

Kenya



# State of land information Kenya

Uncovering Kenya'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.



This work is licensed under the Creative Commons Attribution 4.0 International License. To view a copy of this license, visit <https://creativecommons.org/licenses/by/4.0/>  
Graphic design: Federico Pinci.

# State of land information Kenya

## Uncovering Kenya's land information ecosystem

L. Mey, M. Ochieng Odhiambo,  
C. Tejo-Alonso, M. DeMaria

February 2019

# Contents

<b>Foreword</b>	<b>6</b>	<b>Land Markets &amp; Financing</b>	<b>25</b>
<b>Executive Summary</b>	<b>8</b>	Types of data or information	25
<b>Recommendations</b>	<b>10</b>	Sources of data or information	25
<b>Introduction</b>	<b>12</b>	Timeliness of data and information	26
<b>Methodology</b>	<b>15</b>	<b>Land, Climate Change &amp; Environment</b>	<b>26</b>
<b>Scoping the Land Data Landscape</b>	<b>16</b>	Types of data or information	26
<b>Assessing Accessibility</b>	<b>17</b>	Sources of data or information	27
<b>Availability of Land Data &amp; Information in Kenya</b>	<b>19</b>	Timeliness of data and information	27
<b>Legal, Institutional &amp; Policy Framework</b>	<b>20</b>	<b>Overall Availability of Land Data &amp; Information</b>	<b>27</b>
<b>Land Tenure Data</b>	<b>20</b>	<b>Accessibility of Kenyan Land Information Ecosystem</b>	<b>29</b>
Types of data or information	20	<b>Online</b>	<b>30</b>
Sources of data or information	21	<b>Accessibility</b>	<b>31</b>
Timeliness of data and information	21	<b>Free</b>	<b>31</b>
<b>Land Cover, Use and Management</b>	<b>21</b>	<b>Metadata</b>	<b>32</b>
Types of data or information	21	<b>Standards</b>	<b>33</b>
Sources of data or information	22	<b>Downloadable</b>	<b>34</b>
Timeliness of data and information	22	<b>Open License</b>	<b>34</b>
<b>Land Disputes</b>	<b>23</b>	<b>Machine Readability</b>	<b>35</b>
Types of data or information	23	<b>Overall Accessibility assessment</b>	<b>36</b>
Sources of data or information	23	<b>Conclusions &amp; Recommendations</b>	<b>38</b>
Timeliness of data and information	23	<b>References</b>	<b>42</b>
<b>Human Settlements</b>	<b>24</b>	<b>ANNEX I SCORING CHART</b>	<b>43</b>
Types of data or information	24	<b>Types of Data Criteria</b>	<b>43</b>
Sources of data or information	24	<b>Representation of Types of Sources Criteria</b>	<b>43</b>
Timeliness of data and information	25	<b>Timeliness Criteria</b>	<b>44</b>
		<b>Accessibility Criteria</b>	<b>45</b>
		Overall Accessibility Score	46

# Foreword

By **Laura Meggiolaro**  
Land Portal  
Foundation

Having and using information has always been a powerful force for change, helping to fight corruption, enabling citizens to participate more fully in public life and allowing people from all walks of life to exercise their fundamental human rights. We are living in a time in which paradoxical topics such as 'fake news' and 'big data' are part of our everyday lives. These discussions are symptomatic and reflective of the fact that now, more than ever, we need governments and heads of state, policy makers and change makers, civil society, academia and the private sector alike, to make decisions that are evidence-based, track progress and clearly measure accountability.

The "data revolution", which not only includes an increasing demand for data and its production, but also that it be delivered to the right people and in the right format, is well underway. This revolution has extended to the land sector, yet renowned data initiatives such as the Open Data Barometer and the Open Data Index systematically rank land ownership data as the most closed and inaccessible dataset by governments worldwide. Unfortunately, this trend is not only occurring within government; the entire land sector is lagging behind in talking about and dealing with data. This needs to change, quickly. We need to overcome distrust between actors, competition over funding and fear of misuse of data. In short, we need efforts that are cohesive and we need the right data and information to underpin these efforts.

Within this context, it is with great pleasure that I introduce this first State of Land Information Kenya: Uncovering Kenya's Land Information Ecosystem. We believe that this report is useful and timely in this increasingly data-driven atmosphere in the land sector. An overview of what key information on land is available and exists has never been carried out. For the very first time, we have looked at the entire landscape of data and information related to key land topics in Kenya, assessing over 250 land resources from 60 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.

This report is fundamentally different from other land monitoring initiatives, and also differs from Open Data initiatives such as the Open Data Barometer or Index. The Land Portal considers data useful when there is sufficient information for a user to determine whether the content is reliable, useful or useable for them. We do not pass judgement on its content or meaning. This Report therefore consciously provides an overview of the State of Land Information, not the State of Land in the country. This report also goes further than the Open Data Barometer or Index and analyzes the entire sector, as opposed to only analyzing the 'land ownership' dataset, which the Barometer and Index focus on.

With this inaugural report on Kenya, in partnership with local researchers, we present the information available on key land issues. By assessing how this information is published and how accessible it is, and providing recommendations to improve this, we hope this State of Land Information report will help establish a functioning, inclusive and democratized ecosystem of data in the country, and ultimately, in the entire land sector. We hope that this is of use to you and that initiatives such as these can help the land sector to catch up with other sectors in being part of the data revolution, which in our view is essential to helping secure land tenure and improving land governance worldwide.

# 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 Kenya 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 Kenya, assessing over 250 land resources from 60 different sources. This robust scoping exercise not only allowed us to see trends and gaps when it comes to land data collection, but also prompted us to provide very practical recommendations to improve visibility and usability of data and information, and thus improving the land information ecosystem in the country.

In performing this scoping study on “what is known” or somehow documented about land in Kenya, we considered that it would be an oversight if we only scoped for raw data and statistical indicators. Our expectation was that 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. Our research confirms this expectation, highlighting that **80% of key land resources in Kenya are available as documents, not statistical data.**

**The main information and data providers of key land resources in Kenya are the government** (accounting for 48% of the total resources, though many of those include laws and policies regulating land in Kenya) **and research institutions** (providing 25% of the key land resources). A notable lesser representation is that of national **Civil Society Organizations, accounting for less than 15% of the total key resources** related to land. Our assumption is not that CSOs do not have data, information or other kinds of knowledge to share, but rather that their data and information publishing practices can be improved to increase their discoverability on the web.

Availability of Data and Information							
Key Category	Data available?	Representation of Sources					Data up-to-date?
		Government	Research Institutions	National CSOs	Int. Organizations	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

The knowledge is published online (98%), it is available for free (98%) and largely publicly accessible without requiring registration or identification (94%). The basic access to data and information there seems to be in a very good state in the Kenyan Data and Information Ecosystem. When it comes to more sophisticated accessibility assessments however, the data and information providers score much less high:

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

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

Overall, the health of the Kenyan Land Data and Information Ecosystem is scored with **40/105 points**.

## Recommendations

The key recommendations emerging from this report to data and information providers in Kenya to increase access to and use of their land data and information, as well as to improve the ecosystem in Kenya overall, are the following ones:

- **Establish a platform for policies**, similar to Kenyalaw.org for laws, to allow for a more complete picture of the legal framework that governs land in Kenya;
- **Ensure that datasets and databases are updated** on a regular basis and publication dates are traceable for users;
- **Consider licensing and anonymization** techniques to allow for data publishing without inflicting harm or violating the privacy of data subjects, to allow for better data and information provision on certain key land categories, such as Land Tenure;
- **Support & enforce data and information sharing efforts** by (national) civil society actors to ensure a more inclusive and varied perspective in the Kenyan land data and information ecosystem;
- **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 metadata to promote the usability as well as interoperability of data and information in the Kenyan data & information ecosystem;
- **Enable the possibility to bulk download data** and information from databases to allow for more meaningful and large-scale use and uptake of the 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;
- **Consider the formats in which data (and information) are published**, and specifically consider machine-readable formats to allow for greater discoverability of the information as well as application in technologies;
- **Apply unique identifiers to key elements** of the data to ensure consistency and reference to the data and information, and allows for more efficient exchange within the data ecosystem.

# Introduction

Availability of accurate and up to date data and information on land and different land uses, such as agriculture, forestry, mining, wildlife, water, housing and infrastructure, is critical to effective land governance and crucial for planning and managing the use of land and land-based resources. Public institutions and the government need land data and information for appropriate and timely decision-making; while land users, the general public and other stakeholders need it to effectively monitor and influence those decisions. Land data and information is also critical for effective tracking of land policy implementation processes to inform lesson learning and generate good practices, as well as to ensure sustainable and equitable land investments.

It is an often-repeated rhetoric that there is a lack of land data; either there is no data or the data that exists is unreliable or out of date. Collecting new data is a time-consuming and costly process. Data is collected and captured on a massive scale already, but research shows that of all existing data worldwide, less than 1% is actually analyzed and digested.<sup>1</sup> With increasing digitization of information, increased use of internet in all parts of the world, and continuously growing demand for more data, the risk is that existing data is either purposely cast aside (as the source may be from outside our trusted networks) or simply overlooked. The current reality of land data is that in many parts of the world, data remains inaccessible, fragmented, poorly managed or simply unusable.

Kenya ranks 78th out of 94 countries assessed in the global Open Data Index.<sup>2</sup> The Kenya government scores relatively well when it comes to making certain data, such as government budgets, national statistics and laws, publicly available and for free.<sup>3</sup> The availability and accessibility of land ownership data, however, received the lowest possible score; this crucial piece of land-related information is not publicly accessible, not free, not up-to-date and not available in a way that allows for re-use. However, land issues go much beyond 'simply' land ownership data. There are countless more elements to land, about which data and information are needed, and there are many more possible sources of data than simply government data. An overview of the existence and accessibility of the range of data and information that covers key areas of land beyond ownership, however, simply does not exist.



1 John Gantz & David Reinsel: "IDC's Digital Universe Study", EMC, December 2012.  
2 Open Knowledge International: "Global Open Data Index", 2015.  
3 Open Knowledge International notes however that this data often comes in the format of PDF reports, hidden between lots of text, graphs and tables. To grasp the real data requires a significant amount of effort of a user.

With the **State of Land Information Report**, we seek to provide an overview of existing data and information on key land issues. 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.

### 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.**

# 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 Kenya in 2018 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.



## 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:

- 1 Sustainable Development Goals, “SDGs” (United Nations)<sup>4</sup>;
- 2 Voluntary Guidelines on the Responsible Governance of Tenure, “VGGTs” (FAO)<sup>5</sup>;
- 3 Land Governance Assessment Framework, “LGAF” (World Bank Group)<sup>6</sup>;
- 4 Global Land Indicator Initiative, “GLII” (network facilitated by GLTN/UN-Habitat)<sup>7</sup>;
- 5 Monitoring & Evaluation of Land in Africa, “MELA” (IFPRI & Land Policy Initiative)<sup>8</sup>;
- 6 International Land Coalition Dashboard (facilitated by ILC)<sup>9</sup>;
- 7 Africa Data Revolution Report (facilitated by Open Knowledge International)<sup>10</sup>.

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](#).

It is important to mention that although the scoping study performed has been as rigorous and as targeted as possible with the use of key issues around land, we do not claim to have captured all data and information ever captured or published about land in Kenya. Most particularly, offline resources are difficult to find and it is difficult to know all the possible sources that somehow have collected some form of data, information or knowledge about land issues. Moreover, new perspectives are constantly being collected every day.



4 <https://www.un.org/sustainabledevelopment/sustainable-development-goals/>

5 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.

6 <http://www.worldbank.org/en/programs/land-governance-assessment-framework>

7 <https://gltn.net/global-land-indicators-initiative-glii/>

8 <https://melafrica.wordpress.com/>

9 International Land Coalition, “The Dashboard Indicators”, Rome May 2018.

10 World Web Foundation, “Africa Data Revolution Report 2018. Status and Emerging Impact of Open Data in Africa”, 2018.

The picture of the Data Landscape in Kenya we are presenting in this report is therefore not all-encompassing, but is, rather, a snapshot of a certain moment in time. That said, it is the most comprehensive and definitive resource for land data and information in Kenya available and will be considered a “living” document that we aim at having updated regularly.

### 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.

## 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 Kenya. 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:

- 1 Open Data Index (Open Knowledge International)<sup>11</sup>;
- 2 Open Data Barometer (Web Foundation)<sup>12</sup>;
- 3 5 Stars of Linked Open Data (Tim Berners-Lee)<sup>13</sup>; FAIR principles of Open Research Data<sup>14</sup>;
- 4 Open Data Inventory (Open Data Watch)<sup>15</sup>;
- 5 Africa Data Revolution Report (Open Knowledge International)<sup>16</sup>;



11 Open Knowledge International, “Global Open Data Index. Methodology”, consulted website September 2018: <https://index.okfn.org/methodology/>

12 World Wide Web Foundation, “Open Data Barometer. Methodology”, consulted website September 2018: <https://opendatabarometer.org/leadersedition/methodology/>

13 Berners-Lee, “5 Stars of Linked Open Data”, consulted website September 2018: <https://5stardata.info/en/>

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

15 Open Data Watch, “Open Data Inventory 2017. Methodology Report”, 2018.

16 World Web Foundation, “Africa Data Revolution Report 2018. Status and Emerging Impact of Open Data in Africa”, 2018.

- 6 EU Open Data Maturity Assessment (European Union)<sup>17</sup>;
- 7 OUR Data Index (OECD)<sup>18</sup>.

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.

### 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.<sup>19</sup> It is a common misunderstanding that publishing publications on a website is all you need to do to make the information accessible and useable. 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, make 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](#).

• • •

<sup>17</sup> European Commission, "Open Data Maturity in Europe 2017. Open Data for a European Data Economy", November 2017.

<sup>18</sup> 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.

<sup>19</sup> 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 Kenya

The availability of land data and information in Kenya 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. **✓ indicates a good practice; ! indicates a practice that can be improved; and ✗ indicates a poor practice.** More information about how these scores were allocated can be found in [Annex I-Scoring Chart](#).

## 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 platform Kenyalaw.org, mandated by the Kenyan government, in particular has proven to be a useful tool to access the various laws and legislations that set out the governance framework for land. Unfortunately, no such platform exists for policies that relate to land.

The Ministries that are mandated to govern (certain parts) of land, themselves are not providing information about their work or how they implement their land governance duties. Evidence on the clarity and efficiency of this division of mandates and responsibilities in practice is hard to find. Research Institutions and Civil Society Organizations do provide some evidence, though not on a large scale.

## 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.

### Types of data or information

**Most of the information available was in the form of documents (90%),** consisting mainly of policy and legal instruments as well as research reports on land tenure. The Ministry of Lands and Physical Planning is in the process of digitizing land records, and has reportedly digitized over 3 million land records to date. However, the scoping study was not able to establish whether the digitized records will be accessible for others to view or use.

As regards holders of different natural resources, such as minerals, there is the Mining Cadastre Portal, the data on which is not downloadable nor easily accessed unless one knows the name of a holder of a mineral license or permit or a specific code. To become a member of the Mining Cadastre portal, registration with the Ministry of Petroleum and Mining is required.

Is there data?



### Sources of data or information

**The Government is the main provider of data and information on Land Tenure.** There are research documents as well as research articles by Research Institutions and universities. A noticeable weak source in data or information about land tenure systems in Kenya, is the national Civil Society perspective and International organizations (not operating (only) and/or from Kenya). The overall scoring is given below. More details on the method of scoring can be found in Annex I.

Government	Research Institutions	National Civil Society Organization	International Organization	Other
✓	✓	✗	✗	✗

### Timeliness of data and information

For the timeliness of data and information assessment, we excluded laws and policies because it is not, in principle, the intention of laws and policies to be updated regularly. For the remaining resources, 9% was either not-dated or published before 2010, **83% was published between 2010 and 2017** and, finally, another 9% was published in 2018.

Is the data up to date?



## 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).

### Types of data or information

**All of the information identified for Land Cover, Use and Management is (geo) spatial data.** No offline data or information was identified in this scoping study. For 25% of the data the access is restricted to registration. What is noticeable is that these datasets focus either on Land Cover or Land Use; data on Land Management was not identified in this scoping study.

Is there data?



### Sources of data or information

The main provider of land cover, use and management data are **international organizations** (account for 46% of the data), with governments closely following at 39%. An important element to mention here is that when it comes to land cover and use data, international organizations often act as a data aggregator of governmental data, but are not necessarily the collector of the data. Nonetheless, the main avenue through which this data is accessed is the data portals of the international organizations, so these institutions are considered the source of the data for this exercise. Finally, 8% of the data was accounted for by Research Institutions and 7% by national civil society organizations.

Government	Research Institutions	National Civil Society Organization	International Organization	Other
✓	✗	✗	✓	✗

### Timeliness of data and information

An interesting outcome of assessing the data identified in this category is that for many of the datasets, the dates of publication or collection were unknown: 40% of the datasets did not have a date indicated on the portals on which the data was published. In many instances, this was the case for governmental datasets. Together with datasets that were published before 2010, the unspecified dates for the dataset account for 75% of the total datasets identified for this category. 17% of the datasets were published between 2010 and 2018, with only one dataset (8%) updated in the last year.

Is the data up to date?	✗
-------------------------	---

## 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).

### Types of data or information

The knowledge found during the scoping study consisted **100% of knowledge captured in documents** and no concrete statistical data. Laws documenting the legal framework were found, as well as many research reports on (historical) land disputes in Kenya.

Is there data?	✗
----------------	---

### Sources of data or information

The type of sources for the documentation on land disputes in Kenya is **more dispersed across the various source groups** than the previous categories. From our scoping exercise, 40% of the information providers were Governmental Institutions, another 40% were Research Institutions and finally (national) Civil Society Organizations accounted for 20% of the information.

Government	Research Institutions	National Civil Society Organization	International Organization	Other
✓	✓	!	✗	✗

### Timeliness of data and information

From the documents identified in the scoping exercise, none were as recent as this year (2018). Forty percent of the documents were either not dated or published before 2010. **The majority of the resources however were published between 2010 and 2018 (60%).**

Is the data up to date?	!
-------------------------	---

## 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).

### Types of data or information

In comparison with some of the other categories, there seems to be a lot of data and information that covers this category. The knowledge mostly seems to be captured in documents (88%), but there are also datasets available (22%). This data refers mostly to population and housing, informal settlements, as well as displacement. Most of this data is accessible without any restrictions.

Is there data?



### Sources of data or information

The sources of data and information on human settlements vary enormously; it seems to be a general topic that many types of organizations are involved in. The majority of data and information comes from Research Institutions, accounting for 32% of the total resources identified in the scoping study, followed by International Organizations, which account for 22% of the data and information. Then closely followed are Governmental Institutions, with 19% of data and information, and national Civil Society Organizations providing 16%. The 'Other'-category (mostly news agencies) finally accounts for 11% of the total human settlements resources identified in the scoping.

Government	Research Institutions	National Civil Society Organization	International Organization	Other
!	✓	!	✓	!

### Timeliness of data and information

In order to gain the most realistic perspective of the timeliness of data and information related to human settlements, laws and policies have been excluded from this particular exercise. From the remaining resources, it seems that most data and information (44%) was published between 2010 and 2018. Closely following that is data and information that is either not dated or published before 2010 (40%). The remaining 16% of the data and information is as recent as this year (2018). An interesting finding is that the majority of the resources of which no date for the data or information was indicated came from national Civil Society Organizations.

Is the data up to date?



## 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).

### Types of data or information

**The large majority of knowledge about land markets & financing is captured in documents (83%).** The scoping study identified only one dataset, the Land Matrix, on large-scale land acquisitions.

Is there data?



### Sources of data or information

**The majority of the sources of the scoped data and information for Land Markets & Financing are Governmental Institutions (42%),** though this is mostly relating to land valuation frameworks and laws set out by the government. 'Other' data sources, mostly newspapers, are also a significant source for information on land markets, financing and investments, accounting for 25% of the resources identified under this category. Research Institutions follow with 17%, with national Civil Society Organizations and International Organizations both accounting for 8% of the resources on this topic.

Government	Research Institutions	National Civil Society Organization	International Organization	Other
✓	!	✗	✗	✓

### Timeliness of data and information

Laws and policies have been excluded from the identified resources under this category, to gain the most accurate picture of the timeliness of the data. For the one dataset, the Land Matrix, it was difficult to obtain information on when the last deal was added to the database from the database itself. Database updates need to be found on a separate box on the Land Matrix website. Other than that, the information about the timeliness was available: **most resources were published between 2010 and 2017 (43%), with 29% as recent as this year (2018)**. The other 28% was either updated before 2010 or unable to trace the date. An important note to mention here is that the timeliness of the data refers to when the data or information was published, not the date of the actual valuations or investments.

Is the data up to date?



## 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).

### Types of data or information

**The provision of data and information about Land, Climate Change and Environment is fairly distributed among documents (63%) and data (37%),** especially considering the documents include laws and policies regulating various aspects of climate change and environment. All of the identified datasets are available without any restrictions to access it.

Is there data?



### Sources of data or information

**Governmental Institutions are the majority of information providers in this category (44%),** providing laws and regulations but also national strategies (addressing environmental challenges such as natural disasters). International organizations follow with 31% and finally Research Institutions account for 25% of the resources identified for this category.

Government	Research Institutions	National Civil Society Organization	International Organization	Other
✓	✓	✗	✓	✗

### Timeliness of data and information

Most of the available knowledge on Land, Climate Change and Environment **(55%) is either dated before 2010 or not dated at all**. The Kenya Climate Change Directorate is a crucial database on climate change and environment but it is unclear when it was last updated. Thirty six percent of the resources are dated between 2010 and 2017, with only 9% as recent as this year (2018).

Is the data up to date?

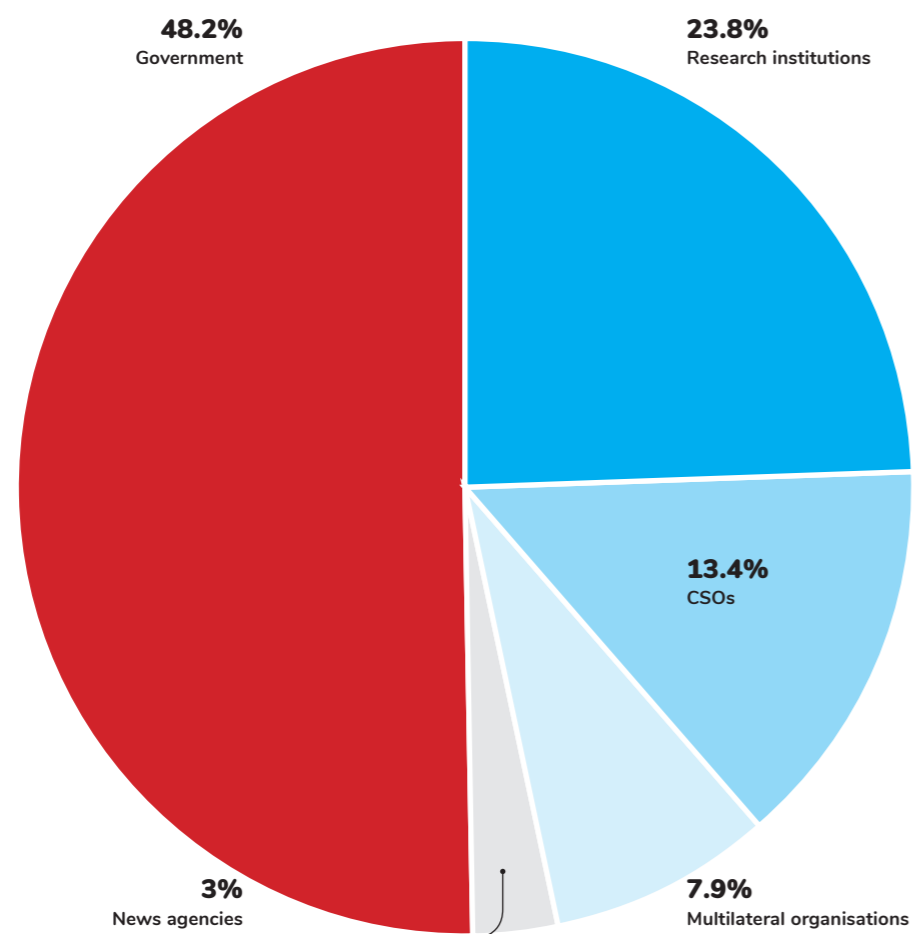


## Overall Availability of Land Data & Information

Overall, we conclude that as regards land Kenya has an information ecosystem and not a data ecosystem. Our findings show that **over 80% of key land information resources in Kenya can be found in documents, not datasets**. Although whether or not there are documents or datasets, highly varies between the type category of land data and information. The documents and data available are predominantly available online (98%), though this is a 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 (excluding laws and policies), 23% of the information is dated from before 2010, while were unable to determine the date of publication or creation for 20% of the information.. This means that **almost half (43%) of the resources that hold key data and information about land is outdated** or lacking details about date of publication or production.

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 Kenya, the division of types of information providers can be grouped as follows:



This figure shows that the **Government is by far the main source of land data and information**. In particular, governmental institutions are the main source of policy and legislative instruments on land. **Research Institutions and Universities** account for almost one-fourth of the data and information identified through the scoping exercise. A noticeable gap in the provision of data and information in this scoping exercise is the (national) **Civil Society Organizations** as an information provider. It is possible the reason CSOs are represented as a smaller group of information providers in this scoping exercise is a reflection more on their dissemination practices than of role and standing as information providers. One needs to be aware of a civil society organization, their work and actively access their website to find the information they are looking for — and even then often their complete body of knowledge is not uploaded online. As a result, potentially important perspectives from the civil society are underrepresented in this comprehensive overview of land data, information and knowledge in Kenya.

## Accessibility of Kenyan 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 criteria, we provide a general score. **✓ indicates a good practice;** **!** indicates a practice that can be improved; and **✗ 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 Kenya's data and information ecosystem.

## Online

A first criteria 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 **98% of the key resources are available online**. A noticeable offline resource is the land tenure data from the National Land Registry, however, the Ministry of Lands & Physical planning, is reportedly in the process of digitizing land records.

### Why is it important data & information are online?

Only 55% of the world's population makes use of the Internet as of June 2018.<sup>20</sup> 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.

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"



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

## Accessibility

The 'Accessibility' criteria 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 criteria, too, the key land resources in Kenya ranked very high with **94% of the resources accessible without any log in barriers**.



The information about original data source is available only to registered users.

SIGN UP NOW

Figure 2. Accessibility barrier experienced during assessment (source: Kenya Data Portal)

Similar to the 'online' criteria, however, the representation of accessible resources may be biased with the scoping exercise, revealing mostly those resources that are more easily accessed.

Overall Score "Accessibility"



## Free

Another important criteria that helps determine the extent to which data and information is inclusive and useful to a wider audience, is whether or not it is available for free (unpaid). Particularly in the academic sector, data and other research findings are often hidden behind publisher paywalls. So how about key land resources in Kenya? Our research findings suggest that the data and information ecosystem overall is freely accessible, with **98% of the data and information available on the web for free**. The same caveat applies as for the two previous criteria, in that the scoping research is more likely to identify freely accessible resources than those behind paywalls.

Overall Score "Free"





## 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.<sup>21</sup> 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 useability 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.

32

From the key land resources identified in this scoping exercise, **only 28% of the data and information came accompanied with metadata**. Of those providers that publish metadata to accompany their data and information, one-fourth provide very limited metadata. Examples of such limitations include lack of a publication date, limited information about the source of the data, and similar missing information that make up vital elements of the metadata that users require.

Overall Score "Metadata"



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

## Standards

The standards criteria 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.

### 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.<sup>22</sup> LandVoc is a part of widely accepted agriculture thesaurus by the Food and Agricultural Organization, AGROVOC.

33

Even with a broad interpretation of the use of 'standards', the land data and information resources in Kenya score incredibly low. Only 38% of the data and information providers that provide metadata use standards. That makes an overall score of **10% of data and information providers using standards in their metadata**. The most commonly identified standard used were the country ISO3 codes, for international data providers.

Overall Score "Standards"



22 <https://www.landvoc.org>

## 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. Only 13% of the data providers actually prevent a user to download the data and restrict its use to their own platform. However, in order to meet the accessibility criteria, being able to download a single file is not sufficient. 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 low on this criteria—**only 9% of key land resources are available to download in bulk.**<sup>19</sup>

Overall Score "Downloadable"



## 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<sup>24</sup> 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.

23 When the key resources referred to a document, the downloadable-criteria was applied to the database this particular document was hosted.

24 <https://opendefinition.org/>

From the key resources on land in Kenya, **31% of the information providers have applied an open license to their resources.** It is worth noting that the large majority of these open licenses apply to policy and legislative instruments, that are licensed under the Public Domain license. Only 6% of the resources with an open license are another type of data. Another important finding in this Accessibility assessment is that while 31% of the resources have an open license—this does not mean that the remaining 69% have a license that does not meet the 'open'-criteria. Rather, **the majority of resources (54%) do not have any license at all.**

Overall Score "License"



## 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.

**Thirty five percent of the key resources related to land in Kenya are published in a machine readable format.** The most commonly used formats for data and information are PDFs (not machine readable<sup>25</sup>), HTML, CSV and XLS (the latter two are common formats for tabular data). An important caveat to mention with this criteria 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 on Kenyalaw.org, for example, are available both in PDF as well as HTML formats. HTML is a machine readable format. The application of this criteria on such documents (which, as mentioned, account for 80% of the key land resources in Kenya) 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 criteria was applied to the metadata of documents, where possible, not the document itself.

Overall Score "Machine Readability"



25 More specifically, PDFs can be read by computers but it is not easily editable or processable by machines.

## (Linked) Data URI

The final criteria 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.<sup>22</sup> 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 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 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”.<sup>23</sup>

Evidently, complying with the principles of the linked web is not a priority for data and information providers on land in Kenya. **A mere 3% provided unique identifiers to classify key elements of the data.** Of those that provide unique identifiers, half linked to other URIs (1.65% of the total resources). These were exclusively global datasets.

Overall Score “(Linked) URIs”



## Overall Accessibility assessment

The Kenyan Land Data and Information Ecosystem scores well with a **basic interpretation of accessibility, namely whether it is online, accessible without registration or other types of barriers, and free.** 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 Kenyan Land Data and Information Ecosystem scores much less high.

A lot of key land data and information on Kenya is available online, can be accessed without needing to register or request access and can be viewed without needing to pay. **The discoverability of the resources within the ecosystem is still considerably poor.** Less than 30% of the data and information were published with metadata and even fewer made consistent use of standards in their metadata. In several instances, a publication date of a particular resource was untraceable.



26 Berners-Lee, “5 Stars of Linked Open Data”, consulted website September 2018: <https://5stardata.info/en/>

27 Idem.

Not only do these weak 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.

**Scores are equally low for other criteria that are intended to promote the use of the data, for whatever purpose.** Only 4% of the data and information are available to download in bulk and only 31% of the data providers apply an open license to their data. What’s even more striking, is that over half (51%) of the data providers do not even specify 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. Another criteria that supports re-use and modification of data and information, machine readability of data and information, met slightly better results: 35% 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, none of the local information providers included this in their data. Of all the key land resources identified in the scoping study, only 3% provided URIs, and those were all global sources.

Overall Score Kenyan “Data & Information Ecosystem”

**40/105!**

# Conclusions & Recommendations

It is an often-repeated rhetoric that there is a lack of land data—that the data is either unavailable or if available 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 Kenya 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 Kenya, assessing over 250 land resources, 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 Kenyan Data & Information Ecosystem on land.

The statement that there is a lack of data is partially accurate: **our scoping exercise shows that 80% key land resources are available as documents, not statistical data.** However, there was no key land category where no knowledge or information was found, suggesting that there is indeed knowledge generated and published, but not (yet) translated to statistical data (where possible). Our research also shows that the knowledge is published online (98%), it is available for free (98%) and largely publicly accessible without requiring registration or identification (94%). The rudimentary access to data and information there seems to be in a very good state in the Kenyan Data and Information Ecosystem, but important to mention is that those resources that are online, free and accessible without barriers are also those most likely to have been identified in our scoping exercise.

Another important caveat to the statement knowledge was found for each key category of land, is that **43% of the resources identified were either dated from before 2010 or were published without a clear publishing date.** 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. The main source of key data and/or information identified in this scoping exercise was the government, accounting for the majority of the resources available. **The government**, in particular, played a significant role in publishing policies and laws on land and land-based resources. Kenylaw.org in particular, is a great resource for accessing laws in Kenya. A similar website for policies does not exist.

**Research Institutions** provided 25% of the total resources identified in the scoping exercise. A notable weaker link in terms of information provision, were the (national) **Civil Society Organizations**, which accounted for less than 15% 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.

Availability of Data and Information							
Key Category	Data available?	Representation of Sources					Data up-to-date?
		Government	Research Institutions	National CSOs	Int. Organizations	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

On accessibility of key land resources, the Kenyan data ecosystem is not quite in a similar good state. As mentioned, on a basic level (available online, for free and without restrictions), the Kenyan information ecosystem performs well. **When it comes to more sophisticated accessibility, however, the state of the ecosystem is still considerably poor.** Less than 30% of the data and information were published with metadata and even fewer made consistent use of standards in their metadata. Not only do these weak data publishing practices make the data and information less discoverable on the web, they also restrict 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.

**For other criteria that are intended to promote the use of the data, for whatever purpose, the scores are very low as well.** Only 4% of the data and information are available to download in bulk and only 31% of the data providers apply an open license to their data. Even more striking is the fact that over half (51%) of the data providers do not even specify 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.

Another criteria that supports re-use and modification of data and information, machine readability, met slightly better results, with 35% of the data and information available in machine readable format. Finally, as regards the criteria of the linked web, namely having unique identifiers (URIs) for key elements of data and/or metadata and linking to other URIs, none of the local information providers included this in their data. Of all the key land resources identified in the scoping study, only 3% provided URIs, and those were all global sources.

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

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

Overall, the health of the Kenyan Land Data and Information Ecosystem is scored with 40/100 points. Recommendations to data and information providers in Kenya to increase access to and use of their land data and information, as well as to improve the ecosystem in Kenya overall, are as follows:

- **Establish a platform for policies**, similar to Kenyalaw.org for laws, to allow for a more complete picture of the legal framework that governs land in Kenya;
- **Ensure that datasets and databases are updated** on a regular basis and publication dates are traceable for users;
- **Consider licensing and anonymization techniques** to allow for data publishing without inflicting harm or violating the privacy of data subjects, to allow for better data and information provision on certain key land categories, such as Land Tenure;
- **Support & enforce data and information sharing efforts** by (national) civil society actors to ensure a more inclusive and varied perspective in the Kenyan land data and information ecosystem;
- **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 metadata to promote the usability as well as interoperability of data and information in the Kenyan data & information ecosystem;
- **Enable the possibility to bulk download data** and information from databases to allow for more meaningful and large-scale use and uptake of the 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;
- **Consider the formats in which data (and information) are published**, and specifically consider machine-readable formats to allow for greater discoverability of the information as well as application in technologies;
- **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.

# References

Berners-Lee, "5 Stars of Linked Open Data", consulted website September 2018:

<https://5stardata.info/en/>

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

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.

International Food and Policy Research Institute, "Monitoring & Evaluation of Land in Africa", website consulted July 2018: <https://melafrica.wordpress.com/>

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

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

Gantz & Reinsel: "IDC's Digital Universe Study", EMC, December 2012.

Global Land Tool Network, "Global Land Indicators Initiative", website consulted July 2018: <https://gltn.net/global-land-indicators-initiative-glii/>

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

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

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

Open Knowledge International: "Global Open Data Index", 2015.

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

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.

United Nations, "About the Sustainable Development Goals", website consulted July 2018: <https://www.un.org/sustainabledevelopment/sustainable-development-goals/>

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

World Bank, "Land Governance Assessment Framework", website consulted July 2018: <http://www.worldbank.org/en/programs/land-governance-assessment-framework>

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

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

# 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: **✓ indicates a good practice; ! indicates a practice that can be improved; and ✗ indicates a poor practice.** 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 data available. Please find below the scoring:

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

## 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:

- 1 Governmental Institutions;
- 2 Research Institutions (including universities);
- 3 (National) Civil Society Organizations;
- 4 International Organizations;
- 5 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 10% and 24% of the total resources per category	!
Group accounted for less than 10% of the total resources per category	✗

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 2017	!
Majority of resources were published since 2018	✓

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 Kenya, the resources were assessed against the following criteria:

- 1 Online;
- 2 Accessible (no registration or other types of barriers);
- 3 Free (unpaid);
- 4 Metadata;
- 5 Standards;
- 6 Downloadable
- 7 Openly Licensed;
- 8 Machine Readable;
- 9 (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 less than 33,33% 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 exactly or more than 66,67% of the total key land resources	✓

### 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<sup>24</sup>, 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:

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

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	



28 World Wide Web Foundation, “Open Data Barometer. Methodology”, consulted website September 2018: <https://opendatabarometer.org/leadersedition/methodology/>







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

[www.landportal.org](http://www.landportal.org)  
[contact@landportal.org](mailto:contact@landportal.org)