The value and challenges of providing and accessing Government open data in developing countries: Kenyan context from a citizen's perspective

Mary Ngure
Aga Khan University, mary.ngure@aku.edu

Soderlund Ochungo
Aga Khan University, soderlund.ochungo@aku.edu

The value and challenges of providing and accessing Government open data in developing countries: Kenyan context from a citizen's perspective

Mary Ngure
Library Department, Aga Khan University, Nairobi, Kenya
mary.ngure@aku.edu

Soderlund Ochungo
ICT Department, Aga Khan University, Nairobi, Kenya
soderlund.ochungo@aku.edu

Abstract:

In the recent years, Governments both in developed and developing countries are embracing open data initiatives as fundament al in facilitating government transparency, accountability, and public participation by making data freely available to the public. In addition, open data serves as an essential cornerstone in supporting technological innovation and economic growth by enabling third parties to develop new kinds of digital applications and services (Gray 2014; Ding et al. 2011; Shadbolt et al. 2012).

Despite the rising uptake of such initiatives, little has been written on the experience as well as the skills and knowledge for citizens in open data and technology environments. This paper seeks to fill this gap by presenting unique lessons learnt from the implementation of Kenya's globally acclaimed Open Data Portal which was launched in July 2011. Kenya forms an interesting study choice as the country was the first developing country in sub-Saharan Africa and the second on the continent after Morocco to develop the portal. The portal, powered by Socrata Inc, aims to make core government developmental, demographic, statistical and expenditure data available for researchers, policymakers, ICT developers and the general public.

The varying technological, economical, and cultural differences in Kenya significantly affect access and usage of the portal as seen in wide inequalities in technical expertise, internet access, and extent of use. In addition, there are various system and management challenges inhibiting the utility and ease of interaction of the Portal. These challenges include empty datasets, broken links, obsolete information, and lack of numerous datasets requested by the public which date back to over two years.

The authors who are Kenyan citizens explore the challenges and best practices learnt from implementation of Kenya Open Data Portal and discuss from a citizen's view point these unique and interesting findings and how they relate and contrast to other countries.

1. Introduction

Open data is based on the idea that some data should be made freely available to the public for use and republishing as needed without copyright, patents or other restrictions and thus anyone can freely access, use, modify, and share for any purpose (Auer et al. 2007). These initiatives have been embraced in recent years by governments both in developing and developed countries as an essential cornerstone in supporting technological innovation and economic growth. Open data helps facilitate government transparency, accountability, and public participation thereby enabling third parties to develop new kinds of digital applications and services (Brito 2008; Gray 2014).

Dataportals.org, an open knowledge project, listed 519 data portals around the world both in developing and developed countries as of May 2016. Some of the countries that have I implemented these initiatives span across the world including Japan, Kenya, United Kingdom and United States to mention but a few.

As the uptake of open data portals increase, an analysis of the skills, knowledge and experience of citizens in open data and technology environments is necessary. This paper interrogates from a citizen's perspective the Kenya Open Data Portal which was launched in July 2011.

2. Research objectives

This paper seeks to explore from a citizen's perspective the challenges and best practices learnt from implementation of Kenya Open Data Portal and how these relate and contrast with other developing as well as developed countries. Further, as there are wide inequalities in technical expertise, internet access, and extent of use across the country; the varying technological, economical, and cultural differences in Kenya and how they significantly affect access and usage of the portal will be discussed in this article with an aim of benchmarking with developed counties.

3. Research methods

The research method employed is an exploratory case study of Kenya Open Data portal conducted using a participatory focused research approach. The case study presents a detailed discussion of the implementation of the portal focusing on the challenges and best practices for successful implementation from a citizen's perspective. The authors are Kenyan citizens and this paper represents their analysis and summary.

4. Findings and Discussion

4.1 Introduction to Kenya Open Data Portal

Kenya Open Data portal has been globally acclaimed as the country was the first developing country to have an open government data portal, the first in sub-Saharan Africa and second on the continent after Morocco. The initiative which was launched in July 2011 has been widely acclaimed globally as one of the most significant steps Kenya has made to improve governance and implement the new Constitution's provisions on access to information. The Kenya Open Data Portal, a project of the ICT Authority, is powered by Socrata Inc and is

available at https://opendata.go.ke/. The portal makes Public Government datasets accessible for free to the public in easy reusable formats, supporting the Government's drive to proactively inform citizens and be accountable. This is in line with the goal of opendata.go.ke which is 'to make core government developmental, demographic, statistical, and expenditure data available in a useful digital format for researchers, policymakers, ICT developers and the general public' through a single online portal. The data made available is key to improving transparency; unlocking social and economic value; and building Government 2.0 in Kenya.

As Kenya citizens have varying technological, economical, and cultural differences; the authors will investigate the effect of these variations on access and usage of the platform. The challenges and best practices from a citizen's perspective and how these relate and contrast with other countries are discussed.

4.2 Analysis of the Kenya Open Data portal from citizen's perspective

This paper will interrogate from a citizen's perspective the Kenya Government open data portal based on Open Government Data Principles which include complete, primary, timely, accessible, machine process able, non-discriminatory, non-proprietary, and license-free. These principles were developed in 2007 during a meeting held in Sebastopol, California, by 30 open government advocates in a view to develop a more robust understanding of why open government data is essential to democracy (opengov.data.org n.d).

Before delving into an analysis of each of these principles based on Kenya open data portal, we briefly discuss a general outlook of the portal from a citizen's perspective. The portal, accessible at https://opendata.go.ke/ as earlier noted has menu items which include Home, Data catalog, County data, Open Budget, Suggest Data and Open data blog which are all easy to interact with.

a) Data Catalogue

This is the first tab after the home page and gives information on different sectors of interest; Agriculture, Economy and Finance, Education, Energy, Environment, Finance, Governance, Governance and Justice, Government accounts, Health and social data, infrastructure, population, water and sanitation. The two main notable miss-outs are politics and religion. A comparison with the US government open data (https://www.data.gov/), they has some other specific data sets; Science and research, Consumer, Public safety and Manufacturing. The one main data sets in the US open data is the applications page that gives the user an opportunity to download an app to help in finding items of interest such us banks, fuel stations, food stores, iCitizen etc.

For quick navigation the data is filtered into recently added, most relevant and mostly accessed. The main items under most relevant are local authority expenditures, county age pyramids, county per capita and county urbanization. It is not clear how these were picked as most relevant but having had the implementation of a new constitution in 2010 creating a devolved government, it is understandable that county news could be the most relevant news as it is a new concept in the country and many web users would like to get information about the counties. In all of the sectors above, there is relevant information uploaded on the site. However, in the Governance and Governance & Justice sectors, one wouldn't know the difference until you look at the type of datasets fed into the section.

The infrastructure sector got more views of over 1000 and mainly concentrating of traffic incidences, road accidents and General Transit Feed Specification (GTFS).

b) County Data

There are forty seven counties in Kenya having been split from the previous eight provinces. The County is headed by a governor who is the custodian of the devolved functions of government and service to the public. On the site, there are only 46 counties indicated. Migori County is missing on the list and it not clear if it was missed during the linking of pages or at the time of creating the web page altogether. The county tabs have almost similar relevant news. About 90% have local authority expenditures, per capital, county urbanization and age pyramid as a standard news item. Apart from Kiambu County that gives more on funds committed, planned expenditure and revenue projections. Nandi County gives more information about water points and Kisumu other details on health facilities, water points and schools. Nairobi County as one would expect to get mote data sets, have only GTFS matatu routes, health facilities and Rift Valley Railway data.

One would expect to find the county government profiles, brief details of the leaders, and the ministries in charge. A link to the county website would held the viewer navigate to the other sites easily.

c) Open Budget

This page purely contains financial figures in terms of the budget. The data contained here are easily understood since they are divided into useful filters. One can either get the allocated budget by Ministry e.g. Ministry of Health or by program e.g. Rail transport, or by sub program e.g. Construction of roads and bridges and lastly by the source of fund e.g. global fund. One can also see which program or sub program have the lion's share of the allocation and who is the main funder of the Kenyan projects.

d) Suggest Data

This is an interactive page where users can suggest what type of data they are looking for, if they didn't find it or if it is in a format they did not like. This is a very important tab both for the user and the administrators. The administrators will be able to know if the users are finding it easy to find data as they will only have suggestions on new datasets rather than having to show where the data resides. The user will be able to gauge if the administrator looks at the website, updates the data as well as work on the suggestions.

For example, a user queried about the most recent list of schools and the response was done within the month;

> *User: An updated version of schools database in Kenya and their details on teacher and pupil numbers. 2007 is 9 years ago and the data in your portal was collected in 2007.*
>
> *Is there any recent version?*
>
> *Admin: this dataset is already available under the education category, please check under https://opendata.go.ke/browse?category=Education. Reach us for further assistance.*
>
> *User: According to the title it is clear that it was collected by Ministry of education during their school mapping database project in 2007.Is there a more recent data because 2007 is 9 year ago?*

Admin: I understand your confusion and want to assure you that the data is from 2007 but was updated with 2014 data early this year. 2014 is the last year the ministry of education carried out a survey of various education indicators, schools being one of them. We are yet to change the name of the dataset because we are struggling with the implications that will have to APIs that use the dataset. Kindly, look at the date created which indicates January 27th, 2016 to clear your doubts. Please reach us for further assistance.

But there are also cases where there are no reply for months. A case in point is a user logged an enquiry about local radio stations mainly those in the slum area on July 11, 2011. By the time this article was being submitted, there is no response to the user. This should be public knowledge data and in coordination with the communication ministry that deals with licensing of broadcasting stations, the information should be availed as requested or at least give a lead to where the user can find the information.

There have been a lot of suggestions in the last 2 years. Most of them are open and with no comments. As indicated on the county data, Migori County is missing. In a means to find how quick the administrator will deal with the issue, the authors have left a suggestion on the page. The system has been set such that the suggestions from users will only appear after approval from the administrator as shown in Figure 1 below. Meaning, the administrators always see and read the suggestions.
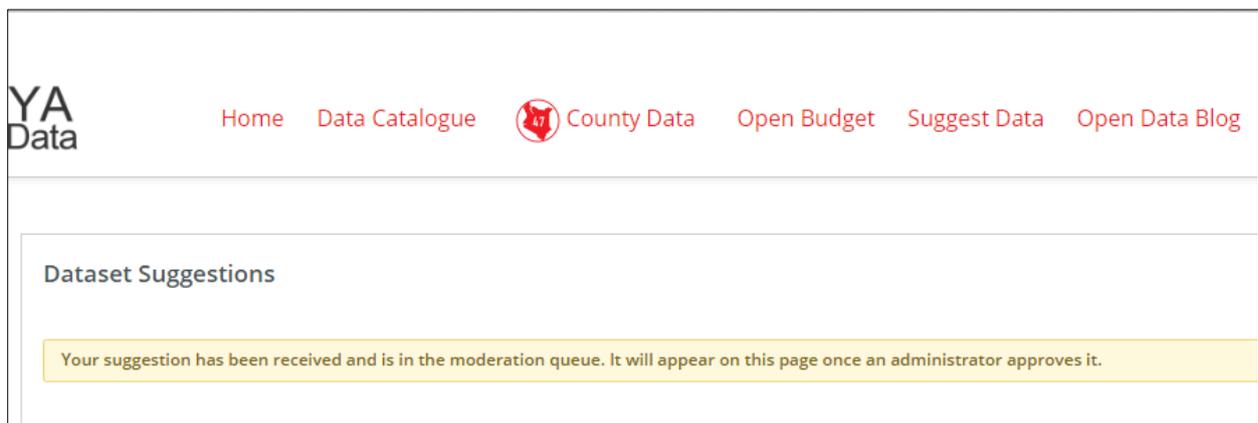
Figure 1: Data Set Suggestion

It is good to note as seen in Figure 2 below that the data request was acknowledged within two hours and that the data set was approved and added within a week.
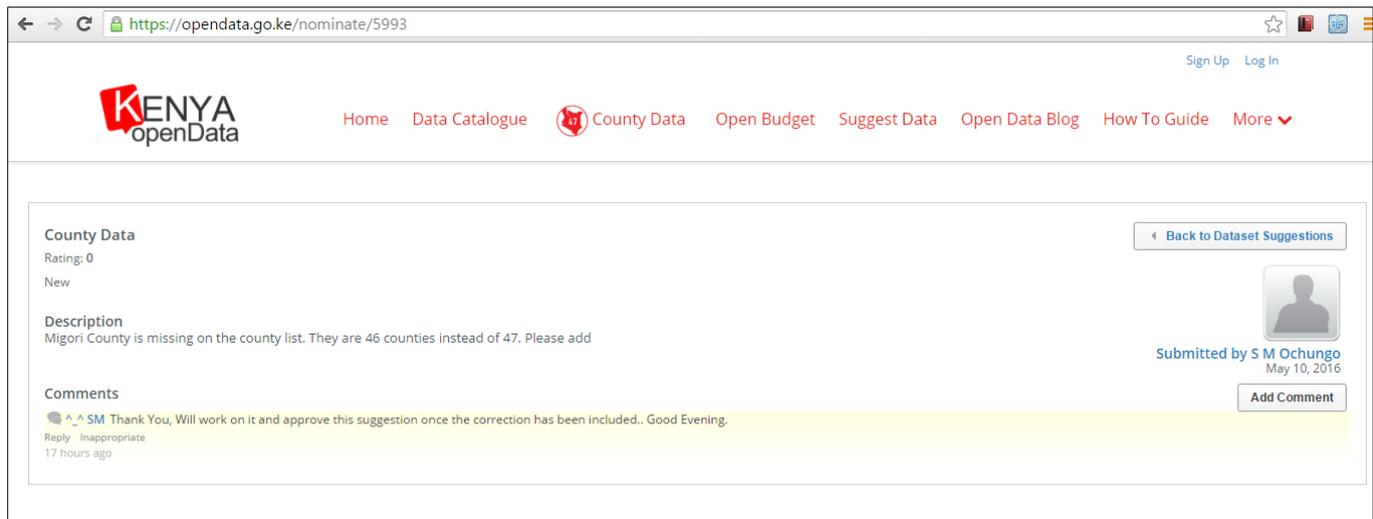
Figure 2: Dataset request acknowledgement

e) Open Data Blog

This page has blogs from the administrator on several topical issues. In discussion, the portal is easy to navigate and the information has been availed plain for easy understanding. The hyperlinks open in another tab ensuring that the user does not lose the navigation trail. The search button makes it easy to search a particular key word in a sector, however, it now not search the whole site for that particular key word. The users have a page to raise queries on data sets they have not found on the site.

The data sets and information currently uploaded on the site and those that are in the process on upload either as new data sets of updates will help individuals, partners and funders to make informed decisions on the projects or initiatives they would like to undertake. For instance, data on Economic survey 2016 makes one understand the economic growth of the sub-Saharan Africa, inflation rates. Employment and such data. Indeed from this survey you could realize the areas that informed the selection of data sets used on the site. Even though not all the data sets of information are available on the portal, there is an avenue of suggesting the required data. For example manufacturing, consumer habits and trends, Security and safety, where to find what etc.

4.3 Analysis of the portal based on open government data principles

For an Open government data initiative to be successful, it is paramount that it meets the laid down fundamental open government data principles (complete, primary, timely, accessible, machine process able, non-discriminatory, non-proprietary, and license-free) without which it would not be effective, transparent, and relevant to citizens. These principles specify the conditions public data should meet to be considered 'open' and in line with this, the challenges and best practices based on these principles are discussed in the sections that follow as they inform key best practices for successful implementation of open government data initiative (opengov.data.org n.d).

4.3.1 Complete

This is based on the premise that '*All public data is made available. Public data is data that is not subject to valid privacy, security or privilege limitations.*'

This principle would therefore help interrogate the utility of the site in terms of availability of all needed data. Four years since the initial launch, the Kenya Government open data portal as noted on the site continues to actively supply the public with datasets from Government ICT Infrastructure in addition to supporting the digitalization of Government records and processes for future upload. However, a look at various topics such as population, education, Constituency Development Fund (CDF), expenditures, and many others indicated 'no results' upon opening the pages. This is despite the counter indicating available data such as population which had indicated '240'. This therefore indicates that the 'complete' principle has not been met.

On the other hand a number of links to external sites such as Kenya Law Reports and Parliament; and Government Tenders and Procurement among others were either broken or had information which was last updated backing date to over 5 years ago. The later for example had both. In order to make sure that all required data by public is available, it is important to ask on what informs the datasets. An additional principle, *designed with public input*, as noted by Association of Government Accountants' Recovery and the Transparency Initiative (Annual CFO Survey) (2009) and Sunlight Foundation Open Data Policy Guidelines (2012) rightly argues that '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' (opengov.data.org n.d). Kenya's portal allows this through 'Suggest data' page, an interactive page where users can suggest what type of data they are looking for and whether it I available in the desired format.

4.3.2 Primary

This principle notes that 'd*ata is as collected at the source, with the highest possible level of granularity, not in aggregate or modified forms.'* As noted on the portal, this has been achieved by providing something for everyone including 'maps to start exploring, interactive charts and tables for a deeper understanding, and *raw data* for technical users to build their own apps and analyses.'

4.3.3 Timely

This principle is based on the argument that 'd*ata is made available as quickly as necessary to preserve the value of the data.'* This would therefore mean the need to avail current data as well as any data as may be requested by the public as quickly as possible as this would citizens in making decisions. A look at the site indicated numerous dataset suggestions from the public but with some having remained open even after over two years of such requests. Where such requests were indicated as approved, they were lacking a link to where such data has since been uploaded or expected date when such would be made available to the public.

It is great to have a blogs page on the portal where the public can engage with the portal administrators. However, there is very little interaction seen on the blogs which then begs the question of whether the topics in the blogs are of interest to the general public and whether citizens do actually know of the existence of the platform and if no, how best to publicize the portal. Further, though there is a suggest data page that allows users to suggest the data they would wish provided, a critical look.

There is need for dedicated portal administrators to help respond to user queries and not only acknowledge requests for data from the public but act on the same to make such data available online. This principle is well summed up in noting that the compliance of these principles must be reviewable in the sense that 'a contact person must be designated to respond to people trying to use the data.'

Indicates numerous dataset suggestions from the public which remain open for over two years of such requests. At the moment, where such requests have been approved the approval note has been used to simply note that such data though currently unavailable will be made available in future. It would be ideal to have the approved note with a link to where such data has been uploaded or details of when such data shall be made available. It is also difficult to gauge the currency of information on the portal as some reports are periodical. From a random check, there are always some uploads on different topics and sectors each month. From the admins responses on user queries, it appears they constantly monitor the portal and presumably upload latest data as soon as they find it. The blogs also show monthly activities from the admin. Since the ICT authority is the host and not a data generator and depends on other arms of government of sectors to provide data, maybe it would have been more appropriate to have the Kenya Bureau of Statistics be the administrator and the ICT authority be the host and facilitators. But like in most scenarios, the initiators of projects tend to hold on to it.

4.3.4 Accessible

This principle is based on the argument that '*data is available to the widest range of users for the widest range of purposes.*' The platform which is a strategic partnership with Socrata I'nc is noted to 'deliver open data to the public through a unique user-friendly platform that allows for visualizations, multi-format downloads and API access for software developers.' Indeed, the architecture of the portal is user-friendly in the sense of good layout and well displayed main menu items which include Home, Data catalog, County data, Open Budget, Suggest Data and Open data blog which are all easy to interact with. Further, the portal is well indexed on search engines such as Google which makes the portal discoverable from data searches. In addition, the portal has data less pages which make data easier to explore and visualize. Integration of Open Budget App (www.budget.opendata.go.ke) for releasing budget data through and further having plans in place to automate data updates and adopt a developer ready data scheme (that will be the default standard of release) to enable better API connectivity and relevance would serve to enhance accessibility of data from the portal once successfully implemented.

Kenya records wide inequalities in technical expertise, internet access, and extent of use across the country with rural areas which have limited or no electricity and internet connections being hugely disadvantaged. The current uptake of mobile phones with internet has however greatly helped in access of such portals from anywhere anytime as citizens who would wish to access such government information are scattered across the country. Data hosted on the portal has been downloaded over half a million times with over 2.5 million views since initial launch in 2011 as noted by the site administration. Nevertheless, in a country where over 10% of the population are on social media, it is of concern to note that the access and the activities on the Kenyan Open data is relatively low. For instance, 2016 Economic Survey data which would be considered quite useful to the public which was uploaded on 9th May 2016 had only one view by the 12th of May 2016. There is therefore

need for vigorous campaign to sensitize the public about the portal, its objective and how it will help the public demand for accountability of the central and the county governments.

As the administration of the portal is currently handled by ICT Authority who may be faced with bureaucratic processes from the data handlers in the counties and central government to intentionally keep some information away from the public, it would perhaps be better if the same were handled by Kenya Bureau of statistics. To increase access, it would also be highly recommended to   have the portal link made available in all Government and county sites for the public to be aware of its existence.

4.3.5 Machine process-able

This is based on the premise that '*data is reasonably structured to allow automated processing.*' The Kenya data portal makes available sufficient data in form of texts, graphics and images, maps, interactive charts and tables, and *raw data* for technical users to build their own apps and analyses.

4.3.6 Non-discriminatory

This requires that '*data is available to anyone, with no requirement of registration*' and this is the case with Kenya government open data portal.

4.3.7 Non-proprietary

This principle argues that *'data is available in a format over which no entity has exclusive control.'* The Kenya data Portal is free and allows for all the government and public sectors to share data that is deemed as a priority public data and can be accessed at any given time. The only challenge is to understand the process and the panel that approves data requested on the portal to be uploaded.

4.3.8 License-free

This principle notes that *data should not be 'subject to any copyright, patent, trademark or trade secret regulation. Reasonable privacy, security and privilege restrictions may be allowed'* and this is the case with Kenya government open data portal.

4   Recommendations and conclusion

This study shared from a citizen's perspective of the authors the experience of use and access of Kenya Government Open Data portal. Owing to the benefits of implementing such initiatives, it is recommended that countries continue to take up the challenge and implement the platforms. As citizens are the end users of these portals, it is important to implement the systems with them in mind so as to ensure the portal is based on the laid down fundamental open government data principles (complete, primary, timely, accessible, machine process able, non-discriminatory, non-proprietary, and license-free) without which it would not be effective, transparent, and relevant to citizens.

It is critical that various challenges be addressed on the Kenya open data portal including empty datasets which indicate 'no results'; broken links; obsolete information; lack of numerous datasets requested by the public dating back to over two years; and the very minimal interactions between the general public and portal administrators as seen in the portal's blog page. As compliance of open data principles must be reviewable, there is need to have a designated contact person respond to public needs. There should be accountability on the portal more so to the administrator responding to users on particular issues. Dataset suggestions need to not only be acknowledged but to be acted upon within the shortest time possible. It is recommended that link to the portal be made available in all government and county sites for the public to be aware of its existence. There is also need for vigorous campaign to sensitize the public about the portal, its objective and how it will help the public demand for accountability of the central and the county governments. Further, inclusion of an application tab where citizens can download all the required apps such as itax, eCitizen, National Transport Service Authority (NTSA), and mobile banking apps would be necessary.

Finally, as countries continue to implement such initiatives, it will be necessary for the library and information science community to especially evaluate and publish their experience so that countries and institutions can learn from one another on challenges and best practices for successful implementation from end user perspective.

References

Ding L, Lebo T, Erickson JS, DiFranzo D, Williams GT, Li X, et al. TWC LOGD (2011) A portal for linked open government data ecosystems. Web Semantics: Science, Services and Agents on the World Wide Web, Vol. 9, issue 3, pp.325-333.

Neville, C. (2007) An Introduction to Research & Research Methods; Effective Learning Service, Bradford University School of Management, pp.1-46

Oengovdata.org (n.d) The Annotated 8 Principles of Open Government Data, accessed on 20/05/2016, available at https://opengovdata.org/

Shadbolt N, O'Hara K, Berners-Lee T, Gibbins N, Glaser H, Hall W. (2012) Linked open government data: Lessons from data. gov. uk. *IEEE Intelligent Systems*, Vol. 27, Issue 3, pp.16-24.

Brito, J. (2008). Hack, mash, & peer: Crowdsourcing government transparency. *The Columbia Science and Technology Law Review, Vol. 9*, pp. 119-157.

Gray, J. (2014). Towards a Genealogy of Open Data. Paper presented at the General Conference of the European Consortium for Political Research in Glasgow.