



HAL
open science

Local Government Open Data (LGOD) Initiatives: Analysis of Trends and Similarities Among Early Adopters

Eric Afful-Dadzie, Anthony Afful-Dadzie

► **To cite this version:**

Eric Afful-Dadzie, Anthony Afful-Dadzie. Local Government Open Data (LGOD) Initiatives: Analysis of Trends and Similarities Among Early Adopters. 18th International Conference on Informatics and Semiotics in Organisations (ICISO), Jul 2018, Reading, United Kingdom. pp.299-308, 10.1007/978-3-319-94541-5_30 . hal-01920725

HAL Id: hal-01920725

<https://inria.hal.science/hal-01920725v1>

Submitted on 13 Nov 2018

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License

Local Government Open Data (LGOD) Initiatives: Analysis of Trends and Similarities among Early Adopters

Eric Afful-Dadzie ^[0000-0003-1514-7825] and Anthony Afful-Dadzie ^[0000-0002-1699-6207]

OMIS Department, University of Ghana Business School
LG 78, Accra-Ghana
{eaafful-dadzie, aafful-dadzie}@ug.edu.gh

Abstract. This paper focuses on developments of open government data (OGD) at city, municipal, county, federal state, regional and provincial levels around the world. This is in line with recent OGD trends where local government authorities, especially in developed countries, are launching separate OGD web portals to complement central governments' efforts at liberating public data. Focusing on early LGOD adopters, an inventory audit of contents and functionalities in use at LGOD web portals was conducted. The data generated was analyzed for trends and (dis) similarities among early LGOD adopters. The results of the study points to a general sense of heterogeneity among LGODs across the world in terms of adherence to OGD web publishing standards. There is also a lack of uniformity in terms of OGD web portal functionalities and contents even among local government authorities within a same country.

Keywords: Open Government Data (OGD), Local Government Open Data (LGOD), Open Data, Early Adopters, Clustering.

1 Introduction

Though relatively new, there is a growing interest in Open Government Data (OGD) from governments, civil society groups, the media, researchers, among others. OGD which primarily seeks the liberation of government controlled data, began with President Obama's open data initiative of 2009 and was subsequently strengthened by the 2013 G8 Open Data Charter [1], [2], [3]. These initiatives sought to encourage countries to make open and for free, public data to the citizenry. To further facilitate the adoption of OGD, the Open Government Partnership (OGP) was created to bring together countries that have affirmed their willingness to provide easy and free access to public data. Typically, countries are admitted as members of OGP when they submit to a number of processes including, formal expression of interest by heads of sovereign states, endorsement of open government declaration, submission of a country action plan and finally, a commitment to an independent reporting mechanism¹. Member countries are then expected to launch an OGD web portal as a public data repository. Both the initial processes together with the opening of OGD web portals, often occur at the instance of

central governments. That is to say that, central government politicians and administrators are often the visible faces in these OGD projects. Such roles played by central governments have partly made central government open data (COGD) the main focus, often at the neglect of other variants.

Recently however, local governments through cities, municipalities, counties, federal states, regions and provinces have been launching their own independently operated OGD web portals. This is welcoming, since it complements central governments' efforts at releasing public data to entrench democracy and spur economic growth and innovation among citizens [4]. The idea of an independently-run local government open data (LGOD) is fast gaining momentum especially in the developed world. For instance, there are as many as 290 local authorities comprising of cities, counties, federal states, regions and provinces that are actively running independent OGD web portals at their local administrative levels. It must be noted that though central governments' role in OGD is crucial, local governments are the real policy actors when it comes to both the supply and demand sides of OGD implementation. This is true since most public datasets are first generated at the level of local authorities or agencies, while a functional OGD system also assists local governments to transform service deliveries through significant cost savings and regular evaluation of local services performance [5], [6]. This helps to actualize the value creation potential of OGD. We therefore argue in this paper, that an equal measure of attention should be given to local government open data (LGOD) as given to central government open data (CGOD). To generate more research and advocacy interests in LGODs, this paper focuses on activities of LGOD early adopters around the world. The paper first audits technical standards in use at LGOD web portals around the world. Data derived from the web content and functionalities audit are analyzed for trends and (dis)similarities among early local government-based OGD adopters.

Though researchers continue to report on OGD activities and initiatives at the level of central governments, very few works have focused on local government open data initiatives. None of the few published works on LGODs comprehensively reviews technical features of web portals for trends and similarities. The closest works in terms of the OGD web audit approach, came from [7] and [8]. The work by Chatfield & Reddick audited open data portals of twenty local governments in some of Australia's large cities. In [8], an audit of the content and functionalities of OGD web portals in seven countries in Africa. The work by [9], looked at the quality of open data web portals but only focused on CGODs. Similarly, [10] also looked at open data at local government levels but focused on determining factors that influence the success or failure of open data initiatives. This paper fills the gap by auditing web portals of 'independent' LGOD early adopters and further analyze the data to glean vital information regarding trends and (dis) similarities among them.

2 Methodology

The methodology is divided into two stages. In the first stage, an inventory audit of LGOD web portals was carried out benchmarked against widely accepted technical

standards for publication of OGD. Two of such technical recommendations come from the World Wide Web Consortium (W3C) authored by [11] and the World Bank’s open data toolkit [12]. In the second part, the data generated from the inventory audit was analyzed using association rules mining and K-Means clustering techniques to determine frequent trends and natural groups among the early LGOD adopters. The following sections present brief introduction to the audit and the techniques used in the analysis.

Table 1. Attribute selection for Inventory strategy.

	Feature	Audit strategy
LGOD web portal content and services	Data format	Open government data requires public organizations to publish data in a form that is easily accessible by data users. In particular, the W3C and the World Bank strongly recommend that public data are published in machine-readable non-proprietary electronic format. Non-proprietary machine-readable data formats ensure easy access and processing by data users. In the audit however, most web portals were found to provide both non-proprietary and some proprietary formats. Some of the common list of data formats audited were RDF, CSV, PDF, XLX(S), JSON, XML, KML, GeoJSON, ODS/ODF etc.
	Metadata	This is one of the essential OGD publication features which requires public organizations to accompany datasets with relevant information that describe the dataset. Some relevant metadata features could be the publication date of data, author attribution, and unit of measurements of data attributes among others. The audit checked for the presence or otherwise of metadata on LGOD web portals by randomly scanning through 5 datasets on each subject (category).
	Data currentness	Here, the audit sought to find the age of the most current data published on the portal. If the latest data is less than 30days old, it is regarded in this research as “current”. However, if the latest dataset is more than 30days old, then it regarded as “not current”. In the data collection, LGOD portals exhibiting current datasets were assigned the code 1 while a 0 for not current data sets.

LGOD web portal functionalities	Data visualization	Another requirement of OGD web portals is the provision of a data visualization tool. Such tools help data users to preview datasets before download. Visualization features come in various forms including tables, graphs, maps etc. The audit strategy checked whether the OGD portals provided this feature in the design or not. If this functionality was provided, the local authority is a 1 in the data collection stage. However, if not available, it is assigned a 0
	Data search	This OGD attribute seeks a data functionality that makes it easier for users to search specific information. Supported by the W3C and the World Bank, OGD web portals are expected to be searched by data types, subject and organization. The audit looked for the presence of this feature and scored a 1 for availability and a 0 for a non-availability of the functionality
	Links to external sites	OGD web portals are additionally required to embed links to external data sources. This serves as a guide to data users to access related data stored on other public servers. If an OGD web portal does not have an ownership right to a data, providing the external link aids data users to access it on a platform that holds the valid license.
	Social media integration	This World Bank Technical Option recommends the integration of social media plugins in OGD web portals. This in line with contemporary ways of engaging with people, ensures that users are able to share their views, experiences and general concerns on the web platform to be subsequently addressed. The feature therefore checks for the inclusion of at least one social media plugin. If this criterion is satisfied, the value 1 is assigned, otherwise a 0 indicates an absence.
	License	The World Bank's technical option on open data stipulates that OGD web portals display license for each dataset. In view of this, this feature checks for the existence of an open data license on the web platform. Typically, most OGD portals are licensed either under Creative Commons or the Open Data License. Some other standards are used by some LGODs.

2.1 Data Collection - LGOD Web Portal Audit Criteria

Attributes used in the study as shown in Table 1 were guided by requirements put forward by the W3C and the World Bank regarding OGD web data publication. Some common standards shared by the two bodies are (1) the need for datasets to be released or published in their raw forms (2) a metadata supplied for each dataset (3) data to be published in machine-readable, non-proprietary electronic formats such as CSV, JSON, XML, KML etc. (4) an open license to be provided by each data web portals and (5) a data visualization tool to guide data users. The criteria (attributes) for the audit were

divided broadly into two categories namely; web portal contents and functionalities and are shown in Table 1. This approach is similar to the one used in [8]. The list of local authorities (cities, counties, municipals, federal states, regions and provinces) were collected from open government data U.S. (<https://www.data.gov/open-gov/>) and Data Portals (<http://dataportals.org/search>). In all, 288 local governments were identified as operating independent OGD web portals different from their national portal. In this paper, local administrations around the world that have shown early interest in LGOD are classified as early adopters. During the web portal audit, about 15 LGODs were found to have nonfunctional Uniform Resource Locators (URLs) and therefore could not be accessed. This brought the total number of LGOD portals studied in this paper to 273. The data collection period was from November, 2016 to March, 2017. A web content analysis was carried out to determine how each local government authority fared on each of the criteria. To this effect, the web portal of each LGOD was examined to identify the features outlined in Table 1. Most of the functions and content of the portals were identified on the home page through visual examination. However, some required a thorough search through all the web pages as well as the html source codes to ascertain the presence or the absence of the feature in question. If a criterion is available on the web portal, the feature was assigned 1 to indicate its presence, otherwise, the feature was assigned a 0 to mean an absence.

2.2 K-Means Clustering

Clustering methods partition data points into homogenous groups called clusters. The K-Means clustering is an unsupervised algorithm that seeks to detect natural groups in unlabeled data. The term unlabeled is used to mean data that does not have pre-defined output [13]. In K-Means, the number of clusters are typically chosen apriori – meaning clustering intends to partition n objects into k clusters in which each object belongs to the cluster with the nearest mean. The goal of this algorithm is to find groups in the data, with the number of groups represented by the variable K . The algorithm works iteratively to assign each data point to one of K groups based on feature similarity. The rationale behind the use of the K-Means algorithm in this study was to find LGOD early adopters who share similar features in web content and functionality and to further determine whether they are in the same country or otherwise. The next section presents research questions to help elicit the right answers from the study.

3 Research Questions

The following research questions were used to guide the study. The questions are further analyzed in the result section.

RQ1. What are the global trends as far as local government open data initiatives are concerned?

RQ2. Which cities, federal states or provinces fall into unique natural groups and why?

RQ3. What do the similarities and dissimilarities among early LGOD adopters say about local government adherence to OGD web publication standards?

4 Results

This section provides answers to research question 1, which sought to understand the global trends as far as local government open data initiatives are concerned. Currently, as of 2016, there are about 27 sovereign countries around the world where local government authorities have launched completely separate OGD portals that are different from what their central governments operate. Conveniently referred to as ‘early adopters’ in this study, majority of them are in the U.S., Canada, Italy, Spain, United Kingdom, France and Australia (See Figure 1). There are however three LGOD adopters in South America (Brazil, Argentina and Chile) and in Asia (Taiwan, South Korea and China). No African city or region is currently implementing an OGD web portal separate from its country-level portal as of the time of this paper. Preliminary audit report as seen in Figure 1 shows that, the U.S. and Canada are currently the leading implementers of LGODs in the world.

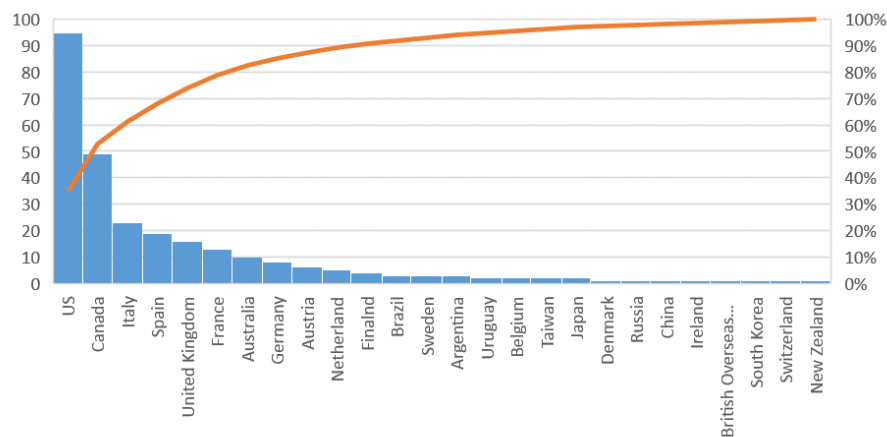


Fig. 1. Country representations of early LGOD adopters.

For instance, out of the 273 early LGOD implementers in the study, the US accounts for 95 of them, representing 34.8% while Canada accounts for 49 representing 17.95% as. The study further classified the early LGOD adopters into two main groups; cities/towns/counties on one hand, and provinces/federal states/regions on

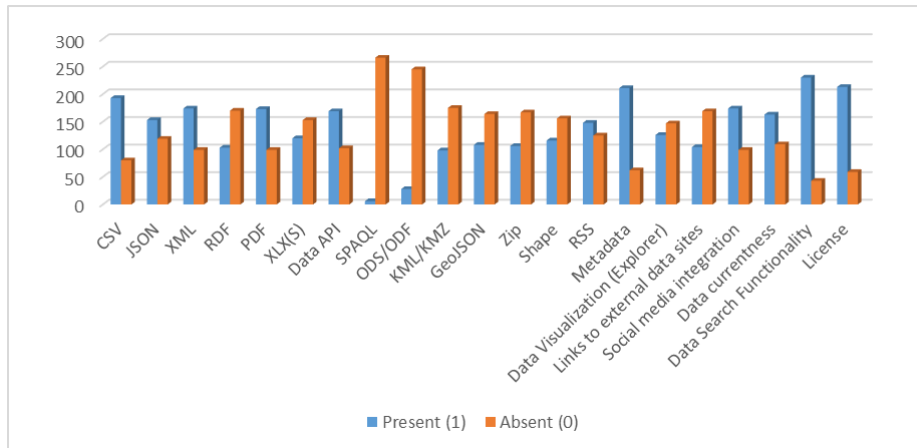


Fig. 2. Frequency count of web content and functionality features on LGOD portals

another. The classification which was simply based on size and level of autonomy of the local administration, found 73 LGODs belonging to the category of provinces/federal states/regions while 195 of them were cities/towns/counties. In terms of content and functionalities provided on LGOD web platforms, the audit data showed that on the whole, most LGOD early adopters are providing adequate OGD web functionalities and content that meet most of the standards put forward by W3C and the World Bank. For instance, in terms of web functionalities, the study observed that most local authorities, representing 84.25% were providing data search features, 78.02% provided open data licenses and 63.7% had social media plugins integrated in their web portals. Similarly, in terms of web content, most LGODs representing 77.3% provided metadata to accompany datasets, 59.7% had up-to-date data sets and majority had a wide range of data formats on offer to help data users to easily access, share and redistribute data. In particular, it was observed that the most used data format on most LGOD web platforms was the comma separated value (CSV), which was present on 70.7% of the LGOD web portals audited. Other notable data formats heavily in use are XML, data APIs and JSON data formats as shown in Figure 2. The use of the non-proprietary data formats is welcoming since it supports some of the ideals of OGD – to make public data progressively free and easily accessible. There are however other data formats that are not being giving much attention. For instance, majority of the LGODs were silent on geographical data formats such as GeoJSON, KML and shape files. The audit data further showed that, most of the geographical data formats were provided by local government authorities in Canada; mostly in the form of either KML of GeoJSON files. There were some equally important OGD web portal features that are not particularly being paid attention to. Some of these are data visualization (previews) and links to external data sites. These requirements are both suggested by the W3C and the World Bank as key to aiding data users in their search of public data. Data visualization for instance ensures that the user previews the data with graphical tools before download. In spite of some of the short comings with LGOD web portals, there is a general satisfiable trend by most of the early adopters to adhere to the international standards for publishing OGD.

Overall, the level of adherence to international standards as far as LGOD's are concerned can be described as high.

The analysis further focused on detecting natural groups or clusters within the LGOD early adopters as stipulated in RQ2. The study sought four groups ($k=4$) in the data, separately for OGD content and OGD functionalities. In all, there were 73 LGODs in cluster 1, 63 in cluster 2, 66 in cluster 3 and 71 in cluster 4. The U.S and Canada dominated in all the clusters contributing a total of 94 and 49 local government authorities out of the 273 in our study.

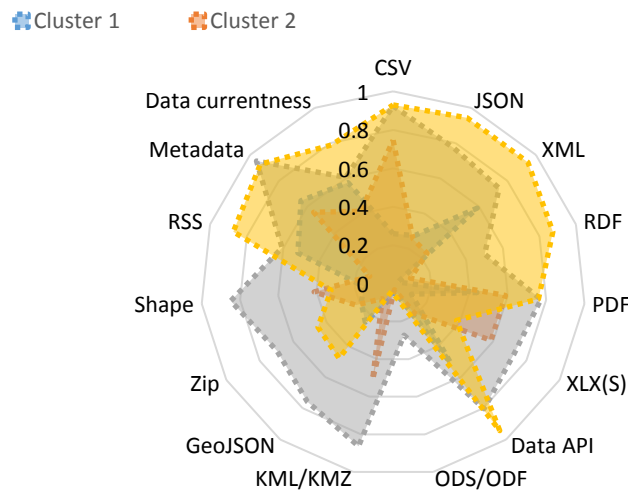


Fig. 3. Strength of LGODs in terms of OGD web contents

Overall, LGODs in clusters 3 and 4 respectively provide the most OGD features in terms of content and functionalities. Contributing to 16.67% of the local authorities in cluster 3, Canada leads in terms of the provision of OGD content benchmarked in the study. The U.S, Spain, Italy and France follow Canada in cluster 3. Similarly, in cluster 4, the U.S contributes to about 63.4% of the total LGODs, and therefore provides far richer OGD functionalities and services than all other countries observed in the study. Canada, Italy, Spain and France follow in that order in terms of OGD functionalities. Specific to OGD contents, the strength of LGODs in cluster 4 as shown in Figure 3, is the provision of features such as metadata, non-proprietary data formats and current datasets. For instance, in terms of data formats, LGODs in cluster 4 provide most of the recommended non-proprietary formats such as CSV, JSON, XML, RDF, as well as RSS and data APIs. However, LGODs in cluster 3 of which Canada dominates, focus more on geographic data formats such as Shape, GoeJSON and KML/KMZ. Geographic data is increasingly becoming one of the sought after data on OGD web portals and therefore adequate provision of appropriate Geo data formats would ease access of such data. Though overall, the clustering shows that, LGODs in clusters 1 and 2 lag

behind their counterparts in clusters 3 and 4, one significant observation seen in cluster 2, is the provision of the XLX(S) data format. This spreadsheet data format is proprietary, and therefore not recommended highly by OGD web publishing standards. In terms of provision of OGD web functionalities, clusters 3 and 4 again provided more OGD functionalities than the other clusters. For instance, the strength of LGODs in cluster 3 can be seen in the provision of open data licenses and links to external data sites. LGODs in cluster 4 on the other hand, provided more of data visualization and data search functionalities. LGODs in cluster 1 urged slightly above clusters 3 and 4 in terms of provision of social media integration plugins.

5 Discussion and Conclusion

The results of the LGOD web portal inventory audit give a general insight into the trends, relationships and similarities among LGOD implementers around the world. In terms of geographical distribution of LGODs, the apparent trend shows early adopters are mostly found in the global north. Significantly, North America represented only by the U.S. and Canada have a combined share of 41.4% of the total LGODs in the world. There are no cities, federal states, provinces or regions in Africa currently implementing the concept of a decentralized open government data. In Asia, one city each in China and South Korea and two in Taiwan are the early LGOD implementers. In South America, Brazil leads in terms of the number of early LGODs, followed by Argentina and Uruguay. Front runners of LGODs in Europe are Italy, Spain, U.K, France, Germany, Austria, Netherland, Finland, Sweden and Belgium in the order of number of local authorities involved. Australia represents the Oceania as the only implementer of LGODs. The LGOD web audit afforded the opportunity to find clusters or natural groups into which LGOD early adopters belong. The clustering scheme though not meant to be a score card of performance of the LGODs, give an indication of the strength of each cluster in terms of which OGD features they provide to data users. This approach is to aid policy makers to support seemingly struggling LGODs in terms of technical support and resources. For instance, local authorities in cluster 1, lag far behind in terms of provision of standard OGD features. Therefore, though a welcoming attempt by local authorities in cluster 1, it is observed that they are not adhering to international OGD publishing standards. Such cities, towns, federal states, provinces etc. would need to be supported to attain the right standards. There is a general lack of uniformity by LGODs in terms of the content and features they publish for their data users. The seeming heterogeneity among LOGD web portals is affirmed by [9] and [14]. For instance, even within the same country, the web template as well as the OGD features used tend to be different at the national, state and city levels. Apart from the U.S. which for the most part provides a common OGD web template for both its national as well as local authorities (federal states, cities, towns and counties), most of the 27 countries have their OGD web architecture different from their local authorities.

References

1. Tauberer, J. (2012). History of the Movement. Open government data: The book. <https://opengovdata.io/2014/history-the-movement/>
2. Attard, J., Orlandi, F., Scerri, S., & Auer, S. (2015). A systematic review of open government data initiatives. *Government Information Quarterly*, 32(4), 399-418.
3. Castro, D., & Korte, T. (2015). Open Data in the G8: A Review of Progress on the Open Data Charter. 2015-04-25]. <http://www.datainnovation.org/2015/03/open-data-in-the-g8>
4. Janssen, M., Charalabidis, Y., & Zuiderwijk, A. (2012). Benefits, adoption barriers and myths of open data and open government. *Information systems management*, 29(4), 258-268.
5. Rahily, L. (2014). *Innovation in Local Government: Open Data and Information Technology*. McKinsey Publications
6. Hammell, R. (2012). *Open data—Driving growth, ingenuity and innovation*. Deloitte White Pap.
7. Chatfield, A. T., & Reddick, C. G. (2017). A longitudinal cross-sector analysis of open data portal service capability: The case of Australian local governments. *Government Information Quarterly*.
8. Afful-Dadzie, E. and Afful-Dadzie, A., 2017. Open Government Data in Africa: A preference elicitation analysis of media practitioners. *Government Information Quarterly*, 34(2), pp.244-255.
9. Sáez Martín, A., Rosario, A. H. D., & Pérez, M. D. C. C. (2016). An international analysis of the quality of open government data portals. *Social Science Computer Review*, 34(3), 298-311.
10. Susha, I., Zuiderwijk, A., Charalabidis, Y., Parycek, P., & Janssen, M. (2015). Critical factors for open data publication and use: A comparison of city-level, regional, and transnational cases. *JeDEM-eJournal of eDemocracy and Open Government*, 7(2), 94-115.
11. Bennett, D., & Harvey, A. (2009). Publishing open government data. World Wide Web Consortium. (<https://www.w3.org/TR/gov-data/>)
12. Herzog, T. (2014). *Technology Options for Open Government Data Platforms*. World Bank Working paper. (<http://opendatatoolkit.worldbank.org/en/technology.html>)
13. Rai, P., & Singh, S. (2010). A survey of clustering techniques. *International Journal of Computer Applications*, 7(12), 1-5.
14. Zuiderwijk, A. and Janssen, M., 2014. Open data policies, their implementation and impact: A framework for comparison. *Government Information Quarterly*, 31(1), pp.17-29.