



Food and Agriculture  
Organization of the  
United Nations



# The role of metadata and open data in the innovation cycle of land administration

April 2023

# Summary

---

**This publication will discuss the importance of open data as a tool for inclusive land governance. It will introduce and describe open data, land governance functions, and metadata. It will provide in-depth looks at AGROVOC, the controlled vocabulary about agriculture and related sciences coordinated by FAO for more than 40 years, and LandVoc, a sub-vocabulary dedicated to the land sector. Using these two examples, the article will explore why structured metadata is integral for open data to support innovation and improvements in land administration.**



# Contents

<b>Introduction to open data in land governance</b>	<b>4</b>
<b>What metadata is and why it is important to open data</b>	<b>8</b>
<b>Data management and the innovation cycle: the power of structured data in the storage, sharing, and use of open data</b>	<b>10</b>
<b>AGROVOC: connecting information resources</b>	<b>12</b>
<b>The role of LandVoc and its purpose in the land governance and agricultural sectors</b>	<b>13</b>
<b>Conclusion and next steps</b>	<b>15</b>
<b>Bibliography</b>	<b>16</b>

## Acronyms

<b>ASFA</b>	Aquatic Sciences and Fisheries Abstracts
<b>CaLAtThe</b>	Cadastre and Land Administration Thesaurus
<b>CAT</b>	Chinese Agricultural Thesaurus
<b>FAO</b>	Food and Agriculture Organization of the United Nations
<b>GEMET</b>	General Multilingual Environmental Thesaurus
<b>GLII</b>	Global Land Indicators Initiative
<b>GLTN</b>	Global Land Tool Network
<b>KOS</b>	Knowledge Organization System
<b>LEGN</b>	Development Law Service of the FAO Legal Office
<b>LuSTRE</b>	Linked Thesaurus framework for Environment
<b>NALT</b>	National Agricultural Library Thesaurus
<b>SDG</b>	Sustainable Development Goal
<b>UNBIS</b>	United Nations Bibliographic Information System Thesaurus
<b>URI</b>	Uniform Resource Identifier



# Introduction to Open Data in Land Governance

---

**“ Open Data is digital data that is made available with the technical and legal characteristics necessary for it to be freely used, re-used and redistributed by anyone, anytime, anywhere.”**

(Open Data Charter, 2022)



**Land governance covers the legal and policy framework for land, as well as traditional practices governing land transactions, inheritance and dispute resolution. In short, it is fundamentally about power and the political economy of land (FAO, 2022a).**

**Land governance data includes data on the core land administration functions of land tenure, use, value and development.**

**These land administration functions are:**

### Land tenure

The processes and institutions related to securing access to land and inventing commodities in land and their allocation, recording and security; cadastral mapping and legal surveys to determine parcel boundaries; creation of new properties or alteration of existing properties; the transfer of property or use from adjudication of doubts and disputes regarding land rights and *parcel boundaries*.



### Land use

The processes and institutions related to the control of land use through adoption of planning policies and land-use regulations at the national, regional, and local level; the enforcement of land-use regulations; and the management and adjudication of land-use conflicts.



### Land value

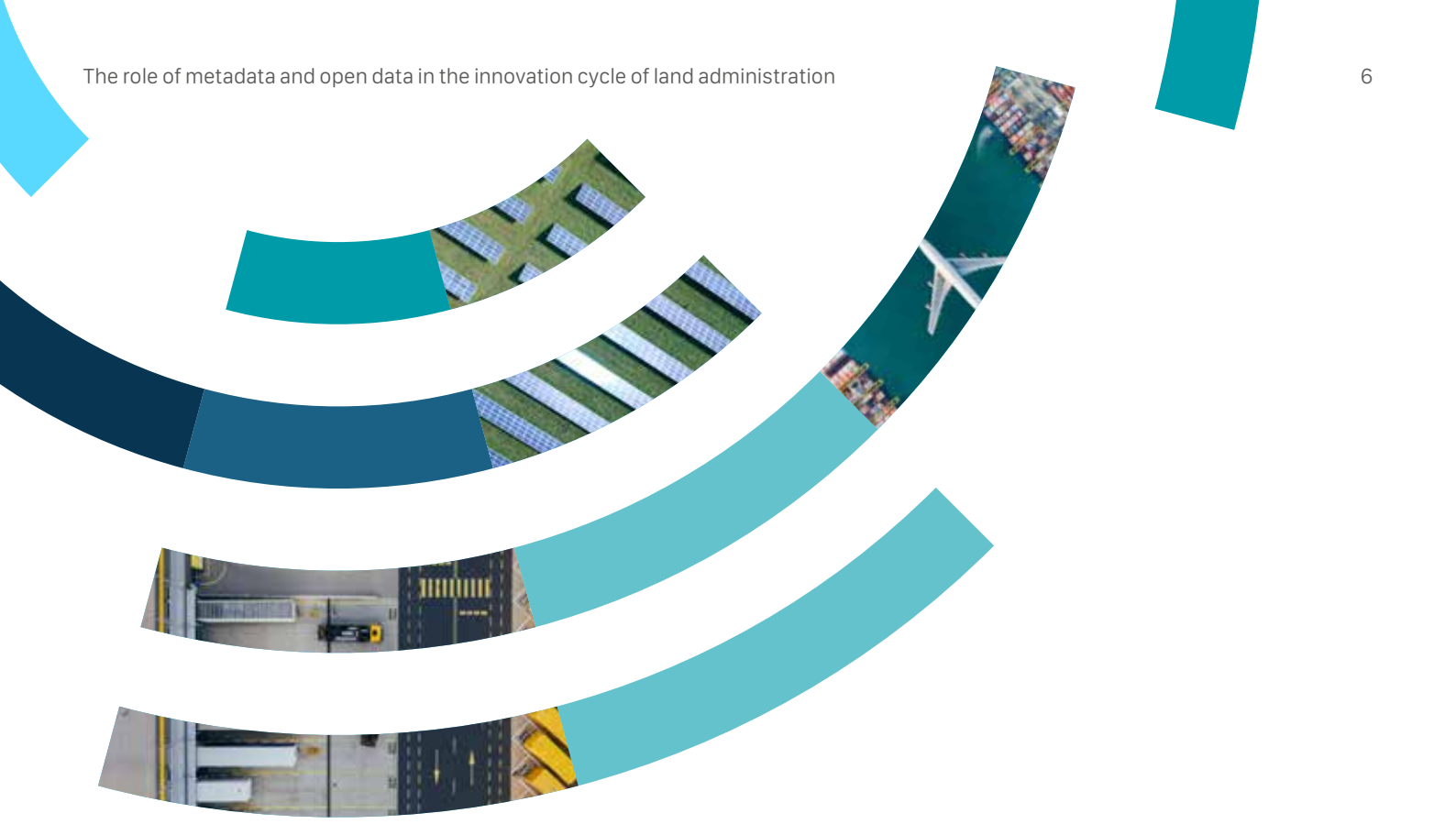
The processes and institutions related to the assessment of the value of land and properties; the calculation and gathering of revenues through taxation; and the management and adjudication of land *valuation and taxation disputes*.



### Land development

The processes and institutions related to building new physical infrastructure and utilities; the implementation of construction planning; public acquisition of land; expropriation; change of land use through granting of planning permissions, building and land-use permits; and the distribution of development costs (Williamson, Enemark, Wallace, and Rajabifard, 2010).





There is little global evidence of openness when it comes to land governance data. Governments are the primary stewards of land data, but they face known barriers to opening up data. Reliable, formal land tenure data sometimes does not exist, or requires time-consuming and expensive digitization. Even though land tenure information is generally not personal data intensive, privacy protection is often cited as one of the reasons for not publishing tenure information. (Davies and Chattapadhyay, 2019).

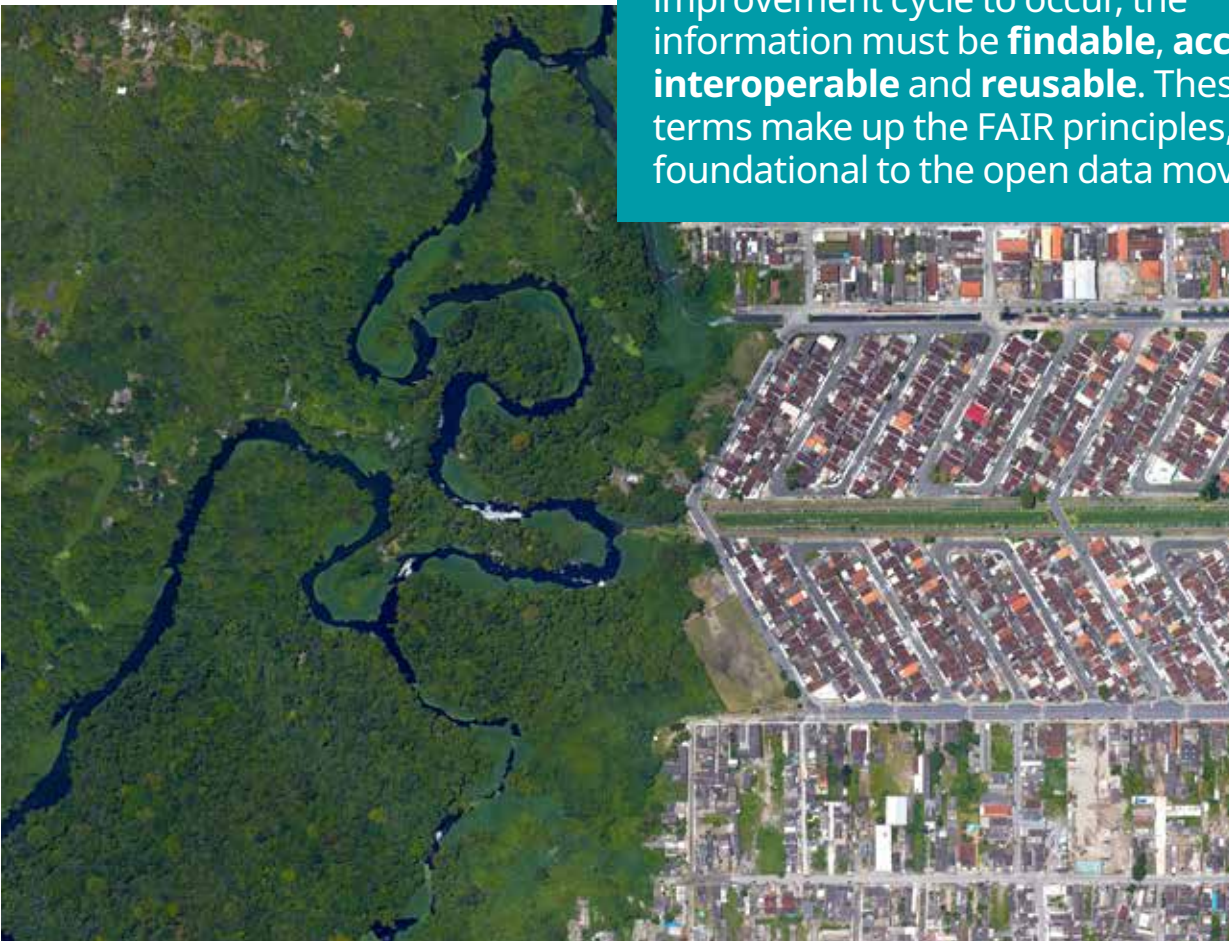
The complexity of tenure rules and regulations resulting in a lack of clarity regarding “ownership” rights is really what precludes easy publication of tenure information. For example, customary rights and obligations are particularly difficult to document in formal systems that are geared towards individual ownership rights. In many cases, there is a lack of trust between government and society where the government may be seen as the primary threat to land rights. In these situations, groups often seek to withhold land data from the government as a means of protection.

Yet land is a critical global asset, and land governance (the exercise of political, economic, and administrative authority over land) should include processes for citizens to participate. Opening up land data provides an opportunity for citizens to participate in the land information ecosystem, which facilitates trust-building and participatory data processes between society and the government. This publication will discuss the importance of open data as a tool for inclusive land governance. It will also delve into metadata, introducing the AGROVOC and LandVoc vocabulary tools, and explore why structured metadata is integral for open data to support innovation and improvements in land administration.

Open access to data is not a panacea to land-related problems, and it is important to refrain from promoting it as an objective in itself. Whether closed or open, high quality data remains high quality, and poor quality data is poor quality. The key difference is that exposing data to the public opens it up to be used, studied, challenged, built upon and improved. In many countries, data is not made available because it is considered incomplete or not up to standard. However, people are able to use even incomplete or imperfect open data to build applications and tools, all the while identifying gaps for improvement. This approach to inclusive data governance can serve to build much-needed trust between government and society as well as provide opportunities for service innovation. Open data is the great facilitator of research, new knowledge and innovation.

Open data can support non-traditional spatial data in the land sector to be harnessed to shore up shortcomings in official data sources. These complementary data sets can serve to broaden the base of information used for land governance decision making. For example, Cadasta Foundation is an organization that works with communities to collect community data on informal settlements or customary lands, which are areas that are not typically included in formal land governance documentation and monitoring. Documenting tenure data from non-traditional sources provides opportunities and insights into future development of land administration systems and the appropriate tenure arrangements that need to be supported.

Opening land data is only the first step towards improving land administration functions. For the open data-innovation-improvement cycle to occur, the information must be **findable, accessible, interoperable** and **reusable**. These four terms make up the FAIR principles, foundational to the open data movement.



# What metadata is and why it is important to open data

**Data can be a critical tool for responding to the biggest social, environmental, and economic challenges. However, data in its purest state is raw material – underground and hard to use. Even after raw data becomes information and knowledge with the help of information scientists and statisticians, it is difficult for humans to absorb it all. Machines are used to call the most relevant data, and in the process, users benefit from the powers of metadata.**

In the simplest terms, metadata is the data that describes data. The descriptions can include author, title, date, geography, abstracts, keywords and much more. These data elements are the primary tool for linking, organizing and connecting data that are generated in different geographies, different languages and across different industries. Metadata has the double function of both describing data and categorizing it logically. For instance, choosing the right keywords for the metadata of an article is important to let people understand what the resource is about. Adding additional machine readable elements (such as Uniform Resource Identifiers, or URIs) makes the data more easily interoperable.

However, it is not enough to tag data and information with keywords freely. Nor is it wise to have unlimited metadata terms with no relationship to each other. When creating metadata, it is important to adhere to standards. Standards facilitate the exchange of data and the harmonization of data among different systems. If no standard is used, it can be very difficult to compare and connect data sets. Standards can apply both to metadata structures (i.e. those elements that need to be univocally referenced to) and their values (these values should point to standardized vocabularies, code lists and such).







Photo: © AdobeStock

To illustrate this:

If someone tags article 1 about slums as **“favela”** and another tags a similar article 2 with **“informal settlements”**, a researcher might not see the second article unless the two terms have some sort of established relationship to each other. This human-controlled list of terms and the relations each term has with other terms is what we call a **“controlled”** vocabulary, an example of a **Knowledge Organization System (KOS)**. When a standard is in place – and used in metadata – a relation is in place to recognize **“slums”** as similar to **“favela”** and **“informal settlements”**. When these terms are exposed as URIs, the machine can make that interpretation as well, which in turn facilitates data retrieval by search engines.

The Land Portal Foundation saw a gap for this type of semantic vocabularies in the land governance sector, and **co-built LandVoc** with the Food and Agriculture Organization of the United Nations (FAO) in 2012.

**“ We value Land Portal’s central commitment to open data and to connecting people and stories on land in the global north and south.”**

**Chris Penrose Buckley**  
Senior Land Policy Lead  
United Kingdom’s Department  
for International Development (DFID)

# Data management and the innovation cycle: the power of structured data in the storage, sharing, and use of open data

**Open land governance data, published in accordance with government laws and regulations, can assist in the provision of efficient and transparent government services. More concretely, open land governance data can be used to support land claims and provide public evidence of land rights, thereby reducing corruption opportunities and offering a route to greater public empowerment. It helps communities monitor whether environmental protections are being upheld, and supports rights claims over geographical areas inhabited for generations. It helps civil society organizations understand patterns of land deals and support environmental and social advocacy.**

Open data systems may also help to bridge the management of data collected by community groups/private individuals, when existing national systems are incomplete or minimal. Open data systems are also able to take advantage of new methods for collecting land data, such as GPS, satellite, drones and mobile phone technologies to help bridge information gaps. Making data interoperable means governments and other data producers use and apply global standards that make the same data work in different contexts (Çağdaş, Meggiolaro and Stubkjær, 2021), (Mey, Zammit and Meggiolaro, 2020).

Critics of open data often circle around issues of privacy, data quality and lack of authoritativeness as reasons why open data is less trusted than closed repositories. These are not problems of open data, but rather problems of a lack of data management, and apply equally well to any data set if it is to be made fit for purpose. The authoritativeness of data derives not from being held in an open or closed repository, but rather from whether good data management practices are applied.

The Land Portal Foundation was established to create, curate and disseminate land governance information by fostering an inclusive and accessible data landscape. The Land Portal Foundation supports data management practices that increase actual accessibility of data and information about land. It supports capacity development activities to improve data management practices at the operational and policy levels through their investments in metadata creation and curation as well as through the creation of the land thesaurus, LandVoc. The Land Portal Foundation also engages in capacity development initiatives such as the Open Up Guide for Land Governance. The Open Up Guide initiative aims to improve data literacy at the government level, document country land information ecosystems, assess the

availability of land data, and guide governments to progressively prioritize country-level data that can be made more open and accessible (Bayer and Booth, 2021).

Evidence shows that very little land governance data is open. There is a poor record of publishing land administration data in support of the core land administration functions (tenure, use value, development) and no data on public land holdings, with Davies et al. (2019) stating that there is no structured open data “on government landholdings, purchases and disposals”.

Land administration has changed from being a parcel data record to an integrated information and decisions support system in pursuit of sustainable social, economic and environmental management (Williamson, Enemark, Wallace, and Rajabifard, 2010), (Zevenbergen, De Vries, and Bennet, 2016). The Land Portal Foundation argues that the land data derived from the functions of modern land administration systems should be opened up in an effort to improve land governance.

**Figure 1.** Open Data for Improved Land Governance



Source: Land Portal Foundation, 2021. Open Data for Improved Land Governance. Rome. Cited 1 February 2023.  
<https://landportal.org/library/resources/open-guide-land-governance-version-20-following-public-comments>



# AGROVOC: connecting information resources

**An invaluable tool to connect information resources is AGROVOC, a controlled vocabulary about agriculture and related sciences, coordinated by FAO for more than 40 years, in partnership with more than 35 institutions in 33 countries. It covers all areas of interest to FAO and it is available in up to 41 languages.**

Adding new translated terms is a contribution to making research and datasets more discoverable and to making national research more visible and accessible. In this way, AGROVOC helps build bridges between datasets and across domains and languages (FAO, 2022b).

Data is an essential component of the decision-making process and the raw material for accountability. New sources of data, if applied responsibly, can enable more agile, efficient and evidence-based decision-making and can better measure progress on the Sustainable Development Goals (SDGs) in a way that is both inclusive and fair (UN, 2022).

Open access to knowledge and information about food and agriculture is crucial to facilitate and promote innovation and leverage the power of data across sources and languages.

Since 2019, FAO has welcomed communities of experts to participate in the maintenance of AGROVOC by including thematic clusters (subvocabularies) such as:

- Land Governance (LandVoc), by the Land Portal Foundation;
- Aquatic Sciences and Fisheries Abstracts (ASFA), by the ASFA secretariat at FAO;
- Legislative and Policy concepts (FAOLEX), by the Development Law Service (LEGN) of the FAO Legal Office;
- FAO Indigenous Peoples;
- One CGIAR. This subvocabulary is the result of an FAO and CGIAR collaboration on increasing interoperability between food and agricultural information systems, and provides an opportunity for CGIAR to suggest new concepts to AGROVOC.

Interested expert communities are welcome to contact [agrovoc@fao.org](mailto:agrovoc@fao.org) for more details.



# The role of LandVoc and its purpose in the land governance and agricultural sector

**LandVoc is a hierarchical set of over 300 carefully selected concepts linked and curated by a community of experts. Hosted within AGROVOC, LandVoc applies to a series of well-known standards used to standardize the creation and maintenance of KOSs. It is available in ten languages: English, French, Spanish, Portuguese, Khmer, Vietnamese, Burmese, Thai, Swahili and Arabic. LandVoc supports the creation of consistent thematic metadata along different types of information systems. Organizations use LandVoc as a glossary and to manually manage their information. Individual researchers use it to make sure they are using appropriate terms in their writing. Across the board, LandVoc supports cross-language information discovery.**

LandVoc began, in 2013, when the Land Portal Foundation started to add new terms and definitions to AGROVOC from existing land glossaries, such as FAO's [Multilingual Land Tenure Thesaurus](#), the Global Land Tool Network (GLTN) and Global Land Indicators Initiative ([GLII](#)) glossary. The commitment to AGROVOC deepened as the Land Portal Foundation started convening the wider land community to validate and translate new additions to LandVoc.

While the editing of LandVoc is done within the AGROVOC database, LandVoc is accessible to end users and is machine readable from its own [website](#), maintained by the Land Portal Foundation. The Land Portal Foundation formed a LandVoc Community of Experts in 2020, which convenes periodically to further develop the thesaurus and provide translations.

LandVoc is the backbone of the Land Portal website. The Land Portal enriches all its content with LandVoc terms. The Land Portal's resources, both originally generated and aggregated from across the internet, range from [academic papers](#) and official reports to [statistical indicators](#), [data stories](#), [blog posts](#) and [news items](#). Aggregating external content and enriching them with new metadata is an especially important service that the Land Portal Foundation provides, as many of the generators of land information do not use [metadata standards](#) themselves. Furthermore, when content is tagged with LandVoc terms, the tags are transparently displayed, making it easy for readers to discover and access additional content. It helps information from Land Portal partners to become more visible and discoverable to those who are seeking it. Over the last ten years, the Land Portal Foundation has created the richest collection of data and publications related to land issues that currently exist for the sector.

Because LandVoc is a tool to facilitate linkages among resources on the Web, it can be used in a variety of different ways beyond the Land Portal's own usage. Firstly, the hundreds of terms contained within LandVoc can be integrated into libraries, repository systems and websites that publish any kind of information related to land governance. This allows information to be classified, indexed and therefore become more discoverable. Secondly, and closely related to this first point, because information is now classified and indexed, those searching for key land governance topics can have access to a wide array of information on their chosen topic matter, including its related terms and translations. LandVoc is meant to be free and re-used by others.

The Land Portal Foundation encourages all partners, contributors and publishers to reach out to learn more about using LandVoc with their own content systems and how to strengthen data and information practices. This is how the land information ecosystem grows more robust.

**“ One of the most significant roles of the Land Portal is its work on open data and building the data ecosystem. They are one of the frontrunners on this topic, which has mushroomed internationally. This is a major added value to the land governance sector.**

**Elke Matthaei**  
Land Governance Advisor  
Gesellschaft für Technische  
Zusammenarbeit (GIZ)

LandVoc goes beyond the sphere of land related matters. It can connect to and exchange information with other databases that are using other vocabularies.

**Since LandVoc is part of FAO's AGROVOC, LandVoc concepts are mapped to other vocabularies such as:**

- [EUROVOC](#)
- Cadastre and Land Administration Thesaurus ([CaLAThe](#))
- Chinese Agricultural Thesaurus ([CAT](#))
- Aquatic Sciences and Fisheries Abstracts ([ASFA](#))
- Linked Thesaurus framework for Environment ([LusTRE](#))
- National Agricultural Library Thesaurus ([NALT](#))
- United Nations Bibliographic Information System Thesaurus ([UNBIS](#))
- General Multilingual Environmental Thesaurus ([GEMET](#))

Connections are being made between concepts with the goal to make information more discoverable and accessible. LandVoc may have an independent hierarchy, but it shares infrastructure with AGROVOC, and consequently, other sub vocabularies. These built-in connections offer high potential for collaboration and AGROVOC is always open to exploring shared opportunities.

The Land Portal Foundation is proud to serve the land community by maintaining – together with FAO – this innovative infrastructure and working with the wider community to promote the growth and expansion of LandVoc by adding new terms, translations, connections and promoting its wider use and application.



# Conclusion and next steps

**Improving land governance requires not only that land data and information be made more open. Opening up of land data is a process for improving the quantity and quality of land data available for decision making and to deal with land governance challenges. The Open Up Guide for Land Governance (using the data set as the unit of observation) allows for the scrutiny, assessment, and subsequent opening up of land governance data at the country level.**

Using LandVoc, country level data and information can be made more visible by linking data and information in a systematic and coherent manner across geographies, including reporting on national development indicators and contributing to international and global indicators for land governance. Together these initiatives from the Land Portal Foundation can play a key role in helping to open up land data and information in a structured manner that addresses the major concerns of stakeholders and improves data quality and accessibility.

From FAO's side, further work on AGROVOC includes improvements to data quality and coverage, such as adding more languages, deepening language coverage, connecting communities, and closing thematic bridges in close collaboration with specific expert communities such as fisheries or land governance. This facilitates knowledge discovery and innovation, data and knowledge integration, and promotion of sharing and reuse of data across the data value chain and across languages. The work of FAO on knowledge sharing contributes to the priorities of FAO to achieve a world without hunger, malnutrition and poverty in a sustainable manner, with a focus on the exchange of knowledge, information and data as a key step towards achieving the SDGs (FAO, 2020).



# References

---

- Bayer, C.-T., and Booth, K. 2021. *Open-up Guide: Open data for improved land governance*. Land Portal, Open Data Charter. Cited 13 May 2022. <https://landportal.org/library/resources/open-data-improved-land-governance>
- Çağdaş V., Meggiolaro L., Stubkjær E. 2021. *Semantic Resources for the Geospatial Domain*. 9th FIG Workshop on the Land Administration Domain Model / 3D Land Administration, FIG e-Working Week 2021, Amsterdam, Netherlands, 20–25 June 2021, pp.1-16
- Davies, T. and Chattapadhyay, S. 2019. Open Data and Land Ownership. In T. Davies, S. Walker, M. Rubinstein, & F. Perini (Eds.), *The State of Open Data: Histories and Horizons*. Cape Town and Ottawa, African Minds and International Development Research Centre.
- FAO. 2020. *Scientific information and digital data on food and agriculture made available, accessible and applicable worldwide*. Rome. [www.fao.org/3/ca9103en/CA9103EN.pdf](http://www.fao.org/3/ca9103en/CA9103EN.pdf)
- FAO. 2022a. Land governance and planning. In: *Land and Water*. Rome. Cited 13 May 2022. [www.fao.org/land-water/land/land-governance/en](http://www.fao.org/land-water/land/land-governance/en)
- FAO. 2022b. *AGROVOC*. Rome. Cited 6 May 2022. [www.fao.org/agrovoc](http://www.fao.org/agrovoc)
- Mey, L., Zammit, S., and Meggiolaro, L. 2020. *Land Governance Lost in Translation – Exploring Semantic Technologies to Increase Discoverability of New Technologies & Data*. FIG. Cited 6 May 2022. [www.fig.net/resources/proceedings/fig\\_proceedings/fig2020/papers/ts06e/TS06E\\_mey\\_zammit\\_et\\_al\\_10648.pdf](http://www.fig.net/resources/proceedings/fig_proceedings/fig2020/papers/ts06e/TS06E_mey_zammit_et_al_10648.pdf)
- Open Data Charter. 2022. Principles. In: *Open Data Charter*. Cited 6 May 2022. <https://opendatacharter.net/principles>
- United Nations. 2022. Big Data for Sustainable Development. In: *Global Issues*. New York, USA, UN. Cited 13 May 2022. [www.un.org/en/global-issues/big-data-for-sustainable-development](http://www.un.org/en/global-issues/big-data-for-sustainable-development)
- Williamson, I., Enemark, S., Wallace, J., & Rajabifard, A. 2010. *Land Administration for sustainable development*. FIG Congress 2010. [www.fig.net/resources/proceedings/fig\\_proceedings/fig2010/papers/ts03a/ts03a\\_williamson\\_enemark\\_et\\_al\\_4103.pdf](http://www.fig.net/resources/proceedings/fig_proceedings/fig2010/papers/ts03a/ts03a_williamson_enemark_et_al_4103.pdf)
- Zevenbergen, J., De Vries, W., & Bennet, R. 2016. *Advances in Responsible Land Administration*. Boca Raton, Taylor and Francis.





**“ As data forms the basis of decision making, policy guidance and resource allocation, the partnership with Land Portal through the ‘Land Governance Programme in Kenya’ opened up an avenue for understanding the importance of data and Open Data platforms and engagement.”**

**Husna Mabarak**  
Land Governance Programme Manager  
and Gender Focal Point  
Food and Agriculture Organization of the  
United Nations, Kenya



**“ Association with Land Portal helped me understand the imperative and nuances of information management for transforming land governance and tenure security. Our partnership has led not only to create a window of India’s land information before the world, but also enabled a better appreciation of India’s land information ecosystem. Importantly, our efforts of connecting the information and actors and arguments for open data, have already started generating good responses, and we look forward to impacting the open land information ecosystem substantially together.”**

**Pranab Choudhury**  
Vice President  
Center for Land Governance  
NR Management Consultants India Pvt. Ltd.

Partnerships and UN  
Collaboration Division –  
Partnerships and Outreach  
[AGROVOC@fao.org](mailto:AGROVOC@fao.org)  
[fao.org/agrovoc](http://fao.org/agrovoc)  
[agrovoc.fao.org](http://agrovoc.fao.org)

**Food and Agriculture  
Organization of the  
United Nations**  
Rome, Italy