



**Ministry of Communication and Information  
Technology (MCIT)**

---

**ETHIOPIAN GOVERNMENT OPEN DATA  
PORTAL**

*Ref: MCIT/EGOV/OD001/2007*

**SITUATION ASSESSMENT AND ANALYSIS OF  
REQUIREMENTS**



September 2015  
Addis Ababa, Ethiopia

Version Control				
Version	Authors	Date	Section Affected	Description
1.0	Solomon Tesfaye, Netsanet Animut, Beza Hailu, Dr. Lishan Adam	September 15, 2015	New	First version of Current Situation Assessment and Requirement Analysis Report

Document Distribution			
Version	Name	Soft Copy	Hard Copy
1.0	MCIT	September 18, 2015	September 18, 2015

## Contents

Abbreviations .....	6
<b>1 Executive Summary .....</b>	<b>7</b>
<b>1.1 Background.....</b>	<b>7</b>
<b>1.2 Key Findings .....</b>	<b>8</b>
<b>1.3 Next Steps .....</b>	<b>10</b>
<b>2 Analysis of Current Situation and Requirements .....</b>	<b>12</b>
<b>2.1 Background.....</b>	<b>12</b>
Purpose of the Assessment .....	15
Methodology.....	16
Data Collection and Analysis .....	16
Desk Research .....	16
Benchmarking visit.....	17
Structure of the Report.....	17
<b>2.2 Overview of International Experiences in Open Data .....</b>	<b>18</b>
Canada .....	18
Indonesia .....	20
Kenya.....	22
United Kingdom.....	23
Lessons Learned from International Experiences.....	27
<b>2.3 Analysis of Policy and Regulatory Environment in Open Data .....</b>	<b>30</b>
General Policy and Regulation on Open Data.....	35
Other Legislations that Provide for Open Data.....	37
Infrastructure Policies that Encourage Access to Open Data.....	38
Commitment to Open Access.....	38
Privacy of Online Data.....	38
Charging for Data .....	39
Arrangement with the Private Sector .....	40
Gap Analysis .....	40
<b>2.4 Analysis of the Economic Impact of Open Data to Ethiopia .....</b>	<b>44</b>
Economic benefit as a driver of open public data.....	44
Cost of Open data .....	45
Quantification of the economic benefit of open data.....	45
Stimulating the economic benefit of open data .....	46
<b>2.5 Analysis of Ethiopian Government Open Data .....</b>	<b>47</b>
Data Availability .....	47
Quality of Ethiopian Government Data .....	53
Ease of Access and Re-Use.....	55
Data Formats.....	56
Metadata for Ethiopian Government Open Data.....	59
Dataset Selection .....	62
Data Catalogue.....	62
Governance of Open Data .....	67
<b>2.6 Analysis of Features for Open Data Portals.....</b>	<b>68</b>
Introduction .....	68
Features in the Terms of Reference of the project.....	68
Feature of open data portals of other countries.....	68
Assessment of the needs of ministries/agencies .....	77
<b>2.7 Analysis of Technology Options.....</b>	<b>81</b>

Background.....	81
Common Features of Data Catalogues.....	81
Types of Data Catalogues.....	83
Open data technologies.....	84
Common Open Data Platforms.....	86
Selection of open data platform for Ethiopia.....	97
Conclusion.....	103
<b>3 Summary of Requirements.....</b>	<b>104</b>
<b>Introduction.....</b>	<b>104</b>
<b>Functional Requirements.....</b>	<b>104</b>
Requirements collected from ministries/agencies during the assessment.....	104
Requirements presented in the TOR of the project.....	108
Recommended site organization of the Ethiopian government ODP.....	110
<b>Non-functional Requirements.....</b>	<b>112</b>
Performance/efficiency requirements.....	112
Security.....	112
Documentation.....	113
Legal Requirement.....	113
Privacy Requirement.....	113
(Data) Quality Requirements.....	113
Language Support.....	113
Interoperability.....	113
Metadata Requirements.....	114
<b>Use case representation.....</b>	<b>115</b>
<b>4 Conclusion and Next Steps.....</b>	<b>120</b>
<b>Conclusions.....</b>	<b>120</b>
<b>Way forward.....</b>	<b>122</b>
<b>References.....</b>	<b>124</b>
<b>Appendix A: List of Institutions.....</b>	<b>126</b>
<b>Appendix B: List of Document Consulted.....</b>	<b>127</b>
<b>Appendix C: Survey Questionnaire.....</b>	<b>128</b>
<b>Appendix D: Summary of Collected Data.....</b>	<b>136</b>

## Figures

Figure 1. Canada's open data portal .....	20
Figure 2. Indonesia Open Data Portal .....	22
Figure 3. UK Open Data Portal .....	24
Figure 4. The Ingredients of Open Data .....	26
Figure 5. Organizational Analysis.....	48
Figure 6. Organizations that Provide Electronic Access to Data Via Different Media .....	49
Figure 7. Available Data Formats .....	54
Figure 8. Analysis of Use of Standards and Other Data Quality Tools .....	55
Figure 9. Data Quality Analysis.....	57
Figure 10. CKAN Interface .....	87
Figure 11. DKAN Interface .....	88
Figure 12. Socrata Interface .....	89
Figure 13. ODS Interface .....	90
Figure 14. OGPL Interface .....	90
Figure 15. Junar Interface .....	91
Figure 16. Swirrl Interface .....	92
Figure 17. OGDl Datalab Interface .....	93
Figure 18. Libre Inteface.....	93
Figure 19. Semantic MediaWIKI Inteface .....	94
Figure 20. fluidOps Interface.....	95
Figure 21. Datahub Inteface.....	96

## Tables

Table 2. Ethiopian Government Datasets by Ministry/Agency.....	50
Table 3. Recommended Dataset Formats.....	58
Table 4. Analysis of Metadata features used by different countries .....	61
Table 6. <i>Analysis of Cataloguing/ browsing features used by different countries.....</i>	63
Table 7. <i>Analysis of Themes/ Sectors used by different countries.....</i>	63
Table 9. Most common open data platforms .....	84
Table 10. Examples of dataset publishing and presentation tools .....	85
Table 11. Licensing and supported deployment methods of different products .....	97
Table 10. <i>Analysis of Portal features of different countries .....</i>	69
Table 10. DCAT-AP Quick Reference of Classes and Properties. Extracted from DCAT-AP Specification Final v1.01.....	114

## Abbreviations

Term	Description
ADB	Address Data Base
CSA	Central Statistical Agency
CT	Consulting Team (eSystems and LogCon)
CKAN	Comprehensive Knowledge Archive Network (CKAN)
CMS	Content Management System
DHCP	Dynamic Host Configuration Protocol
DNS	Domain Name Server
GDP	Gross Domestic Product
GIS	Geographic Information System
GOE	Government of Ethiopia
ICT	Information and Communication Technologies
MCIT	Ministry of Communication and Information Technology
OG	Open Government
OGD	Open Government Data
ODP	Open Data Portal
OPGL	Open Government Platform
UAT	User Acceptance and Testing
UN	United Nations

# 1 Executive Summary

## 1.1 Background

The Ministry of Communication and Information Technology (MCIT) initiated the development of national open government data portal as part of its eGovernment implementation. A large amount of government data will be made available to citizens and the general public through the portal. The data is expected to be available in a flexible and user-friendly platform that allows data users to view and download different types of data sets at national and regional levels, compare data sets, and create maps and other visualizations. The open data portal is expected to be robust and secured that will be used to find, download, query and visualize data sets from government ministries and agencies. It will also help implement policies of open data and establish relevant organizational processes to prepare high quality data sets to be published on their respective data portals. The open data portal will centrally hold data sets published by public administrators at all levels.

eSystems Africa has been contracted to develop Open Data Portal for the government of Ethiopia. One component of this assignment is to conduct the current situation assessment of the legal/policy environment for open data, identify the various data sets that exist at various ministries and agencies and recommend appropriate classification and metadata, recommend open data portal features, and study and recommend platform for the development of the portal.

Accordingly, 21 government ministries and agencies were included in the assessment. Questionnaires, interviews, and focus group discussions were conducted at various government ministries/agencies to collect primary data. Desk research and benchmarking visits were also conducted to learn from experiences of other countries. Various legal and policy documents, including proclamations, policies, standards, and guidelines relevant to open data were also collected and analyzed.

## 1.2 Key Findings

### Legal and Policy Framework

These country experiences suggest that Ethiopia can move ahead with its open data by consolidating and publishing existing open data sets through data.gov.et. In parallel, it needs to launch a promotional scheme to increase data supply in priority areas like agriculture, health, education, economy and financial management, geographic information, law and others.

This can be achieved through mandatory policy requiring government departments and agencies to maximize the release of data and information of business value subject to applicable restrictions related to privacy, confidentiality, and security. Eligible data and information should be released in standardized, open formats, free of charge, and without restrictions on reuse.

In addition the government needs to work on capacity building initiatives that improve data availability, quality and metadata. This demands training of data suppliers in government open licensing techniques and presentation of data in open machine-readable format. It also involves training of decision makers on data-driven analysis and arguments.

At the same time, it is important to identify at an early stage actual or perceived “blockers” in order that policy or legal change can be initiated early. The current policies, legislative frameworks like the National Data Protection Law, e-commerce law, computer misuses and cyber crime law and the e-signature law can be revised for inclusion of clauses that cater for open data.

The government needs to establish a legal framework for proactive online disclosure of public data. A subsidiary legislation that builds on the Mass Media and Freedom of Information Law needs to be enacted in order to addresses electronic access to raw data and provide for the proactive release of public information before individuals requests it. The law will update and improve access to information and specify that new data sets and records to be collected and published. This would eliminate legal barriers to access to online public information. The law would also include mandates on anonymous requests and regulation on fees and fee waivers.



In addition, Ethiopia needs to work on the requirements to manage risks of data security and privacy. Breaches of data privacy and security needs to be anticipated and managed based on international regulation. The data protection regulation needs to be enforced together with the accountability of entities managing public data.

## **Data and Metadata**

The country survey shows that Ethiopia is at the early stages of development of open data. From the 20 surveyed government ministries and agencies, a total of about 84 datasets are identified as data that can be available for the public. The fear for releasing data in open reusable format abounds. The current data is largely aggregated and only available through individual institutions. Data is not harmonized and has not been presented in open standard format. There is a mismatch between data supply and demand. The current data sets are not commensurate with the needs of the population. The survey also show that no dataset is openly licensed, and there is limited understanding of what open licenses entails.

The assessment recommended the use of open data formats such as XML and RDF. It is also recommended to adopt the DCAT metadata dictionary nationally and ensure its full implementation when publishing Open Data.

## **Technology Options**

There are a number of open data technologies in the market. The products may be open source or commercial, self-managed or cloud-based. The products may be complete platforms or supporting tools designed for purposes of facilitating specific functions of the data publishing or presentation process.

The team undertook a two-phase evaluation and selection of a suitable open data platform for Ethiopia. The first phase of the selection process involves evaluation of the basic needs of the ministry, which is the product has to be free and open source (FOSS) and should be deployable at its premises. The second phase of selection involves a detailed analysis of the selected platforms in regards to their suitability for implementing the Ethiopian ODP. The main criteria used for

evaluating the candidate platforms include ease of installation and configuration, availability of documentation, existence of strong support community, feature sets of the platform and adoption by other national-level open data portals.

The study team analysed about 12 different open data platforms. Accordingly, two platforms (CKAN and DKAN) were shortlisted after the two phases of assessment based on pre-defined criteria.

Out of the free open source and on-premise deployable platforms, the best platform for national level government open data portals was found to be the CKAN platform. This is the most widely used platform for government open data portals. It has rich documentation and a strong support community. Most of the requirements presented in the terms of reference of the project can be met with the CKAN platform. Hence, the team recommends CKAN as the best platform for the Ethiopian government open data portal implementation. For any content management system requirements that cannot be provided by CKAN, the team recommends Liferay portal, the most widely used portal platform in Ethiopian government organizations.

The team recommends DKAN platform as a secondary option. DKAN is a PHP based platform that is easy to deploy and configure. It is compatible to CKAN and runs both on a Windows server and Linux platforms. Most of the requirements in the terms of reference can be met with DKAN. It has rich documentation and support community.

## 1.3 Next Steps

From Policy/legal and Organizational perspective it is essential to secure high level political commitment, create awareness on the benefits of open data at all levels, develop and launch the coordination mechanism for open data in order to make it operational and sustainable. In providing data, cooperation is more important than going it alone. Building a national open data platform is crucial. All relevant government ministries/agencies should be brought on board in open data activities.

Capacity building around open data is also an important component. It is important to systematically promote data skills – reducing shortages of skilled workers in government ministries and agencies and narrowing the digital divide

In terms of datasets, it is important to adopt the DCAT metadata dictionary nationally and ensure its full implementation. MCIT should issue guidelines on recommended licenses for datasets. The guideline should encourage the use of open licenses. Under the Open Data Initiative, government ministries/agencies should publish their data in the most open way possible. One way to measure the openness of the formats used is through the 5-star deployment scheme for Open Data.

MCIT in consultation with the assessment team should discuss the recommended platforms (CKAN and DKAN) and select one suitable platform for development of Ethiopian government open data portal.

## 2 Analysis of Current Situation and Requirements

### 2.1 Background

The Ministry of Communication and Information Technology (MCIT) of the Federal Democratic Republic of Ethiopia (FDRE) has a mission to develop, deploy and use ICTs to improve the livelihood of Ethiopians and contribute to the development of the nation. To this end, the ministry has undertaken a number of e-government initiatives that improve the delivery of services to citizens and enhance internal efficiencies within government organizations. One of these initiatives is the dissemination of open data to the public.

It is part of a bigger open government (OG) initiative that many governments around the world are currently undertaking to deliver better services to their constituencies using the Open Government (OG) approach. Open Government is an innovation in governance and administration that instigates a rethinking and updating of administrative procedures to meet citizens' demands and expectations<sup>1</sup>. It entails cultural, organizational, procedural and attitudinal changes in how public servants relate to citizens.

OG initiatives intend to increase transparency of governments, boost a number of services to people, open up new business opportunities and jobs for application and service developers, and develop new synergies between government, public administration, and civil society organizations.

The most important benefits generated by OG include:

- Increased transparency, engagement and accountability of governments
- Increased number of services to people resulting from increased base of potential service providers
- New innovative business opportunities and jobs for application and service developers
- New synergies between government, public administration and civil society organizations
- Increased citizen participation and inclusion through extended offers of services closer to

people's needs

- Interoperability and openness of government information

Open Government Data (OGD) is the pillar of OG strategies where public institutions including state and local government agencies upload their raw data on the Web in reusable, machine-readable and open formats. OGD facilitates access to public information stored by governments on behalf of the citizens such as meteorological, legal, traffic, financial, economic, population and demographic data as well as maps; excluding sensitive or individual information.

The Ministry of Communication and Information Technology (MCIT) initiated a project for implementation of a national open government data portal as a key OG initiative that will make a large amount of government data available to citizens and the general public. The data is expected to be available in a flexible and user-friendly platform that allows data users to view and download different types of data sets at national and regional levels, compare data sets, and create maps and other visualizations.

- Developing a robust and secured open data portal that will make government data sets available to general public. The open data portal will be used to find, download, query and visualize data sets from government ministries and agencies.
- It will help implement policies of open data and establish relevant organizational processes to prepare high quality data sets to be published on their respective data portals.
- The open data portal will centrally hold data sets published by public administrators at all levels.

eSystems Africa has been contracted to develop Open Data Portal for the government of Ethiopia.

The scope of the assignment includes:-

- Assess data sets that exist at various ministries and agencies and recommend appropriate classification and metadata. The recommendation shall consider the National Data Set under study.
- Produce a report of the current status of datasets in the country,
- Developing, launching, operating and maintaining a large-scale web-based Ethiopian

Government open data portal system to act as a single point of access to (open) datasets produced in the country;

- Gathering, analyzing, designing, implementing, configuring, commissioning, testing, deploying and providing training on the portal system.
- Installing, configuring and implementing the software system and test with actual data at the working environment of the ministry.
- Defining and documenting each section of the portal system in system development life cycle.
- Examining the sustainability of open data portal systems
- Providing evidence on the economic impact of publishing data assets by public administrations
- Developing web applications that can be used to analyze data sets and generate reports
- Develop sample mobile apps that can be used to view datasets and reports offline using mobile devices
- Collect data sets from selected ministries and agencies and integrate them in the open data portal and provide a mechanism for searching datasets in the portal
- Enable locating of data sets using maps,
- Enable mechanisms for integrating content in the national data portal with other similar portals,
- Provide local language support of the open data portal,
- Open data portal management training for 5 peoples; Note: the training could be carried out locally or abroad and all the necessary expenses required for the training should be covered by the firm and should be clearly seen in the financial document.

- Preparing international experience sharing visit on national open data portal for 3-4 peoples at least for duration of one week. Note: all the necessary expenses required for the visit should be entirely covered by the firm and should be clearly seen in the financial document.

### **Purpose of the Assessment**

eSystems Africa began the project with a survey that was designed to assess available data sets and public policies and standards. We have also carried out the assessment of a technologies and platforms currently in use for delivery of data. This was followed by the assessment of current data portal of public institutions.

This study builds on the methodology and experience of a prior studies conducted by Grail Consulting Services for the National Dataset Master Plan Development. The objectives of the OGP current situation assessment in Ethiopian Government Ministries and Agencies include the following:

- Describe legal and policy environment, identify gaps and recommend strategies to bridge the identified gaps
- Assess the various data sets that exist at government ministries and agencies.
- Study the availability and use of data and metadata standards by government ministries and agencies and recommend a national metadata standard for OGP
- Assess the open data portals of other governments to learn how their data sets have been classified and presented in their respective open data portals. It includes an assessment of the needs of MCIT, major data owners, major stakeholders and potential users of the data sets.
- Recommend appropriate categorization and classification scheme for the data sets, which will be used to present data sets on the open data portal
- Assess the design requirements of the open data portal to be developed, including assessment of similar portals of other governments to find out the features and functionalities that are commonly available on similar national level government data portals
- Study the various platforms and technologies used for implementation of open data portals and related components, which include a data catalogue, data visualization, web application, mobile application, location based presentation/GIS, etc.
- Conduct detailed assessment of the platforms and technologies used in implementation of

open data portals of other governments

## Methodology

The assessment team utilized a participatory approach that involved all relevant stakeholders. To understand the current situation of open government data provision in Ethiopia, the assessment team identified and examined four essential dimensions of open data implementation: legal and policy framework, datasets and metadata standards, portal features, and technology choices. To answer the assessment questions, various methods were used, including document reviews, in-depth interviews, questionnaires and observation. Upon gathering the relevant data, analysis, inferences and recommendations were made.

## Data Collection and Analysis

To gather primary data, questionnaires were distributed to 21 government ministries and agencies jointly identified by the assessment team and the counterpart team at MCIT (See Annex A for the list of government ministries and agencies consulted). In addition, in-depth, face-to-face key informant interviews were conducted with relevant staff(team) at 7 ministries (33%). In most of the ministries relevant personnel involved in data management activities were participated during the interviews. The data collection instrument can be seen in Annex C.

The collected primary and secondary data was analyzed using qualitative and quantitative techniques. The quantitative data were analyzed and presented as tables and graphs. The qualitative data were analyzed using thematic analysis based on themes identified based on the coding of the data, objectives of the assessment, and international standards.

## Desk Research

The assessment team collected and rigorously analyzed important legal and policy documents, including proclamations, policies, standards, and guidelines relevant to open data (See Annex B for the list of documents reviewed). The team also assessed portals, e-services, e-participation and ministerial websites. The websites are assessed for content, features, and level of web content accessibility. Similarly, open access portals of other countries were also assessed to learn from the experiences of other countries and important lessons were drawn from their experiences.



## Benchmarking visit

A team composed of key personnel from MCIT and eSystems Africa conducted a week long benchmarking visit to Indonesia to learn from the experiences of implementation of Open Data Initiatives from a country in similar socio-economic development with Ethiopia. During their stay the team visited the Open Data Lab, the SmartCity project of Jakarta, the Open Data project at the President’s Office, and other important stakeholders. The assessment team conducted brainstorming sessions with the site visit team to draw lessons from the Indonesian experience. (The site visit report is available at Annex E)

## Structure of the Report

The report has two parts. Part One is the Current Situation Assessment Report and the subsequent sections that follow this background chapter are as follows:

- Section 2 discusses the international experience
- Section 3 presents the status of open data in Ethiopia
- Section 5 provides analysis of the various available technological options for development of open data portal
- Section 6 provides an overall gap analysis in the current open data offering
- Section 7 presents the next step based on the preceding sections

Part two covers the functional and non functional requirements of the Ethiopian government open data portal.

## 2.2 Overview of International Experiences in Open Data

There is a significant experience in the implementation of open data and the development of government open data policies. Over seventy countries have initiated open data over the last decade. Albania, Azerbaijan, Brazil, Canada, Czech Rep, Denmark, Dominica, Estonia, Finland, Germany, Ghana, Greece, Israel, Italy, Jordan, Kenya, Korea, Mexico, New Zealand, Netherlands, Nigeria, Norway, Sweden Tunisia, United Kingdom and the United States are among those that have implemented a wide range of policies on open access. A global collaboration on open data under the Open Government Partnership (OGP) that aims to promote innovative approaches to transparency throughout these countries has also been established.

Countries in the European Union have been the leaders in open government data system. The European Union has issued several directives to expand residents' access to public sector information. The EU adopted the Public Sector Information ("PSI") Directive in 2003, which regulates the obligations of public sector bodies in the EU when re-using public sector data. The EU's Data Protection Directive protects individuals' fundamental right to privacy regarding personal data usage, broadly defined to include personal information encompassed in public sector data.

Canada, Kenya, Indonesia and the United Kingdom are among the many government that have active on going open data initiatives that aim to make government information available online for free to the public. Australia, Netherlands, Hungary and Poland have enshrined citizens' right to access to information in their constitution. Additional provisions enhancing access to information are included in laws in areas such as environment consumer protection and health. This section presents examples from four continents.

### Canada

Canada' is one of the countries that has made a considerable progress with open data. Canada's freedom of information laws dates in 1983. The Access to Information Act gave Canadians the ability to access a range of government documents and other information and specified a time period in which they could expect to receive that information. The act applies to federal agencies.


Canada has a Privacy Act that grants individuals access to certain types of their own information that is held by the government, and forbids the government from disclosing individuals' personal information or using that information for purposes other than that for which the information was originally obtained, without the individuals' consent.

The Public Servants Disclosure Protection Act of 2007 amended the Access to Information Act, the Privacy Act, and the Personal Information Protection and Electronic Documents Act to narrow the timeframes under which agencies could deny information requests.

Canada has committed to the principle of open by default. While the country does not have an “open by default” law on its books, it has affirmed its commitment to the concept in forward-looking plans. The government declared its intention to promote an “open by default” approach in a forthcoming Directive on Open Government, which will be created by the federal Treasury Board but will apply to the entire federal government. The directive will issue guidance to federal agencies on publishing open data and establishing data standards, publication priorities, and other information on publication.

Canada's Open.canada.ca data portal has published over 200,000 data sets, more than any other nation. All but approximately 8,000 of these are geospatial datasets, such as the geospatial vector data used in geographic information system software. The federal data portal suffers some quality and completeness issues.

Français

 **Government of Canada** **Gouvernement du Canada**







Search Canada.ca

Jobs ▾ Immigration ▾ Travel ▾ Business ▾ Benefits ▾ Health ▾ Taxes ▾ More services ▾

Home → Open Government

## Open Data

Search open data that is relevant to Canadians, learn how to work with datasets, and see what people have done with open data across the country.

Follow:      

**Search Through our Open Data Portal**

Looking for data about Government of Canada services, financials, national demographic information or high resolution maps? Discover that and more through our Open Data portal, your one-stop shop for Government of Canada open datasets.

**Working with Data and Application Programming Interfaces**

Learn more about how to use a variety of data formats, including Application Programming Interfaces (APIs).

**Apps Gallery**

Browse and download apps that were built both by the Canadian public and the Government of Canada that use open data.

**Open Data 101**

Learn what open data is, how to get the most out of it, and explore the principles that are guiding the Government of Canada's Open Data initiatives.

**Suggest a Dataset**

Can't find the dataset that you are looking for? Let us know!

**Submit Your App**

Have you made an application using GC Open Data or Open Information? If so, let us know and we will share it with Canadians in our [Apps Gallery](#).

**G8 Open Data Charter** [Canada's G8 Open Data Action Plan](#)

Figure 1. Canada's open data portal

While many of the available data sets have been improved by Open.canada.ca's efforts to standardize metadata, some data sets, such as collections of geospatial data are still underway.

## Indonesia

Indonesia is one of the developing countries that have been active in Open Government Data initiative. It launched the policy to open public data and the Freedom of Information (FoI) law for opening data in Indonesia. The FoI law is considered a big step towards a more transparent government. It gives citizens' rights to information related to public policy making, encourages active participation of the citizens, and improves the managing of public services.

The Open Government Indonesia (OGI) movement started in September 2011. In the effort to enforce the open government movement in Indonesia, a specific unit was appointed under the President called the UKP4. This unit is tasked with all the open government duties, also those related to the global Open Government Partnership of which Indonesia is co-chair in 2013. The ‘Open Government Indonesia’ (OGI) website ([opengovindonesia.org](http://opengovindonesia.org)) was launched in 2012 and contains information about the OGI initiative and its actions plans. The goal of the Indonesian government in launching this program was to initiate a change in the bureaucracy between sectors and institutions, to improve transparency for the public which in turn leads to identification of ineffective and inefficient processes or individuals, and also to improve the overall service that is provided to the public.

The Indonesia Open Government Action Plan was approved in 2014. It emphasizes the importance of increasing access to data to 1) Improve Public Services; 2) Increase Public Integrity; and 3) manage Public Resources effectively. The government launched a data portal (<http://data.id>) in September 2014 to encourage more data transparency in the country. Launched simultaneously with the Open Data Movement in Indonesia, the data portal was the result of greater collaboration, specifically the President’s Delivery Unit, and civil society groups.

**Data** Visualisasi Aplikasi Komunitas Tentang

## Data Indonesia dalam satu portal

Temukan data berbagai sektor instansi pemerintahan yang dapat anda gunakan secara bebas

Cari Data, Tema, Instansi...

**TOPIK**

- Ekonomi dan Keuangan
- Pendidikan
- Kependudukan dan Ketenagakerjaan
- Energi dan Sumber Daya Alam
- Kesehatan
- Pengadaan

**Blog**

- Pantau Harga Sembako, Jokowi Gelar Lomba IT >>
- Kantor Staf Presiden Fasilitas Komunitas IT Pantau Harga Komoditas >>

**Featured**

- Kementerian Pekerjaan Umum dan Perumahan Rakyat**
- Data Bongkar dan Muat Barang melalui Jalur Transportasi Laut**
- Pemerintah Provinsi DKI Jakarta: 318 dataset**

**Statistik**

- 1006 dataset**
- 29 instansi**
- 18 grup**

## Figure 2. Indonesia Open Data Portal

By 2015, the portal had 1006 data sets available in standard CSV format for download and reuse. The open data is the basis of the Indonesia Open Data lab that was launched by the World Wide Web foundation with the aim to develop innovative, home-grown approaches to open data — creating powerful tools to tackle practical problems and improve lives. The data lab Using a combination of research, incubation, training and engagement, to accelerate progress and ensure that open data rapidly becomes a vital tool to tackle practical problem in Indonesia.

### Kenya

Kenya emerged as one of the countries that had traditionally the most restrictive access to information to one that champion open data. Under former President Daniel arapMoi, between years 1978-2002, the government restricted the free flow of information and hindered other forms of media networks. This was held under the Official Secrets Act that had existed from the colonial era. Under the influence of such an Act, it is understandable that the government culture was built to restrict access to information from the public and to closely guard all categories of information.

In 2005, Dr. Bitange Ndemo became the Permanent Secretary of Kenya’s Ministry of Information and Communications and brought changes within the government that would enable a more open government. On 8 July 2011, President Kibaki officially launched [opendata.go.ke](http://opendata.go.ke) with 200 datasets that were categorized into education, energy, health, population, poverty, and water and sanitation. The launching of the website also launched the Kenya Open Data Initiative which was an important step for the country.

When the portal launched, Ndemo wanted to avoid confrontation with public officials on publishing non-public data by launching the portal with data that was already categorized for the public but not yet published. However, no policies were enacted to ensure the government opened their non-published data. In 2010 there was an addition to the constitution, which called for the government to “publish and publicize important information affecting the nation”. This constitution amendment was the anchor to the open data efforts instead of waiting for the Freedom of Information law. Another document that contains statements about open data is the Vision 2030 Plan, which is a

long-term development blue print launched in 2008. This Plan provides the means towards a more open government through ICT infrastructural developments.

However, because of the lack of Freedom of Information law there is no legal background or formal policies that enforce the government to open the data. The Freedom of Information law is currently still tied up in parliament and has yet to be announced as a new law. This law may be one of the turning points needed for the Kenya Open Data Initiative to be completely in action.

On the global scale, Kenya has also made a statement to the world by joining the Open Government Partnership. Through the OGP, Kenya has formulated an action plan that is committed towards addressing certain areas of open government. The areas include the improving of public services, increasing public integrity, and more effectively managing public resources. Kenya has also been active in stimulating applications that promote open government and access to information by encouraging data journalism and open coding.

## **United Kingdom**

The United Kingdom is one of the top countries in the development and use of open data. It claimed the top spot on the World Wide Web Foundation's 2015 Open Data Barometer. The UK passed the Freedom of Information Act (FOIA) in 2000, which establishes the public right to access paper or electronic copies of government-held information from all public authorities, excluding personal information about individuals and information that relates to national security or that is held by intelligence departments.

In 2012, the Cabinet Office, an agency charged with supporting the Prime Minister and Cabinet, released a white paper outlining its vision for open data, including a schedule for openly releasing high-value data sets from various departments. The Cabinet Office also published its 2012 Public Data Principles, a set of guidelines for how national government agencies should publish open data, such as using standard features like machine-readability, open formats, and open licenses.

The UK has created four non-binding commitments to open by default. First, the 2012 Public Data Principles state that public data should be open without caveat or exception.

Second, the Secretary of State’s 2013 Code of Practice, which provides guidance for public authorities on handling Freedom of Information Access requests, reinforced this sentiment, noting that authorities releasing data under the act must, as far as is reasonably practical, provide it in a machine-readable format, with the UK Open Government License as the default licensing model.

As of January 2015, the UK’s data portal contained approximately 20,000 published and over 4,000 unpublished data sets, and government agencies publishing data through Data.gov.uk have generally been proactive in releasing large amounts of high quality data.

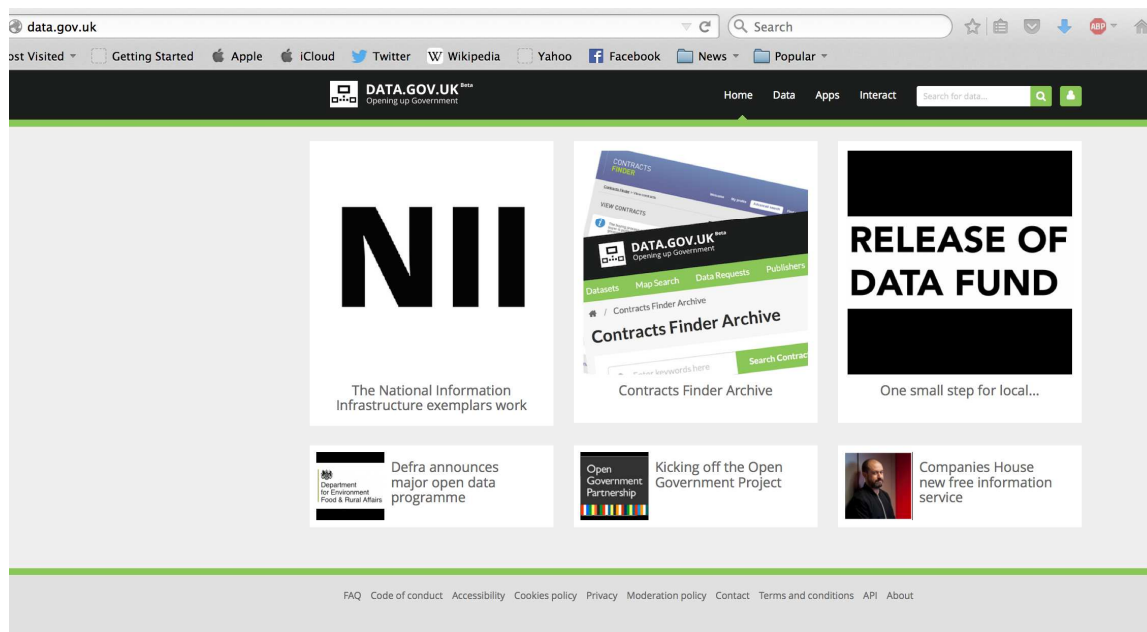


Figure 3.UK Open Data Portal

The portal offers several unique components, including a joint effort from the Cabinet Office and Data.gov.uk developers to identify and fix broken links in data sets and a feature that lets users view certain metadata of those data sets that agencies have included in their data inventories but have not yet published.

The UK has also been a global leader in ensuring its public data is usable by all, in part for having completed its commitments around open licensing ahead of schedule. The Open Government



License, first published in 2010, permits free and unlimited reuse with modification, subject only to attribution.

The UK has made considerable progress in releasing data for innovation, convening private-sector stakeholders to prioritize datasets for release, offering funding for data-driven start ups, and hosting a variety of hackathons and other events to promote data use.

The Cabinet Office has instituted feedback mechanisms with private sector stakeholders through the Open Data User Group (ODUG), an independent advisory group representing data users that provides recommendations on funding for open data initiatives and that conducts public consultations to identify and prioritize data sets for open release. The ODUG has also released a set of fifty case studies of businesses that make extensive use of open data, in an effort to demonstrate data. The non-profit Open Data Institute (ODI) also works to promote open data release within the government and innovation in the private sector, convening experts, publishing open data use case studies, and offering training courses. Regarded as a successful model for encouraging open data innovation, the Open Data Institute has expanded to include eleven branches in other cities around the world, and it has been replicated in the United States as the U.S. Open Data Institute.

Public and private sector efforts have been made to engage the developer community directly. The Department for Business, Innovation, and Skills has tried to spur open data innovation through its Technology Strategy Board. The board operates an open data innovation voucher program, whereby startups can receive grants of up to £5,000 (\$8,300) for open data consulting and services. The Open Data Challenge Series, administered by the Open Data Institute and innovation charity Nesta, hosts challenges to encourage open data innovation in the private sector and offers £400,000 (\$660,000) in prizes annually.

Local and regional governments in the UK have also been proactive in engaging with developer communities to promote open data innovation. Agencies have hosted hackathons and challenges around issues such as mental health care, flooding, and food security, and new events are regularly announced.

The experiences of these countries show that to be successful, the open data initiative should be accompanied by a series of policy effort including the Freedom of Information Act legislation. In addition countries need to move the data online as quickly as possible in a standard machine readable format. Portals need to be managed in a continuous basis, with data sets being added and updated to ensure open access to public sector information.



Figure 4. The Ingredients of Open Data

Countries have also moved ahead in publishing open data. However, there are many gaps to overcome before open data availability can lead to widespread effective use and impact. These include data availability, data quality and meta data and policy decisions that improve on the non-technical aspects of open data like improving connectivity, capacity and resources.

Experience in the United Kingdom has already shown a governance framework plays a critical role in enhancing access to open data. The governments needs to set up independent advisory group representing data users that provides recommendations on funding for open data initiatives and that conducts public consultations to identify and prioritize data sets for open release. Government should also encourage the private sector and civil society to develop applications around open data and to offer training courses.

## Lessons Learned from International Experiences

The following is a summary of some of the lessons we learned from the desk research that we have conducted on the open data initiatives of other countries and the benchmarking visit conducted in Indonesia.

### Political/Legal

- Top-level political leadership essential “Greater transparency will enable the public to hold politicians and public bodies to account” “Public information does not belong to Government, it belongs to the public
- 

### Organizational

- Raise awareness: being Open Data a recent topic is still not well known and it is important to spend significant time on educating people, which proved to be very useful to address several of the issues and concerns raised by people about open data.
- Resistance to change is a delicate issue. Government people tend to avoid risks (so it's better to keep them as low as possible). Much internal outreach is needed and it is essential to show all sorts of social and economic benefits and identify champions, as they are key.
- Start small and grow steadily
  - Start with just a few ministries/agencies and grow the base over time
  - Involve as many people from as many agencies as possible
- It is important to not focus on technology
  - Technology represents a small percentage
  - Do not have an Open Data catalog portal as the target

- An Open Data initiative is much more than a website, the website should be only a deliverable and, in many cases, the easiest one to do

### **Economic**

- Beware of the economic growth argument
  - Even in developed world data re-use is not that high just yet
- Improved government efficiency
  - In several cases the initial beneficiary of the project is the government itself. As an example, consolidated information from several databases into a single dataset that didn't exist before. The use of common vocabularies has improved interoperability across agencies
- Licensing schemes
  - It's better to just give the information away for free
  - Promote open licenses
  - The Open Government License is a great example and step forward
  - Avoid the use of different licensing schemes in different projects even when applied to the same type of information.
- M&E
  - Measure achievement
  - Measure delivery and conformance

### **Technical**

- Don't try to do everything perfect the first time but try to do it fast
  - Users are much more interested in getting the information soon than in getting the information in the best shape.
  - So what we recommend is: raw data now!... and better data afterwards
    - Improve over time (step by step)
- Do not try to enforce an specific architecture
  - Try to adapt to existing systems and build on top of them as a start
  - Use some innovative approaches from the beginning though
- Create Sustainable Applications
  - Sustainable apps come from social entrepreneurs
  - Do things that matter to people

## Datasets

- Release data people want and care about
- Clear, common, licensing approach
- Make sure data is re-usable 11
- Data Publishing
  - Put your data on the Web with an Open License (any format)
  - Make it available as structured data (e.g. Excel, CSV, instead of PDF)
  - Use open, standard formats (e.g. XML, RDF)
  - Use URLs to identify things (so people and machines can point at your data)
  - Link your data to other people's data
- Ensure Privacy of Personal Data
- It's not just about new data. Scope for "Open Data" also includes data previously "published" but in non-reusable format with restricted license.
- Recognize that not all Government data is accurate 22

## 2.3 Analysis of Policy and Regulatory Environment in Open Data

Open data is information that is accessible to everyone in machine-readable format, offered online at no or very minimum cost, and with no limits on reuse and redistribution. The free re-use of public sector information raises a multitude of policy and legal issues.

Policy informs the key motivation and objectives for opening data in machine readable form with no limit on reuse and influences the infrastructure and services that are available for accessing to public data. Policies ensure that citizens receive the information they sought. The policy must answer key questions like:

- Which public sector information should be available for citizens?
- How to safeguard privacy of citizens?
- What is the practical response time to receive public information?
- Whether to or how much to charge?
- How to manage information?
- How to present information?

Policy issues for opening data range from those that encourage proactive release of government information (opening data by default) and create the context to addressing technical issues like data quality and standards. Legislations that ensure the privacy of individuals and the security and integrity of data are also among the key aspects of government policies for open data. Open data policies also embrace aspects of ownership and custodian of public sector data and those that limit fees to be charged for accessing to public sector information. Policies may also cover technical aspects such as data presentation and retention, presentation channels such as portals, data formats, standards, quality and metadata.

Open data is underpinned by policies on disclosure of government information. The Freedom of Information (FOI) has been the main legislative document for disclosure of information produced by governments in many countries. While the Freedom of Information or media laws allow individuals (mainly via a request) to access the information that the governments hold, Open

Government Data policies proactively provides people with data in reusable digital formats for them to use and reuse in the way they prefer. The utility derived from the reuse of open government data depends on certain rights of use being granted, such as the right to freely reuse, freely share and freely alter the data. Freedom of Information tend to focus on print and traditional media like radio and TV, while the Open Government data focuses on row data available online from the government for reuse and manipulation.

Operational policies that cover more aspects of the open data process such as quality of the published data, machine readability of the data and the accessibility of the datasets on an open data platform. The policy also creates an open data support system through organizational and infrastructural changes. Organizational policies include those that enable the designation of open data to a specific agency centrally or specific organizational entity within each data publisher’s organization that is responsible for the published data. The main open data policy and regulatory issues are listed in Table 1:

Issue	Description	Key questions
General policy and legislation on open data and data reuse	Government needs to ensure access to data either through Freedom of Information legislation or open data policy- it needs to release all public data in open and machine-readable formats unless there is a compelling reason not do so, such as national security or privacy concerns.	Is there any policy on Open Data or (re-use) of public sector information? Is there a “right to data” or “open by default” legislation? If so, how is this enforced?
Other legislations that may promote or deter open data	The various legislation on access to information, copy right, intellectual property rights, freedom of speech and media laws need to be	What are policies/laws on government secrecy and access to information? What official policy/law/regulations exist on privacy, official secrecy, copyright, intellectual property, freedom

	<p>modernized to promote open access</p>	<p>of information, information security, electronic transactions or other information-related issues?</p> <p>What policies/laws help or hinder the use of information by public and civil society?</p> <ul style="list-style-type: none"> <li>•Freedom of speech law?</li> <li>•Internet access and freedom?</li> <li>•Press and media law?</li> </ul>
<p>Infrastructure policies</p>	<p>Government policies on universal access to communication and broadband plans can accelerate the use of open data.</p>	<p>Does the government have infrastructure and access policies that promote open access?</p> <p>Does the law encourage tax free computer ownership?</p> <p>Are there polices on universal access that encourage access to information?</p>
<p>Commitment to open access and the use of open licenses such as creative common</p>	<p>Public institutions need to commit open licensing, where data sets are available for reuse under a Creative Commons license</p>	<p>Does government (or any agency or local authority) use any license with respect to release/use of its data?</p> <p>Have recognized Open Licenses, such as Creative Commons By Attribution, been transposed into local form, or could the imported forms be effective under the local legal code?</p> <p>What government secrecy policies, laws and obligations apply on public agencies and their employees? Who has authority to authorize release data (and override civil service secrecy obligations, etc.)?</p> <p>What approvals need to be addressed</p>



		before publishing data?
Privacy	<p>Digitized information can be rapidly copied, mined, matched, and used for a broad range of purposes that many would consider invasive of privacy citizens, therefore legislation and guidelines around privacy needs to be in place</p> <p>Open data should ensure that third party rights are respected, thus data should generally be anonymous in terms of individual, business data and other third party</p>	<p>How are concerns around privacy/harm handled?</p> <p>What privacy safeguards exist for data about citizens?</p> <p>What laws exist about respecting anonymity/third party rights in government data (e.g. company registers, business data underlying official statistics), and how would these impact an Open Data Initiative?</p>
Ownership of government data	Central government needs to coordinate the licensing of public sector information (e.g. under creative commons, Government Open data license); ministries, agencies and departments may own data in the sense of updating it.	Who “owns” government data in a legal/copyright sense - is it government as a whole or individual ministries? Who is able to set/agree a license?
Charging for data	There should be minimum charge for accessing public data. Citizens should normally access to public data free of charge. If exist,	What is the policy (if any) on charging for government information? Who is responsible for setting this policy? Does this policy allow the commercial use of data? Do any/many agencies sell their

	fees should justify the effort of making specific data available	data? Should citizens, individuals and business pay for information defined as public by the law?
Arrangements with the private sector to use and reuse public data	Access to data by the private sector should be non-discriminatory. SMEs need to be encouraged to develop innovative applications around open data to increase its use	Does government have any exclusive arrangements with any companies with respect to any datasets? Conversely, does the government have any national or international obligations to be non-discriminatory?
Liability of misuse or accuracy of data	Civil servants should normally be liable for misuse of data	Can the government avoid liability for misuse and/or accuracy?
Availability of official statistics	Government should have explicit policies for making statistical data available for citizens	When was the Statistics law, if it exists, last updated? Would its provisions help or hinder the release of open data?
Governance of open data	There should be a national body for spearheading open data initiative. A participatory process needs to be put in place to encourage open data.	Who has a role in setting policy on the release of information? What is the role of judicial review in open government information? Are there any other independent bodies that review or approve release of government information?

The Government of Federal Republic of Ethiopia has embraced open government agenda. A series of initiatives have been underway including a reform the public service through Business Process Re-engineering, Balancing Score Card and Citizen Charter. This has created opportunities for improved disclosure and increased access to public information.

## General Policy and Regulation on Open Data

The findings of the survey indicate that access to information is guaranteed by the constitution and different ministries, departments and agencies have made effort to make data online. Article 29 of the Federal Democratic Republic of Ethiopia guarantees a right to obtain information about the activities of state organs and organs of local administration.

The government has also enacted its first Freedom of Information law - the “Proclamation to provide for Freedom of the mass media and access to Information”. The Proclamation provides that all persons have the right to seek, obtain and communicate any information held by public bodies, except exempted information therein. Under the proclamation, citizens have a right of “access, [to] receive and import information held by public bodies, subject to justifiable limits based on overriding public and private interests.

The mass media and freedom information law in article 13 (1) states that government institutions shall publish information concerning:

- its organizational structure, main duties and responsibilities;
- the power and responsibilities of the officials as well as decision-making procedures ;
- description of the services it provides for the public;
- brief descriptions of the complaint hearing mechanism available to the public and the public body's response to frequently asked questions by the public;
- a description of the type of the records its possession, a brief description of the contents of its records and the detailed explanation of the procedures to be followed by persons who wish to access this information;
- a description of its regulations, directives, policies, guidelines and manuals, which govern the operation and activities of its various organs, along with a description of any amendment or repeal of such provisions;
- its directives, regulations, guidelines and other documents which governs the activities of the employees of the organization;
- the name and address of the public relation officer; and

- other particulars.

The article does not cover aspect of the raw data that is a pre-requisite for open Government data. The absence of raw data from the list also means data is not “open by default.”

However article 13 (3) a states that a public body shall “publish all relevant facts concerning important decisions and policies that affect the public not later than the time at which such decisions and policies are announced” this implies that the data that is a basis for decision may as well need to be published in conjunction with the outcome of a given decision.

The law provides for protection of privacy of information. The exempted information from disclosure is personal information. The exempted information include:

- Information relating to third party
- Commercial information of a third party
- Protection of confidential information of third party
- Third party notification and intervention
- Protection of safety of individuals and property
- Protection of proceedings of law enforcement and legal investigation
- Protection of records privileged from production in legal proceedings
- Defence, security and international relation
- Cabinet documents
- Economic interests and financial welfare of the country and commercial activities of public bodies
- Operations of public bodies
- Requests that are too broad, likely to divert resources or pertaining to information to be published in the future

The survey shows that there is a general apprehension that data may be misused and misreported, if available in a raw form. There are also fears that business or citizens may use raw data irresponsibly and an incorrect analysis of data may lead to misconceptions and may hurt business interests or

larger public interests. Hence, many data suppliers are cautious in releasing any data in the public domain.

The value of the open data can only be realized when public sector raw data is available freely for reuse through proactive disclosure. Proactive disclosure ensures that information seekers get immediate access to public information and avoid the costs of filing a request or engaging in administrative procedures.

In sum, while the Mass Media and the Freedom of Information proclamation provides for overall guarantee for access to information and privacy, it does not provide for publications of raw data and electronic information. The law is a vehicle for reactive disclosure where a question/request has to be filed before an answer is given and that public information must be requested before it is disclosed.

A subsidiary legislation that builds on the Mass Media and Freedom of Information law needs to be enacted in order to address access to electronic and raw data and provide for the proactive release of public information before individuals request it. The law will update and improve access to information and specify that new data sets and records to be collected and published.

### **Other Legislations that Provide for Open Data**

The Government of the Federal Democratic Republic of Ethiopia has initiatives towards enacting laws and legislations to promote access to online information. This includes the National Data Protection Law, e-commerce law, computer misuses and cyber crime law and the e-signature law. It has also been working towards the establishment of a national Public Key Infrastructure (PKI) framework. The government has also adopted an e-government enterprise architecture and interoperability framework. The enactment of these laws will provide for availability of safeguards for data in open format including the protection of privacy and security of institutions and individuals.

It is evident from the survey that respondent did not know the existence of these laws and their implications to open data in Ethiopia.

## Infrastructure Policies that Encourage Access to Open Data

There is no broadband master plan and universal access policy in Ethiopia. However, the Ministry of Communications and Information Technology has made a deliberate effort to encourage the operator the Ethio Telecom to expand access to broadband services. Ethio telecom has over 16,000 Km fiber backbone that extends throughout the country. Effort has also been made to extend access to broadband mobile services that is now readily available in major towns. This has encouraged access to the Internet. There are about 2.5 million Internet users of which about 1.5 million are face book users. This implies, at least about 2% of the population, that represents the core user of open data can access to open data.

Increase access to the Internet will encourage further use of open data. The figures shows that the government needs to promote access to intermediary institutions like schools, universities, libraries and kebeles.

## Commitment to Open Access

The survey finds that ministries department and agencies do not use any license when making data available to the public information -even where they make data available in machine-readable format. Creative commons is largely unknown; therefore all ministries present information on their websites under full copyright protection. A review of the web sites indicates that there is no mention of free distribution of information on the web sites in any media and on the Internet without any conditions except citation of source. The extent to which “open” licenses conform to the Ethiopian legal framework including copyright laws is unclear, and there is no provision for creative commons to date.

## Privacy of Online Data

Privacy establishes the basic protections for personal information. Processing personal data usually requires prior consent of individual whose data will be processed. Consent is specially important when information is collected in digital format and can easily be copied and exchanged. The review indicates that Ethiopia has a more comprehensive privacy safeguard that is introduced by the constitution. The constitution protects privacy of persons, their home and correspondences, which

is informed by the privacy provisions of the UDHR and the International Covenant on Civil Political Rights (ICCPR) to which Ethiopia is state party.

According to the constitution (1) Everyone has the right to privacy. This right shall include the right not to be subjected to searches of his home, person or property, or the seizure of any property under his personal possession. (2) Everyone has the right to inviolability of his notes and correspondence including postal letters, and communications made by means of telephone, telecommunications and electronic devices. (3) Public officials shall respect and protect these rights.

The Constitution requires public officials not only to refrain themselves from interferences with individual privacy, but also to prevent private persons or entities that would impair the right. Limitation to the right to privacy is allowed only when three important elements are satisfied together: (1) there must be compelling circumstances; (2) restriction must be in accordance with specific laws whose purposes shall be the safeguarding of national security or public peace, the prevention of crimes or the protection of health, public morality or the rights and freedoms of others; and (3) there must be legitimate aims.

Although the constitution provides safeguards for privacy, and a clear understanding of the importance of privacy, it is not clear how institutions protect the privacy of individuals in the online environment.

The review has also shown that there is limited knowledge on the anonymization of data (stripping off personal data from row data). Anonymisation is regarded as a good strategy to keep the benefits of open data and mitigating the risks. Once a dataset is truly anonymised and individuals are no longer identifiable, although the creation of a truly anonymous dataset from a rich set of personal data, whilst retaining as much of the underlying information as required for the task, is not that simple.

### **Charging for Data**

There is no established system for charging fees for online information. The National Metrological Agency charges nominal fees for its data. Almost all users are expected to fill in the form in person

and pay the fee at the agency. The absence of an online payment system has also been the main constraint.

The Mass Media and Freedom of Information proclamation sets the conditions for fees to access to information, although this was not meant for print information. Article 14 (12) states that “ fees payable for access to information ...shall not exceed the actual cost of searching, collecting and duplicating the record containing the information provided that no fee shall be charged for requests of personal information or for requests from those who cannot afford to pay.”

However, there is no a general provision for online purchase of data. This needs to be established through a new law that establishes the legal framework for online data.

### **Arrangement with the Private Sector**

The survey finds that there is no specific arrangement to work with private sector in the area of open data. There is no application that has been developed to improve the visualization and use of open data.

### **Gap Analysis**

The country survey shows that Ethiopia is at the early stages of development of open data. The fear for releasing data in open reusable format abounds. The current data is largely aggregated and only available through individual institutions. Data is not harmonized and has not been presented in open standard format. There is a mismatch between data supply and demand. The current data sets are not commensurate with the needs of the population. The survey also show that no dataset is openly licensed, and there is limited understanding of what open licenses entails.

It is imperative to migrate from a few scattered data supply approach to a more balanced comprehensive data demand perspectives that would benefit from reuse and value creation. This requires widening data offer from government institutions based on users needs in areas such as health, education, agriculture, environmental and geographic data.



These country experiences suggest that Ethiopia can move ahead with its open data by consolidating and publishing existing open data sets through data.gov.et. In parallel, it needs to launch a promotional scheme to increase data supply in priority areas like agriculture, health, education, economy and financial management, geographic information, law and others.

This can be achieved through mandatory policy requiring government departments and agencies to maximize the release of data and information of business value subject to applicable restrictions related to privacy, confidentiality, and security. Eligible data and information should be released in standardized, open formats, free of charge, and without restrictions on reuse.

In addition the government needs to work on capacity building initiatives that improve data availability, quality and meta data. This demands training of data suppliers in government open licensing techniques and presentation of data in open machine-readable format. It also involves training of decision makers on data-driven analysis and arguments.

Ethiopia has progressive information access policies, when compared to other African countries. Thus, it can leverage of existing polices and legal frameworks to promote open data. It needs to complement open data policies with existing legislation for publication of information. Building on the basis of these precedents can help to strengthen open data policies and identify new niches where such policies can be developed. This can greatly reduces the legal/policy impediments and lead times.

At the same time, it is important to identify at an early stage actual or perceived “blockers” in order that policy or legal change can be initiated early. The current policies, legislative frameworks like the National Data Protection Law, e-commerce law, computer misuses and cyber crime law and the e-signature law can be revised for inclusion of clauses that cater for open data.

The government needs to establish a legal framework for proactive online disclosure of public data. A subsidiary legislation that builds on the Mass Media and Freedom of Information Law needs to be enacted in order to addresses electronic access to raw data and provide for the proactive release of public information before individuals requests it. The law will update and improve access to information and specify that new data sets and records to be collected and published. This would

eliminate legal barriers to access to online public information. The law would also include mandates on anonymous requests and regulation on fees and fee waivers.

In addition, Ethiopia needs to work on the requirements to manage risks of data security and privacy. Breaches of data privacy and security needs to be anticipated and managed based on international regulation. The data protection regulation needs to be enforced together with the accountability of entities managing public data.

The government needs to promote debate on open data laws. These will provide an opportunity to update and improve access to information that is already being published, and make it possible to specify those new datasets and records that should be published

The experience with Open Data in other countries has shown that demonstrating the possibilities for data-drive applications and promoting reuse among developers is important. At the same time, building the capacity of public officials to use data (from their own agency and others) and leverage new applications is crucial. This implies that Ethiopia needs to strength the capacities of the local ICT industry in order to develop application on open data.

The other important challenges identified on the course of this assessment include:

At the level of technical infrastructure:

- The availability of government data in open and machine-readable format and the anticipated high cost of digitization and data conversion to required format
- Although most data available in electronic format, it is not well structured, complete, and properly archived in a required format
- Ethiopian government metadata standard is still at the draft stage and there is no practice of using metadata standards for data catalogue
- Lack of proper IT infrastructure in some institutions. Challenges related to connectivity have been raised by almost all ministries/agencies covered by the survey
- Varied capacity and efficiency to process data among government ministries and agencies
- Inadequate data management staffing

At the institutional, administrative level

- Ambiguity in the legal environment regulations on what data to open and how to do it. Although the Freedom of the Mass Media and Access to Information Proclamation gives adequate legal framework to implement open data initiatives, detailed policy guidelines need to be developed for the proper implementation and management of Ethiopian Government open data.
- Lack of awareness on potentials and opportunities of implementing OGD in some ministries
- Lack of human resources and capacity to implement open data in most of the ministries
- Lack of proper collaboration mechanism among government ministries and agencies

At the level of demand

- Lack of culture of opening and consuming data
- Often data is mis-interpreted
- Limited access to technologies
- Low capacity to make use of open data (analysis, providing feedback, etc.)

## 2.4 Analysis of the Economic Impact of Open Data to Ethiopia

### Economic benefit as a driver of open public data

There are a number of political, ethical and economic incentives for opening data. The activities that the government undertakes and information that is produced in the course of these activities is a public good. Public sector information, which is a public good should be available for citizens at no cost because the public has already paid for production of such information during government activities.

The economic impact of open data is another driver for promoting access to data. The government needs to make its data available, because it supports economic growth, spurs innovation and promotes job creation. The economic value from increased transparency and accountability, and improvement in the efficiency and operations of public services is another incentive for opening government data. The direct and indirect economic benefits of opening data include:

- Public sector performance improvement,
- Development of new markets especially by information and data companies,
- Enterprises efficiency through open data,
- Efficiency gain at individual levels through access to open data, and
- Impact on research and innovation

The availability of open data will improve transparency, education and research and support personal decision making capabilities that in turn contribute to economic growth.

The government can accrue economic benefits from opening its own data. The benefits can be financial, for instance in terms of effective policy making, resource allocation or operational management and optimisation; or they may be non-financial, for instance in meeting legislative and other requirements for transparency and accountability. The opening of data attracts new types of re-users and enterprises that add value to it and resell products and services that depend on open data. Individual benefits from open data can also be high. For example, open cadastral information

contributes to authoritative system of land ownership and proof of title, which has a direct economic benefit for individuals.

It is important to note that access to data by itself may not bring the desired benefits. A raw data is a bare bone; it needs the meat that makes it live. A combination of data with innovative ideas and new ways of delivery including smart phones are important to expand the real benefits of open data.

### **Cost of Open data**

Costs that are incurred in opening data include consultancy for the development of open data portal (like this one), ongoing maintenance of the portal and information gathering, data extraction from internal databases and integration. Some of the open data costs are one-off expenditures and the amount of these is often small compared to the on-going costs for collection and administration of data. The cost of collecting and administering of data can be cut by encouraging decentralization of management and collection of data at its source.

### **Quantification of the economic benefit of open data**

A lot of data, once published in an open format, can appreciate economically and new value can arise from its re-use. However, it is difficult to quantify the exact benefits of open data. Firstly, it is not easy to separate data-dependent elements from the rest of information economy. The limited availability and quantity of information and inadequate accounting of the usage of information at institutional levels poses a serious challenge to quantification.

Besides, the benefits of open data take time to emerge. Innovators, enterprises or decision makers may act as a result of spotting an opportunity and then seek the data needed to build service or system to address it. Furthermore, in the cases of developing countries like Ethiopia, it is difficult to carryout cost-benefit analysis for open data that does not exist.

Current economic analysis of open data relies on benchmarking the experiences of developing countries, where open data exists and the use of open data is wide spread. Most of the economic studies measure the overall contribution of open data to the GDP or specific sectors. A study by Graham Vickery for example, shows that the direct and indirect economic benefits of open data for

the European Union' Economy could be of order of 200 billion Euro in 2008 - 1.7% of the GDP of the Union. A 2013 in the United Kingdom indicates that the over all direct and indirect impact of open data could be around £6.8 billion, with direct impact being estimated to be £1.8 billion. In the UK case, the benefit of open data accounts for about 0.4% of the economy in 2013.

Factoring the low level of development, the indirect benefit of open data can be around 0.1% of the GDP in Ethiopia at the beginning. This is very high compared to the effort that is needed to make the data available. The direct and indirect benefit will increase dramatically, once the data become accessible and re-usable. It is expected that the indirect and direct benefits will be to raise to about 1% of the GDP or more within five to ten years.

### **Stimulating the economic benefit of open data**

Ethiopia is at the early stages of open data development. The economic impact of open data can only be achieved through concerted effort that increases the quality and quantity of data sets available for access and re-use. Prioritization of the release of valuable data sets can also stimulate the economic benefits of data. Widely used data sets such as weather data, road and transport data, education, health, official registers, address data and cadaster can stimulate usage and the economic impact of open data.

At the same time, there is a need for improving the availability of open data through different platforms including mobile handsets. Multiple platforms will increase usage of open data by enterprises and citizens.

Furthermore, there is a need for marketing of available data sources raising awareness of the availability of open data. It is also important to convince developers to use open data and build applications that stimulate use.

## 2.5 Analysis of Ethiopian Government Open Data

Availability of government generated data and accessibility to the public is the basis of any open government data initiative. In order for stakeholders outside of government to be able to use, re-use and disseminate the data it first needs to be collected, stored and managed by government institutions – ideally in digital form and through a standardized process. Making government data available irrespective of quality, form, or potential for reusability does not serve its purpose. For its re-use value to emerge, the data needs to be of sufficiently high quality, and legal and technical gaps discussed in the previous section to access and use the data have to be bridged.

### Data Availability

In Ethiopia, as shown in the graph below, almost all ministries and agencies have either dedicated data management units or in some cases the IT department is responsible for processing and managing data. Although most of the institutions claim that they have some data processing and management capacity, there is still a big gap that should be bridged in order to successfully implement the open data initiative in Ethiopia. The survey shows that considerable variations exist between government institutions in terms of their staffing and capacity to manage open government data in their institution.

The Ethiopian Statistics Agency (ESA) is the key provider of statistical data. The agency generates a large amount of data through its own portal, even though the data is aggregated. ESA has made preliminary efforts toward releasing large quantities of data including the results of 1994 census and the projections in aggregated format. The current proclamation to establish the central statistics authority, proclamation No. 442/2005 provides for disclosure of information and access to reports and publications is not constrained. Hard and soft copies of reports are sold on a cost recovery basis. Basic results of the surveys are also available on the web. However, the law does not provide for raw data. Access to the micro data is administered by a data release protocol, which describes the principles and costs of users accessing the unit record information.

The survey indicates that although efforts are made by many organizations to provide access to data, these do not use common standards. Much of the data has not been updated with the latest data

series, hampering sustainable and on going data re-use. There is much work to do to increase the availability of open data.

An additional problem is lack of engagement of important stakeholders in the process of developing and managing open data. As the graph below shows, only 10% of the surveyed organization claimed to have engaged stakeholders in their selection of data to make publicly available. One of the eight principles of Open Government Data clearly states the importance of public engagement during the process of opening up government data. “The public is in the best position to determine what information technologies will be best suited for the applications the public intends to create for itself. Public input is therefore crucial to disseminating information in such a way that it has value.”<sup>1</sup> The survey finds that there is no specific arrangement to work with private sector in the area of open data. There is no application that has been developed to improve the visualization and use of open data.

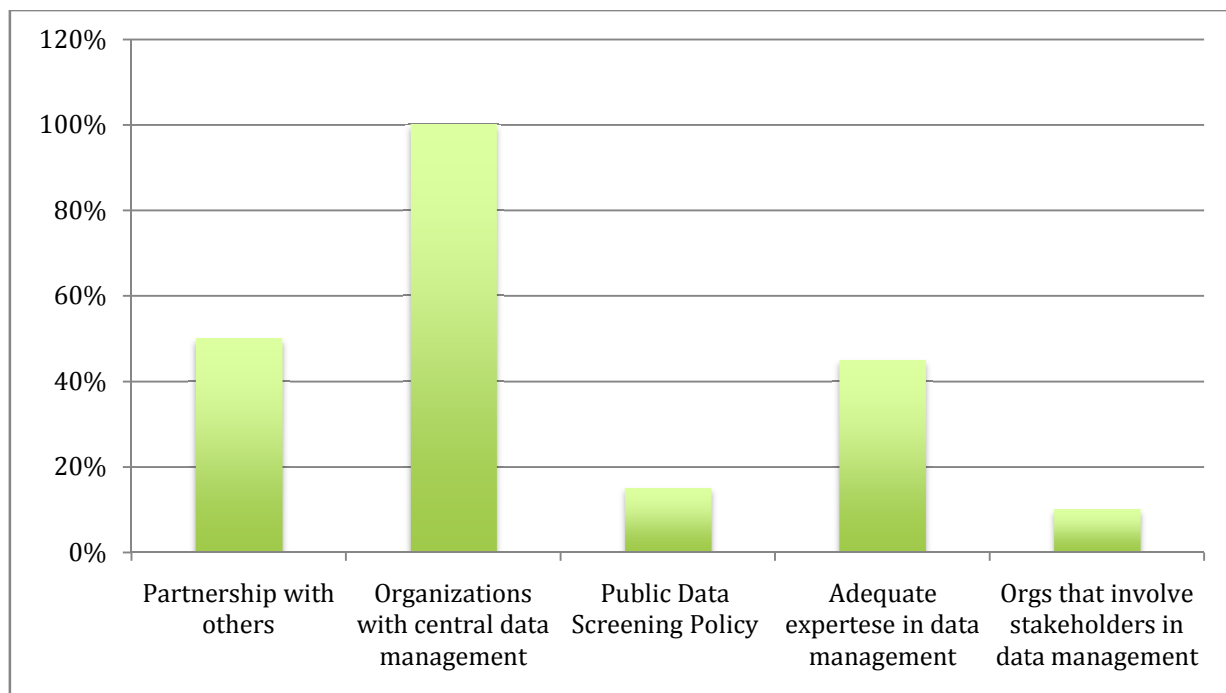


Figure 5. Organizational Analysis

During the survey, several ministries and agencies acknowledged that they have considerable quantities of not well-shared datasets. It was noted that despite best efforts of the

<sup>1</sup>The Annotated 8 Principles of Open Government Data. <http://opengovdata.org>



ministries/agencies, public availability of data still remains a significant challenge. The survey results also support this view. As shown in the graph below, although more than 85% of government ministries/agencies provide electronic access to selected few data through their portal, almost 98% of the ministries use offline media such as CD-ROMs and flash drive to provide access to data upon request. However, access through mobile devices, APIs, and direct downloads is available only in <20% of the surveyed ministries and agencies.

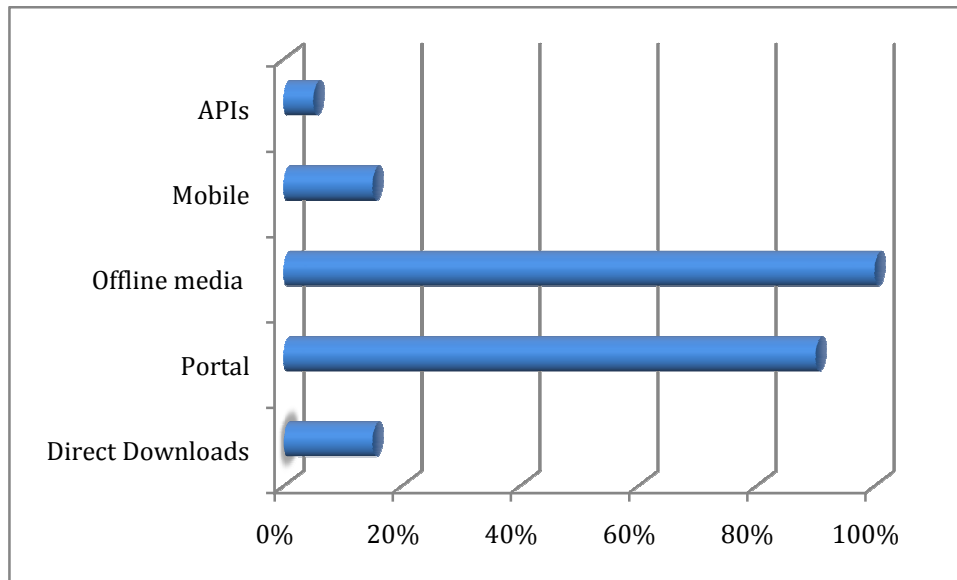


Figure 6. Organizations that Provide Electronic Access to Data Via Different Media

Yet, it is important to point out that there are several initiatives by Ethiopian government ministries and agencies to open up their data. For example, several sources of useful and reliable data, such as the Population and Housing Census Report data from the Central Statistics Agency and the Daily and Seasonal weather forecasts from Ethiopian Meteorological Agency, which are two of the most progressive institutions in terms of openness and data provision, are made available to the public through their organizational portal. There are also similar initiatives in other governmental organizations and their data can be found in several places. Some examples are the annual abstracts that provide detail information on different educational variables of the last 15 years, which are available from the ministry's web portal. Additionally, data related to import and exports are available in the ministry of trade portal.

There is no established system for charging fees. The National Metrological Agency charges nominal fees for its data. Almost all users are expected to fill in the form in person and pay the fee at the agency. Thus there is no provision for online purchase of data.

From the 20 surveyed government ministries and agencies, a total of about 84 datasets are identified as data that can be available for the public. The table below shows summary of the datasets and their description by Ministry/Agency.

**Table 1. Ethiopian Government Datasets by Ministry/Agency**

<b>Ministry/Agency</b>	<b>Dataset</b>
Ministry of Finance and Economic Development	<ul style="list-style-type: none"> <li>• National budget, cash flow, disbursement and related data</li> <li>• Civil service payroll data</li> <li>• National economic and social development including plans, projects, programs, indicators and achievements</li> <li>• List and profiles of public budget organizations. Budget requests, budget ceilings, budget utilization performance</li> <li>• List and profile of internal and external auditors and auditors and audit reports.</li> <li>• Data on accounts of ongoing and completed projects</li> <li>• Data on federal budget assistance to regions and to wards</li> <li>• Registration and analysis of bank statement and budget ceiling for federal payment</li> </ul>
Ministry of Education	<ul style="list-style-type: none"> <li>• Student enrolment, dropout, repetitions</li> <li>• Teachers, principals and supervisors</li> <li>• Buildings and teaching leasing materials</li> <li>• School infrastructure</li> </ul>

	<ul style="list-style-type: none"> <li>• Financial data/revenue and expenditure</li> </ul>
Ministry of Health	<ul style="list-style-type: none"> <li>• Health and health related indicators</li> </ul>
Ministry of Transport	<ul style="list-style-type: none"> <li>• Policies and strategies</li> <li>• Proclamations</li> <li>• Regulations</li> </ul>
Ministry of Trade	<ul style="list-style-type: none"> <li>• Trade Registration</li> <li>• Trade License</li> <li>• Company Name</li> <li>• Trade Name</li> </ul>
Ministry of Communication and Information Technology	<ul style="list-style-type: none"> <li>• Telecom and IT statistical abstract</li> </ul>
Ministry of Agriculture	<ul style="list-style-type: none"> <li>• Land Leased</li> <li>• Investment incentive</li> <li>• List of registered pesticides</li> <li>• List of registered fertilizers</li> </ul>
Ethiopian Metrological Agency	<ul style="list-style-type: none"> <li>• Rainfall (daily, decade, monthly, seasonal, annual, rainfall intensity, etc....)</li> <li>• Temperature (daily, decae, monthly, seasonal, annual, soil temperature, etc....)</li> <li>• Relative humidity</li> <li>• Wind direction and speed</li> <li>• Pressure</li> <li>• Evaporation</li> <li>• Global radiation</li> </ul>
Ethiopian Mapping Agency	<ul style="list-style-type: none"> <li>• Thematic map</li> <li>• Topographic map</li> <li>• Areal photographs</li> </ul>
Ethiopian Anti-Corruption Commission	<ul style="list-style-type: none"> <li>• Tips</li> <li>• Investigated tips off.</li> </ul>

	<ul style="list-style-type: none"> <li>• Prosecution and court decision</li> <li>• Asset recovery</li> <li>• Corruption prevention</li> <li>• Ethics education and communication</li> <li>• Coordination of ethics infrastructure</li> </ul>
Ethiopian Chamber of Commerce	<ul style="list-style-type: none"> <li>• Membership data</li> <li>• Import/export data (by commodities, by destination)</li> <li>• Investment (by commodity, sector, region, and investor)</li> <li>• Business related proclamations</li> </ul>
Ethiopian Revenue and Customs Authority	<ul style="list-style-type: none"> <li>• Import/export dataset</li> <li>• Annual report</li> <li>• Proclamations</li> <li>• Regulations</li> <li>• Directives</li> </ul>
Ethiopian Commodity Exchange Authority	<ul style="list-style-type: none"> <li>• Price information of coffee, sesame etc</li> <li>• Import and export information</li> <li>• Exchange/commodity status and historical background</li> <li>• ECX regulatory results /exchange violation and manipulation etc</li> <li>• Exchange regulation, rules and standards</li> <li>• Trade rules regulation and standards</li> </ul>
Document Authentication and Registration Office	<ul style="list-style-type: none"> <li>• No public dataset</li> </ul>
Central Statistics Agency	<ul style="list-style-type: none"> <li>• Census in different areas</li> </ul>
Parliament	<ul style="list-style-type: none"> <li>• Bill</li> <li>• Proclamation</li> <li>• Regulation</li> <li>• Directive</li> </ul>
National Electoral Board	<ul style="list-style-type: none"> <li>• Number of voters</li> </ul>

	<ul style="list-style-type: none"> <li>• Number of candidate</li> <li>• Number of political parties' contesting</li> <li>• Number of winner candidate and parties</li> </ul>
Supreme Court	<ul style="list-style-type: none"> <li>• Case status</li> <li>• Daily court list</li> <li>• Report (#of cases opened, #of cases closed, decisions by type)</li> <li>• Cassation decision</li> </ul>
National Bank of Ethiopia	<ul style="list-style-type: none"> <li>• Daily exchange rate</li> <li>• Commercial banks rates for major currencies against birr</li> <li>• International commodity price</li> <li>• Gold purchasing rate</li> <li>• Annual and quarterly bulletins, monthly macroeconomic statics</li> <li>• Directive, circulars, proclamation, regulation, guideline</li> <li>• Foreign exchange rate</li> </ul>
Agricultural Transformation Agency	<ul style="list-style-type: none"> <li>• Agricultural best practices</li> <li>• Agricultural inputs (fertilizers and seeds)</li> <li>• Agricultural planning data</li> <li>• Soil information</li> <li>• Fertilizer recommendation</li> </ul>

### Quality of Ethiopian Government Data

Data quality is one of the biggest challenges for countries that start open data initiatives. In Ethiopia, issues related to quality of data are well identified with clear evidence from survey results as well as from interviews. The most common data quality issues identified by the survey are irregularities in datasets, the use of different standards or no standard at all, and a high level of aggregation. In addition, data is at times incomplete, inaccurate, not frequently updated, and difficult to process and understand. It is also noted that machine-readable formats are few and far between.

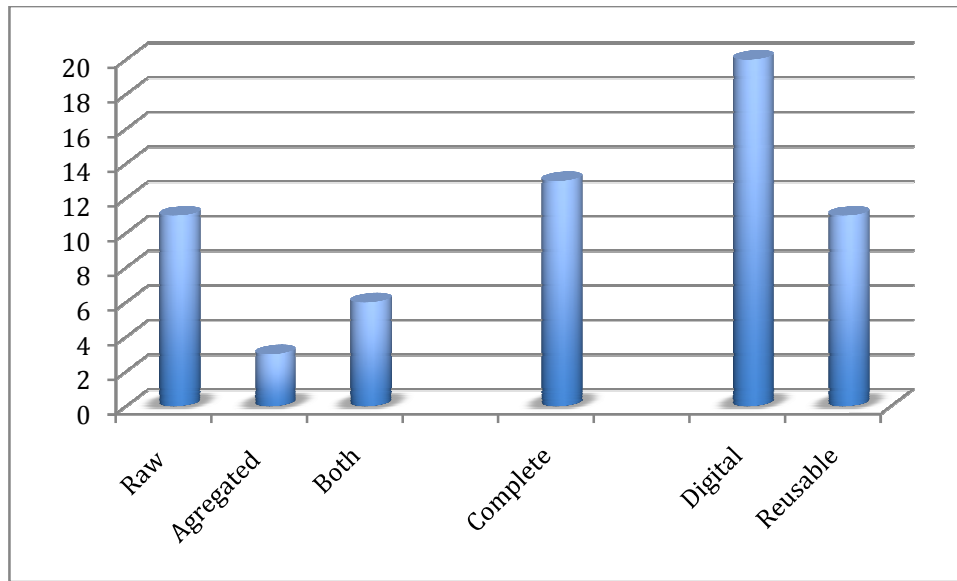


Figure 7. Available Data Formats

As the above graph shows, most of the Ethiopian government data are in digital format but a big proportion of the data is not accessible in reusable format. More than half of the government organizations have made raw data available to the public. However, there is still big portion of data available in only aggregated format.

Interviewees cited two main factors to explain most of the quality-related issues. First, data collection and management systems are largely inefficient and not linked. Second, a general lack of capacity and skills in terms of data processing and management exists in a number of ministries and agencies. 65% of the government ministries and agencies believe that the dataset made available to the public are complete and in good quality. However, sample datasets accessed by the assessment team show some of the data in government ministries and agencies are incomplete and not frequently updated.

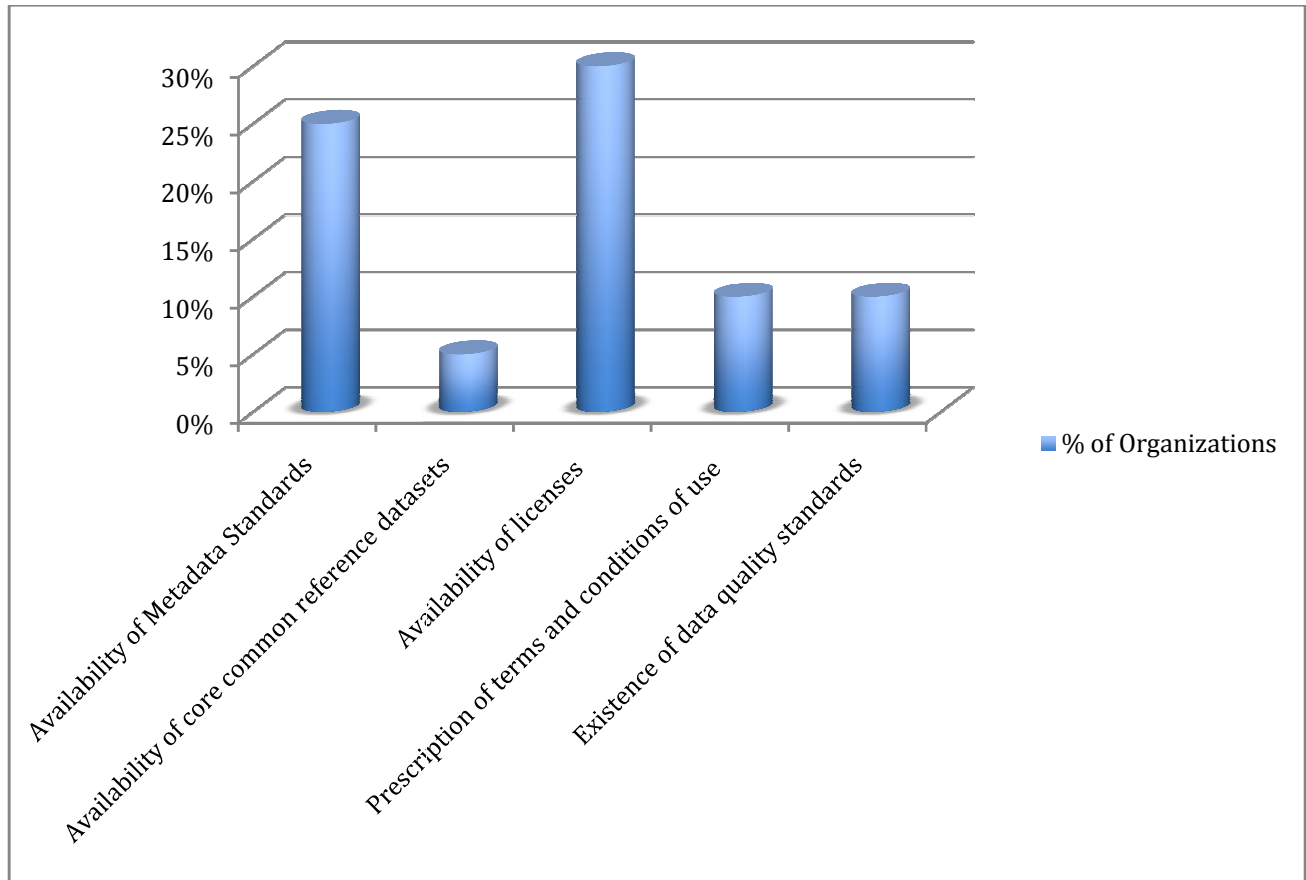


Figure 8. Analysis of Use of Standards and Other Data Quality Tools

### Ease of Access and Re-Use

Even though the Freedom of the Mass Media and Access to Information Proclamation has been in place for six years and while some ministries and agencies have made data available online, it is often difficult to obtain information and data from government institutions.

In most of the government ministries/agencies access to data requires written approval at the higher level. The data access approval process is very bureaucratic that it often requires official documents to be submitted and in person appearance at government offices, presentation of reference letters, and meeting various public officials etc.

Another barrier to accessing data is lack of clear policy to screening sensitive data. Among the surveyed ministries and agencies, only 16% have public data screening policy or mechanisms. The

rest of the organizations depend on the case-by-case decisions of higher officials to screen public data.

Issues related to copyright and licensing are also identified as barriers for many government institutions to putting data on their websites. Only 28% of the Ethiopian government ministries/agencies claimed to have different degrees of access when they provide data. For example, the Ethiopian Mapping Authority has different categories of access based on sensitivity of data.

- Restricted data – access is provided for the data by signing a pledge form to not pass the data to third party
- Confidential data – not public. Only selected few government organizations have access for this type of data
- Data in the public domain – accessible to the public freely or fee based, depending on the nature of the data

Others also have relatively similar categories of licenses for usage of government data. However, the majority of the Ethiopian government ministries/agencies do not have any well-defined licensing mechanism when they provide public data. Only 8% of the Ethiopian government ministries/agencies clearly prescribe terms and conditions of use of government data when they provide access.

## Data Formats

It is recommended that data has to be published in open format and it should be machine-readable.

Though there are many formats suitable to different category of data.

The survey indicated that there is no national standard on formats and procedures for government internal data processing and there is no common framework that is followed across government ministries and agencies. As the table in Annex-D shows there is a big variation in the format(s) of existing data currently, the level of granularity, and frequency of data updates.

The following graph shows the status of quality of Ethiopian government data as per Tim Berners-Lee's classification of data into single star to five star categories based on formats of data.



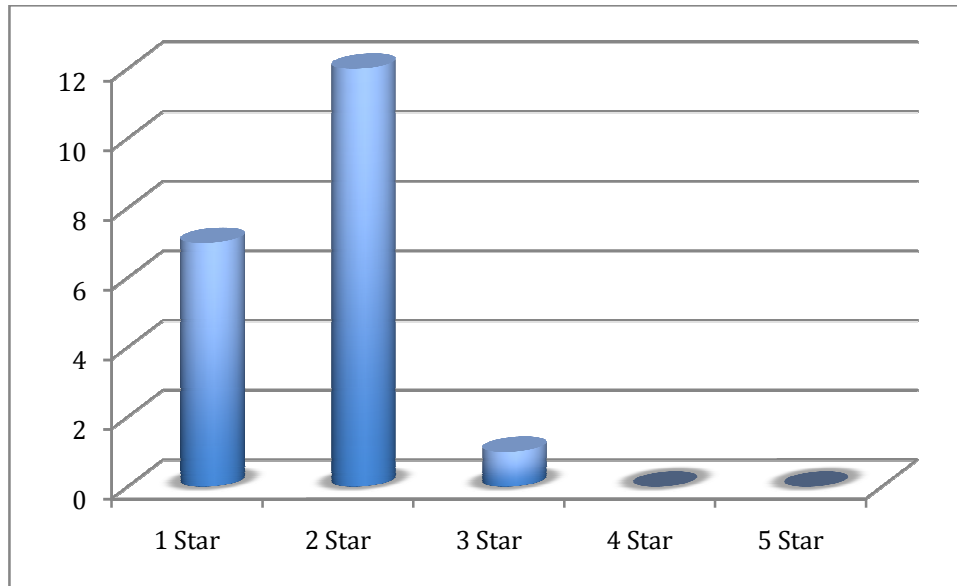


Figure 9. Data Quality Analysis

The stars denotes as follows. <sup>2</sup>

- ★ make your stuff available on the Web (whatever format) under an open license
- ★★ make it available as structured data (e.g., Excel instead of image scan of a table)
- ★★★ use non-proprietary formats (e.g., CSV instead of Excel)
- ★★★★ use URIs to denote things
- ★★★★★ link your data to other data to provide context

The above graph shows almost all government ministries and agencies, except CSA, fall under one and two Star levels. Data at the one star level allow users to easily look, print, store locally, and share the data and it is easier to publish whereas data at star two allows additional feature of directly processing it with proprietary software to aggregate it, perform calculations, visualize it, etc. The three star format data, which is available at CSA, allows manipulation of the data in any format without being confined by the capabilities of any particular software.

<sup>2</sup><http://5stardata.info>

However, all the datasets held by the Ethiopian government ministries and agencies do not receive the benefits of 4 and 5 star web data, such as data reuse, ability to reuse existing tools and libraries, even if they only understand parts of the pattern the publisher used, combine the data safely with other data, access to granular control over the data items and can optimize their access (load balancing, caching, etc.), can allow other data publishers to link into your data.

Therefore, it is important to develop a data format standard and enforce it across ministries. Government ministries and agencies shall strive to promote their datasets to 4 and 5 star by building their capacity and control the quality of their data. For inclusion in the Open Data Portal, public bodies must publish data at a minimum of 3 Star Open Data, such as CSV, JSON or XML. However it is encouraged to publish datasets in multiple formats, for example, 1 Star (e.g. PDF), 2 Star (e.g. Microsoft Excel) in addition to the required 3 Star (e.g. CSV).

### Recommendations

- Based on current analysis of data formats prevalent in Ethiopian Government, it is proposed that data should be published in any of the following formats:

Table 2. Recommended Dataset Formats

General	Geospatial
CSV	GeoJSON
JSON	GML
XML	KML
ODF	WKT
RDF	LAS
	IFC
	Shapefile
	ASCII Grid

Please note that this list is subject to review and updating as new formats are developed due to technological developments.

- It is also possible to publish datasets in multiple formats.
- Datasets not yet available in one of the recommended open formats should have a clear timeframe when it will be available in an open format. The publication of data in open formats should be built into data publication processes of all government ministries and agencies.

### **Metadata for Ethiopian Government Open Data**

Metadata is structured information that describes, explains, locates, or otherwise makes it easier to retrieve, use, or manage an information resource. Metadata can range from basic to advanced, from allowing one to discover the mere fact that a certain data asset exists and is about a general subject all the way to providing detailed information documenting the structure, processing history, quality, relationships, and other properties of a dataset. Making metadata machine-readable greatly increases its utility, but requires more detailed standardization, defining not only field names, but also how information is encoded in the metadata fields.

In order to help realize the benefits of Open Data, Ethiopian government ministries/agencies should make their data more searchable and usable. To achieve this, ministries and agencies should provide precise descriptors about their datasets to help in the identification, location and retrieval of online resources by data-users. Besides facilitating easy access to datasets, this shall be extremely useful in the future for federation/integration of data catalogues.

Typically data publishers provide some form of metadata to describe their datasets. Currently there are a wide variety of forms and formats for such metadata, with varying degrees of machine readability. To support easier discovery and use of open data as the number of datasets and number of open data catalogues expands, interoperable machine-readable metadata becomes steadily more important. Interoperability tends to require some degree of standardization.

There are a number of metadata vocabularies in common use: Dublin Core, VoiD, DCAT. More recently schema.org has developed RDF-compatible vocabularies that provide a number of relevant

terms and may be of particular benefit in assisting generic search engines to find and correctly categorize the data.

Countries use different metadata standards. Although there are variations as well as pros and cons of using a certain vocabulary, there is also commonality on the metadata features. The table below summarizes metadata features of open data portals from few countries.

Table 3. Analysis of Metadata features used by different countries<sup>3</sup>

	General												Categorization		Data Access			Others	Characteristics	Data Collection	
	Title	Description	Publisher	Frequency	Release date	Update date	coverage	coverage	License	Data Dictionary	Granularity	Metadata update	Themes	Tags/Keywords	Dataset URL	Format	Size	Citations			
USA	Y	Y	Y	Y	Y	Y	Y	Y	N	Y	Y	N	Y	Y	Y	Y	Y	Y	Y	Y	
UK	Y	Y	Y	Y	Y	Y	Y	Y	Y	N	Y	N	Y	Y	Y	Y	N	Y	Y	N	
New Zealand	Y	Y	Y	N	Y	N	N	N	Y	N	N	N	Y	Y	Y	Y	N	N	N	N	
Australia	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	N	Y
Ireland	Y	Y	Y	Y	Y	Y	Y	Y	N	Y	Y	Y	Y	Y	Y	Y	N	Y	Y	Y	

<sup>3</sup><https://hal.archives-ouvertes.fr/hal-01056576/document>

## Dataset Selection

For the purposes of prioritizing public datasets, Ethiopian government ministries and agencies should consider whether information embodied in the public dataset:

- Increases the ministry/agency’s accountability and responsiveness;
- Improves public knowledge of the ministry/agency’s and its operations;
- Responds to a need or demand identified by the public;
- Furthers the mission of the ministry/agency;
- Fosters ministerial/inter-ministerial efficiency; or
- Creates economic opportunity

All the datasets identified by government ministries and agencies are important for the public, including businesses. However, it is possible to phase out the datasets based on their formats. It will be easier to start with datasets in two star and above and gradually convert one-star datasets to machine-readable format.

## Recommendations

- Follow a phased approach in uploading the data. Start by targeting two and three star datasets for the launching of the portal
- Build capacity of Ethiopian government ministries and agencies to retrospectively convert their data to reusable formats

## Data Catalogue

In order to develop a data catalogue for the Ethiopian government open data portal, the assessment team analyzed portals of different developed and developing countries. Table 6 summarizes the various browsing features available in different portals by country. The more features are utilized for indexing the discoverability of datasets will increase.

Table 4. *Analysis of Cataloguing/ browsing features used by different countries*

Country	India	UK	Australia	Kenya	Ghana
Catalogue	<ul style="list-style-type: none"> <li>• Ministry/ Agency</li> <li>• State</li> <li>• Sector</li> <li>• File Format</li> <li>• Resource Category</li> <li>• Frequency</li> <li>• Keyword</li> </ul>	<ul style="list-style-type: none"> <li>• Publication status</li> <li>• License</li> <li>• Themes</li> <li>• Resource format</li> <li>• Publisher</li> <li>• Openness score</li> </ul>	<ul style="list-style-type: none"> <li>Organization</li> <li>Jurisdictions</li> <li>Groups</li> <li>Tags</li> <li>Formats</li> <li>Licenses</li> <li>Locations</li> </ul>	<ul style="list-style-type: none"> <li>• Categories</li> <li>• Topics</li> </ul>	<ul style="list-style-type: none"> <li>• Ministry/ Department</li> <li>• Resource Format</li> </ul>

The sector or theme classification system of different countries has also been analysed. The table below shows the different variations in the classification of the themes and the commonalities have been identified.

Table 5. *Analysis of Themes/ Sectors used by different countries*

Country	India	UK	Australia	Germany	Ghana
Themes/ Sectors	<ul style="list-style-type: none"> <li>• Water Resources</li> <li>• Agriculture</li> <li>• Health and Family welfare</li> <li>• Governance and Administration</li> <li>• Parliament</li> <li>• Finance</li> <li>• Education</li> <li>• Transport</li> <li>• Home Affairs and</li> </ul>	<ul style="list-style-type: none"> <li>• Environment</li> <li>• Government</li> <li>• Location</li> <li>• Health</li> <li>• Society</li> <li>• Transportation</li> <li>• Education</li> <li>• Finance</li> <li>• Administration</li> <li>• Policy</li> <li>• Spending data</li> <li>• Linked data</li> </ul>	<ul style="list-style-type: none"> <li>• Community Services</li> <li>• Business Support and Regulation</li> <li>• Science</li> <li>• Environment</li> <li>• Sport and Recreation</li> <li>• Finance Management</li> <li>• Health Care</li> <li>• Civic</li> </ul>	<ul style="list-style-type: none"> <li>• Agriculture</li> <li>• Counties</li> <li>• Economy and Finance</li> <li>• Education</li> <li>• Energy</li> <li>• Environment</li> <li>• Financial Sector</li> <li>• Governance</li> <li>• Government Accounts (National &amp;</li> </ul>	No classification by theme

	<p>Enforcement</p> <ul style="list-style-type: none"> <li>• Power and Energy</li> <li>• Environment and Forest</li> <li>• Industries</li> <li>• Census and Surveys</li> <li>• Urban</li> <li>• Information and Communications</li> <li>• Rural</li> <li>• Labor and Employment</li> <li>• Statistics Science and Technology</li> <li>• Infrastructure</li> <li>• Commerce</li> <li>• Social Development</li> <li>• Economy</li> <li>• Biotechnology</li> </ul>	<ul style="list-style-type: none"> <li>• Defence</li> </ul>	<p>Infrastructure</p> <ul style="list-style-type: none"> <li>• Cultural Affairs</li> <li>• Communications</li> <li>• Transport</li> <li>• Employment</li> <li>• Education and Training</li> <li>• GovHack</li> <li>• Governance</li> <li>• Government</li> <li>• Emergencies</li> <li>• Tourism</li> <li>• Society</li> <li>• Indigenous Affairs</li> <li>• Safety</li> <li>• Primary Industries</li> <li>• News</li> <li>• Law</li> <li>• Geography</li> <li>• Technology</li> <li>• Planning</li> <li>• Natural Resources</li> <li>• Information-communication Technology</li> <li>• Immigration</li> </ul>	<p>County)</p> <ul style="list-style-type: none"> <li>• Health and Social Data</li> <li>• Infrastructure</li> <li>• Population</li> <li>• Water and Sanitation</li> </ul>	
--	--	---	---	---	--



### *Recommendations*

- The Ethiopian government open data portal should provide as many options as possible for browsing of the datasets. Accordingly, the datasets should be indexed using the following categories:
  - By ministry/Agency
  - By regional state
  - By sector/theme
  - By file Format (CSV, JSON, XML, ODF, RDF, etc...)
  - By resource Category (dataset or app)
  - By data frequency (Daily, weekly, monthly, quarterly, annual...)
  - By keyword
- The following categories shall be used for ministry/agency and there should be flexibility to modify if there is a change in the name and/or structure of government ministries/agencies.

#### *Legislators*

- House of Peoples Representative
- House of Federation

#### *Ministries*

- Ministry of Foreign Affairs
- Ministry of Education
- Ministry of Health
- Ministry of Trade
- Ministry of Finance and Economic Development
- Ministry of Culture and Tourism
- Ministry of Agriculture
- Ministry of Defence
- Ministry of Communication and Information Technology
- Ministry of Justice
- Ministry of Federal Affairs
- Ministry of Construction and Urban Development
- Ministry of Transport
- Minister of Revenues

- Ministry of Water Resource
- Ministry of Mines and Energy
- Ministry of Labour and Social Affairs
- Ministry of Youth, Sport
- Ministry of Women's Affairs

*Agencies and Authorities*

- Disaster Prevention and Preparedness Agency
- Ethiopian Mapping Authority
- Privatization and Public Enterprises Supervising Authority
- Social Security Agency
- Ethiopian Insurance Corporation
- Ethiopian Science and Technology Commission
- Ethiopian Telecommunication Agency
- Micro and Small Enterprises Development Agency
- Maritime and Transit Service Enterprise
- National Bank of Ethiopia
- Commercial Bank of Ethiopia
- Transport Construction Design Enterprise
- Ethiopian Investment Agency
- Ethiopian Airlines
- FDRE Supreme Court
- Ethiopia Electric Power Corporation
- Ethiopian Revenues and Customs Authority
- National Archives and Library Agency
- Ethiopian Export Promotion Agency
- Environmental Protection Authority
- National Agriculture Input Authority

***Data Catalogue by Sector***

The following list of sectors is recommended as themes for cataloguing purposes. The sectors are identified after analysing the sector classification schemes of five countries from Europe, Australia, and Africa. The proposed scheme is customized to reflect the various sectors of the Ethiopian economy and it is hoped that it will

provide comprehensive list of sectors during classifying datasets. The selection of classification of sectors has also keeps in mind the relevance of each sector to the citizens as well as the government's emphasis in the socio-economic development of the country.

- Agriculture
- Commerce & Industries
- Communication & IT
- Defence
- Education
- Environment & Forest
- Finance
- Health
- Infrastructure
- Power and Energy
- Rural Development
- Science & Technology
- Transport
- Water Resources

### **Governance of Open Data**

There is no known governance framework for open data in Ethiopia. However, efforts are underway to coordinate the common data sets of the government. The open data initiative aims to address the governance issues. The open data initiative will create national body that spearhead open data initiative and create a participatory process to encourage the publication of open data across the government.

## 2.6 Analysis of Features for Open Data Portals

### Introduction

One of the key objectives of the project is the development and implementation of the Ethiopian national open data portal. The development of the portal requires identifying the features that will be included in the portal.

The project team used the following sources to identify the features of the portal:

- Features requested in the Terms of Reference of the project
- Features of open data portals of other countries
- Requirements collected from ministries and agencies

The finding of the assessment has been presented as follows:

### Features in the Terms of Reference of the project

The Terms of Reference (TOR) prepared by the Ministry of Communication and Information Technology (MCIT) contain the features that are required to be made available in the Ethiopian government open data portal. Moreover, it contains the standard features that need to exist in national level government open data portals. The TOR has been attached in the annex section.

### Feature of open data portals of other countries

In order to identify the features that exist in a typical government open data portals, the project team reviewed the national open data portals of selected countries. The most common features will be made available in the Ethiopian government open data portal.

The findings of the study have been summarized in the following table.

Table 6. *Features of open data portals of different countries*

No	Country	Portal	Features
1	USA	<a href="https://www.data.gov/">https://www.data.gov/</a>	<ul style="list-style-type: none"> <li>• Simple design with minimal graphics</li> </ul> <p><i>Main menu</i></p> <ul style="list-style-type: none"> <li>• Data</li> <li>• Topics</li> <li>• Impact</li> <li>• Applications</li> <li>• Developers</li> <li>• Contact</li> </ul> <p><i>Middle section</i></p> <ul style="list-style-type: none"> <li>• Simple description about the portal</li> <li>• Search</li> <li>• Browse Topics using thematic icons</li> <li>• Highlights</li> <li>• Updates</li> </ul> <p><i>Footer</i></p> <ul style="list-style-type: none"> <li>• About</li> <li>• Open Government</li> <li>• FAQ</li> <li>• Media</li> <li>• Glossary</li> <li>• Participation</li> <li>• Accessibility</li> <li>• Data Policy</li> <li>• Privacy Policy</li> <li>• Performance</li> <li>• Link to national government portal</li> </ul>

			<ul style="list-style-type: none"> <li>• Link to Twitter</li> <li>• Link to Github</li> <li>• Link to Login</li> </ul> <p><i>Data section</i></p> <ul style="list-style-type: none"> <li>• Organization</li> <li>• Search <ul style="list-style-type: none"> <li>○ Search</li> <li>○ Search results sorting</li> </ul> </li> <li>• Filter by location</li> <li>• Topics</li> <li>• Topic categories</li> <li>• Datasets type</li> <li>• Tags</li> <li>• Formats</li> <li>• Organization types</li> <li>• Organizations</li> <li>• Publisher</li> </ul>
2	UK Government	<a href="http://data.gov.uk">http://data.gov.uk</a>	<p><i>Main Menu</i></p> <ul style="list-style-type: none"> <li>• Data</li> <li>• Apps <ul style="list-style-type: none"> <li>○ By category</li> <li>○ By sector</li> </ul> </li> <li>• Interact</li> <li>• Search</li> </ul> <p><i>Middle section</i></p> <ul style="list-style-type: none"> <li>• Graphical presentation of featured links</li> </ul> <p><i>Footer</i></p>

			<ul style="list-style-type: none"> <li>• FAQ</li> <li>• Code of conduct</li> <li>• Accessibility</li> <li>• Cookies policy</li> <li>• Privacy</li> <li>• Moderation policy</li> <li>• Contact</li> <li>• Terms and conditions</li> <li>• About</li> </ul> <p><i>Data section</i></p> <ul style="list-style-type: none"> <li>• Datasets</li> <li>• Published status</li> <li>• License</li> <li>• Theme</li> <li>• Formats</li> <li>• Publisher</li> <li>• Vocabulary</li> <li>• Code lists</li> <li>• Openness score</li> <li>• Broken links</li> <li>• UK location data set type</li> </ul> <p><i>Datasets section</i></p> <ul style="list-style-type: none"> <li>• Search <ul style="list-style-type: none"> <li>○ Search results sorting</li> <li>○ RSS feed</li> </ul> </li> <li>• Map search</li> <li>• Data requests</li> <li>• Publishers</li> <li>• Organograms</li> </ul>
--	--	--	--

			<ul style="list-style-type: none"> <li>• Spend reports</li> <li>• Site Analytics</li> <li>• Reports</li> <li>• Contracts</li> </ul>
3	Australia	<a href="http://data.gov.au/">http://data.gov.au/</a>	<p><i>Main Menu</i></p> <ul style="list-style-type: none"> <li>• Datasets</li> <li>• Organizations</li> <li>• About</li> <li>• Site Statistics</li> <li>• Use cases</li> <li>• National map</li> <li>• Request data</li> <li>• Toolkit</li> </ul> <p><i>Middle section</i></p> <ul style="list-style-type: none"> <li>• Search</li> <li>• Introduction</li> <li>• Statistics</li> <li>• News</li> <li>• Usage reports</li> </ul> <p><i>Footer</i></p> <ul style="list-style-type: none"> <li>• About</li> <li>• Links to CKAN API</li> <li>• Open knowledge foundation</li> </ul> <p><i>Datasets section</i></p> <ul style="list-style-type: none"> <li>• Search <ul style="list-style-type: none"> <li>○ Search</li> <li>○ Search results sorting</li> </ul> </li> </ul>



			<ul style="list-style-type: none"> <li>• Organizations</li> <li>• Jurisdiction</li> <li>• Groups</li> <li>• Tags</li> <li>• Formats</li> <li>• Licenses</li> <li>• Filter by location</li> </ul>
4	Canada	<a href="http://open.canada.ca/data/en/dataset">http://open.canada.ca/data/en/dataset</a>	<p>The open data portal is a section of the national government portal.</p> <p><i>Major categories</i></p> <ul style="list-style-type: none"> <li>• Search Datasets by category</li> <li>• Organization</li> <li>• Data Type</li> <li>• Subject</li> <li>• File Format</li> <li>• Suggest data set</li> <li>• Apps and visualization</li> <li>• License</li> <li>• About Open Government</li> <li>• Bilingual</li> <li>• Integration with government portal</li> </ul>
5	Ireland	<a href="https://data.gov.ie">https://data.gov.ie</a>	<p><i>Main Menu</i></p> <ul style="list-style-type: none"> <li>• Datasets</li> <li>• Publishers</li> <li>• Suggest a dataset</li> <li>• News and Events</li> <li>• Developers</li> </ul>

			<p><i>Middle section</i></p> <ul style="list-style-type: none"> <li>• Search</li> <li>• Introduction</li> <li>• Simple graphical presentation</li> <li>• Dataset Categories</li> <li>• News and Events</li> <li>• Open Data License</li> <li>• Open Consultations</li> </ul> <p><i>Footer</i></p> <ul style="list-style-type: none"> <li>• Guide to publishers</li> <li>• Open data technical framework</li> <li>• Open data license</li> <li>• Resources</li> <li>• About</li> <li>• Link to twitter, e-mail, telephone, portal owner department information</li> <li>• Powered by CKAN link</li> </ul> <p><i>Datasets section</i></p> <ul style="list-style-type: none"> <li>• Search <ul style="list-style-type: none"> <li>○ Search</li> <li>○ Search results sorting</li> <li>○ RSS feed</li> </ul> </li> <li>• Theme</li> <li>• Resource Format</li> <li>• Publisher</li> <li>• Licenses</li> </ul>
--	--	--	---

The features of the open data portals of the countries presented in the above table and other open data portals of other countries show that most open data portals have similar sections. Accordingly, the portals have the following sections and features:

A header section containing the following:

#### Header

- Logo
- Main menus

#### Middle section

- Brief or detailed description about the data portal, in some case small graphics
- Thematic search using graphical icons for themes or topics

#### Footer

- Containing link that display at the bottom of every page such as About, FAQ, privacy, link to social media, contacts, etc.

Moreover, the Data or Datasets link exists on the main menu of all portals reviewed. The Data/Datasets section has a separate page that contains a number of links related to datasets including the following Search, Search results sorting, Organizations, Themes, Formats, Licenses and Search by location.

In addition to the Data/Datasets menu item, there are a few other menus that exist in the header of most open data portals. Some of these are Topics or Organizations/Publishers, Applications or apps, Developers, Contact, News and Events, Maps, Analytics/Statistics, About, Suggest or Request a Dataset, etc.

Accordingly, the project team recommends the following organization of the site to be used as an initial list to be used for development of the prototype of the Ethiopian government open data portal:

**Header**

- Logo
- Main menu
  - o Datasets
  - o Organizations
  - o Applications
  - o Statistics
  - o Suggest data
  - o Contact

**Middle Section**

- Brief description about the open data portal
- Search bar
- Icons for the various thematic topics

**Footer section**

- About
- FAQ
- Participation
- Accessibility
- Data Policy
- Privacy Policy
- Link to Ethiopia.gov.et
- Link to Social media
- Link to GKAN or data portal framework

**Data/Datasets section**

- Search
  - o Search
  - o Search results sorting
  - o RSS feed
- Theme
- Organizations
- Formats

### Assessment of the needs of ministries/agencies

The requirements of the various ministries/agencies that will be using the national data portal with regard to the features of the data portal to be developed was assessed using a questionnaire. The responses have been summarized in the following table:

No	Question	Response
Participation in the National Portal		
1	Are you willing to participate in the national portal where your organization data will be presented for access to the public? If not, why not?	All respondents are willing to participate in the national portal.
What kind of features would you like to see in the Open Data Portal?		
2	Tools to publish and find datasets Publish datasets a. via import b. Through a web interface c. Both	All respondent except ECX and document authentication and registration office select both features to publish and find dataset, ECX and DARO selected only the web interface.
3	Search by: a. keyword b. Filter by tags c. See dataset information at a glance d. All	All respondent selected all search features to be incorporated in the portal.
4	Tools to store & manage data: a. Store the raw data b. Store metadata c. Both	The respondent from Supreme court, Ministry of Trade and Ministry of Agriculture selected “ <i>store the raw data</i> ” tool. The respondent from National Election Board didn’t select any option. The respondent from Metrology selected “ <i>Store meta data</i> ”. The rest of the respondents selected both options.

5	<p>Visualize structured data with interactive:</p> <ul style="list-style-type: none"> <li>a. Tables</li> <li>b. Graphs</li> <li>c. Maps</li> <li>d. All</li> </ul>	<p>The respondent from Supreme court selected “Tables and Graphs” features to visualize structured data</p> <p>The rest of the respondents selected all three options.</p>
6	<p>Get statistics and usage metrics for your datasets?</p>	<p>Except National Election Board, all respondents wanted to have statistics and usage metrics.</p>
7	<p>Search geospatial data on a map by area?</p>	<p>Except Supreme Court, National Election Board, Ministry of Finance and Economic Development and Ethiopian Commodity Exchange, all other respondents wanted to have the geospatial search feature.</p>
8	<p>Theme the portal with CSS or integrate with a CMS?</p>	<p>Except Supreme Court, National Election Board, Metrology, Ethiopian Commodity Exchange and Ethiopia Mapping Agency, all other respondents’ wanted to have the CSS or integrate with a CMS.</p>
9	<p>Build a community with extensions that allow users to comment on and follow datasets</p>	<p>Except National Election Board and Ethiopian Commodity Exchange, all other respondents selected the feature.</p>
10	<p>What other features do you want in the portal?</p>	<p>Supreme court</p> <ul style="list-style-type: none"> <li>- Online recommendation of datasets by users.</li> </ul> <p>Ministry of Transport</p> <ul style="list-style-type: none"> <li>- Our website to join application web services (SOA) to applications of organizations like CBE (Commercial Bank of Ethiopia), Face book, etc.</li> </ul> <p>Ministry of Education</p> <ul style="list-style-type: none"> <li>- Enable users to manage, analyze and generate customized reports with different possible parameters online.</li> <li>- Control and record users interacting to the system.</li> <li>- Direct integration to already implemented data</li> </ul>

		<p>collection tools in our organization.</p> <p>Ethiopian Commodity Exchange</p> <ul style="list-style-type: none"> <li>- Video uploading features.</li> <li>- Trading manuals uploading features.</li> <li>- Audio uploading features.</li> </ul> <p>Chamber of Commerce</p> <ul style="list-style-type: none"> <li>- Mobile apps should be developed for accessing the data.</li> </ul> <p>Central Statistical agency</p> <ul style="list-style-type: none"> <li>- Online data processing and generation indicators is good if you incorporate it in the portal as well as make it as comprehensive as possible</li> </ul> <p>Federal Ethics and Anti-Corruption Commission</p> <ul style="list-style-type: none"> <li>- The best future and currently used features that you think is required for the portal.</li> </ul>
--	--	--

From the above table, we can see that the suggestion of most ministries and agencies is similar. The project team analyzed the selected options of the ministries and agencies and produced the aggregated requirements presented below. The additional features requested under item number 10 have been analyzed and only the relevant suggestions have been included.

- Tools need to be provided to publish and find datasets Publish datasets. The tools shall support publishing via import an through a web interface.
- Search feature shall support searching using keywords, filter by tags and shall allow viewing dataset information at a glance
- Tools used to store and manage data shall support storing of raw data and metadata
- The system shall support visualization of structured data with interactive tables, graphs and maps whenever possible.
- The system shall support getting statistics and usage metrics for datasets

- The system shall support searching geospatial data on a map by area
- The system shall support managing the look and feel of the user interface using Themes with CSS or integrated themes and CSS of a CMS.
- The system shall have a feature to build a community with extensions that allow users to comment on and follow datasets.
- The system shall have a feature to allow online recommendation of datasets by users
- The system shall support integration with applications using Web service technology.
- The system shall have a feature that enables users to manage, analyze and generate customized views with different possible parameters/indicators online, when possible.
- The system shall have logs of users visits
- The system shall have a feature that allows direct integration to already implemented data collection tools in organizations. (This is only possible if the data collection tools have been developed using the Ethiopian eGIF standard, or allows web services technology.)
- The system shall allow multimedia uploading features
- The system shall allow uploading of relevant manuals or documents with the dataset
- The system shall have Mobile apps for selected datasets to enable accessing the data using mobile devices.
- The system shall be scalable to enable inclusion of future requirements



## 2.7 Analysis of Technology Options

### Background

Open data portal is a type portal that provides public access to datasets provided by government ministries and agencies under Open Data terms of use through a web based portal interface. Open data portals consist of data catalogue, content management system and data management tools.

The main component of open data portals is the data catalogue, which is a web based tool that provides public access to datasets provided by organizations under Open Data terms of use. Open data catalogs allow users to easily search for and download available datasets in open data formats, display terms of use, and provide attribution, metadata and other information about the data sets.

Implementation of open data portals requires a suite of technologies that are used to provide a data catalogue along with other information and services such as an online forum, a feedback mechanism, a blog, which are usually provided by the data catalogue provider or implemented using a separate content management system or a portal solution that has a content management system as one of its components.

### Common Features of Data Catalogues

Data catalogues may vary from one another in terms of the features they provide. The common features of most data catalogues are the following:

#### *Data Access and Storage*

Data catalogues provide a feature to store and access publicly available data. Catalogs may have the ability to store data in its own management system, or manage metadata along with links to the datasets stored on other servers.

Data catalogues provide data in open, non-proprietary formats such as CSV, XML and JSON. They allow metadata of data sets to be stored with the datasets. The data sets can be stored within the data catalogue system, or in another separate system/server, in which

case the link to the data sets along with its metadata is stored in the data catalogue. Data catalogues have the facility to publish data sets and a workflow that may be used for approval purposes.

### *User Services*

Data catalogues provide a number of features to enable access to the data sets in a convenient way. Some of these features include searching datasets, online data preview, data visualization using charts and maps, commenting on dataset, suggestion of new data set, linking data sets to social media, and accessing data sets using mobile devices.

### *Customization*

Data portals may allow users to customize the look and feel of default user interfaces by modifying the layouts and color schemes of the data catalogue pages using custom made themes. This feature may not be available in some commercial products that use cloud based hosting of the data catalogue.

### *Integration*

Data catalogues may offer integration with external content management systems, portals or websites using a variety of integration technologies. If the open data portal requires extensive content management features and the data catalogue does not have the required features, it should allow for integration of an external content management system or a portal that can provide the required features.

### *Analytics*

Data catalogues may have analytics feature that provides information on page access, views and downloads. External analytic tools such as Google Analytics can also be integrated with the data catalogues. However, for datasets specific information, the built-in analytic features of data catalogues are required.

### *Application Programming Interface (API)*

APIs allow developers to access different components of the data catalogue programmatically in order to add additional features to the data catalogues, update the data via external systems or use the data catalogues built-in functionalities in other custom applications.

### *Extensibility*

The data catalogue should allow extending the built-in feature of data catalogues through incorporation of additional modules or through development of custom application using the API.

## **Types of Data Catalogues**

There are a number of ways of classifying data catalogues. Two of these classification methods have been presented below, the first one is based on openness of the source code of the data catalogue and cost and the second one is based on deployment method.

### *Open Source versus proprietary Data Catalogues*

Data catalogues may be either open source and free or commercial and proprietary. Open source products are free. They may be acquired via download for no cost and may be modified or customized without restriction or licensing fees. Commercial products incur cost for the software licenses or in the case of cloud based solutions for hosting services, typically for a setup charge and recurring monthly or annual fee.

### *Self-managed versus Cloud based solutions*

Data catalogues may be self-managed or cloud based. In the case of self-managed solutions, the software will be installed in the server of the client at its premises. Cloud based products, which are usually in the form of SaaS model, provide software and hosting services on the cloud through the Internet. The vendor of the software is responsible for software maintenance, server availability and reliability and performance monitoring. The

vendor may provide training and customization service at cost. However, the degree of customization may be limited to what the vendor is capable of and willing to provide.

### Open data technologies

There are a number of open data technologies in the market. The products may be open source or commercial, self-managed or cloud-based. The products may be complete platforms or supporting tools designed for purposes of facilitating specific functions of the data publishing or presentation process.

The following list contains the most common open data platforms:

Table 7. Most common open data platforms

No.	Product	URL
1	CKAN	<a href="http://ckan.org/">http://ckan.org/</a>
2	DKAN	<a href="http://nucivic.com/dkan">http://nucivic.com/dkan</a>
3	Socrata	<a href="http://www.socrata.com/">http://www.socrata.com/</a>
4	Open Data Soft	<a href="http://www.opendatasoft.com/">http://www.opendatasoft.com/</a>
5	Open Government Platform (OGPL)	<a href="http://ogpl.gov.in/">http://ogpl.gov.in/</a>
6	Junar	<a href="http://junar.com/">http://junar.com/</a>
7	Swirrl	<a href="http://www.swirrl.com/">http://www.swirrl.com/</a>
8	Information Workbench (FluidOPS)	<a href="http://www.fluidops.com/en/">http://www.fluidops.com/en/</a>
9	OGDI	<a href="http://ogdi.codeplex.com/">http://ogdi.codeplex.com/</a>
10	LIBRE	<a href="http://libre.readthedocs.org/en/latest/index.html">http://libre.readthedocs.org/en/latest/index.html</a>
11	Datahub	<a href="http://datahub.io/">http://datahub.io/</a>
12	Semantic media wiki	<a href="http://semantic-mediawiki.org/">http://semantic-mediawiki.org/</a>

A brief description of some of the common open data platforms will be provided in the next section.

The following table contains examples of tools that have been designed to undertake specific functions in the publishing and presentation of datasets.

**Table 8. Examples of dataset publishing and presentation tools**

No.	Product	Feature and URL
1	Data Visualization wizard	The Data Visualization Wizard provides a fast way to get data visualizations online using Drupal. It provides the following visualizations after a quick upload of a spreadsheet: Geo-spatial map, Pie and bar graphs, Summary listings, Detail pages, Searchable and filterable listings, Downloadable data <b>and</b> Lightweight API. <a href="https://www.drupal.org/project/datavizwiz">https://www.drupal.org/project/datavizwiz</a>
2	Geo Network	GeoNetwork is a catalog application to manage spatially referenced resources. It provides powerful metadata editing and search functions as well as an interactive web map viewer. It is currently used in numerous Spatial Data Infrastructure initiatives across the world. It has been developed to connect spatial information communities and their data using a modern architecture, which is at the same time powerful and low cost, based on the principles of Free and Open Source Software (FOSS) and International and Open Standards for services and protocols. <a href="http://geonetwork-opensource.org/">http://geonetwork-opensource.org/</a>
3	CSV to API	Dynamically generate RESTful APIs from static CSVs. Provides JSON, XML, and HTML. <a href="https://github.com/project-open-data/csv-to-api">https://github.com/project-open-data/csv-to-api</a>
4	Database to API	Dynamically generate RESTful APIs from the contents of a database table. Provides JSON, XML, and HTML. Supports most popular databases. <a href="https://github.com/project-open-data/db-to-api">https://github.com/project-open-data/db-to-api</a>

## Common Open Data Platforms

### CKAN (the Comprehensive Knowledge Archive Network)

CKAN is an open source data portal platform developed by the Open Knowledge Foundation, a non-profit organization. The platform is aimed at government agencies, organizations and companies who want to publish and share open data. CKAN is the most widely used platform in government organizations around the world.

CKAN can be installed on any Linux server, including cloud-hosted configurations. The Open Knowledge Foundation also offers hosting services for a monthly fee. CKAN is written in the Python programming language, and is designed for publishing and managing data either through a user interface or through an API. CKAN has a modular architecture through which additional or custom features may be added.

The following list contains a few of national-level government data portals developed using CKAN:

- United States: [Http://www.data.gov](http://www.data.gov)
- United Kingdom: [Http://www.data.gov.uk](http://www.data.gov.uk)
- Australia: [Http://www.data.gov.au](http://www.data.gov.au)
- Canada: [Http://www.data.gc.ca](http://www.data.gc.ca)
- Ireland: <http://www.data.gov.ie>
- Germany: <https://www.govdata.de/>
- Romania: [Http://www.data.gov.ro](http://www.data.gov.ro)
- Slovakia: [Http://www.data.gov.sk](http://www.data.gov.sk)
- Austria: <https://www.opendataportal.at>
- Netherlands: [Http://www.data.overheid.nl](http://www.data.overheid.nl)
- Russia: <http://hubofdata.ru/>
- Italy: <http://www.dati.gov.it>
- Norway: [Http://www.data.norge.no](http://www.data.norge.no)
- Belgium: <http://portal.openbelgium.be/>
- Switzerland: <http://opendata.admin.ch/>
- Hungary: <http://opendata.hu/>

- Greek: <http://open-data.okfn.gr/>
- Sweden: <http://oppnadata.se/>
- Indonesia; <http://data.go.id/>
- Japan: <Http://www.data.go.jp>
- Brazil: <http://www.dados.gov.br>
- Ecuador: <http://www.datasabiertos.ec>
- European-Union: <http://open-data.europa.eu>

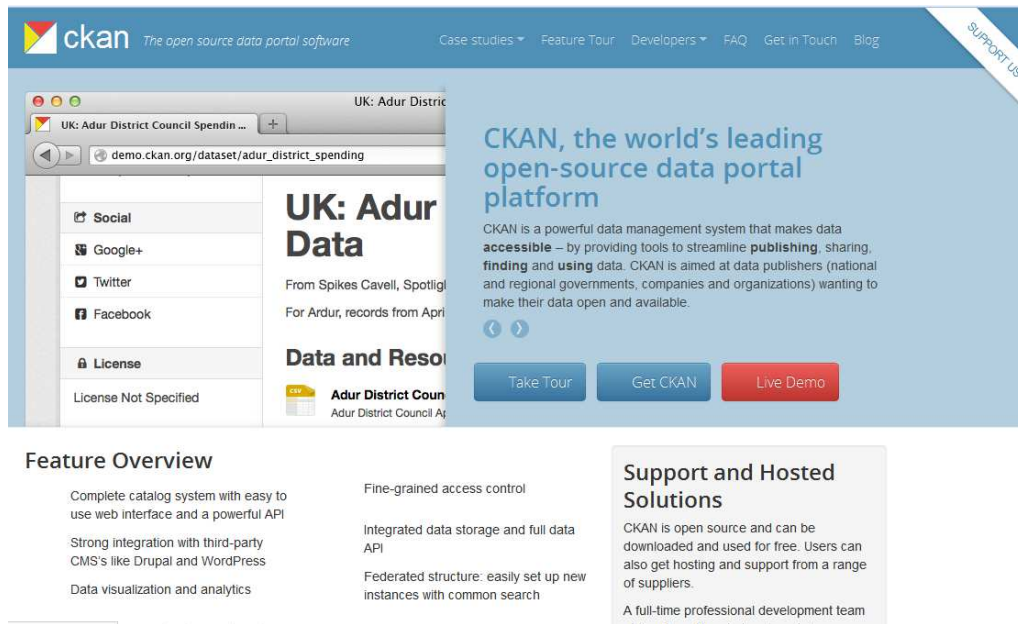


Figure 10. CKAN Interface

## DKAN

DKAN is a Drupal-based open data tool with a full suite of cataloging, publishing and visualization features that allows governments, nonprofits and universities to easily publish data to the public. DKAN is maintained by NuCivic, <http://nucivic.com/dkan>. It is an open source distribution of Drupal delivering a full featured open data publishing portal as well as the full suite of Drupal content management (CMS) functionality, to deliver a richer open data portal experience for users.

DKAN is designed to be feature compatible with CKAN and its underlying API is identical. DKAN is also open-source, but is based on Drupal, a popular content management system written in PHP instead of Python. Drupal has its own modular architecture with thousands of modules available for download and a large developer community.

The Cologne city data portal in Germany <http://www.offenedaten-koeln.de> uses DKAN in the implementation of its open data portal.

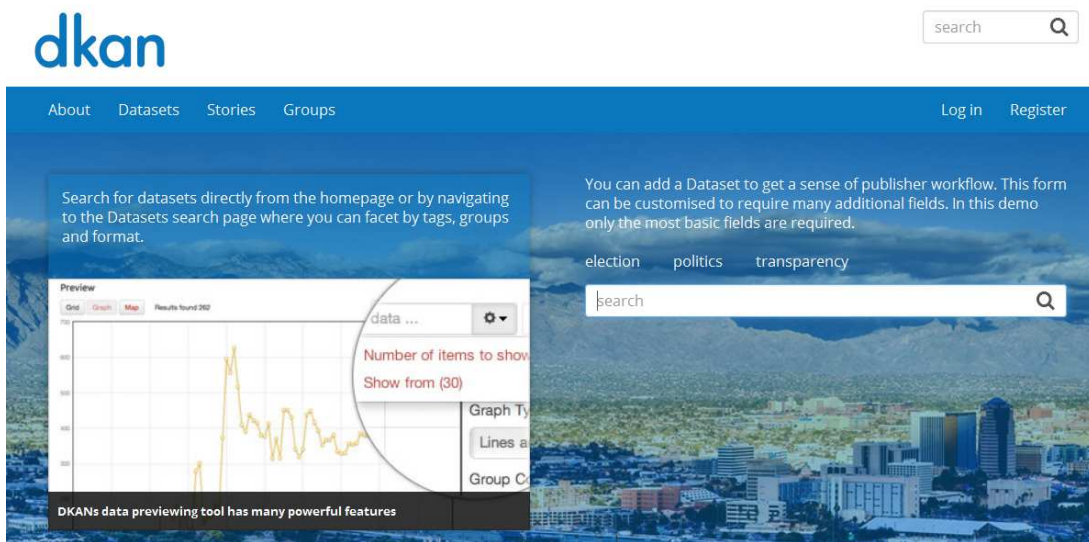


Figure 11. DKAN Interface

### Socrata Open Data Portal

Socrata is a commercial cloud-based SaaS open data catalog platform that provides API, catalog, and data manipulation tools. It is hosted by Socrata Inc. for publishing, browsing, comparing and visualizing different datasets. Socrata offers a free version of its API server, which is just one component of the complete platform. A community edition of the Socrata platform, which is open source, is under ongoing development. The Kenyan national open data portal uses Socrata platform. Examples of other government departments that use the Socrata platform include the Chicago and San Francisco city open data portal. <http://www.socrata.com>



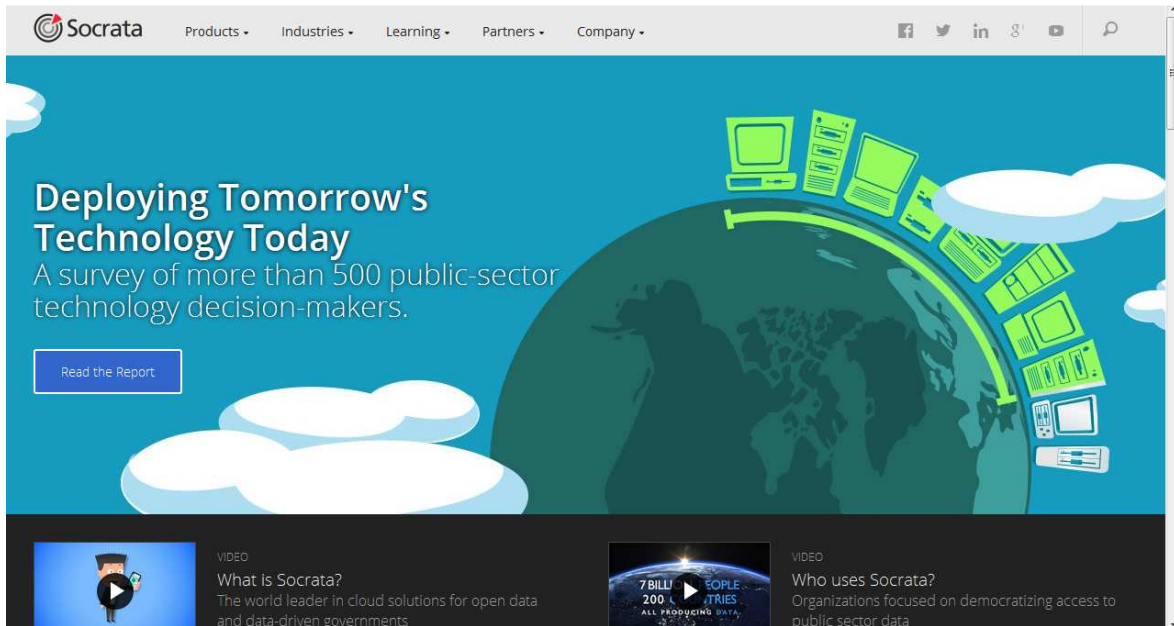


Figure 12.Socrata Interface

### Open Data Soft (ODS)

Open Data Soft is a commercial cloud-based SaaS platform that offers a comprehensive suite of Open Data and visualization tools. ODS is written in Python with Django as the web framework, uses Exalead for search functionality, Hadoop for data processing and MongoDB as its data store. It enables organizations to quickly develop and launch open data portals as it has many built-in rich features. The platform supports common Open Data formats such as CSV, JSON and XML, as well as many geospatial formats such as KML, OSM and SHP. Search functionality is very straightforward and easy to use.

It uses commercial cloud hosting providers such as Amazon cloud services, Microsoft Azure, eBRC and Numergy. ODS' business model is a pay-as-you-use subscription fee based on data volume, usage and SLAs. They charge for platform setup, data hosting, access to software, maintenance and support. Among the earliest adopters of ODS are the city of Brussels, Belgium, the city of Paris, France, and the company Veolia Environment. eSystems Africa's team observed that ODS has its largest number of users in France.

<http://www.opendatasoft.com/>



# MAKE THE MOST OF YOUR DATA

TRANSFORM DATA INTO UNDERSTANDABLE INFORMATION AND **INNOVATIVE APPLICATIONS**



Figure 13. ODS Interface

## Open Government Platform (OGPL)

OGPL is an open-source Drupal-based data catalog. It was not designed to be CKAN-compatible at the API level. OGPL is a joint product of the government of India and the United States government to promote transparency and greater citizen engagement by making more government data, documents, tools and processes publicly available. OGPL is an open source platform. <http://ogpl.gov.in>. The government of and Ghana used OGPL platform for implementing their respective national data portals.



Figure 14. OGPL Interface

## Junar

Junar is a cloud-based SaaS open data platform. Data is typically managed within Junar's infrastructure. Junar can either provide a complete data catalog or can provide data via an API to a separate user catalog. eSystem Africa's team observed that Junar has its largest number of users in South American region. Some examples are the data portals of Lima (Peru), Chile and Costa Rica. <http://www.junar.com>.

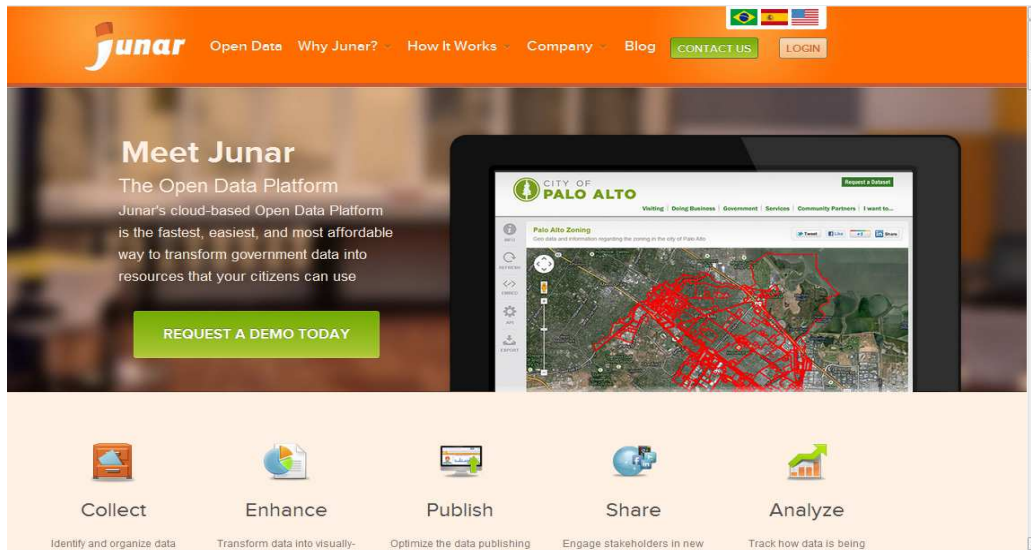


Figure 15. Junar Interface

## Swirrl

Swirrl is a cloud based open data portal platform developed by a small company based in Manchester and Stirling in the UK to make government data available for the public. It was built on linked data technologies. Documentation on its website and on the Internet shows that its adoption could not go beyond a few UK government departments such as Open Data Scotland and Hampshire City Council, not even in the national UK open data portal.

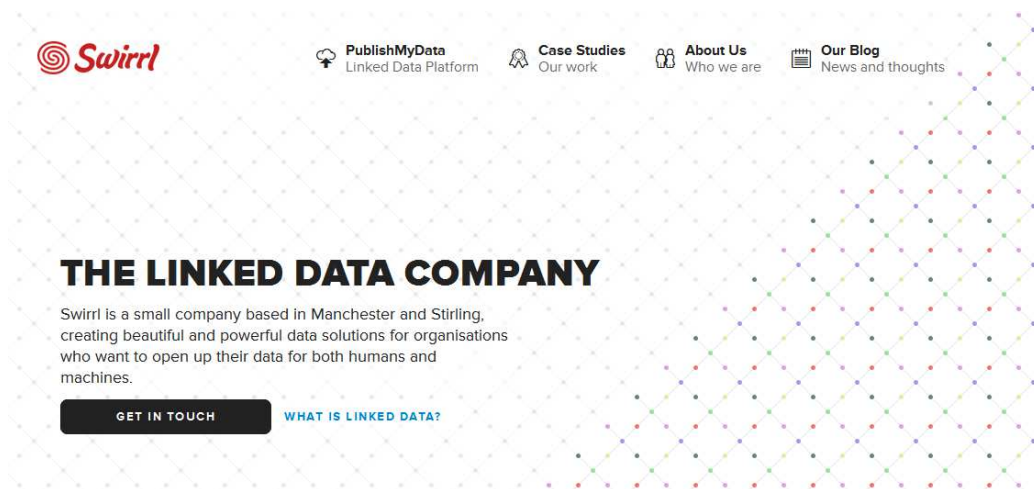


Figure 16.Swirrl Interface

## OGDI (Open Government Data Initiative) DataLab

OGDI DataLab is Microsoft contribution to the global open data initiative and was initially made available on Codeplex, Microsoft's open source code portal at <http://ogdi.codeplex.com/>. It was later moved to Github. OGDI is an open source platform written using Microsoft development technologies C#/.NET. It is developed to run on Windows Azure, a Microsoft cloud solution. OGDI has three main components: Data Service, Data Loader and Data Browser. OGDI is being used by different organizations such as City of Medicine Hat, Canada and City of Regina, Canada.

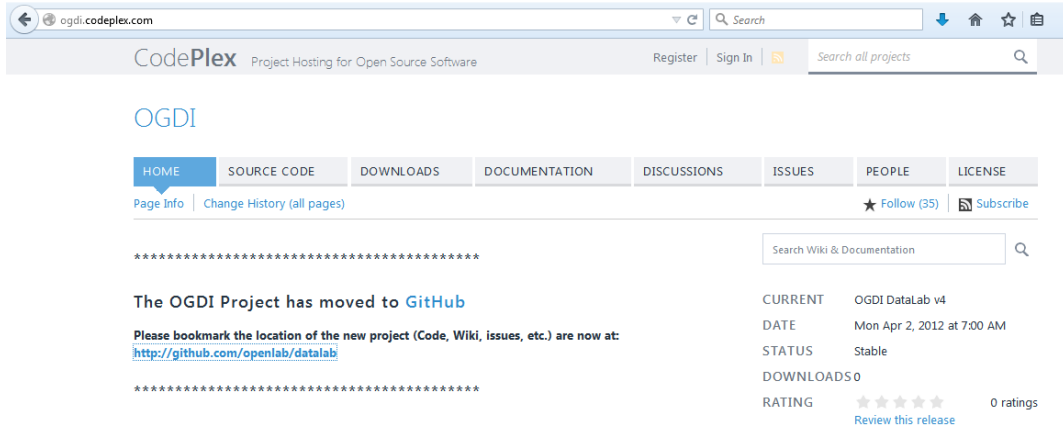


Figure 17. OGDIDataLab Interface

### LIBRE (Large Information Batch Restructuring Engine)

Libre is an open source platform written in Python and based on the web framework Django. The platform was created by the chief information officer of the Commonwealth of Puerto Rico to publish open government data and it was made available as Free software with the hopes that other countries and individuals may benefit from it too. <http://libre.readthedocs.org/>

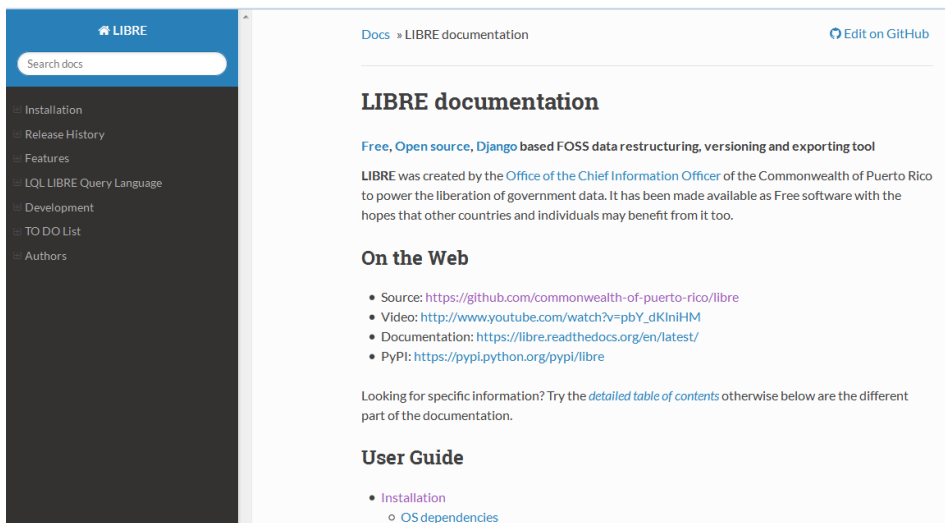


Figure 18. Libre Interface

## Semantic MediaWiki

Semantic MediaWiki was created as an extension of the wiki application best known for powering Wikipedia called MediaWiki. While traditional wikis contain only text, Semantic MediaWiki adds semantic annotations that allow a wiki to function as a collaborative database and data catalog. Semantic MediaWiki is an RDF implementation, meaning that both data and metadata are stored as linked data and are accessible via linked data interfaces such as SPARQL.

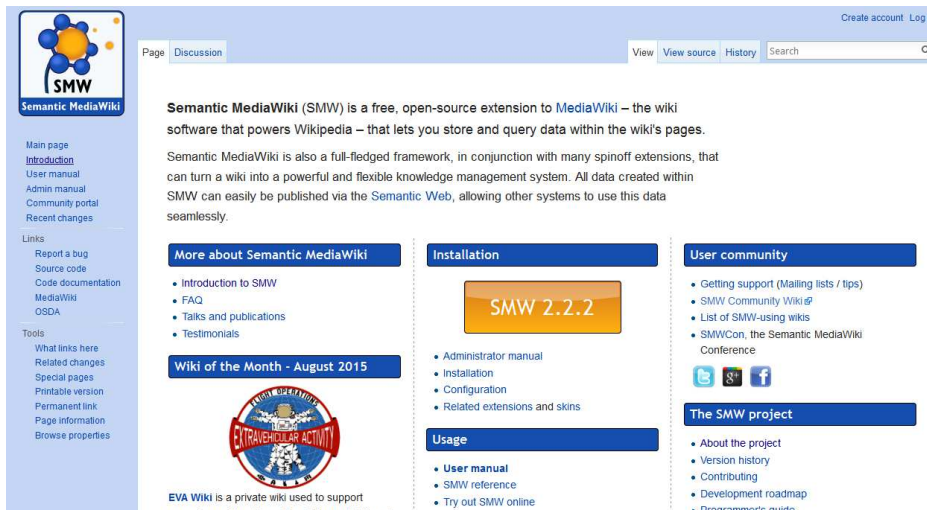


Figure 19. Semantic MediaWiki Interface

## Information Workbench (fluidOps)

Fluid Operations (fluidOps) provides innovative solutions in semantic technology through its flexible and powerful semantic integration platform called the Information Workbench. Information Workbench unifies semantic data management with innovative cloud management tools and links it with best-in-class data center technologies. The following are the main features of Information Workbench:

- Automatically analyze and use data from different data sources:
- Access, use and combine data flexibly:
- Use data to proactively drive your business:
- Orchestrate data, systems, workflows and processes as well as the underlying IT and infrastructure

Main benefits of the technology are to connect data silos in an open, extensible platform, gain an overview of enterprise processes and structures, increase data quality, create and leverage actionable data, turn data into smart data, extract new insights from smart data, make well-founded decisions more easily, automate processes and workflows, save time and costs.

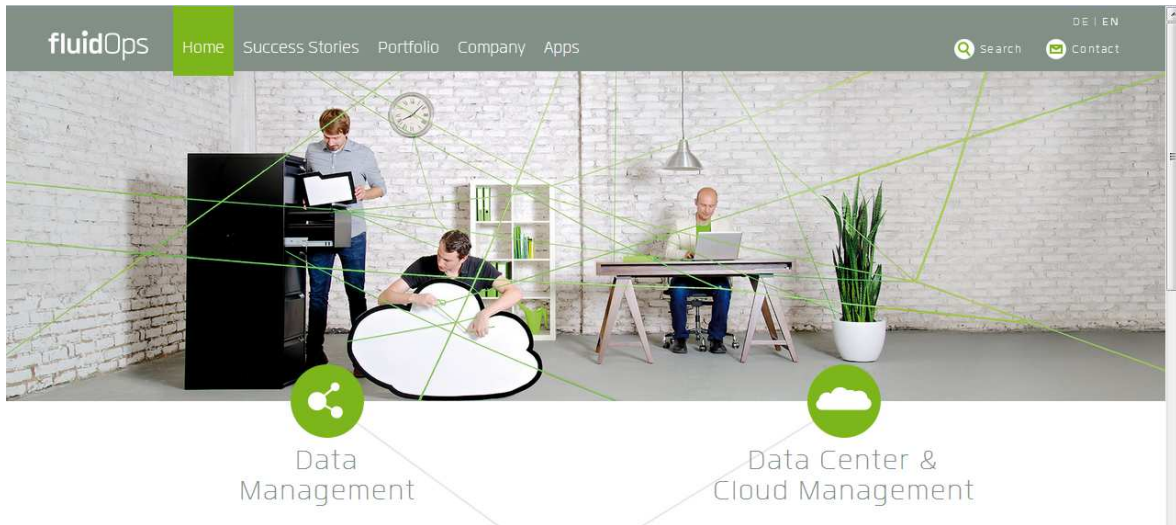


Figure 20.fluidOps Interface

## Datahub

Datahub is a hosted free and powerful data management platform from the Open knowledge foundation, the organization responsible for the development of the CKAN platform. Data hub is a hosted solution and allows anyone to publish data. It runs on a CKAN platform. It has around 9800 datasets hosted on it from different organizations and countries around the world. It is a good demonstration of the power of CKAN in action.

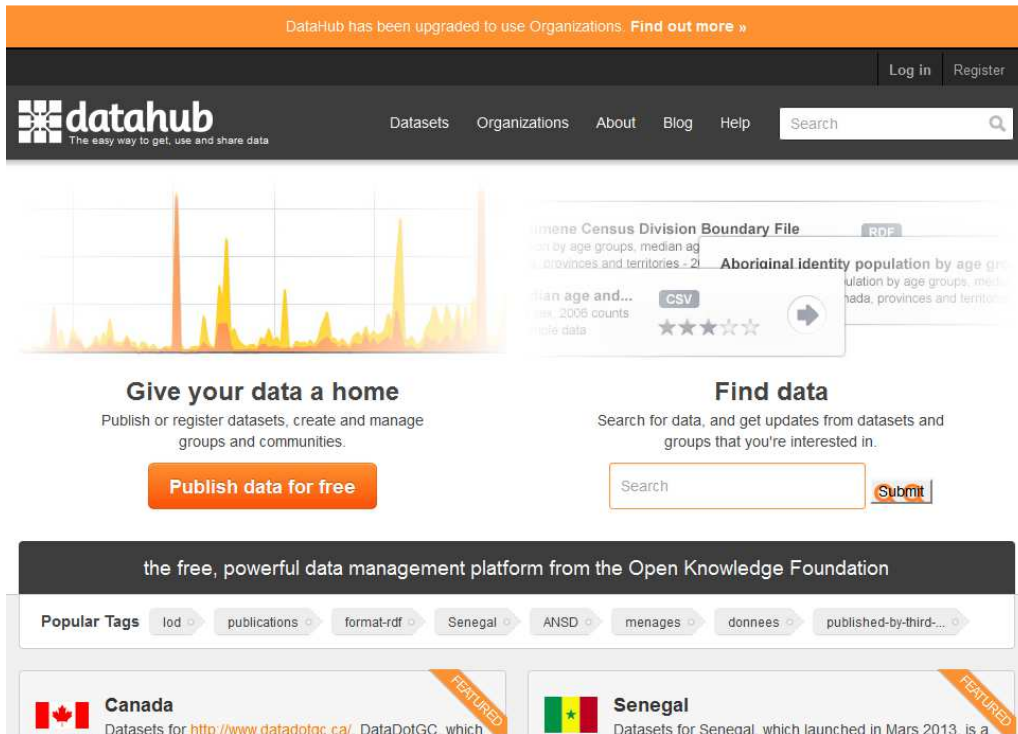


Figure 21. Datahub Interface



## Selection of open data platform for Ethiopia

The team undertook a two-phase evaluation and selection of a suitable open data platform for Ethiopia. The first phase of the selection process involves evaluation of the basic needs of the ministry, which is the product has to be free and open source (FOSS) and should be deployable at its premises.

The second phase of selection involves a detailed analysis of the selected platforms in regards to their suitability for implementing the Ethiopian ODP. The main criteria used for evaluating the candidate platforms include ease of installation and configuration, availability of documentation, existence of strong support community, feature sets of the platform and adoption by other national-level open data portals.

The following table contains the list of products with their licensing and supported deployment methods.

**Table 9. Licensing and supported deployment methods of different products**

No.	Product	Licensing	Deployment
1	CKAN	Open Source	On-premise or Cloud
2	DKAN	Open Source	On-premise or Cloud
3	Socrata	Commercial, Open Source community edition under development	Cloud, On-premise possible for the API server
4	Open Data Soft	Commercial	Cloud
5	Open Government Platform (OGPL)	Open Source	On-premise
6	Junar	Commercial	Cloud
7	Swirrl	Commercial	Cloud
8	Information Workbench (FluidOPS)	Enterprise edition is Commercial, Community edition is free.	Cloud based solution, the community version is a self-managed version,
9	OGDI	Open Source	On-premise, data stored on Azure cloud

10	LIBRE	Open Source	On-premise
11	Datahub	Open Source, runs on CKAN	Hosted solution
12	Semantic media wiki	Open Source	On-premise

As shown in the above table, the following products are either “free or open source” or have a free community version, and support on premise deployment:

1. CKAN
2. DKAN
3. OGPL
4. Libre
5. Information Workbench
6. Semantic Media-wiki

The project team further investigated these products to determine their suitability for the Ethiopian government open data portal. The team used the following as key factors for determining suitability for the Ethiopian ODP:

1. Ease of installation and configuration
2. Documentation
3. Support community
4. Feature sets
5. Adoption by other national-level open data portals

The team used internet based research to determine the suitability of the platforms, and complemented it with actual installation and configuration of the platforms at the lab facilities of e-System Africa. The result of the study has been presented as follows:

## **CKAN**

The installation files as well as the installation and configuration instructions are available on the CKAN website. The installation process was straightforward but the configuration process encountered a few challenges. The platform requires installation and configuration of a Linux operating system, of which Ubuntu 12.0 is recommended as the best by CKAN. The team installed the operating system and the associated files of the CKAN platform. It took a little bit of effort to do the configuration and get the main features to work properly. There is sufficient documentation on the CKAN website. The documentation was very helpful to successfully carry out the installation, configuration and troubleshooting of the errors. There is also a strong CKAN support community.

With regard to adoption by national-level data portals, the team found out that the platform is the most widely used platform by national-level as well as department/ministry level government data portals. Most countries that implemented open data portals used the CKAN platform. Hence, this shows that CKAN is the most popular platform for government data portals.

The team tried to evaluate the features of CKAN against the requirements presented in the terms of reference of the project and found out that most of the requirements will be met by the CKAN platform. Moreover, any requirements that may not be met by the default features of CKAN can be met using one or more extensions developed by the CKAN community. These extensions are accessible on the CKAN website.

## **DKAN**

The installation files as well as the installation and configuration instructions are available on the DKAN website. The installation and configuration process was straightforward. The platform requires installation and configuration of either Linux or Windows operating system. There is also a need to install the Drupal content management system, which is the key component of the platform.

The team installed the operating system and the associated files of the DKAN platform. There is sufficient documentation on the DKAN website to do the installation and configuration. There are two main sites that provide the installation files and the documentation. The first one is [www.getdkan.org](http://www.getdkan.org) and the second one is the drupal website [www.drupal.org](http://www.drupal.org).

The team observed that there is no specialized website specifically designed for DKAN. It looks like the DKAN and Drupal CMS content on the Drupal website are mixed up. The documentation as well as the support community is presented on the same website; hence, it is a little bit confusing to see the distinction between DKAN related documentation and that of Drupal CMS. Hence, unlike CKAN, DKAN does not have a specialized website, documentation and support base.

Documentation shows that DKAN is a PHP based platform that was created by re-writing the CKAN Python code using PHP. Hence, any core functionality written on the CKAN platform needs to be continuously monitored and re-written using PHP by Drupal developers. This makes the sustainability of DKAN questionable, though currently it seems there are no such issues as there is a large support community. It even claims there are over 1800 extensions of DKAN.

Regarding adoption by national level open data portals, the team could not find any national level portals developed using DKAN. Even if they exist, their numbers will not be as many as those developed using CKAN. Some governments such as the government of Ireland tried to use CKAN and Drupal together, Drupal to serve the CMS features required in the open data portal. However, they learnt that the integration effort was very complicated and resorted back to using CKAN and the default CMS features available in CKAN.

## **OGPL**

The team's effort to carry out a local installation of the OGPL platform could not succeed due to unavailability of enough documentation on the official website. The team tried to search for the necessary documentation in other websites on the Internet, but it could not

find any. There are a few product brochures on the OGPL website, but there is no documentation that contains installation and configuration instructions. There are also no extension development instructions; hence, it will be difficult to add additional features on the ODP portal if required. According to the desk research results, it is only India and Ghana that implemented OGPL. When Ghana implemented its ODP, it used unstable version of the OGPL platform, and was assisted by the Indian national institute of information and communication technology.

## **OGDI**

OGDI is Microsoft Corporation's contribution to the global open government initiative. Microsoft developed OGDI to be used on a Microsoft platform to run open data portals. The product was developed using Microsoft's development technologies C#.NET. The installation files and the documentation have been placed on the Microsoft open source portal [www.codeplex.com](http://www.codeplex.com). The site states that the installation files have been moved to the Github website. This indicates that the sustainability of the support from Microsoft for the platform is questionable.

The project team downloaded the installation file and tried to install and configure the data portal. However, the installation could not be completed since it requires a cloud-based storage, specifically on Microsoft's Azure, in order to proceed. It means the platform can only be used if the implementer buys Microsoft's Windows Azure cloud hosting service. This will create dependency and vendor lock-in. Moreover, it is not possible to put the storage in the premises of the national data center. This is against the basic requirements of MCIT as presented in the Terms of Reference of the project. The project team could not find any national level open data portals that used OGDI as their platform.

## **Information Workbench (fluidOps)**

Information workbench (IWB) is one of the best open data portal platform technologies in terms of out-of-the-box features and the ease and speed of implementation. It is a commercial solution, with the availability of a community edition that is open source and free.

The team downloaded the installation files of the community edition of IWB from the fluidops website [www.fluidops.com](http://www.fluidops.com). Installation of the platform was straight forward and the easiest of all. However, when the team tried to configure the different features of the platform, most of them could not be completed, as the features have not been made available in the community edition of the software.

### **LIBRE (Large Information Batch Restructuring Engine)**

Libre is an initiative of the chief information officer of the Commonwealth of Puerto Rico to publish open government data. Documentation on the Libre is available on its website <http://libre.readthedocs.org/>.

The documentation on the Libre website and further study on the Internet shows that even though Libre has very good features of a typical open data portals, there is no activity in further development of the platform after December 2013. According to the documentation on website, the main development was undertaken by a few individuals but could not get acceptance by a large community. It looks like the platform has been used for the government of Puerto Rico and could not be move further.

### **Semantic MediaWiki (SMW)**

Semantic MediaWiki was created as an extension of the wiki application best known for powering Wikipedia called MediaWiki. While traditional wikis contain only text, Semantic MediaWiki adds semantic annotations that allow a wiki to function as a collaborative database and data catalog. Semantic MediaWiki is an RDF implementation, meaning that both data and metadata are stored as linked data and are accessible via linked data interfaces such as SPARQL.

The team found out that SMW is an extension of the wiki application to handle the publishing of national data sets into open data portals. However, it is not a specialized open data platform. Hence, it does not have a large community of users nor a support base to be used as a reliable open data portal. The team tried to study if there are any national

level government open data portals that use the platform and found none. The team could not also find many non-governmental organizations and businesses that use SMW.

## Conclusion

There are a number of open data platforms used for implementing open data portals for government organizations, private companies and non-governmental organizations. The platforms could be free and open source or commercial. There are also those that are commercial but have a free community version. The delivery model of the platforms can either be cloud based or on premise. The platforms vary in their feature sets, ease of implementation, availability of documentation, support community and degree of adoption by governments in the implementation of national level open data portals.

The project team undertook an extensive study of the various products available for implementing open data portals, and carried out testing of selected products by actual installation at the lab of eSystems Africa. The result shows that many of the platforms use cloud based delivery model. There are also many that are commercial. Many of the cloud-based platforms are commercial products.

Out of the free open source and on-premise deployable platforms, the best platform for national level government open data portals was found to be the CKAN platform. This is the most widely used platform for government open data portals. It has rich documentation and a strong support community. Most of the requirements presented in the terms of reference of the project can be met with the CKAN platform. Hence, the team recommends CKAN as the best platform for the Ethiopian government open data portal implementation. For any content management system requirements that cannot be provided by CKAN, the team recommends Liferay portal, the most widely used portal platform in Ethiopian government organizations.

The team recommends DKAN platform as a secondary option. DKAN is a PHP based platform that is easy to deploy and configure. It is compatible to CKAN and runs both on a Windows server and Linux platforms. Most of the requirements in the terms of reference can be met with DKAN. It has rich documentation and support community.

## 3 Summary of Requirements

### Introduction

This section presents a summary of the requirements for the Ethiopian government open data portal. The requirements in the terms of reference of the project, the requirements collected during the assessment of the government ministries/agencies, and the requirements identified in the desk research have been used to produce this consolidated requirements for implementation of the Ethiopian government open data portal.

### Functional Requirements

#### Requirements collected from ministries/agencies during the assessment

1. The Ethiopian government open data portal should provide tools to publish datasets via different mechanism. The portal should enable users to add and edit data in many ways, including:
  - a. Directly via the web interface
  - b. Using API
  - c. Via custom spreadsheet importer, and other options
2. The portal should have customizable harvesting mechanism, which can fetch and import records from many different repository sources of government ministries and agencies that already have their data in repositories.
3. The Ethiopian government open data portal should provide as many options as possible for browsing of the datasets. Accordingly, the datasets should be indexed using the following categories:
  - By ministry/Agency
  - By regional state
  - By sector/theme
  - By file Format (CSV, JSON, XML, ODF, RDF, etc...)
  - By resource Category (dataset or app)



- By data frequency (Daily, weekly, monthly, quarterly, annual...)
  - By keyword
4. The following categories shall be used for ministry/agency and there should be flexibility to modify if there is a change in the name and/or structure of government ministries/agencies.

*Legislators*

- House of Peoples Representative
- House of Federation

*Ministries*

- Ministry of Foreign Affairs
- Ministry of Education
- Ministry of Health
- Ministry of Trade
- Ministry of Finance and Economic Development
- Ministry of Culture and Tourism
- Ministry of Agriculture
- Ministry of Defence
- Ministry of Communication and Information Technology
- Ministry of Justice
- Ministry of Federal Affairs
- Ministry of Construction and Urban Development
- Ministry of Transport
- Minister of Revenues
- Ministry of Water Resource
- Ministry of Mines and Energy
- Ministry of Labour and Social Affairs
- Ministry of Youth, Sport
- Ministry of Women's Affairs

*Agencies and Authorities*

- Disaster Prevention and Preparedness Agency
- Ethiopian Mapping Authority
- Privatization and Public Enterprises Supervising Authority

- Social Security Agency
- Ethiopian Insurance Corporation
- Ethiopian Science and Technology Commission
- Ethiopian Telecommunication Agency
- Micro and Small Enterprises Development Agency
- Maritime and Transit Service Enterprise
- National Bank of Ethiopia
- Commercial Bank of Ethiopia
- Transport Construction Design Enterprise
- Ethiopian Investment Agency
- Ethiopian Airlines
- FDRE Supreme Court
- Ethiopia Electric Power Corporation
- Ethiopian Revenues and Customs Authority
- National Archives and Library Agency
- Ethiopian Export Promotion Agency
- Environmental Protection Authority
- National Agriculture Input Authority

5. The following categories shall be used as themes/sectors for cataloguing purposes

- Agriculture
- Census and Surveys
- Commerce and Trade
- Economy
- Education
- Enforcement
- Environment and Forest
- Federal Affairs
- Finance
- Governance and Administration
- Health
- Industries
- Information and Communications

- Infrastructure
  - Labor and Employment
  - Parliament
  - Power and Energy
  - Rural
  - Social Development
  - Transport
  - Urban development
  - Water Resources
  - Governance of Open Data
6. Tools need to be provided to publish and find datasets Publish datasets. The tools shall support publishing via import and through a web interface.
  7. Search feature shall support searching using keywords, filter by tags and shall allow viewing dataset information at a glance
  8. Tools used to store and manage data shall support storing of raw data and metadata
  9. The system shall support visualization of structured data with interactive tables, graphs and maps whenever possible.
  10. The system shall support getting statistics and usage metrics for datasets
  11. The system shall support searching geospatial data on a map by area
  12. The system shall support managing the look and feel of the user interface using Themes with CSS or integrated themes and CSS of a CMS.
  13. The system shall have a feature to build a community with extensions that allow users to comment on and follow datasets.
  14. The system shall have a feature to allow online recommendation of datasets by users
  15. The system shall support integration with applications using Web service technology.

16. The system shall have a feature that enables users to manage, analyze and generate customized views with different possible parameters/indicators online, when possible.
17. The system shall have logs of users visits
18. The system shall have a feature that allows direct integration to already implemented data collection tools in organizations. (This is only possible if the data collection tools have been developed using the Ethiopian eGIF standard, or allows web services technology.)
19. The system shall allow multimedia uploading features
20. The system shall allow uploading of relevant manuals or documents with the dataset
21. The system shall have Mobile apps for selected datasets to enable accessing the data using mobile devices.
22. The system shall be scalable to enable inclusion of future requirements

### **Requirements presented in the TOR of the project**

23. The portal shall provide government ministries and agencies capability to manage members, datasets, approve datasets to be public, manage harvest sources all from each ministry/agency's admin page.
24. The portal shall provide government ministries and agencies capability to create ministry/agency specific forms that pre-fill certain fields or have additional required fields to fit individual requirements.
25. The portal should allow a workflow to categorize datasets as public or private. If they are private they are only visible to the logged in members of their owner ministry/agency.

26. The portal should enable users to search datasets using keywords and browse the dataset collection using different tags. It should also enable users to be able to see dataset information at a glance with their formats and license type.
27. The Ethiopian government open data portal should provide tools to store & manage raw and metadata in a pre-specified format. As well as holding metadata and links to the offsite data, the portal should provide secure storage for the data itself. When creating the dataset or resource, it should provide options to either link to data hosted elsewhere, or upload it in the same action as registering the metadata.
28. The portal should allow storage of data in any format.
29. When structured data, e.g. when a spreadsheet is uploaded, the system should provide an API for the data itself, allowing users to query, retrieve and use data instantly from datasets without needing to download or process it first.
30. The portal should provide its own visualization tools to display data previews, graphs and visualizations of the data on the dataset resource page.
31. The portal should provide users the ability to download public datasets for mobile applications
32. The portal should provide users the ability to visualize structured data with interactive tables, graphs, and maps. The data previewing tool of the portal should have a host of features for previewing data stored as follow:
  - **Table view:** The system should allow a range of ways to view and process structured data and display it as a table. It should allow users to sort the data on particular columns, filter or facet by values, or hide columns entirely.
  - **Graphing data:** It should also allow displaying the data on a graph, choosing the variables on the axes and comparing a number of variables by graphing them together on the same y-axis.
  - **Mapping data:** the system should allow plotting data points in a map if the table has columns with latitude and longitude. Users should be able to pan drag and

zoom a particular point in the map. It should also enable displaying of all the field values in the corresponding row by selecting a data point.

- **Image data:** data previewing should not be restricted to tabular data. Common image formats should be able to display.
33. The portal should provide tools to generate statistics and usage metrics for datasets
  34. The portal should search geospatial data on a map by area
  35. The system should allow instant view of whether the data is available under an open license or not. This makes it clear to users whether they have the rights to use, change and re-distribute the data.
  36. The system should allow see what labels the dataset in question belongs to. Tags also allow for browsing between similarly tagged datasets in addition to enabling better discoverability through tag search and faceting by tags.
  37. The system should allow publishing datasets in multiple formats
  38. The system should provide API key that allows access every metadata field of the dataset and ability to change the data if you have the relevant permissions via API.
  39. The portal should have a feature that enables the formation of a community which will allow users to comment on and follow datasets
  40. It should enable users to manage, analyse, generate customized reports with different possible parameters online; control and record users interaction to the system; direct integration to already implemented data collection tools in ministry/agency
  41. The Ethiopian government open data portal should provide tools for online data processing and generating indicators
  42. The Ethiopian government open data portal should provide tools online for recommendation of datasets by users

### **Recommended site organization of the Ethiopian government ODP**



## Non-functional Requirements

The following are the non-functional requirements suggested by eSystems Africa’s project team based on analysis of requirements and past experience in developing similar portals for the government organizations. These requirements make the open data portal more reliable, secure, stable and dependable.

### Performance/efficiency requirements

**Availability:** measures to track the availability of data sources.

**Device Independent:** measures to enable device independence.

**Feedback from Stakeholders:** a mechanism to collect recommendations on how to improve the platform from all stakeholders (e.g., a repository of ideas and applications that rely on government data).

**Scalability:** The platform must be federated across several administrative boundaries and across several data repositories.

**Performance:** The system shall be fast so that users will be able to get information easily without wasting much time. The portal solution shall have the facility to tune for optimal performance under load.

**Portability:** Since the system is going to be developed on open source software it shall be platform independent and it shall be designed and tested on different browsers like Mozilla fire-fox, Internet Explorer, Opera, etc.

**Stability:** The portal solution shall be stable. The portal product used shall be a stable version and should not have frequent failures or require a lot of bugs to fix during operation of the portal.

**Usability:** The portal shall be easy to use and navigate from page to page. The user interfaces shall be user friendly.

**Reliability:** The portal solution shall have support for redundancy, failover and load balancing to ensure maximum uptime of the portal.

### Security

The system shall have an access control system and facilities on the critical part of the portal so that only authorized users can access some parts of the portal like membership management, participation on discussion forums, updating data, etc.



## Documentation

**Platform User Manual:** a user manual that addresses the needs of data contributors, data consumers, and platform administrators.

**Platform Architectural Design Document:** a platform architectural design document to provide details suitable for platform engineers, who aim to customize and/or further extend the capabilities of the platform.

**Platform Software Design Document:** a platform software design document, in order to provide support for platform engineers, who aim to customize and/or further extend the capabilities of the platform. In particular, listing all software tools (e.g., name, version, and additional reference information) and their dependencies that make up the platform should be provided.

**Document Source Code** For any source code produced, proper documentation (according to an established standard) should be included to ease maintainability of the code.

**Listing of Available APIs:** Provide a list of available APIs

## Legal Requirement

**Disclosure of Licensing Terms:** All data sets, should, preferably must, include associated licenses that clearly specify the licensing terms that govern the usage of the data, by third parties.

## Privacy Requirement

**Ensure/Preserve User Privacy:** The platform must preserve user privacy.

## (Data) Quality Requirements

**Cleanliness:** The data should (ideally, must) be clean in order for it to be usable and reliable.

**Timeliness:** The data should be timely for it to be suitable in critical situations.

**Feedback Mechanism:** The platform should provide a feedback function to facilitate the collection of user recommendations for improvements.

## Language Support

The system shall have multilingual support.

## Interoperability

The portal to be implemented shall be interoperable with standard portal products. It should be compatible with the Ethiopian EGIF (Government Interoperability) standard.

## Metadata Requirements

The Ethiopian government Open Data portal must use standardized metadata. Accordingly, the adoption of Data Catalog Vocabulary (DCAT) as a standard Metadata Schema by Ethiopian government ministries and agencies is recommended. The rationale for the recommendation of DCAT includes:

- DCAT is one of the most widely used Semantic Web vocabularies for describing datasets and data catalogues. It is being used in a number of Open Data portals in developed and developing countries including the US, Europe and many Asian countries.
- Although the standard DCAT is lacking geospatial metadata coverage, the EU DCAT-AP Working Group has identified the need to describe geospatial datasets, data series, and services. As a result, that Group has developed GeoDCAT, an extension of DCAT. It provides an RDF syntax binding for the union of metadata elements defined in the core profile of ISO 19115:2003.
- It incorporates most of the major metadata features. This Metadata Schema includes the following three categories of metadata:
  - Mandatory
  - Recommended
  - Optional
- The following is a list of tools that support the export of data using DCAT: CKAN, EsriGeoportal Server, Neologism, and DKAN. CKAN and ESRI map their metadata to DCAT, whereas Neologism exports its data (vocabularies descriptions) in ADMS (an application profile of DCAT). DKAN also embeds DCAT metadata in HTML pages.

An extracted Reference Guide to DCAT, which is recommended as a standard for Ethiopian government open data portal, is available in the table below

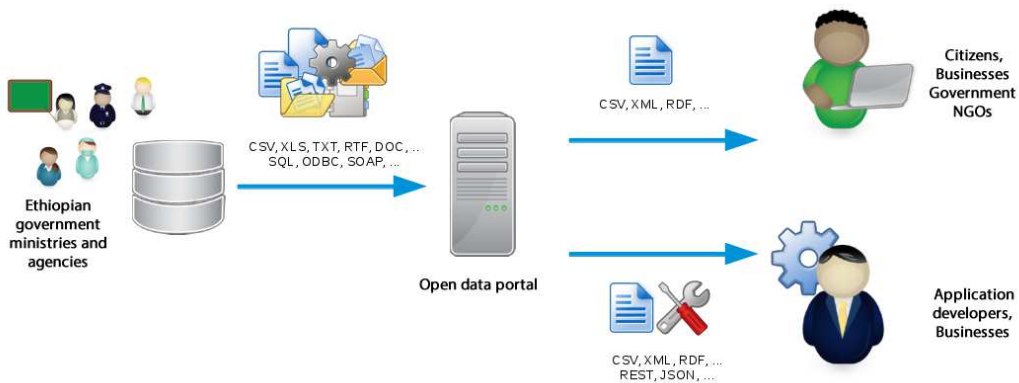
**Table 10.DCAT-AP Quick Reference of Classes and Properties. Extracted from DCAT-AP Specification Final v1.01**

Class	Class URI	Mandatory properties	Recommended properties	Optional properties
Catalogue	dcat:Catalog	dcat:datas etdct:desc riptiondct :publisher dct:title	dct:issueddct:language dct:licensedct:modified dcat:themeTaxon omyfoaf:homepage	dcat:recorddct:rightsdct:sp atial
Dataset	dcat:Dataset	dct:descri ptiondct:t itle	adms:contactPointdc at:distributiondcat:ke yworddcat:themedct: publisher	adms:identifiernadms:versio n adms:versionNotesdcat:l andingPagedct:accrualPeri odicitydct:conformsTo dct:identifierdct:issueddct:lang uagedct:modifieddct:spatia ldct:temporal
Distribution	dcat:Distribut ion	dcat:acces sURL	dct:description dct:formatdct:license	adms:statusdcat:byteSized cat:downloadURLdcat:me diaTypedct:issueddct:modi fieddct:rightsdct:title

In addition, DCAT-AP extended version should be used for Geospatial metadata elements to be included in the Ethiopian government open data portal.

## Use case representation

In this section, the most important requirements of the data portal have been represented using a Use case diagram. In the use case diagram, the main users of the system and functions that are required to be available in the system have been presented.



The main users of the open data portals have been categorized into the following four groups of users:

- *Data Consumers*

These are users of the open data portal that consume the datasets for some purpose, including research, analysis, reporting, decision making, developing apps, etc.

Data consumers include citizens, government, non-governmental organizations, businesses, app developers, etc. Each user may have its own

- *Data publishers*

These are Ethiopian government ministries and agencies that will be responsible for publishing government data on the open data portal

- *Organizational Data Administrators*

These are users that are responsible for configuring the data portal and administering the open data portal including managing users and organizations. The data administrators is also responsible for modifying the portal and creating content on the portal pages.

- *System Administrators*

These are IT staff of the national data centre that are responsible for installation, maintenance and configuration of the open data portals, create system users, produce statistics and reports.

The actions that will be carried out by the actors on the Ethiopian government open data portal are the following:

*Data consumers*

- Browse dataset
- View dataset
- Search dataset
- Download dataset
- Suggest dataset
- Request dataset

*Data publishers*

- Publish dataset
- Modify dataset

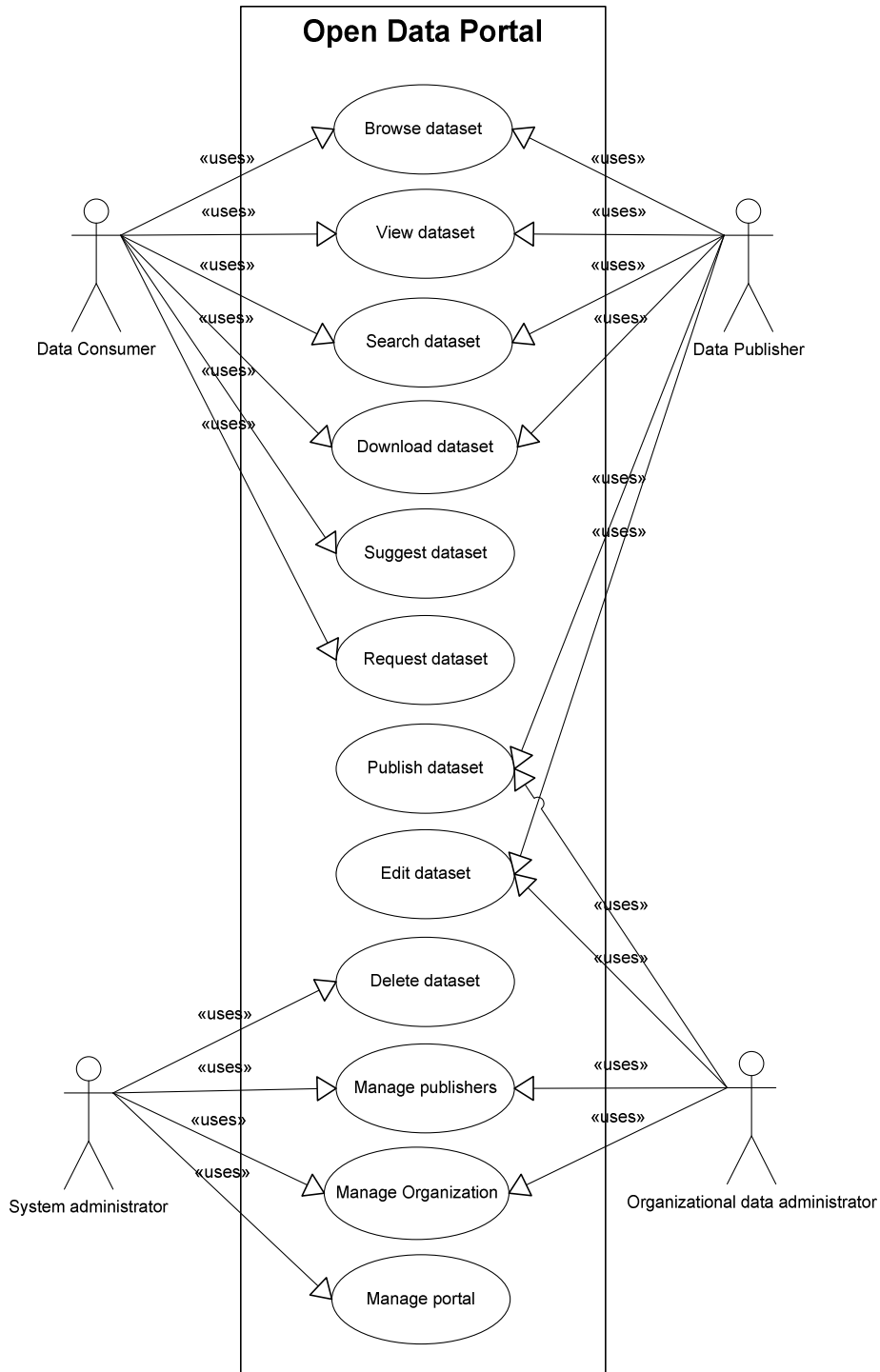
*System Administrators*

- Delete dataset
- Manage data publishers
- Manage organizations
- Manage portal content
- Configure portal
- 

*Organizational Data Administrators*

-

## Use Case Representation



The following table provides description of the use cases.

No.	Use case	Short Description
1	Browse dataset	Browsing the datasets in the open data portal.
2	View dataset	Viewing the dataset in the open data portal using various tools such as data explorer, visualization, apps, applications, maps, etc.
3	Search dataset	Searching required datasets in the portal including searching on a map
4	Download dataset	Downloading the datasets existing in the data portal in available formats such as MS Excel, CSV, XML, JSON, RDF, etc.
5	Suggest dataset	Suggesting a dataset that does not exist in the open data portal
6	Request dataset	Request a dataset from the open data portal so the system administrator communicates with the requester with the response.
7	Publish dataset	Publishing a dataset including uploading the data or specifying links, specifying metadata, attaching related files, etc.
8	Edit dataset	Editing the dataset including changing the data, the metadata, attaching additional resources, deleting dataset, etc.
9	Delete dataset	Permanently deleting the dataset from the open data portal.
10	Manage Publishers	Creating, editing or deleting dataset publishers, changing the roles of data publishers.
11	Manage Organization	Creating, editing or deleting organizations
12	Manage Portal	Creating content in the portal, changing portal features, moving datasets, creating statistics and reports, etc.

## 4 Conclusion and Next Steps

### Conclusions

As one of the major components of its e-government strategy, the government of Ethiopia pursues to expand access to government data, ensure transparency and accountability, and strengthen citizen engagement in the activities of government. The government expressed its commitment to transparency through its constitution and passing a legislation of Freedom of the Mass Media and Access to Information (590/2008).

The government of Ethiopia recognizes the social and economic value of free and open access to government data to the public while respecting privacy, security, and confidentiality restrictions. Opening government data represents an essential change from the traditional way of government functions, which requires a high-level commitment, legal and policy support, standardization of data formats and metadata across ministries and agencies, and data management capacity building at each ministry.

The purpose of the current situation assessment was to investigate the legal/policy environment for open data and identify high value datasets in Ethiopian government ministries and agencies to be published for the public. Current standards for metadata and options for cataloguing were investigated. The technology options to develop the open data portal were also rigorously investigated.

The assessment revealed that overall there is an inappropriate legal environment for open data. Other than the constitution, the government enacted the “Proclamation to provide for Freedom of the mass media and access to Information” that provides all persons the right to seek, obtain and communicate any information held by public bodies, except exempted information therein. It also provides details on the rights to access information, duty to publish government data/information, and operations of public bodies in the process of providing information to the public. However, detailed open data policy or guideline needs to be developed for the smooth implementation of the open data project and it will be a key enabler for openness and transparency.



However, it does not provide for proactively publications of raw data and electronic information. The law is a vehicle for reactive disclosure where a question/request has to be filed before an answer is given and that public information must be requested before it is disclosed. Therefore, a subsidiary legislation that builds on the Mass Media and Freedom of Information law needs to be enacted in order to addresses access to electronic and raw data and provide for the proactive release of public information before individuals requests it. The law will update and improve access to information and specify that new data sets and records to be collected and published.

Regarding ownership, there is no clearly defined ownership of government data. There is no a general government view and each agency is holding its own data and regards it as a strategic resource for its existence. All ministry websites include a copyright notice by that individual ministry, and not the federal government. The absence of an open government license means, institutions continue to “own data” and license it individually. There is no experience in metadata and a variety of “core common reference datasets.”

The implementation of a Government Enterprise Architecture and open data portal is expected to alleviate the problem of silos and encourage a holistic view of government ownership of data, while leaving individual organization to maintain specific information on behalf of the government.

The Ethiopian Statistics Agency (ESA) is the key provider of statistical data. The agency generates a large amount of data through its own portal, even though the data is aggregated. ESA has made preliminary efforts toward releasing large quantities of data including the results of 1994 census and the projections in aggregated format. The current proclamation to establish the central statistics authority, proclamation No. 442/2005 provides for disclosure of information and access to reports and publications is not constrained. Hard and soft copies of reports are sold on a cost recovery basis. Basic results of the surveys are also available on the web. However, the law does not provide for raw data. Access to the micro data is administered by a data release protocol, which describes the principles and costs of users accessing the unit record information.

The survey indicates that although efforts are made by many organizations to provide access to data, these do not use common standards. Much of the data has not been updated with the latest data series, hampering sustainable and ongoing data re-use. There is much work to do to increase the availability of open data.

There is no known governance framework for open data in Ethiopia. However, efforts are underway to coordinate the common data sets of the government. The open data initiative aims to address the governance issues. The open data initiative will create national body that spearhead open data initiative and create a participatory process to encourage the publication of open data across the government.

## Way forward

### Policy/legal and Organizational Recommendations

1. Secure high level political commitment
2. Create clear roadmap and have everyone agree on it
3. Develop and launch the coordination mechanism – make it operational and sustainable. In providing data, cooperation is more important than going it alone. Building a national open data platform is crucial.
4. Start small make it big in the end ...
5. Strategize new objectives of open data and raise awareness
  - a. The government should work rigorously to create awareness and communicating open government data to specific target groups – motivating data providers, demonstrating usefulness to the public
  - b. All relevant government ministries/agencies should be brought on board in open data activities. A series of awareness raising workshops needs to be organized.
6. It is important to systematically promote data skills – reducing shortages of skilled workers in government ministries and agencies and narrowing the digital divide
7. It is vital to actively offer open data and paying attention to demand of citizens, business, and the civil society

8. Follow a genuine Open-by-default approach for implementation across government ministries and agencies will need to start embedding Open Data in every day's tasks, moving to a model in which Open Data is not just the final result of a process but the very way processes are run.

### **Datasets and Metadata**

- When publishing Open Data, adopt the DCAT metadata dictionary nationally and ensure its full implementation
- For a dataset to be considered as Open Data, it must be published under an Open License. MCIT should issue guidelines on recommended licenses for datasets. The guideline should encourage the use of open licenses.
- While data published in any format can be considered Open Data if associated with an Open License, the type of data format used can have significant implications for the usability of the data. Under the Open Data Initiative, government ministries/agencies should publish their data in the most open way possible. One way to measure the openness of the formats used is through the 5-star deployment scheme for Open Data. The greater the number of stars, the more reusable the data.

### **Technology**

- The assessment team recommended both CKAN and DKAN as a platform for development of Ethiopian government open data portal. A discussion needs to be conducted with the project counterpart team at MCIT and select the platform by openly discussing the merits and demerits of both platforms.

## References

- Bloomberg, M. R., & Merchant, R. N. (2012). Open Data Policy and Technical Standards Manual, (September).
- Capgemini Consulting. (2013). The Open Data Economy Unlocking Economic Value by Opening Government and Public Data Only Few Governments are Leveraging Open Data for Economic Benefits.
- Cygniak, R., Maali, F., & Peristeras, V. (2009). dcat: An RDF vocabulary for interoperability of data catalogues.
- Farhan, H., Alonso, J., Davies, T., Tennison, J., Heath, T., & Berners-lee, T. (2013). Open Data Barometer, 1–45.
- Foundation, O. K. (2012). Open Data Handbook Documentation.
- Germany, G. (n.d.). White Paper Framework for Open Government Data platforms Open Government Data, 1–12.
- Government, M. C. (2014). Montgomery County Government Open Data Operations Manual.
- Government, T. S. (n.d.). Open Data Consultancy Final Report, (November 2013).
- Government, U. (2012). *Open Data White Paper Unleashing the Potential Unleashing the Potential Open Data White Paper*.
- Hohn, N., & Day, N. (2014). EU Open Data Portal, (January).
- Hoppin, A., Byrnes, A., Couch, A., & Ideas, N. A. (2013). Open-Source Open Data Platforms The Proprietary SaaS Competition ...
- Indonesia, R. A. (n.d.). Open Government Data.
- Lee, A. D., Cygniak, R., & Decker, S. (2014a). Open Data Ireland: Best Practice Handbook, (May).

Lee, A. D., Cyganiak, R., & Decker, S. (2014b). Open Data Ireland: Open Data Publication Handbook, (May).

Makers, D. (n.d.). *Linked Open Data: The Essentials*.

Opare, E. (2013). Accra Hackathon – Lessons Learned.

Plan, M., & Services, G. C. (n.d.). National Data Set-Design Document Master Plan, (April 2015).

Sharing, N. D. (2014). Open Data Implementation Guidelines, (February).

Ubaldi, B. (2013). OECD Working Papers on Public, (22).

## Appendix A: List of Institutions

1. Ministry of Finance and Economic Development
2. Ministry of Education
3. Ministry of Health
4. Ministry of Transport
5. Ministry of Trade
6. Ministry of Communication and Information Technology
7. Ministry of Agriculture
8. Ethiopian Metrological Agency
9. Ethiopian Mapping Agency
10. Ethiopian Anti-Corruption Commission
11. Ethiopian Chamber of Commerce
12. Ethiopian Revenue and Customs Authority
13. Ethiopian Commodity Exchange Authority
14. Document Authentication and Registration Office
15. Central Statistics Agency
16. Parliament
17. National Electoral Board
18. Supreme Court
19. National Bank of Ethiopia
20. Agricultural Transformation Agency
21. Ethiopian Institute of Ombudsperson

## Appendix B: List of Document Consulted

1. Federal Democratic Republic of Ethiopia Freedom of the Mass Media and Access to Information Proclamation No. 590/2008
2. Federal Democratic Republic of Ethiopia. “Central Statistics Agency Establishment Proclamation No.442/ 2005”.
3. Central Statistics Agency. DIRECTIVE No. 1 /2004. “DIRECTIVE ISSUED TO ESTABLISH PROCEDURES FOR ACCESSING RAW DATA TO USERS”
4. Central Statistics Agency. Form for requesting access to raw data.
5. MCIT: e-Government Strategy
6. MCIT: e-Government Strategy and Implementation Plan – Report January 2011
7. MCIT: The National Information and Communication Technology Policy and Strategy
8. FDRE-MCIT: Information Exchange Standard
9. MCIT: Government Information Systems Interoperability Standard (e-GIF)
10. MCIT: National Dataset Design Document
11. MCIT: National Dataset Design Current Situation Assessment Report

## Appendix C: Survey Questionnaire



Open Data Portal Development Project  
Current Situation Assessment  
Data Collection Tool

### General Information

Name: \_\_\_\_\_

Organization : \_\_\_\_\_

Title: \_\_\_\_\_

### 1. Policy and Legal Framework for Open Data

1. Do you know of any laws of the government in provision of online data? For example
- Does the government have a governance process for information and data security?
  - What policies or standards exist on the provenance, accuracy and quality control of data?
  - What policies/laws exist on use of languages? Do all government publications and data have to be in multiple languages?
  - Is there a policy for data archiving?



2. If the policies and laws do not exist, what do you think the best approach will be for ensuring data availability through policy and legislation (enacting legislation, new guidelines, policies, etc.)? Who should take the initiative for open data policy (political, executive, ministries, etc.)?

3. How should the government address public or institutional concerns for privacy and security? How did you address these concerns within your organization?

5. Are there established metadata standards, and are data holdings described by accurate metadata records? Are there established “core common reference datasets” which are used across government (e.g. organization codes, address register)?

## 2. Organizational Content

Do you provide data sets and information to public?

List all the data sets below:

---

—

---

—

---

---

---

---

<i>State of the data sets</i>	
In general, are the data sets in digital and reusable form?	
What is the format of the data? Is the data provided in open formats? e.g. XLS, XLSX, ODS, CSV, etc. Is the data structured?	
Are the data available as a complete set (eg raw data download) or aggregated?	
Are metadata –formal description of the data– provided along with the data?	
How often are the data sets updated?	

### Technical approach to data

1. If you currently provide data for the public, what is the technical platform that supports your open data service? Why did you choose that platform?

2. Are the data sets available online? Please provide the URL:

---

3. Are the technical services provided to develop and support the open data initiative in house or by third parties?

4. Is there a license to grant the right to use, download, and reproduce the your organization's publicly available data? Are there any restrictions imposed by the license? Are there any other usage requirements and do you see them as barriers to use?

5. Do you prescribe terms and conditions of use for the online data? Can you share a copy of the terms of use of your online data sets?

6. How is the quality of the data guaranteed, in terms of consistency, permanency, and updating? What are the processes to guarantee the quality of the data? Is there any explicitly stated quality agreement with the users?

7. What channels are provided for the users to access the information –i.e. portal, mobile apps, direct download, APIs, etc.? Can bulk and raw data be downloaded directly using machine-to-machine protocols?

### **3. Organizational Analysis**

1. Do you work with any other organizations, public or private, on your open data measures?

2. What different departments/areas of expertise do you have within your organization? What are the responsibilities of each with regards to open data?

3. Who provides the open data content within your organization? Is there a central section to which the data provided and uploaded by different departments?

4. Do you have organizational policy and mechanism for screening data that should be available to public? What degree of oversight do you have in regard to what data is published/made available, etc.?

5. Do you have experts in data management? What relevant skills, capacity and experience does your organization has in data management? If you do not have that how are you intend to upgrade your skills?

#### **4. Financing Data Availability**

1. Do you have resources for making data open? Are you responsible for obtaining funding from sources such as grants? What percentage of funding comes from public and other sources?

#### **5. Stakeholder Engagement**

1. Are citizens, businesses, academia, nonprofits, and civil servants involved in the open data initiative? What actions were taken to involve them?

2. Has your organization taken steps to make open data accessible to diverse populations (i.e. providing materials in languages other than English?)

3. What about those with disabilities, those who live in remote areas with little access to the Internet or technology, those who cannot afford to pay for internet connection, or those who do not know how to use the internet?

4. Have you created secondary policies to include diverse groups of the population? If yes, which type of policies are you implementing or planning to implement?

6. Assessment of the Benefits of Open Data

1. What are the main benefits your organization has experienced as a result of the making data available to users? Have you conducted any evaluation of those benefits?

2. Some experts advise that government efficiency increases with data transparency. Do you feel that has been the case in your organization? If yes, would you provide a concrete example?

3. Have you noticed an increase in the number of citizens' requests, letters, e-mails or complaints?

4. Are you aware of companies or organizations that use your datasets to develop new businesses or applications? Have you received requests from the private sector or any other organizations that would like to use those datasets? Would you provide details of a concrete case?

## **7. Possible Challenges for Opening Data**

1. What have been the main challenges to making data available for public? How has your office tried to solve them?

2. Do you anticipate future challenges? How are you planning to address them?

3. What would be your recommendation for the government in trying to implement open data portal?

## **8. Participation in the National Portal:**

1. Are you willing to participate in the national portal where your organization data will be presented for access to the public? If not, why not?

2. What kind of features would you like to see in the Open Data Portal?

1.1. Tools to publish and find datasets Publish datasets

- a. via import ( )
- b. through a web interface ( )
- c. Both ( )

1.2. Search by:

- e. keyword ( )
- f. filter by tags ( )
- g. See dataset information at a glance ( )
- h. All ( )

1.3. Tools to store & manage data:

- d. Store the raw data ( )
- e. Store metadata ( )
- f. Both ( )

1.4. Visualize structured data with interactive:

- e. tables ( )
- f. graphs ( )
- g. maps ( )
- h. All ( )

1.5. Get statistics and usage metrics for your datasets ( )

1.6. Search geospatial data on a map by area ( )

1.7. Theme the portal with CSS or integrate with a CMS ( )

1.8. Build a community with extensions that allow users to comment on and follow datasets ( )

1.9. What other features do you want in the portal?

---

---

## Appendix D: Summary of Collected Data

Data 1: Technical Approach

	Ministry of Finance and Economic Development	Ministry of Education	Ministry of Health	Ministry of Transport	Ministry of Communication and Technology	Ministry of Agriculture	Ethiopian Metrological Agency	Ethiopian Mapping Agency	Ethiopian Anti-Corruption Commission	Ethiopian Chamber of Commerce	Ethiopian Revenue and Customs Authority	Ethiopian Commodity Exchange Authority	Document Authentication and Registration Office	Central Statistics Agency	Parliament	National Electoral Board	Supreme Court	National Bank of Ethiopia	Transformation Agency	Agricultural
Metadata Standards	No	Yes	No	No	Yes	Yes	Yes (GeoNet)	Yes	No	No	No	No	No	Yes (DDI, RI)	No	No	No	No	No	No
Availability of core common reference datasets	No	Yes (OC)	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No
Technical platform		Php				Joomla	Java and por.Joomla	Mojo (open source) and Joomla						NADA, ETHioinfo ai Joomla	Java and SQL server					
URL	www.mof.gov.et	www.moe.gov.et	http://www.mot.gov.et	www.EAP.gov.et	www.emapc.gov.et	www.feac.gov.et	www.ethiopianchamber.com	www.ecx.gov.et	www.daro.gov.et					www.hwww.electicwww.fswww.nbe.gov.et						
technical services development and support	Third part	Inhouse	Third party	Inhouse	Third party	Both	Third party	Both	Inhouse	Both	Both	Third party	Both	Inhouse	Third party	Both	Inhouse	Third party	Both	Inhouse
Availability of licenses	No	No	No	Yes (Open)	No	No	Yes	No	Yes	No	Yes	No	No	Yes (6 differ)	Yes (op)	No	No	Yes	No	No
Prescription of terms and conditions of use	No	No	No	No	No	No	Yes	No	No	No	No	No	No	Yes	No	No	No	No	No	No
Existence of data quality standards	No	No	No	No	No	No	Yes	No	No	No	No	No	No	Yes	No	No	No	No	No	No
Access Channels to Users	Portal	Portal	Direct dc	Portal	Portal	Portal, Radio, Portal, CD, D	Portal, CD, D	Portal	Portal	Portal	Website, Portal	Portal, web	Portal	Portal	Portal	portal a	Mobile, dire	Portal, Mo		



Data 2: Organizational Context

	Ministry of Finance and Economic Development	Ministry of Transport	Ministry of Trade	Ministry of Communication and Information Technology	Ministry of Agriculture	Ethiopian Metrological Agency	Ethiopian Mapping Agency	Ethiopian Anti-Corruption Commission	Ethiopian Chamber of Commerce	Ethiopian Revenue and Customs Authority	Ethiopian Commodity Exchange Authority	Document Authentication and Registration Office	Central Statistics Agency	Parliament	National Electoral Board	Supreme Court	National Bank of Ethiopia	Agricultural Transformation Agency
Partnership with other gov organizations on Open Data	Yes	No			Yes	Yes		Yes	Yes		No	No	Yes	No	No	No	Yes	Yes
Who provides the open data content	Economics, Department	from trade registra	Each dir	Met	theoretical	c	Different de	Other govt orgs		Department	Departments				Different depa	Differen	Different pr	
Is there a central section to which the data provided and uploaded by different departments	Yes	Yes	Yes		Yes	Yes		Yes	Yes		Yes	Yes	Yes	Infor	Yes	branc	informal	Yes
Availability of organizational policy and mechanism for screening data that should be available to public	No	No	No		No	No		Yes	No		Yes	No	Yes	No	No	Yes	No	No
Availability of experts in data management	No	No	Yes		No	Yes		No	Yes		Yes	No	Yes	No	Yes	No	Yes	
Involvement of citizens, businesses, academia, nonprofits, and civil servants in the open data initiative	No	No	No		No	No		No	No		No	No	Yes	No	Yes	No	No	No

Data 3: Dataset of Government Ministries and Agencies

	Ministry of Economic Development	Ministry of Education	Ministry of Health	Ministry of Transport	Ministry of Trade	Ministry of Communication and Information Technology	Ministry of Agriculture	Metrological Agency	Ethiopian Mapping Agency	Ethiopian Anti-Corruption Commission	Ethiopian Chamber of Commerce	Ethiopian Revenue and Tax Authority	Ethiopian Commodity Exchange Authority	Ethiopian Authentication and Registration Office	Central Statistics Agency	Parliament	National Electoral Board	Supreme Court	National Bank of Ethiopia	Agricultural Transformation Agency
No. of available Datasets	8	6	2	4	4	0	4	3	7	3	6	7	5	4	4	4	7	3		
Is it in digital and reusable form?	Yes	Yes	No	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
What format(s) is the data currently	Word, excel, pdf	Excel and PDF	PDF	PDF and Excel	PDF	Excel, PDF	PDF	Excel	Excel	Excel	PDF	PDF and PDF	Excel, PDF	Excel, PDF	Excel, PDF	Excel, PDF	Excel, PDF	Excel, PDF	Excel, PDF	Excel, PDF
Level of granularity	Raw	Raw	Aggregated	Raw	Raw	Raw	Both	Raw	Both	Both	Aggregated	Raw	Raw	Raw	Raw	Raw	Raw	Raw	Raw	Raw
Frequency of updates	Daily, Monthly	Annual	Annual	Monthly	Weekly	Daily, Monthly, Annual	Annual	Annual	Regularly	Weekly	Yearly, C	Monthly	Annual	Weekly	Annual	Weekly	Annual	Weekly	Annual	Regularly
Is the data available as a complete s	Yes	Yes	No	No	Yes	Yes	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No	Yes/No	Yes/No
Funding of data production and dist	Government	External funding	Government	Government	Government	External funding	External Government	Internal and External	Government	Government	Government	Government	Agency f	Government	External	Government	External	Government	External	Government
Is data free or charged-for?	Free	Free	Free	Free	Free	Free and paid	Free	Free and paid	Free	Free	Free	Free	Free	Free	Free	Free	Free	Free	Free	Free

Data 4: Portal Features

	Ministry of Finance and Economic Development	Ministry of Education	Ministry of Health	Ministry of Transport	Ministry of Trade	Ministry of Communication and Information Technology	Ministry of Agriculture	Ethiopian Metrological Agency	Ethiopian Mapping Agency	Ethiopian Anti-Corruption Commission	Ethiopian Chamber of Commerce	Ethiopian Revenue and Customs Authority	Exchange Authority	Ethiopian Commodity	Authentication and Registration Office	Document	Central Statistics Agency	Parliament	National Electoral Board	Supreme Court	National Bank of Ethiopia	Agricultural Transformation Agency
Tools to publish and find datasets																						
Publish datasets																						
a. via import	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	N	N	Y	Y	Y	Y	Y	Y	Y	Y
b. through a web interface	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Search by:																						
a. keyword	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
b. filter by tags	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
c. See dataset information at a glance	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Tools to store & manage data:																						
a. Store the raw data	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
b. Store metadata	Y	Y	Y	Y	N	Y	N	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	N	Y	Y	Y
2.4. Visualize structured data with interactive:																						
a. tables	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
b. graphs	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
c. maps	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	N	Y	Y	Y
Get statistics and usage metrics for your datasets																						
Search geospatial data on a map by area	No	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	N	Y	Y	Y	Y	Y	N	Y	Y	Y
Theme the portal with CSS or integrate with a CMS	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	N	Y	Y	Y
Build a community with extensions that allow users to comment on and follow datasets	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Other portal features																						
Able to users to manage, analyze, generate customized reports with different possib Mobile apps should t Uploading audio, video, Online data processing Online recommendation of																						