), thereby allowing state agencies to coordinate their plans and interventions, and to start to achieve "spatially targeted investment" (SA Government, 2018)[3], and land reform and redistribution which are both seen as high priorities in reversing some of the lasting effects of past systems;
- the state, with the help of agencies in the private sector, academia and civil society, being able to understand and respond to emerging patterns of urban growth and migration thus allowing planning in advance;
- citizens being able to hold the state accountable by monitoring (for instance) the satisfaction of the rights related to access to, and use of, land for a variety of uses;
- both government and civil society being better able to understand and monitor change in who is succeeding in accessing better-located land in the context of a sometimes very competitive land market[4];
- investors (whether small or large) being able to make decisions about future investments and improving the levels of predictability that foster investor confidence;
- private-sector agencies being able to build on to state datasets and adding to the utility of such datasets, while also deriving commercial value[5].

2 Manona's paper, written for the LandNESS network, has a more extensive discussion of the constitutional basis for transparency and open data.

3 A term that is used to imply coordinated government capital expenditure in concentrated geographic areas identified as high priority.

4 Refer Urban Land Markets Programme work on addressing competition in the urban land market **www.urbanlandmark.org**

5 Although, we should add that this benefit to private sector agencies of open access to state data should not cause the state to abrogate its responsibility to ensure that high-quality data is freely available.

There are many other real and potential benefits to open access to data and obviously more specific advantages can be identified for the many types of data that are commonly used in different sectors such as the variety of urban land uses, and agriculture, forestry, and mining.

While the principle of open access to land data and information is a sound one, there can be an argument of course for some limits to open access to all data.

For instance, when it comes to the state (typically municipalities) releasing spatially-specific information about future priorities for infrastructure investment, there is a case for thinking about the timing of the release of certain types of data and information. There are examples where municipalities or provinces have disclosed future plans prematurely, before taking steps to do their own planning sufficiently, which have led to less positive forms of speculation by the private sector to the detriment of less powerful sectors in society.

A typical example has been the release of information about the proposed location of public rail and road routes, and stations, in South African cities without a prior effort by the state to secure land for lower cost housing and government services in what then become prime locations as a result of the state transport and other infrastructure investment.

There is of course a distinction between arguing the benefits of access to more basic forms of data collected using public funds, and this case of the premature release of public planning information affecting intended, future infrastructure investment by the state.

On the flip side, prior to its release, government employees who have privileged access to information about land need to exercise care in the proper use of this information, as it can be used for the common good or for pursuing more individual agendas (National Treasury, 2018).

There are many other generic benefits to open government data that would apply in most or all countries (as discussed by Manona, 2019 and many others). But one positive feature of the South African situation, as we shall see, is that the legal system ensures that information about who owns what land is not regarded as private information[6]. Although the information is not always accurate or fully up to date (or free), the legally-enforceable access to information about land ownership is a significant transparency advantage in a country that is consciously trying to move away from its unequal past. Transparency in land ownership information does help to measure progress away from the situation where a privileged few had access to the benefits of the land and exclusive rights over property, and many of the benefits derived from the most productive and well-located land, much to the detriment of the majority, and towards a situation of greater distribution of resources and opportunities.

6 A point expanded on in the section addressing the legal, policy and institutional environment.

Methodology

The State of Land Information methodology consists of two consecutive phases, namely the scoping research, followed by an accessibility assessment of the identified datasets and other sources of information. These aspects together provide a snapshot of the state of the land data ecosystem in South Africa in 2019 and the only true, comprehensive reference point for available land data and information in the country. We intend this to be a "living" document to be updated regularly and through an open process.

2.1 Scoping the Land Data Landscape

The parameters for the scoping study were set on the basis of key land issues identified by the Land Portal. The mantra of "building on rather than duplicating" that underlies the entire effort of this study has also been applied to the process of identifying the key land issues. We drew from key land indicators and guidelines from several global and regional land monitoring initiatives. The Land Portal team assessed overlaps and availability of information based on indicators identified in the following initiatives:

- Sustainable Development Goals, "SDGs" (United Nations)[7];
- Voluntary Guidelines on the Responsible Governance of Tenure, "VGGTs" (FAO)[8];
- Land Governance Assessment Framework, "LGAF" (World Bank Group)[9];
- Global Land Indicator Initiative, "GLII" (network facilitated by GLTN/UN-Habitat)[10];
- Monitoring & Evaluation of Land in Africa, "MELA" (IFPRI & Land Policy Initiative)[11];
- International Land Coalition Dashboard (facilitated by ILC)[12];
- Africa Data Revolution Report (facilitated by Open Knowledge International)[13].

Based on the categories, indicators and principles included in these international land data monitoring and governance guidelines and frameworks, the Land Portal has grouped overlapping indicators and principles into the following seven categories: Legal, Policy & Institutional Framework; Land Tenure data; Land Cover, Use and Management Data; Land Disputes; Human Settlements; Land Markets & Financing; and Land, Climate Change & Environment. For each of those categories, associated key information has been identified based on the principles and indicators identified in the initiatives above. The full methodology can be accessed through the online and open **State of Land Information Research Guide**.

7 **https://www.un.org/sustainabledevelopment/sustainable-development-goals/**

8 Food and Agriculture Organization of the United Nations, "Voluntary Guidelines on the Responsible Governance of Tenure of Land, Fisheries and Forests in the Context of National Food Security", Rome 2012.

9 **http://www.worldbank.org/en/programs/land-governance-assessment-framework**

10 **https://gltn.net/global-land-indicators-initiative-glii/**

11 **https://melafrica.wordpress.com/**

12 International Land Coalition, "The Dashboard Indicators", Rome May 2018.

13 World Web Foundation, "Africa Data Revolution Report 2018. Status and Emerging Impact of Open Data in Africa", 2018.

The observation unit for the purpose of this scoping exercise is the dataset. In the context of this research, a dataset is defined as "a structured collection of information, including (numerical) data, publications and multimedia contents". This may be a statistical dataset that contains land cover data (for example), but it can also include a database of publications that contains information about (a certain topic of) land governance, for example.

Data or Information?

You will notice we use data and information almost interchangeably, purposely so. When we perform a scoping study on "what is known" or somehow documented about land in a country, it would be a major oversight if we only include raw data and statistical indicators. Much of what is known, particularly at the grassroots level, is not captured in an indicator, but rather in a publication or news article, for example. In this scoping exercise, we therefore very purposely talk about both data and information.

An important disclaimer is that although the scoping study performed has been as rigorous and as targeted as possible with the use of key issues around land, this scan is not exhaustive because of the size of the land information ecosystem, because of the broad scope of the scan, because of the many actors (public, private, civil) who are the developers and custodians of datasets, platforms and decision support systems, because of the local nature of (for example) land use planning data (hence the plethora of local datasets over the 278 municipalities in the country), and because of the fluid legal and institutional context that will be described in more detail in this report. Added to this, a number of online repositories identified and documented in another scan only five years ago no longer exist or their content has not been maintained beyond a certain date. Therefore, the picture of the Data landscape in South Africa we are presenting in this report is therefore not all-encompassing, but is, rather, a (reasonably current) snapshot of a certain moment in time.[14] That said, it is the most comprehensive and definitive resource for land data and information in South Africa available and will be considered a "living" document that we aim at having updated.

14 When a dataset is still available online but content has not been updated, this is noted in the SoLI Matrix.

2.2 Assessing Accessibility

Following the scoping research, the study focuses on a rigorous assessment based on the accessibility of the identified sources of data and information on key land issues in South Africa. Similar to the scoping study, accessibility of the data and information was assessed on the basis of key criteria, guidelines and principles that have generally been accepted to define "accessible" and "open" data. The following frameworks and initiatives have been used to identify the criteria:

- Open Data Index (Open Knowledge International)[15];
- Open Data Barometer (Web Foundation)[16];
- 5 Stars of Linked Open Data (Tim Berners-Lee)[17];
- FAIR principles of Open Research Data[18];
- Open Data Inventory (Open Data Watch)[19];
- Africa Data Revolution Report (Open Knowledge International)[20];
- EU Open Data Maturity Assessment (European Union)[21];
- OUR Data Index (OECD)[22].

The Land Portal identified 18 criteria against which every information item identified during the scoping study has been assessed. This has been done on the basis of extensive studying of the available data and information online, as well as contacting data owners with additional questions and clarifications to gain as much information about the particular data or information source as possible.

15 Open Knowledge International, "Global Open Data Index. Methodology", consulted website September 2018: **https://index.okfn.org/methodology/**

16 World Wide Web Foundation, "Open Data Barometer. Methodology", consulted website September 2018: **https://opendatabarometer.org/leadersedition/methodology/**

17 Berners-Lee, "5 Stars of Linked Open Data", consulted website September 2018: **https://5stardata.info/en/**

18 Wilkinson, Dumontier et al, "The FAIR Guiding Principles for scientific data management and stewardship", Scientific Data No 3, March 2016.

19 Open Data Watch, "Open Data Inventory 2017. Methodology Report", 2018.

20 World Web Foundation, "Africa Data Revolution Report 2018. Status and Emerging Impact of Open Data in Africa", 2018.

21 European Commission, "Open Data Maturity in Europe 2017. Open Data for a European Data Economy", November 2017.

22 Ubaldi, B., "Open Government Data: Towards Empirical Analysis of Open Government Data Initiatives", OECD Working Papers on Public Governance, No. 22, OECD Publishing, Paris 2013.

Why does Open Data matter?

Open Data principles are critical to bring a perspective to data that makes it more useful, more democratic and less harmful.[23] It is a common misunderstanding that publishing publications on a website is all you need to do to make the information accessible and usable. Data that is published according to Open Data principles is much more visible on the web than a single PDF on a website, and, perhaps more importantly, makes it possible for anyone to use, re-use and build upon the data for innovations, thereby empowering citizens and fostering transparency and accountability. Open Data empowers, democratizes and enables large-scale impact!

An important caveat to this research is that the above-mentioned criteria and initiatives are based on assessing datasets, whereas this study focuses on documents and other types of information as well. This means that the application of the criteria from the above-mentioned initiative are therefore not always (completely) performed in the way they were intended. To understand how we interpreted those criteria when it comes to documents and other sources of information than data, please refer to our public Open Data Assessment methodology.

23 Joel Gurin, "Big data and open data: what's what and why does it matter?", The Guardian, April 15th 2014.

Availability of land data & information in South Africa

The availability of land data and information in South Africa was assessed with reference to: types of data or information, representation of sources of data and information and finally, timeliness of the resources (are they up to date). This is done for all key categories with the exception of the first category, Legal, Institutional and Policy framework, as this category mostly covers the availability of laws and policies, and therefore less suitable to assess based on the aforementioned criteria. For each criteria, a general score is given. A green score indicates a good practice; an orange score indicates a practice that can be improved; and, finally, a red score indicates a poor practice. More information about how these scores were allocated can be found in Annex I-Scoring Chart.

3.1 Legal, Institutional & Policy Framework

The first category of key land issues is the Legal, Institutional and Policy Framework. The scoping research aimed to uncover whether the legislative and policy framework could be identified with the accessible data and information, as well as use the framework as the basis to find possible data and information providers from the government based on their respective mandates. Naturally, this category lends itself to mostly documents and other types of information, rather than (statistical) data.

The Constitution of the Republic of South Africa

The Constitution of the Republic of South Africa, 1996, is central to the county's economic, social and spatial transformation, and the reform of its legal system. Given the well-known history of the country, access to land and shelter is one of the strongest themes in the Constitution. The text of the Constitution was well disseminated and popularised after its release, remains widely available offline (in print format) and online (e.g. from **https://www.gov.za/documents/constitution-republic-south-africa-1996**, Matrix Ref 1), and is often quoted in government and wider debates and discussions.

There is a large body of available evidence and commentary on how this has played out in the Courts and in society since 1994. The rulings are available online (Land Claims Court Database–Matrix Ref 59).

Institutional ecosystem

The questions posed in this scan about the ecosystem, focus mainly on state actors. The quality of data and the level of analysis in some government datasets are such that they create opportunities for private operators (whether for profit, or not for profit) to build upon and add value to the base datasets that are held by the state. This is not a matter of principle but one of degree–there is state data, it is in most cases available, but it is often poorly captured, coordinated, collated, maintained and presented[24].

24 This is an important concluding comment for the whole exercise.

In some cases, as with the example of land use data, it is very localized and so far the systems are not consistent with one another. This is the data and information gap or niche that is filled by others downstream of government, often for commercial gain. However, for the purposes of this exercise, the description of the land governance ecosystem focuses on the state actors and their respective mandates.

In answering whether there is clarity of mandates in practice, that is quite a large topic, with some overlap between provincial and municipal level functions (especially in planning and housing). There was a great deal more contestation around the overlap of mandates around spatial planning, and planning and building approvals, with a variety of planning routes (e.g. the Development Facilitation Act (67 of 1995), the Less Formal Township Establishment Act (113 of 1991) etc.) up until 2013 (South African Cities Network, 2012) when this was partly clarified by the newly promulgated Spatial Planning and Land Use management Act (SPLUMA). However, because that is framework legislation those overlaps are still being resolved as regulations are being developed. This is quite an expert area and this report is not hazarding an opinion about progress in resolving overlapping mandates for land governance and information provision, nor the sum total of jurisprudence to date in advancing the (land-related) rights and principles enshrined in the 1996 Constitution.

Governance of Spatial data

Another key dimension of the South African land ecosystem is who is responsible for various aspects of spatial data. A large part of the SA Matrix relates to online spatial datasets that are available for free or for a fee. Many of the base datasets that stand behind much of this part of the ecosystem are controlled by the government under specific arrangements and as determined by the South African Spatial Data Infrastructure Act, 54 of 2003. Here we give a brief overview of this dimension of the ecosystem, and some of the recent developments in this area. Again, this area is a work in progress especially around the establishment of a national Spatial Planning Data Repository.

When it comes to spatial data, the key legislation as identified by South African Geomatics Council (SAGC, previously PLATO[25]), the legal body with which all geographic informatics system (GIS) and surveyor professionals must register, filtered by relevance to this report, includes[26]: the Land Survey Act 8 of 1997; Deeds Registries Act 47 of 1937; Constitution of the Republic of South Africa 1996; Expropriation Act 63 of 1975 (1 Apr 1994); Extension of Security of Tenure Act 62 of 1997 (16 May 2011); Interim Protection of Informal Land Rights Act 31 of 1996 (16 May 2011); National Environmental Management Act 107–1998 (18 Dec 2014); Promotion of Access to Information Act 2 of 2000 (2 Aug 2017); Sectional Titles Schemes Management Act 8 of 2011 (7 Oct 2016); Spatial Data

25 https://sagc.org.za/6legislation.php

26 The date following the original promulgation date indicates the latest amendment date, although some of these may already have changed.

Infrastructure Act 54 of 2003 (16 May 2011); Regulations in terms of Act No 54 of 2003 (The Spatial Data Infrastructure Act); Spatial Planning and Land Use Management Act 16 of 2013 (1 Jul 2015); Subdivision of Agricultural Land–Act 70 of 1970 (4 Oct 1996); and Electronic Deeds Registration Systems Act 19 of 2019.

The international movement GSDI (Global Spatial Infrastructure Association) was a key influence behind the South African Spatial Data Infrastructure (SDI) Act 2003. GSDI motivated for every country to have its own spatial data infrastructure system through every vital dataset being gazetted and documented, and having a legal, state custodian. This vision is now being fulfilled by the Committee for Spatial Information (CSI) (as required in terms of the Spatial Data Infrastructure Act No. 54 of 2003) and the Department of Agriculture, Land Reform and Rural Development is the overall body responsible for the implementation of the SDI Act. However, this has become a contested space due to issues regarding custodianship, responsibilities and capabilities.

It is interesting to note that in government documents and state-commissioned research about the national spatial data infrastructure, the legislation that is cited is not only all of that that applies in some way to land, but also many of the pieces of legislation that have to do with access to information (as in the list quoted above).

The SDI Act provides for building South Africa's "Spatial Data Infrastructure" (then referred to as SASDI) overseen by a Committee for Spatial Information. This includes capturing and publishing metadata in an electronic metadata catalogue, and determining standards and prescriptions for the sharing of geospatial information (Schwabe & Govender, 2012). According to Schwabe and Govender:

> *"The [SDI] Act is not explicit in defining either criteria for core geospatial datasets or who data custodians should be. Where it is explicit is in defining what a base dataset and data custodian is. A base dataset is those themes of geospatial information which have been captured or collected by a data custodian. Thus, the emphasis is on the defining of the data custodians rather than the core geospatial datasets themselves.*
>
> *[...] The Act describes a data custodian as an organ of state or an independent contractor or person engaged in the exercise of a public power which captures, maintains, manages, integrates, distributes or uses geospatial information. The emphasis is therefore on organs of state that have a legislated responsibility of providing geospatial data. These data custodians are to form part of the CSI [Committee for Spatial Information]."*
>
> (Schwabe & Govender, 2012, p. 30 - emphasis in original)

The report by Schwabe and Govender captures a set of discussions about who should take custodianship of various datasets. Where this is known and has been resolved, it is reflected in the SoLI Matrix for South Africa. And this is the significance of referring to certain state institutions (whether departments or state-owned entities) as the 'custodians' of certain datasets.

The report goes into great detail in defining and naming the state's 'base' datasets[27]. In this context, they are referring to a subset of this SoLI's scope of datasets, in that they are trying to identify (for the purposes of implementing the SDI Act), South Africa's core geospatial datasets.

After collecting the views of stakeholders and researching the situation in a variety of other countries, they settled on this provisional definition:

> *"Core geospatial datasets are those identified as the minimum[28] set of essential[29] datasets that are widely used as a reference base at various administrative levels to accomplish South Africa's national and international priorities.[30]"* (Schwabe & Govender, 2012, p. 44)

From their research they listed what stakeholders and survey respondents identified as base datasets. Including repetition and duplication where it arose, the datasets identified by respondents included cadastral, geodetic, rectified imagery (satellite, aerial and raster topographical maps), administration boundaries, census data, hydrology, rivers, dams, catchments, roads, streets, road centre lines, land cover, land use, physical infrastructure, addresses, government services, health facilities, place names, topography, digital elevation models, human settlements, transport, vegetation, agricultural potential etc.

They then worked this into a complex set of classified datasets that runs over five pages, but is a very useful reference list (Schwabe & Govender, 2012, pp. 48-52). Because the legal process of establishing each agency as a custodian is complex and not resolved in all cases (Schwabe & Govender, 2012, p. 61), a more simplified version will need to suffice. The following diagram is a frequent element of public presentations by the Department of Agriculture, Land Reform and Rural Development, and indicates a simplified version of state datasets and their respective custodians.

One of the key areas of activity (and debate) has been the establishment of a National Spatial Planning Data Repository (NSPDR). This seems to have arisen indirectly from some of the requirements of the SDI Act (2003). The earlier emergence of the idea can be traced back to the National Spatial information Framework (NSIF) that was established in 1997 to drive the implementation of South Africa's National Spatial Data Infrastructure, including developing a metadata repository and portal. The origins and intentions of the national repository are further described at this website of the NSDPR (Matrix Ref 14).

27 "Datasets, which may be used for many different purposes and in many different applications, are often referred to as base data, core data, fundamental data or reference data" – Schwabe and Govender quoting United Nations Economic Commission for Africa's 2003 Committee on Development Information.

28 "Minimum refers to the smallest number or set of geospatial datasets."

29 "Essential geospatial datasets are those that are absolutely necessary or extremely important."

30 "It is assumed that core geospatial datasets will be those that are systematically or programmatically maintained to appropriate standards. It is also refers predominantly to those core geospatial datasets that are mandated through some form of legislation, policy document or cabinet decision."

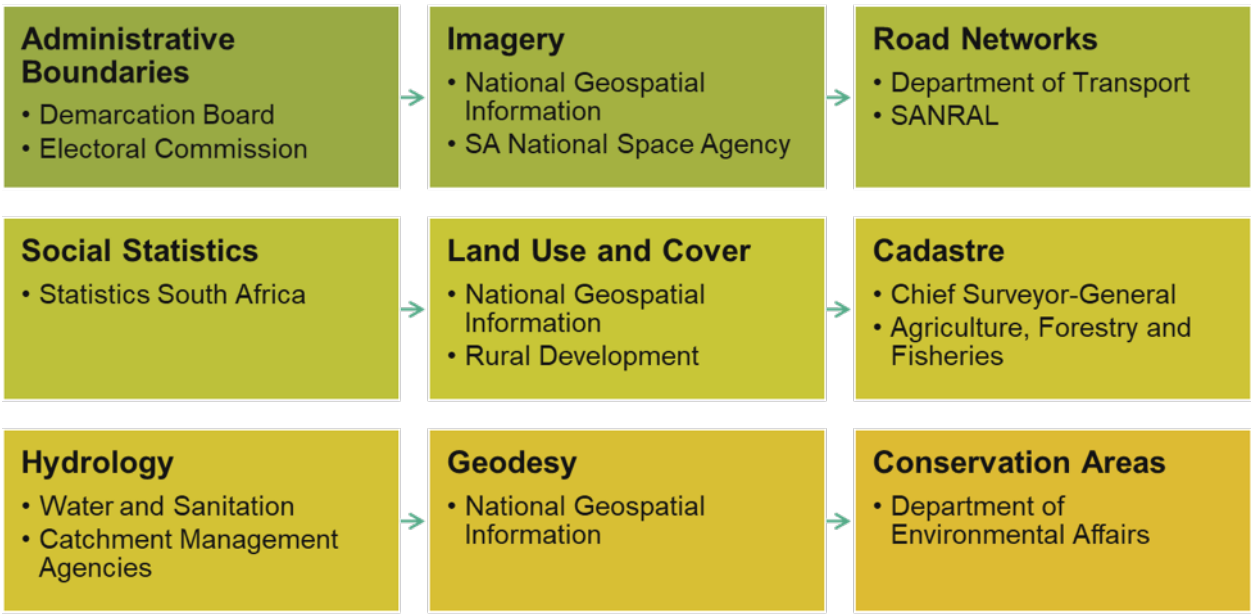


Figure 1. SA government base datasets and data custodians (Department of Rural Development and Land Reform, 2018)

Although the plans and stated intentions are in place, it is unclear when this national spatial data repository will be operational. From the website about the NSPDR, the following commitment is made:

> *"The DRDLR has already completed the User Requirements Specification (URS), Functional Requirements Specification (FRS) and Technical Specification making it possible to start with the construction of the NSPDR Ecosystem modules. The Esri South Africa and Agizo Solutions Joint Venture (JV) will develop and maintain a central repository that utilises spatial data for spatial planning, monitoring, evaluation and coordination purposes over a period of five years."*[31]

Conclusions

Government is generally well-intentioned and effective in making data and information available. We have written elsewhere (Napier, Sebake, & Rajab, 2018) however, that more work needs to be done in establishing more consistent protocols and effective practices that make information reliably available over longer time periods, more interoperable (shareable across platforms), and with less duplication of effort.

As we have seen, the move is in the right direction for the sharing of spatial data. Hopefully we can see this also advance in the more general sharing and accessibility of, for example, data from the range of departments that deal with analysis, plans and delivery, so that the state does effectively become more open, transparent, and accountable[32].

31 (http://nspdr.info/docs/about.html)accessed online Nov 2019

32 Useful as a concluding comment.

As far as whether the policy framework is well known to actors dealing with land, the best summary of expert opinion on this would be the 2013 Land Governance Assessment Framework. The summary of strengths and weaknesses in the system is also useful to bear in mind when assessing the datasets in the various categories covered in the remainder of this report.

3.2 Land Tenure Data

When scoping for land tenure data, the researchers scoped for cadastral data (of mining, forestry or agriculture cadasters) and/or land registry data (are there individual or community land records available; are these disaggregated by urban/rural areas, by gender or rights holder, by indigenous and non-indigenous peoples or communities?). The scoping research also focused on whether any evidence existed on whether or not the land registry data is contested.

Availability of data

Availability of Cadastre data

Cadastre data is usually captured in the form of a GIS map (Ref 3 Matrix). The map contains the following elements:

- Boundaries (i.e. Provincial Boundaries, Magisterial District Boundaries, Local Authority Boundaries, Farm Parcel Boundaries, Urban Parcel Boundaries, General Plan Boundaries);
- Farm Names;
- Urban Parcel Numbers;
- Places of Interest;
- Servitudes;
- Cadastral land parcel information (such as Number, name, legal extent, registered owner); and
- Third party rights (including Type of right, owner of right).

With regards to types of cadastre, there are several departments within the South African government that are charged with maintaining a specific cadastre. The Department of Agriculture, Forestry and Fisheries maintains **maps of forests** per province, including government-owned commercial forests, leased plantations, government-owned plantations, etc. (Ref 9). The South African Cadastral Spatial Data of the Chief Surveyor-General stores all **surveyed properties and registered (in the Deeds Offices) rights**, like ownership, bonds (mortgages) and servitudes (easements) in South Africa. Information on any specific erf or portion can be supplied by the Chief Surveyor General upon request or via their online map viewer. This information provides the parcel boundaries (also in the form of shapefiles), the parcel locations and the registered owner (Matrix Ref 2 & 3).

The South African Mineral Resources Administration System (Matrix Ref 52) provides a space to lodge applications for **prospecting and mining** as well as a space where anyone can view them. It allows the user to view the locality of applications, permits and rights made or held in terms of the Mineral and Petroleum Resources Development Act.

Availability of registry data

For each parcel (or 'erf'), the National Deeds Registry has records of Property details: province, diagram number, extent, owners, endorsements (bonds, etc), history (previous owners and title deeds), contract details, etc. Note that because of the nature of the registry, it is not a registry of all land rights.

Once purchased (Matrix Ref 4), individual land records can be aggregated by:

- by **urban or rural** – an approximate proxy for urban or rural would be land use (i.e. areas zoned as residential, commercial, retail, industrial etc.). Now that South Africa has a municipal boundary system which covers the whole country, whether an area is municipal or not is not indicative of urban or rural. Built up areas are identified by geo-informatics professionals using remote sensing data. This then includes places which are built up but have not been zoned (e.g. informal settlements, built up rural areas on communal land);
- by **gender of rights holder** – in the case of individual ownership or joint ownership by partners this could be derived from the identity number of property owners, but of course only covers land where a title deed over the land parcel exists. And more general information about gender of heads of household and tenure can be derived from the National Census (Matrix Ref 69);
- by **indigenous and non-indigenous individuals** – see next section, but note that the meaning of indigenous in South Africa is special (du Plessis, 2011).

Community land records are not held in the National Deeds Register other than as broad trust lands, leading to a call for land law and administration reform, for example in (Kingwill, 2019; Mahlati et al., 2019).

Many commercial companies make a business of purchasing the deeds registry data (usually at quarterly intervals), cleaning the data and making it available in various forms. The data around property transactions is particularly valuable to estate agents, investors and financiers. Many research organisations and even government departments purchase their deeds registry data from commercial organisations. There are many such companies that specialise in different areas, three examples being Lesis WinDeed (Matrix Ref 38), Lightstone Property[33] and Knowledge Factory[34]. Others specialize in linking the cadastre to the deeds registry, or the cadastre and registry to more up to date land use data (see land use section). Because the ecosystem of such companies is so large we have not tried to scan all of these.

33 https://lightstoneproperty.co.za/

34 http://www.knowledgefactory.co.za/

It appears that the South African Deeds Registry is well organized and easily available online (for a price). However the system as a whole is expensive to access for people wanting to change title deeds to reflect reality (e.g. changes in names on title deeds because of inheritance, sale, divorce, etc.), or to do research.

Availability of prevalence of state-owned land

On the issue of public land holdings, information about land owned by the national, provincial and local governments and by state-owned entities (so-called SOEs) should be provided by the Chief Registrar of Deeds with the Chief Surveyor General (Matrix Ref 51). There have been several attempts by the Chief Surveyor General and the Chief Registrar of Deeds to compile a register of all State-owned land, but these have only partly succeeded (in the form of once-off land audits) because of poor record-keeping: when land is transferred between public entities, the land is not always taken off the records of the former owner and/or added to the records for the new owner.

Land ownership audits

At Matrix Ref 85 we reference one land audit exercise in 2017[35] that was set up to gather information to answer for each land parcel:

- Who is the owner;
- Who is the occupant/user;
- The rights to the land;
- Current usage of the land;
- What buildings and improvements exist on it.

The report on the results of the 2017 land audit (covering 2014 to 2017), included information on land ownership in South Africa, private land ownership, farms and agricultural holdings ownership by race, farms and agricultural holdings ownership by gender, farms and agricultural holdings ownership by nationality, 'erven' ownership by race, 'erven' ownership by gender, 'erven' ownership by nationality, sectional title ownership, sectional title ownership by race, sectional title ownership by nationality, sectional title ownership by gender.

These land audits tend to be ad hoc exercises because the information must be gathered from a variety of sources and datasets, collated and analysed, not least spatially, and so are expensive to undertake.

The findings emerging from land audits around trends in land ownership patterns also tend to be contested by, for example, private agricultural land holders. Overlapping with the state-led land audit just mentioned, Agri-SA released the findings of their own land audit (AGRI SA, 2017). The ensuing debate was certainly well covered in the media.

35 http://www.ruraldevelopment.gov.za/phocadownload/Cadastral-Survey-management/Booklet/land%20audit%20booklet.pdf

Data and information on customary land

The Constitution does recognize customary land, however, there is a gap when it comes to recording more individual or household level rights over land in traditional areas.

There is no individualized recording of community land rights as the registry is limited to deeds (Deeds of Grant and Deeds of Sale) which only apply to 'formal' properties. This type of registry needs reform as, according to the book "Untitled", 59.7% of the South African population in 2011 were holding land outside the formal property system, with 32.8% of this being in the form of communal property (Hornby, Kingwill, Royston, & Cousins, 2018, p. 8).

There is limited data available, however, on the existence and localization of customary land. The Chief-Directorate provides a layer of "Tribal Authority Areas in South Africa" on the Mapable Viewer by the National Geo-Spatial Information (NGI). This layer shows the boundary information of tribal/customary land areas in South Africa over the land area (Matrix Ref 112). The Redistribution and Restitution in South Africa 2018 report (Matrix Ref 103) details key statistics for redistribution and restitution in the country. It also describes the beneficiaries by gender and age. The South African Legal Information Institute (SAFLII) is an online legal repository that gives the user access to a number of court cases, including ones to do with land restitution and land claims. (Matrix Ref 60)

Gaps in data or information

When it comes to information gaps for this key category of land governance "Land Tenure", some of those are in the area of aggregating who owns what, or who lives where (by gender and ethnicity), are addressed from Census data. The information gaps mainly have to do with:

- the focus of the registry on deeds, to the exclusion of other forms of tenure;
- the expense of accessing the deeds records system (affordable for small quantities of records but not for larger)
- the excessive expense and complexity of making changes to deeds (e.g. changes in who owns a property)
- the links between the cadastre and the deeds registry (a complex topic, but it seems that the links between the systems have some challenges)
- the ongoing issue around individual land rights on customary land and that this is in most cases not surveyed (no formally recorded boundaries) nor formally recorded. The powers of traditional leaders over land allocation and ongoing security of tenure is a complex, somewhat sensitive and ongoing debate.

Types of data or information

While the description above clearly highlights there is data and information available on this key category–whether or not it is disputed–from the materials collected in the Matrix, it is evident that while there is a **great prevalence of (statistical or geospatial) data over documents** (69% of the observed data units were in fact data, not documents), only 2 data units were available without needing to register or identify oneself to view the data.

Is there data?	✘

✔ = good practice ! = room for improvement ✘ = poor practice

Sources of data or information

Government institutions were the source of the large majority of data units identified in this scoping exercise, accounting for almost 70% of the total resources. Private sector actors (categorized here under 'other') account for 23% of the data units and finally Research institutions for only 7%. No data was identified from Civil Society Organizations or International Organizations on this scoping exercise.

Government	Research Institutions	National Civil Society Organization	International Organization	Other
✔	✘	✘	✘	!

✔ = good practice ! = room for improvement ✘ = poor practice

Timeliness of data and information

Cadastres are said to be updated daily (private land surveyors generally do the work, which is then validated by the Surveyors General and then lodged), and indeed, **the majority of the data units recorded in this scoping exercise (46%) cover data until 1/1/2019**. This is just within the limit of our 'green'-scoring for timeliness of data, however, it is worth noting that 38% of the resources were dated from before 2010 or unknown.

Is the data up to date?	✔

✔ = good practice ! = room for improvement ✘ = poor practice

3.3 Land Cover, Use and Management

For the Land Cover, Use and Management category, we sought to establish whether there was any land cover data or information (i.e. land surface data, soil type data) or data or information on land management (such as land consolidation, exchanges or other approaches for the readjustment of parcels or holdings).

Availability of data or information

With regards to **land surface data**, there are a number of datasets that cover this in South Africa. The South African National Space Agency has an extensive online map viewer that highlights vegetation density, NDVI & Leaf Area Indices and forest cover densities. It also shows water distribution in the country as well as human settlements, formal and informal (Ref 114). The South African National Land Cover dataset showcases the 72 different land cover classes in the country (Ref 7). The SANBI BGIS portal also offers access to some land surface datasets – for example, the KwaZulu-Natal Systematic Conservation Plan: Vegetation Types (Ref 96). Finally, access to the various satellites, such as those found at Landsat (Ref 20), the SPOT versions (Ref 21) and the MODIS, AQUA and TERRA (Ref 23), can assist with determining land surface information in the country.

With regards to **soil type data** the International Soil Reference and Information Centre's Soil and Terrain (SOTER) Database for South Africa provides information about South Africa's soil properties as well as the underlying lithology and landforms (Ref 118). The current legal responsibility lies with the Council of Geo-Science (**http://www.geoscience.org.za/**).

The terms Land 'use' and 'cover' are often conflated, unfortunately. As a result, many datasets encompass several land use, cover or management elements. The South African Land Observatory (SALO) is a civil society, university-based initiative whose objective is to make land data and documents available to improve decision making by providing access to evidence and information (Matrix Ref 13). The South African Cities Open Data Almanac (SCODA) aims to support South African cities planning, managing, monitoring and reporting needs. It provides users with information such as the State of the City Reports by the South African Cities Network (SACN) that details many issues regarding cities in South Africa. For example, the number of formal, informal and traditional dwelling types in South Africa, the open-space in the country and the population size (Matrix Ref 40).

Maintaining land use records is a local function, and as a result, in this scoping study we limited ourselves to documenting a few examples. The City of Cape Town, as a first example, provides a shapefile dataset on the undeveloped public open space in the city (Matrix. Ref 80). Many datasets can also be found at the eThekwini Municipality GIS portal relating to land use, such as the housing plan and the informal settlement programme, both of which adhere to the SPLUMA (Matrix Ref 88).

The intention is that the National Spatial Planning Data Repository (Matrix Ref 14) discussed above will also have land use data and that this will be the national platform for that, although this platform is still under development.

Gaps in data or information

Land use data is locally collected and maintained by municipalities on a plethora of platforms, and until the promised National Spatial Planning Data Repository or NSPDR is in place the absence of a national dataset on land uses based on a shared zoning system remains one of the larger gaps (or weaknesses) in the country's land information system.

The absence of regularly updated and available data on the proportion of public to private land in the country is another gap, as discussed, although we did find a number of datasets that highlighted public land holdings.

Following on from this, there is yet to be a consistent National system that regulates land use and land cover information in the country. At the moment, it is mostly at the municipal level where the data is of varying standards. The adoption and roll out of SPLUMA gives some hope, and the building of the National Spatial Planning Data Repository.

Types of data or information

A majority of the information on land cover, use and management that were identified are in the form of geospatial or statistical data (62%), mostly published by governmental or research institutions. Of all the geospatial or statistical datasets, most of them (68%) were accessible without any login barriers or other types of requirements to identify oneself prior to accessing the data. Closed datasets are inherently more difficult to assess and review, but mostly referred to data from satellite imagery, for example.

Is there data?	✔

✔ = good practice ! = room for improvement ✘ = poor practice

Sources of data or information

Similar to the previous key category, **Governmental institutions are the main source of information (54%)** for this category (for this particular scoping exercise). The Department of Cooperative Governance and Traditional Affairs (CoGTA) is responsible for ensuring all municipalities adhere to and uphold their respective mandates. Because of this department and the South African Local Government Association (SALGA) overseeing the progress, much of the information on land use, cover and management originates from governmental institutions. The private or research-based institutions tend to then refine, develop and analyse the data from the government for their own research, commercial or other purposes. Other than government, Private Sector actors (represented here under 'Other') accounted for nearly 12% of the resources identified under this category.

Government	Research Institutions	National Civil Society Organization	International Organization	Other
✔	✘	✘	✘	!

✔ = good practice ! = room for improvement ✘ = poor practice

Timeliness of data and information

For many of the data units identified under this key category, the dates of publication are unknown. This resulted in a red score. Twenty seven percent of the identified resources were published between 2010–2018, with a mere **8% covering a period from 2019 or more recent**.

Is the data up to date?	✘

✔ = good practice ! = room for improvement ✘ = poor practice

3.4 Land Disputes

For land disputes, the scoping research focused specifically on (historical) data and information. We also looked for data or information on the legal framework for land disputes resolution and specifically evidence on the effectiveness of this framework. Finally, we looked for data on concrete disputes, such as share of land affected by disputes (possibly disaggregated by type of land: agricultural, forest, urban), the number of people affected by land disputes (possibly disaggregated by type of people, indigenous/gender).

It is worth noting that the history of the country and the place of land rights in the Constitution and supporting legislation are all very core to the framework in which land disputes are settled. The topic is potentially very broad, stretching all the way from large land redistribution and restitution programmes initiated by the state (to correct historical injustices), including the system of historical land claims (and the disputes arising from that), and all the way to how legal experts sort out personal disputes over property ownership as properties are transferred from rental to ownership under programmes like the Township Discount Benefit Scheme, and allocation of deeds under the current housing subsidy scheme.

There are a wide range of dispute resolution institutions from rental tribunals, planning tribunals, once-off provincial dispute resolution panels. Customary authorities also play major roles in dispute resolution on community land held in trust.

Availability of data or information

Information on land disputes is often found in the form of case law, where disputes are addressed and (hopefully) settled. In Urban Landmark's Land Governance in South Africa report, they state that land disputes in the formal court system are less than 10% of the total court cases. However, the process is slow and expensive.

According to the over performance of land dispute resolutions using seven indicators, South Africa performs fairly poorly. (Ovens, 2012)

We do not know of a broad measurement of the numbers of all types of land disputes in the country (urban, rural, eviction, transfer of properties, expropriation, etc.) but there are few specific datasets included in the Matrix (e.g. Ref 12, AFRA's repository of farm evictions), that provide some insight.

Another type of event that may fall under the umbrella of disputes is community protests. Municipal IQ is a web-based data and intelligence service that monitors and assesses all of South Africa's municipalities, covering local protests and whether these are linked to service provision, housing or land, amongst other things (Ref 68). SAFLII also allows the user to access court cases on land disputes and could give an indication of how many people are affected.

Gaps in data or information

Because of its history, land in South Africa and how to manage it, is a highly charged issue. Despite this, there is little high level information on land disputes and land dispute resolutions in the country. The reason could possibly be due to the informality of many dispute resolutions or the expense and lengthiness of the formal system that deters citizens from following through with issues in the formal system.

According to Urban LandMark's Land Governance in South Africa, a key challenge in South Africa is that municipalities often do not recognize local 'informal' land dispute resolutions as legal. They, therefore, often overlook community-based mechanisms which ultimately delegitimizes these dispute resolutions.

Types of data or information

In this scoping study, **statistical (and/or geospatial) data was identified, making up 43% of the total resources identified under this key category**. Thirty three percent of the data sources (not documents) were accessible without needing to register or log in. An important caveat, however, is that, as described above, most of the information about land disputes are recorded in court cases, which are usually recorded in documents. For the purpose of this exercise, a repository of court cases (website, database) is recorded as one data unit. Therefore the picture shown of 'data vs documents' is somewhat skewed. Another important element to mention is that although there may be data available, this by no means indicated whether or not this captures all land disputes in the country.

Is there data?	✔

✔ = good practice ! = room for improvement ✘ = poor practice

Sources of data or information

When it comes to main information providers, **this category showed a great diversity of sources compared to the other key categories**. Government again provided the majority of the resources (43%), but followed by research institutions (29%) and National Civil Society organizations (29%). The private sector was also represented (here under 'Other'), providing 14% of the total resources identified under this key category. When it comes to specific information providers, legal repositories like SAFLII are useful resources. Besides that, Urban LandMark's work in addressing land governance in the country is extensive and indispensable when it comes to understanding the status of land reform and, with it, land disputes in the country. Groups like SERI and LandNESS (and many others) have continued with this process of elevating awareness of onto land disputes (and pursuing litigation in some cases).

Government	Research Institutions	National Civil Society Organization	International Organization	Other
✔	✔	✔	✘	!

✔ = good practice ! = room for improvement ✘ = poor practice

Timeliness of data and information

The Land Dispute information identified in this scoping study was largely published recently, either between 2010–2018 or in 2019 or more recent. A notable exception to most of the key categories, none of the identified data units dated from before 2010.

Is the data up to date?	✔

✔ = good practice ! = room for improvement ✘ = poor practice

3.5 Human Settlements

For Human Settlements, the scoping study focused on whether or not there is any data or information about the number of people without a registered address (possibly disaggregated by women, indigenous peoples, youth and other marginalized groups); legal frameworks on (social) housing provisions and any evidence of their effectiveness in practice; data on informal settlements (such as the number of people living in informal settlements; data and information about their access to basic services within informal settlements); laws and policies on regularization of tenure in informal settlements; and any evidence on the implementation and effectiveness of these policies in practice.

In addition, the scoping exercise focused on displacement and eviction information (such as the number of displaced people (possibly disaggregated by gender, youth, indigenous/non-indigenous peoples), statistics or other information about the cause of displacement (such as conflict/violence, natural disasters, development, or others) and finally, expropriation data (such as the number of expropriations, statistics or information on the provided compensation for the people that were expropriated, etc).

The topic of human settlements is fairly extensive both because of the country's history, because of the large state-funded housing programme since 1994, and because of complex dynamics between that intervention and the ongoing growth of informal housing and settlements. There are a range of active urban and rural housing NGOs and CBOs that form another part of the country's institutional ecosystem, and they use information for their programmes (advocacy, intervention, etc.) and are themselves generators of data. There are also several urban and housing under- and postgraduate programmes and centres across the country. These also actively research the topic and generate a great deal of useful knowledge and information. The levels and layers of state departments dealing with housing and human settlements have been described above, along with the legal framework (e.g. in the section on 'The institutional ecosystem' p). In some cases there are partnerships between government and universities, such as the case of the Gauteng City Regional Observatory, that deal with detailed housing, planning and quality of life information.

There is fairly comprehensive information about the housing situation in the country, and about the supply of housing through state-funded programmes, by the private sector, and through unassisted self-build.

Availability of data or information

Many cities have housing (or human settlement) departments and so also maintain localized housing data, along with Provinces. To narrow the potentially enormous scope of this research, we have only used Cape Town as an example. Generally information about housing in the country, and who lives where, is fairly comprehensive. There may be some possible exceptions such as collecting information about the numbers of homeless people in Metro's and other municipalities across the country, although there was some attempt to do this as part of the last census (2011).

The National Census[36] (Ref 69) under the custodianship of the government agency Statistics South Africa (Stats SA), is conducted every ten years. The Census method is based on a questionnaire addressed to a household living in a housing unit. The Census provides a lot of information pertaining to housing, such as dwelling type, house materials, and tenure. Intermediate Stats SA surveys also provide housing information such as the Community Survey (2017) and others. Trends analyses for demographics, migration, urbanisation and changing housing conditions between censuses is also common[37].

36 Matrix Ref 69

37 The CSIR and the Human Sciences Research Council, along with many other research and academic agencies are very active in this space.

As noted in the Matrix (Ref 47) there is not an official online **national street address dataset** other than those that appear in other datasets for different purposes (elections, post, municipal billing, etc.). Several private companies maintain and sell their own address datasets, such as AfriGIS and credit bureaus. According to Untitled: Securing Land Tenure in Urban and Rural South Africa on page 8, 59.7% of the South African population were holding land outside the formal property system in 2011. Thus, more than half of the population in South Africa do not have a registered address.

With regards to data or information on **informal settlements**, there are many laws and policies in South Africa (including but not limited to the Housing Act), indicating a responsibility for the South African government to prioritize the needs of the poor in respect of housing developments. The Housing Act also led to the Upgrading of Informal Settlement Programme (UISP) that allows municipalities to apply for government funding to redevelop informal settlements incrementally. This includes securing tenure.[38]In this space of informal settlements, there is an active movement of Civil Society. There are many initiatives undertaken by civil society to assist in garnering information on informal settlements and monitoring state interventions. Know Your City is one of these initiatives and with their informal settlement enumerations (Ref 29) they are running community-led censuses whereby a socio-economic and demographic profile is generated. The tenure status, level of services and development aspirations of each household are also documented.

Finally, with regards to data or information **expropriation**, we are not aware of a national dataset that quantifies or locates land expropriations, other than searching through court proceedings. Researches note that during the LGAF investigation process (2011/12), there was word of the existence of some kind of government record of land expropriations, but we were not able to get access to it at the time. When it comes to information about compensation provided in cases of expropriation, Kitchin & Ovens (2013) wrote:

> *"Compensation occurs within a year for between 70% and 90% of expropriated land owners. However, complaints against expropriation need to follow an expensive process, which means that the poor have little recourse. There are capacity shortfalls (financial and staffing) around management of public land. Problems around expropriation include the fact that displaced households do not have comparable assets despite receiving compensation, and that unregistered rights, such as grazing, are not usually compensated. Major problems include the long period of time taken to resolve complaints about expropriation and the fact that disposal of public land has generally not been transparent."*

38 https://www.ohchr.org/Documents/Issues/Housing/InformalSettlements/SERI.pdf

Gaps in data or information

Access to official state delivery figures has, over the last ten years, become slightly more obscure. Expenditure on state housing programmes is fairly clear from national Treasury reports, but the information about the types of 'housing opportunities' that government is funding is less clear than it used to be.

The seeming absence of a dataset (or register) of information about state land expropriations (if we are correct about this) is a gap (other than searchable information about expropriation cases in http://www.saflii.org/). Land claims, evictions and disputes information seems more accessible. It is possible that we just have not located a dataset on expropriations.

Types of data or information

Most of the information identified in this scoping study is **captured in data (75%)**. Some of the resources were available without a requirement to register (46%), but the majority of data did require some form of identification prior to accessing the data.

Is there data?	✔

✔ = good practice ! = room for improvement ✘ = poor practice

Sources of data or information

As we have seen with the other key categories, the Government remains a main source of data and information, also on the human settlements topic. Government accounts for 67% of the data units identified in this scoping exercise. Main sources within government are Stats SA, the Department of Human Settlements and the National Treasury. Notably, civil society organizations also provided a fair amount of data units for this category, accounting for 18% of the total data units identified.

Government	Research Institutions	National Civil Society Organization	International Organization	Other
✔	✘	!	✘	✘

✔ = good practice ! = room for improvement ✘ = poor practice

Timeliness of data and information

With regards to timeliness of data and information, the majority of resources were published at an unknown time or published before 2010 (39%). This means a red score needs to be given to this key category. However, it is important to note that 32% of the resources were published in or after 2019.

Is the data up to date?	✘

✔ = good practice ! = room for improvement ✘ = poor practice

3.6 Land Markets & Financing

The scoping exercise focused on land valuation information as well as land transaction data and information, such as market transaction data (disaggregated by sale and lease), market transaction data of indigenous and community lands, any information on land investments (if possible, disaggregated by public/private investments, disaggregated by scale of land areas, disaggregated by indigenous and non-indigenous lands, or foreign and domestic investments), as well as data or information on national government's foreign land investments (in other countries).

Availability of data or information

Land valuation is a key subject in a number of different university degree programmes in the country. It is of great interest of course to the private sector (e.g. the SA Property Owners' Association). There is not a national, coordinated system of property valuation as in some countries. Land valuations are recorded mainly in municipal valuation rolls as already described and documented.
The City of Cape Town provides two relevant datasets through their data portal, Cape Town Residential Property Valuations (Ref 76) and Cape Town Valuations Property Bands (Ref 77). The former details the median property valuations for every suburb in the City of Cape Town and the latter shows the number of properties falling within a certain valuation property band grouped by suburb. Municipal Money is an online tool started by the National Treasury that showcases extensive municipal financial data in order to increase transparency of government expenditure. Included in the information provided is property tax values. (Ref 19).

Outside of the private sector property sector and commercial data around property values and finance datasets that are available for purchase, we have not found a great deal of accessible online data. There are some exceptions (e.g. farm land transactions). The state's role in land parcel valuation may, ironically, become more clear and systematically applied if new legislation around expropriation comes into effect, depending on the final nature of that legislation, and the regulations that are then developed around valuations.

In terms of **land transaction price data** across the whole country (not just large scale or rural), it is publically available from the Deeds Registry office and online portal (Ref 4), however the process of acquiring the information is somewhat lengthy and costly. Other private companies, such as Lightstone, add additional value to the Deeds Registry information and it comes at an even higher cost. CityMark does a lot of work in providing the public with easily accessible data on property transactions and valuations in the eight major metros of South Africa – the two most relevant are the Interactive Housing Markets Insights 2017 Dashboard (Ref 35) and the Mortgage Lending in South Africa 2017 Dashboard (Ref 36). The Land traded in South Africa (Ref 98) dataset provides information on land traded per ward, municipality and district in South Africa by province. It also includes the average price.

When it comes to other **land investment data**, The Land Matrix provides the user with particularly useful information on land acquisitions between countries, including by other countries in South Africa and by SA in the rest of the world (Ref 41). The Sustainable Human Settlements Investment Potential Atlas (Ref 31) is a publication based on spatial analysis that was aimed at guiding the locality of investments in housing and settlements by various stakeholders.

Private sector agencies and publicly-owned investment entities (such as the Public Investment Corporation) work with a range of property value datasets that are available commercially. Apart from their own property holdings data, these can include:

- MSCI – Investment Property Database **https://www.msci.com/real-estate**
- the Rode Research Report – a quarterly panel measuring investor confidence **http://www.rode.co.za/**
- the SA Property Owners' Association datasets, e.g. office and industry vacancy reports
- Global Trade Data from Harvard Centre for International Development.

Gaps in data or information

There is a general need for more accessible and legible datasets and information about land transactions and investments. There is a fair degree of transparency in this area because all transaction data are available, but with limitations (i.e. relating to data from the Deeds Registry). Certainly there was a gap in knowledge around the extent of so-called foreign land owners (from high end coastal properties, to large agricultural holdings) as was evident when the Panel on Foreign Land Ownership was doing its work in 2007 (Department of Agriculture and Land Affairs, 2007).

Types of data or information

The large majority of knowledge about land markets & financing is captured in a statistical or geospatial dataset (89%). This is also the only key category for which all of the data was available without any sort of log in barrier or requirement of registration.

Is there data?	✔

✔ = good practice ! = room for improvement ✘ = poor practice

Sources of data or information

For the Land Markets & Financing category, Governmental institutions again account for most of the resources (44%) identified in the scoping study. Specific state actors providing information on this key category are Municipalities, the Deeds Registry and CityMark. However, this category is notably more diverse in its type of information providers than others. International organizations (such as Land Matrix), national Civil Society Organization as well as Private Sector (represented under the 'Other'-category), account for 11%, 22% and 22% respectively of the data units identified in this study.

Government	Research Institutions	National Civil Society Organization	International Organization	Other
✔	✘	!	!	!

✔ = good practice ! = room for improvement ✘ = poor practice

Timeliness of data and information

When it comes to key resources for Land Markets & Financing, one third of the key resources identified were published in 2019 or after that. The majority of the resources (55%) was published between 2010-2018 however. Notably there were no data units identified in this key category for which the publication dates could not be ascertained.

Is the data up to date?	!

✔ = good practice ! = room for improvement ✘ = poor practice

3.7 Land, Climate Change & Environment

The data and information that was scoped for under the Land, Climate Change & Environment category, was land degradation information (data on proportion of degraded land over total land area, data or information on causes of land degradation), data and information on protected areas (proportion of protected areas over total land area, data or information on existing restrictions of land use or access with regards to protected areas and evidence of possible non-compliance with restrictions) and data and information on natural disasters (data on number of natural disasters per year including disaggregation by type of natural disasters, and number of displacements due to natural disasters).

As with many of the other areas in this scoping study, this is a specialist area and more datasets would be revealed where different types of specialists begin to share their own knowledge. However, there is a lot of environmental information available online for South Africa both because of domestic and international efforts to make it so.

On the legal side, climate change mitigation requirements are quite a specialist field and one which is evolving. Based on experience, there is often a gap between what commitments are on paper and what is done in practice. However, the many departments involved in service delivery are obliged to report in terms of National Environmental management Act and development projects are subject to Environmental Impact Assessments.

Availability of data or information

One of the first categories the researchers were asked to explore is **data or information on land degradation**. In partnership with the Land Degradation Assessment in Drylands (LADA) and the International Soil Reference and Information Centre (ISRIC), the Global Assessment of Land Degradation and Improvement (GLADA) Report 2008/01 stated that South Africa has 351,555 km2 of degrading area which makes up 28.82% of the total territory (Ref 49). South Africa is one of the GLADA partner countries.

Regarding data on **natural disasters and subsequent displacement**. The Emergency Events Database (EMDAT) highlights the number and type of natural disasters per year(s), the number of people affected and deaths due to the disaster, as well as the total economic cost of damage (Ref 86). The Internal Displacement Monitoring Centre (IDMC) (Ref 119) details the number of internal displacements in South Africa each year due to conflict and environmental disasters. It uses graphs that show each year, the type of disaster and how many people were displaced because of it.

Gaps in data or information

Finding datasets that wholly answered a subcategory was the greatest challenge as often datasets only partially gave the required information. For example, the SANBI BGIS datasets would have terrestrial protected areas datasets for six out of the nine provinces' but would not have anything on the remaining three.

Or the IDMC would detail the number of people in a nation displaced by a natural disaster but not tell you how many natural disasters occurred in a given year.

It was also challenging to always find up-to-date information. Often, data portals would not be regularly updated and datasets would be exactly what you were looking for but they were outdated.

Types of data or information

In this scoping study, **the resources identified under this key category were mostly (statistical or geospatial) data (95%)**. From these data sources, only 40% was available without needing to register or provide an email address.

Is there data?	✔

✔ = good practice ! = room for improvement ✘ = poor practice

Sources of data or information

Once again, the large majority of sources of data units identified for this key category is **Government, accounting for no less than 81% of the total resources identified in this scoping study**. Research institutions follow with just shy of 20% — for the other categories of sources, no information sources were identified in this scoping study.

Government	Research Institutions	National Civil Society Organization	International Organization	Other
✔	!	✘	✘	✘

✔ = good practice ! = room for improvement ✘ = poor practice

Timeliness of data and information

The majority of Land, climate Change and Environment **resources identified in this scoping study were published between 2010-2018, namely 57%**. Twenty three percent of the resources identified were published in 2019 or later, including disaster risk profiles that date until 2050.

Is the data up to date?	!

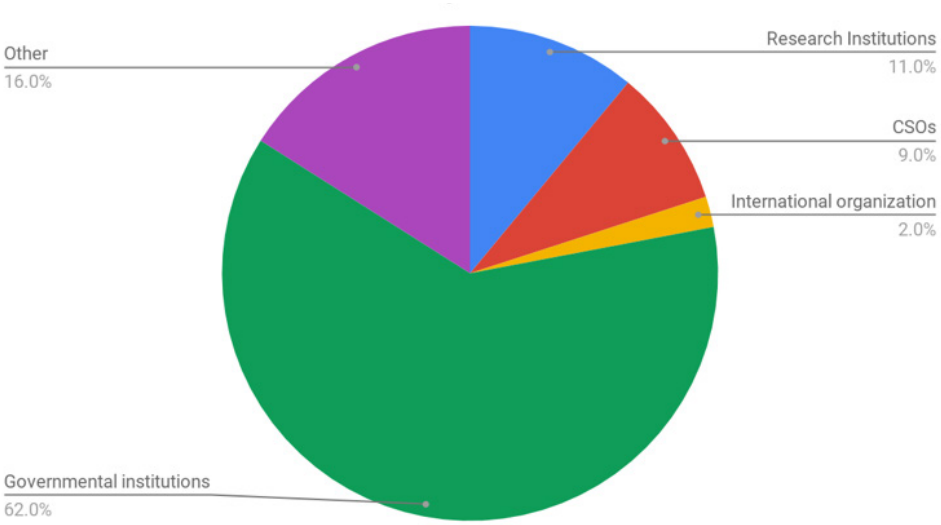
✔ = good practice ! = room for improvement ✘ = poor practice

3.8 Overall availability of land data and information

Overall, we conclude that as regards land South Africa has largely a data ecosystem and not an information ecosystem, as we have seen in our previous studies on four countries in East Africa. However, this is of course depending on the scoping exercise performed. Our findings show that **over 66% of key land information resources in South Africa can be found in documents, not datasets**. The documents and data available are all predominantly available online (97%), though this is a very skewed picture considering that offline data and information are harder to access and therefore easily missed in scoping the information landscape.

As regards availability of up-to-date information, only 6% of the information is dated from before 2010, which is remarkable considering the comparison with the East African countries researched (with the only notable exception of South Sudan, which is a relatively young state). We were unable to determine the date of publication or creation for 11% of the information. Altogether, this means that 17% of the total resources are outdated or lack details about date of publication or production. Most of the information identified was published between 2010 and 2018 (43%), with 39% of the resources identified in this scoping study published in 2019 or later.

A data ecosystem is defined not only by the type and coverage of the information it contains, but also by its data and information providers. The source of data and information is almost as important as its content. As consumers of data and information, our judgment of the accuracy and reliability of the data is, to a large extent, based on our perception of the trustworthiness of the source. From the identified datasets and other resources on key categories of information on land governance in South Africa, the division of types of information providers can be grouped as follows:



This is a considerably different picture than the study in Kenya, Tanzania, Uganda and South Sudan have shown. Clearly, the government in South Africa publishes a lot of data, accounting for well over half of the total resources identified (62%). A clear gap here is though the data governance, and the need for a clear overview of custodianship over certain key governmental datasets. Private sector (represented under the 'Other'-category) is also notably more present than in the four East African countries. One of the most pressing conclusions from this scoping exercise is the difficulty of accessing data and information from National Civil Society Organizations. Indeed, it is well known that CSOs collect and publish a lot of data and information, and there is a vibrant CSO-collective working on land in the country. Their data and information is notably missing from this scoping exercise, partly due to the difficult and cumbersome nature of attaining it (often published on individual websites and largely dependent on a personal network or connection). The South Africa Land Observatory (SALO), hosted at the University of Pretoria, used to fulfill a role here, however, this initiative has been dormant for several years. Initiatives such as Urban Knowledge Exchange for Southern Africa (UKESA) or Knowledgebase. LAND have stepped in, but do not have the same focus or scope as the SALO database once had(i.e. UKESA and KBL do not focus on statistical or geospatial data, whereas this scoping study proves a lot of such data exist and is accessible). An important recommendation for the next iteration of this scoping study however, is to look more closely and intensively at this category of sources.

Accessibility of South Africa land information ecosystem

Having mapped the information ecosystem based on availability, type and relevancy of the data and information, the study subsequently focuses on the accessibility of the data and information. The criteria to assess the accessibility are based on Open Data principles as laid out in the initiatives highlighted in Chapter 2 of this report. The final criteria against which each document or dataset was assessed against are: 1) Online; 2) Accessible; 3) Free; 4) Metadata; 5) Standards; 6) Downloadable; 7) Open License; 8) Machine Readability; and 9) (Linked) Data URI for key elements of the data.

In this chapter we highlight, per criteria, how the various data and information sources on key land issues are ranked. For each criterion, we provide a general score. Green indicates a good practice; orange indicates a practice that can be improved; and red indicates a poor practice. More details on how those scores are allocated can be found in Annex I–Scoring Chart. The chapter concludes with an overall assessment of these criteria combined to provide one measurement for the state of South Africa's data and information ecosystem.

4.1 Online

A first criterion to assess the accessibility of key land resources is whether or not the information is available online or offline. The findings of the scoping exercise are positive and indicate **97% of the key resources are available online**.

Why is it important that data and information are online?

Only 55% of the world's population makes use of the Internet as of June 2018.[39] A valid question therefore is why data or information being online is one of the criteria used to define accessibility. There has been an exponential increase in Internet users in the last few years, particularly in the global South. Another undeniable advantage of the Internet is that knowledge can reach a great audience at an unequalled speed and scale than any other medium. The potential of knowledge being put into practice in other parts of the world, is endless. To ensure maximum reach and impact of data or information, making it available online is essential.

39 Internet World Stats, "Internet Users in the World by Regions", June 30, 2018, Miniwatts Marketing Group.

The representation of online materials through this scoping research may be skewed, considering the scoping research was largely a desktop study and offline materials are more difficult to scope. A potential recommendation for continued or future expansion of this scoping research could be to apply scoping methods to allow for more inclusion of offline sources and resources.

Overall Score "Online"	✔

4.2 Accessibility

The 'Accessibility' criterion looks into the ease with which the resource may be accessed. We studied whether users are required to register, log in or perhaps request access, to be able to study the complete resource of key land information. For this criterion, **only 55% of the key land resources in South Africa were available without any login or identification barriers**. This is a notable difference from the scores in the scoping studies of the East African countries.

Overall Score "Accessibility"	!

4.3 Free

Another important criterion that helps determine the extent to which data and information are accessible and useful to a wider audience, is whether or not they are available for free (unpaid). Particularly in the academic or the private sector, data and other research findings are often hidden behind (publisher) paywalls. So how about key land resources in South Africa? Our research findings suggest that the data and information ecosystem overall is freely accessible, with **75% of the data and information available on the web for free.** Just shy of 11% of the resources are behind a paywall and there were also some data services that provided part of the data for free, but for more details one needed to pay.

Overall Score "Free"	✔

4.4 Metadata

Crucial to the accessibility of data and information is being able to find it on the web. Metadata, or information about the data or information, is key to catalogue data and information in databases or repositories.

What is metadata and why does it matter?

Metadata, or 'data about data', explains a dataset or information resource and allows for data providers as well as users to understand what the data or information resource is about at a later time.[40] Metadata provides information on the source of the data, the date of publication and other important characteristics of the data. Metadata therefore plays an important role in the usability of the data or information resource. But it is not only that, metadata also plays a key role in discoverability of data and information resources on the web, playing a key role in cataloguing of resources in databases and for search engine optimization.

From the key land resources identified in this scoping exercise, **57% of the data and information came accompanied with metadata**. For 23% of the resources we were unable to verify the availability of metadata, considering the researchers of this study were unable to assess those resources behind paywalls.

Overall Score "Metadata"	!

4.5 Standards

The standards criterion is based on the FAIR-principles and is arguably one of the more subjective criteria to assess accessibility of key land data and information in this study. The importance of standards in accessibility of data is largely uncontested, the qualification of whether something is a 'standard' or not is mostly subjective. The approach taken here is to assess whether any kind of standard is used, whether that is a standard way to classify geographical or topical coverage, or the type of metadata fields. South Africans have been active participants for over 20 years in the development of standards by ISO/TC 211, Geographic information/Geomatics, including leading several projects developing standards, and some of these standards are being adopted by the CSI.

Sixteen percent of the data and information providers make use of standards in their data or their metadata. **Of the total resources, including those that did not provide metadata, 39% of the key resources were published using standards either within the data or its metadata.** Commonly found standards were ISO3 codes as well as DOI identifiers, when publishing data. Again, this is a notable difference with countries assessed in East Africa.

40 GODAN Action, "Open Data in Agriculture & Nutrition: Making Data Open", November 2017.

Overall Score "Standards"	!

Potential of a Standard Vocabulary for Land

Land is a topic which is debated across the world, in many natural languages and in a variety of different (academic) disciplines. Having a common and standard vocabulary to classify data and information to ensure no perspective is lost, is therefore very important. When a grassroots NGO wants to spread its good practice on mapping land boundaries in a "favela" in Rio de Janeiro, it would be a missed opportunity if this could not be applied in a "township" in Johannesburg, simply due to a linguistic difference in describing an issue–and therefore the right connections are not being made. To accommodate for the fact that no vocabulary standard for land existed, the Land Portal helped facilitate the establishment of LandVoc, the Linked Land Governance Thesaurus.[41] LandVoc is a part of widely accepted agriculture thesaurus by the Food and Agricultural Organization, AGROVOC.

4.6 Downloadable

A measure of accessibility that is crucial for the usability of the data and information, is whether or not the data or information can be downloaded by the user. Downloading the data allows a user to perform more rigorous data analysis and application for their particular use; it is also important to be able to reach offline communities and make the data or information useful to them.

In principle, many of the key land resources are downloadable by the user. **About 8% of the data providers actually prevent a user from downloading the data** and restrict its use to their own platform (think of data visualizations on a website that do show the data and information, but do not allow for downloading the raw data).

In order to meet the accessibility criteria, the data and information should be downloadable in bulk and/or queried in bulk through an API or other access protocol. The data and information providers scored fairly well in this criterion, with 68% of the resources being downloadable, many of which in bulk and/or API.

Overall Score "Downloadable"	✔

41 https://www.landvoc.org

4.7 Open License

A license regulates the manner in which data and information can be used. It is one of the cornerstones of Open Data, because the Open Definition[42] specifies that open data should be allowed to be used, re-used and modified by anyone and for any purposes. This includes commercial purposes, thereby allowing a data user to make a profit out of the use and application of another party's data.

Why does a License matter?

When it comes to data and information about land, privacy and safety concerns are always important topics to consider. They are common incentives for data and information providers not to publish their data at all. Paradoxically, if this data is opened up by using an open license, it can protect because the license facilitates a controlled and steered way in which the data can be used. An open license allows for the best of both worlds: safe and controlled publishing as well as increased awareness and (controlled) use of the dataset. An open license is a key element for a democratized data and information ecosystem.

From the key resources on land in South Africa, **55% of the information providers have applied a license to their resources**. Forty four percent did not provide any type of license when publishing the data. **Of the total resources, only 16% provided an Open License**.

Overall Score "Open License"	✗

4.8 Machine Readability

The criteria of machine readability is a common criteria used to assess compliance with (linked) open data principles. As mentioned, the Open Definition includes that data and information should be able to be re-used and modified by anyone for whatever purposes. For users to be able to modify, re-use and build on existing data–for example by designing innovations or technologies based on the data–the data needs to be in a machine readable format. A machine readable format means that a machine (a computer) can easily process the data.

42 https://opendefinition.org/

Fifty percent of the key resources related to land in South Africa are published in a machine readable format, which is incredibly high compared to the countries researched in East Africa. The most commonly used formats for data and information are PDFs (not machine readable[43]) and shapefiles. An important caveat to mention with this criterion is that machine readability in the Open Data assessment tools on which these Accessibility criteria are based, really applies to raw, numerical data–not documents. The laws and legislations various websites, for example, are available both in PDF as well as HTML formats. HTML is a machine readable format. The application of this criterion on such documents (which, as mentioned, account for 85% of the key land resources in South Africa) needs to be interpreted carefully; having an HTML page through which a computer could process the contents, does not mean that the raw HTML code allows for 'clean' data exchange or application in technologies without any manual intervention. To mitigate this, the criterion was applied to the metadata of documents, where possible, not the document itself.

Overall Score "Machine Readability" !

4.9 (Linked) Data URI

The final criterion in our Open Data-compliance assessment is investigating whether the key land resources can be awarded the fourth star of the famous "Five Stars" of Linked Open Data.[44] This fourth star is awarded to a dataset if it contains URIs: a Uniform Resource Identifier. The URI was invented by Sir Tim Berners-Lee as a protocol to provide a unique 'identifier' to a resource, a piece of data. This unique identifier is usually in the form of a code that should not change in the future; it is an ever-fixed reference point in the World Wide Web, completely unique for this one resource. Each indicator, piece of data and overall dataset should have a URI to comply with the fifth star of Linked Open Data. If that URI refers to (links) to other URIs, we create what Sir Berners-Lee called the "linked web".[45]

South Africa attains a very low score when it comes to this criterion. **Only one of the key resources contained unique identifiers to classify key elements or the data or metadata.**

Overall Score "(Linked) URIs" ✗

43 More specifically, PDFs can be read by computers but are not easily processable by machines.

44 Berners-Lee, "5 Stars of Linked Open Data", consulted website September 2018: https://5stardata.info/en/.

45 Idem.

4.10 Overall Accessibility assessment

The **South Africa Land Data and Information Ecosystem scores relatively well with a basic interpretation of accessibility, namely whether it is online and free.** Notably, many resources do require some sort of login or identification of the user prior to being able to access the data. However, true accessibility of data goes much beyond these three criteria. True accessibility of data and information means that any person is free to use, re-use and modify the data and information for any possible purpose and that the data and information is published in such a way that allows for effective and unrestricted flow across websites and to and from people. For these latter accessibility criteria, the South Africa Land Data and Information Ecosystem scores less high, but notably much better than countries previously researched in East Africa: Kenya, Tanzania, Uganda and South Sudan.

The discoverability of the resources within the ecosystem can be improved, but is not in bad shape. Well over half of the data and information were published with metadata, fewer have made consistent use of standards in their metadata. In several instances, a publication date of a particular resource was untraceable. Not only do such data publishing practices make the data and information less discoverable on the web (metadata and standards strengthen the (relevant) cataloguing in databases and the web in general), but it also restricts the possible use of the resource — metadata often contains vital information for a user to determine whether or not the resource is of relevance or of sufficient quality and reliability for them to use. This means that there is definitely still room for improvement.

Equally scores for other criteria that are intended to promote the use of the data, for whatever purpose, are a significant improvement from the previous SoLI studies. No less than 68% of the data and information are available to download in bulk (with many others simply presenting data in a visualization format on a website, without the ability to download). Where there is room for improvement is in the use of licensing. Only 16% of the data providers apply an open license to their data. What's even more striking, is that almost half (44%) of the resources were published without specifying a license! These criteria are at the very core of the Open Definition. Using, re-applying and building on data and information has an enormous potential and can increase the impact of the knowledge considerably. A more positive score for another criteria that supports re-use and modification of data and information, machine readability of data and information: 49% of the data and information are made available in a machine readable format.

Finally, as regards having unique identifiers (URIs) for key elements of data and/or metadata and linking to other URIs, only one key resource provided this in their data.

Overall Score South Africa "Data & Information Ecosystem"	! 60/105

Conclusions and recommendations

It is an often-repeated rhetoric that there is a lack of land data–that the data is either unavailable, or if it is available, that it is unreliable and/or out of date. With this State of Land Information Report we seek to provide an overview of existing data and information on key land issues. Our aim was to uncover the many different sources of land data and information in South Africa and thus provide a basis to substantiate, refute or nuance the rhetoric that no land data exists. For the very first time, we looked at the entire landscape of data and information related to land in South Africa, assessing 104 land resources from 59 different sources, to see trends and gaps when it comes to data collection as well as how accessible it is on the world wide web. Ultimately, we hope to improve the overall health of the South Africa Data & Information Ecosystem on land.

The statement that there is a lack of data can partially be refuted: **our scoping exercise shows that 67% key land resources are available as statistical or geospatial data, not documents**, as was the case in our studies in East Africa. However, we did not assess how complete or accurate this data is, which remains an area of further work. Another important caveat to saying that there is "no lack of data", is that 60% of the key resources were either from before 2019 or the publication date was unidentifiable. This is a significant constraint for these resources to be useful or used.

Our research shows that the knowledge is published online (97%) and it is mostly available for free (75%). The rudimentary access to data and information there seems to be in a good state in the South Africa Data and Information Ecosystem, with a notable exception of the fact that for many of the key resources identified (40%) there was still a login barrier or some kind of requirement to identify oneself prior to accessing the data.

Another important aspect that defines the usability of a resource for a user, is knowing the source of the data or information. **In this particular scoping study, the government of South Africa was identified as the main provider of data (over 60%)**. This is very different from the four countries in East Africa. Where South Africa does not deviate from the other four East African countries in terms of weakest link in information provision are national **Civil Society Organizations**, who accounted for less than 9% of the total resources identified and provided little information for almost each key category. This is not necessarily because CSOs do not have data, information or knowledge to share, and may well reflect on their poor information sharing practices, and demonstrate the need to improve the discoverability of their perspectives online. The South Africa Land Observatory (SALO), hosted at the University of Pretoria, used to fulfil a role here, however, this initiative has been dormant for several years. Initiatives such as Urban Knowledge Exchange for Southern Africa (UKESA) or Knowledgebase. LAND have stepped in, but do not have the same focus or scope as the SALO database once had (i.e. UKESA and KBL do not focus on statistical or geospatial data, whereas this scoping study proves a lot of such data exist and is accessible).

Availability of Data and Information							
Key Category	Data available?	Representation of Sources					Data up-to-date?
		Government	Research Institutions	National CSOs	International Organisations	Other	
Land Tenure Data	✗	✔	✗	✗	✗	!	✔
Land Cover, Use & Management	✔	✔	✗	✗	✗	!	✗
Land Disputes	✔	✔	✗	✔	✗	!	✔
Human Settlements	✔	✔	✗	!	✗	✗	✗
Land Markets & Financing	✔	✔	✗	!	!	!	!
Land, Climate Change & Environment	✔	✔	!	✗	✗	✗	!

✔ = good practice ! = room for improvement ✗ = poor practice

As mentioned, on a basic level (available online and for free), the South Africa information ecosystem performs relatively well. **When it comes to more sophisticated accessibility, the country scores much better than the previously researched East African countries**. Almost half of the key data units identified in this scoping study were available to download in bulk (through direct download or APIs), which leaves a lot of room for improvement, but still is considerably higher than we've seen in Kenya, Tanzania, Uganda or South Sudan. When it comes to providing metadata, which is an important practice to make data and information more discoverable on the web, the identified South African information providers perform a lot better than in the East African countries as well. Where in those countries the provision of metadata averaged around 30% of the resources identified, in South Africa we've found that almost 60% of the data was provided with some form of metadata. Almost 70% of those that provided metadata, also used standards when publishing their data (country codes or DOI identifiers were used most frequently). Another notable difference with the SoLI reports in Kenya, Tanzania, South Sudan and Uganda, was the prevalence of resources that were provided in a machine readable format. Half of the resources were provided in a machine readable format, which increases the usability and discoverability considerably.

However, **there is still room for improvement**. Many of the resources identified required some sort of login or identification prior to accessing the data. And even though the majority of the key resources were identified with a license, only 30% of the resources were actually licensed openly, which is an essential criteria that promotes the user of data, for whatever purpose. South Africa does not perform much better than the other four East African countries when it comes to resources where no license was found: no less than 44% of the key resources identified did not specify any kind of license. This restricts the possible use of the resource, and thus its impact, considerably. Finally, the last 'star' of the Linked Open Data 5-star system, namely having unique identifiers (URIs) for key elements of data and/or metadata and linking to other URIs, only one of the key resources had this included in their data.

Accessibility of Data and Information								
Online	No (log in) barriers	Free (unpaid)	Metadata	Standards	Downloadable	Open License	Machine-readability	(Linked) Data URIs
✔	!	✔	!	!	✔	✘	!	✘

✔ = good practice ! = room for improvement ✘ = poor practice

Overall, the health of the South Africa Land Data and Information Ecosystem is scored with 60/105 points. This score is considerably better than what we've seen in Kenya, Tanzania, Uganda or South Sudan. This does not mean there is not considerable room for improvement. An important caveat we want to reiterate is that this is the result of the assessment of the resources identified during this study; there is a limited view there on what is not present. Another caveat is that this study only covers availability and accessibility of data; there is no assessment or judgment on the completeness or accuracy of the data.

The following recommendations are made to improve the state of the Land Data and Information Ecosystem in South Africa:

1. **Continue the process of clarifying custodianship of governmental datasets** within government to ensure proper maintenance of the respective databases;
2. **Consider facilitation of more equitable access to data by removing login requirements or payment barriers wherever legally possible**, at least for certain groups, and/or certain elements of the data;
3. Support & enforce data publishing practices to **include a minimum set of metadata** with each publication, dataset or other type of information published by any type of information providers;
4. **Support & enforce the use of standards** when publishing data and metadata to promote the usability as well as interoperability of data and information in the South African data & information ecosystem;

5. **Institute a system of publishing downloadable 'raw data' alongside tableaus and visualization** methods published online to enable more re-use of data and information;
6. **Apply open licenses to published data and information** to allow for more meaningful and in depth use, re-use and modification of data and information to increase its impact, and most importantly, consider licensing and publish it along with the data and information;
7. **Apply unique identifiers** to key elements of the data to ensure consistent and reference to the data and information, and allows for more efficient exchange within the data ecosystem;
8. **Commission specific research and action into availability of data and information from civil society organizations** or NGOs, to gain a further understanding in their data and information supplies as well as sharing practices.

References

AGRI SA. (2017). Land Audit: A Transactions Approach. (November), 1–34. Retrieved from **https://www.agrisa.co.za/wp-content/uploads/2017/11/AgriSA_Land-Audit_November-2017.pdf**

Berners-Lee, T. 5 Stars of Linked Open Data. Consulted September 2018: **https://5stardata.info/en/**.

Department of Agriculture and Land Affairs. (2007). Report and Recommendations by the Panel of Experts on the Development of Policy Regarding Land Ownership by Foreigners in South Africa. In Department of Agriculture and Land Affairs (Vol. 46). Retrieved from **http://www.ruraldevelopment.gov.za/phocadownload/Foreign-land-owners/PLOF_Report.pdf**

du Plessis, J. (2011). LGAF definitions adapted for South Africa.

European Commission. (2017). Open Data Maturity in Europe 2017. Open Data for a European Data Economy. .

Food and Agriculture Organization of the United Nations. (2012). Voluntary Guidelines on the Responsible Governance of Tenure of Land, Fisheries and Forests in the Context of National Food Security. Rome.

Gantz & Reinsel. (2012). IDC's Digital Universe Study. EMC.

Global Land Tool Network. Global Land Indicators Initiative. Consulted July 2018: **https://gltn.net/global-land-indicators-initiative-glii/**

GODAN Action (2017). Open Data in Agriculture & Nutrition: Making Data Open.

Gurin, J. (2014). Big data and open data: what's what and why does it matter?. The Guardian.

Hornby, D., Kingwill, R., Royston, L., & Cousins, B. (2018). Untitled: Securing Land Tenure in Urban and Rural South Africa. Pietermaritzburg: University of KwaZulu-Natal Press.

International Food and Policy Research Institute. Monitoring & Evaluation of Land in Africa. Consulted July 2018: **https://melafrica.wordpress.com/**

International Land Coalition. (2018). The Dashboard Indicators. Rome.

Internet World Stats. (2018) Internet Users in the World by Regions. Miniwatts Marketing Group.

Kingwill, R. (2019). Discussion document for LandNNES Inclusive Land Administration in the context of People Centred Land Governance.

Kitchin, F., & Ovens, W. (2013). Land Governance in South Africa: Implementing the Land Governance Assessment Framework: booklet.

Matfess & Smith (2018). Argument: Africa's Attack on Internet Freedom. Foreign Policy.

Mahlati, V., Hall, R., Karaan, M., Kriek, D., Mabasa, B., Moagi, T., ... Sihlobo, W. (2019). Final Report of the Presidential Advisory Panel on Land Reform and Agriculture. [online]. Republic of South Africa. [Viewed 20 March 2020]. Available at: **https://www.gov.za/documents/final-report-presidential-advisory-panel-land-reform-and-agriculture-28-jul-2019-0000**

Manona, S. (2019). Discussion document for LandNNES: An overview of South Africa's Land Data Ecosystem for a People Centered Land Governance system.

Napier, M., Sebake, N., & Rajab, A. (2018). Policy Brief: Guidelines for government departments and agencies to commission web-based, integrated, long term knowledge platforms of built environment know-how in partnership with state and non-state actors.

National Treasury. (2018). Catalytic Land Development Guideline. Retrieved from **https://csp.treasury.gov.za/Projectdocuments/01 CLD Guideline Aug 18 PDF.pdf**

Open Data Watch. (2018). Open Data Inventory 2017. Methodology Report.

Open Knowledge International. (2015). Global Open Data Index.

Open Knowledge International. Global Open Data Index. Methodology. Consulted September 2018: **https://index.okfn.org/methodology/**

Ovens, W. (2012). Improving Land Sector Governance in South Africa–Country Report.

Parliament of the Republic of South Africa. (2019). Current Composition of the Newly Sworn-In 6th Parliament [online]. Parliament of the Republic of South Africa. [Viewed 20 March 2020]. Available from: **https://www.parliament.gov.za/press-releases/current-composition-newly-sworn-6th-parliament**

Parliament of the Republic of South Africa. (2019). Establishment of the 6th Democratic Parliament [online]. Parliament of the Republic of South Africa. [Viewed 20 March 2020].
Available from: **https://www.parliament.gov.za/project-event-details/43**

Platzky, L., & Walker, C. (1985). The surplus people: Forced removals in South Africa. Johannesburg: Ravan Press.

Republic of South Africa. Constitution of the Republic of South Africa. , Government Gazette § (1996).

Right of Access to Information Act, Republic of South Africa, 2013, consulted website March 2019: **http://www.ssbalaw.org/downloads/**

SA Government. (1996). Constitution of the Republic of South Africa.

SA Government. (2018). Draft National Spatial Development Framework.

Samasuwo, N. (2004). Gentrification, foreign land-ownership and market-led land reforms in South Africa. Global Insight, Issue No 38, 11. Retrieved from **http://www.igd.org.za/jdownloads/Global%20Insight/gi_38_part_2.pdf**

Schwabe, C., & Govender, S. (2012). Stakeholder survey on defining the criteria and identifying core geospatial datasets and data custodians in South Africa. Retrieved from **https://erln.gtac.gov.za/images/jevents/54575518073572.20358747.pdf**

South African Cities Network, Abrahams, G., & Berrisford, S. (2012). Addressing the crisis of planning law reform in South Africa. Retrieved from **http://sacitiesnetwork.co.za/**

Ubaldi, B. (2013). Open Government Data: Towards Empirical Analysis of Open Government Data Initiatives. OECD Working Papers on Public Governance, No. 22, OECD Publishing, Paris.

United Nations. About the Sustainable Development Goals. Consulted July 2018: **https://www.un.org/sustainabledevelopment/sustainable-development-goals/**

Wilkinson, Dumontier et al. (2016). The FAIR Guiding Principles for scientific data management and stewardship. Scientific Data No 3.

World Bank. Land Governance Assessment Framework. Consulted July 2018: **http://www.worldbank.org/en/programs/land-governance-assessment-framework**

World Wide Web Foundation. Open Data Barometer. Methodology. Consulted September 2018: **https://opendatabarometer.org/leadersedition/methodology/**.

World Web Foundation. (2018). Africa Data Revolution Report 2018. Status and Emerging Impact of Open Data in Africa.

Annex I
Scoring Chart

For ease of reference and understanding, the various criteria used in availability and accessibility assessments in this study have been collated into three scoring categories highlighted through colors: green (good); orange (good, but room for improvement; and red (poor). This Scoring Chart highlights for each individual assessment, how a certain scoring category was determined and allocated.

Types of Data Criteria

We assessed per key land category whether or not there is statistical and/or geospatial data available. Please find below the scoring:

Criteria	Scoring Chart
Statistical and/or geospatial data is available and accessible, with 33% or fewer of the datasets accessible only after registering or identifying yourself.	✔
Statistical and/or geospatial data, but more than 33% of the datasets are not accessible without having to register or identify yourself.	!
Statistical data is not available	✘

✔ = good practice ! = room for improvement ✘ = poor practice

Representation of Types of Sources Criteria

Per key category of land issues, we highlighted the groups of sources and assessed their contribution to the key resources identified for each respective category. The following types of data and information providers were grouped together:

- Governmental Institutions;
- Research Institutions (including universities);
- (National) Civil Society Organizations;
- International Organizations;
- Other.

Whenever a data source was an international research institution or international civil society organization, these were grouped under ‘international organizations’, in order to highlight as much as possible whether a perspective was ‘local’ or not.

The classification of the representation of these groups for a particular category was done as follows:

Criteria	Scoring Chart
Group accounted for more than 25% of the total resources per category	✔
Group accounted for between 11% and 24% of the total resources per category	!
Group accounted for less than 10% of the total resources per category	✘

✔ = good practice ! = room for improvement ✘ = poor practice

The threshold of 25% for the green score was chosen relatively low to avoid misrepresentation of perspectives and reduce the chance that the scoring of one group is too heavily dependent on the actions of another group. For example, in the event many different groups provided a similar amount of resources per category, the respective percentages of the total would automatically be on the lower side (if all provided the same amount, all would account for 20% of the resources for a category). Similarly, if one group of information providers simply provided an extremely large volume in comparison with the other groups, other groups–even though they might also provide a fair amount of data and information–would rank lower simply because another group increased the total significantly. To allocate absolute number-thresholds was not possible either because that would have been heavily dependent on each category and differ per each country.

Timeliness Criteria

For each key category of information, we assigned a red, orange or green score indicating whether or not the key resources are up-to-date. The scoring based on the findings was done as follows:

Criteria	Scoring Chart
Majority of resources were either not-dated or published before 2010	✘
Majority of resources were published between 2010 and 2019	!
Majority of resources were published since 2019	✔

✔ = good practice ! = room for improvement ✘ = poor practice

Laws, policies and other legal documentation were purposely left out of this assessment, as it is not in the nature of legal documents to be regularly updated.

Accessibility Criteria

To determine the accessibility of the key land resources in South Africa, the resources were assessed against the following criteria:

- Online;
- Accessible (no registration or other types of barriers);
- Free (unpaid);
- Metadata;
- Standards;
- Downloadable
- Openly Licensed;
- Machine Readable;
- (Linked) data URIs.

We allocated one score (red, orange or green) for each category, assessing all the key resources identified. The scoring was based on the following criteria:

Criteria	Scoring Chart
Accessibility criteria is met by 33,32% or fewer of the total key land resources	✘
Accessibility criteria is met by between 33,33% and 66,66% of the total key land resources	!
Accessibility criteria is met by or more than 66,67% of the total key land resources	✔

✔ = good practice ! = room for improvement ✘ = poor practice

Overall Accessibility Score

Not each of the nine accessibility criteria is generally considered of equal importance. Therefore, to accommodate for that fact and provide a general assessment for ease of reference and understanding, an "overall accessibility" score has been given to assess the overall "health of the Data and Information Ecosystem in the country.

Following the Open Data Barometer methodology[46], particular weight is given to the criteria Free (3), Downloadable (6), Openly Licensed (7) and Machine Readable (8). Points per criteria along with their associated weight have been incorporated as follows:

46 World Wide Web Foundation, "Open Data Barometer. Methodology", consulted website September 2018: **https://opendatabarometer.org/leadersedition/methodology/**.

Accessibility Criteria	Points if red score	Points if orange score	Points if green score
Online	0	5	10
Accessible	0	5	10
Free	0	5	15
Metadata	0	5	10
Standards	0	5	10
Downloadable	0	5	15
Openly Licensed	0	5	15
Machine Readable	0	5	15
(Linked) Data URIs	0	2	5

✔ = good practice ! = room for improvement ✘ = poor practice

The total score (if all green scores are given) can be 105 points. Based on the scoring per country of the overall accessibility, a subsequent green, orange or red score will be given to the "overall accessibility" of the information ecosystem. This ranking is allocated as follows:

Criteria	Scoring Chart
Total points below 35	✘
Total points between 35 and 65	!
Total points of 65 and higher	✔

✔ = good practice ! = room for improvement ✘ = poor practice

Notes

Notes

LAND PORTAL FOUNDATION
Postbus 716, 9700 AS
Groningen, the Netherlands

www.landportal.org
contact@landportal.org