Open Data Portals in Africa: An Analysis of Open Government Data Initiatives

Olayiwola Bello, Victor Akinwande Oluwatoyosi Jolayemi and **Ahmed Ibrahim**
Faculty of Communication and Information Sciences, University of Ilorin,
Ilorin, Nigeria.
laibello@gmail.com, toyosi.jolayemi@gmail.com
victour19@gmail.com, ahmedolaibrahim@gmail.com

Abstract

African countries like other developed nations are beginning to open up data towards attaining transparency and accountability. A better understanding by citizens of what a government does and the level of performance achieved can be better understood by making public sector data open. Citizens can therefore hold their government accountable for not meeting up with stated goals and misconduct. This study employs a survey through content analysis to evaluate the extent of implementation of open data portals. Variables used for the evaluation are Berners-Lee 5 star of open data, implementation technology, data formats, licensing, major data sets and functionality across the African continent. A total of twenty-two (22) data portals from seventeen (17) different countries were assessed. Seven (7) of the data portals representing 32% of the total number of the data portals were implemented using Drupal tool (Dkan). About 60% of the portals under investigation are national initiatives and Nigeria has the only regional/state initiative and the only two independent organisation data portals. The only two university related initiatives as well as the only city based data portal were from South Africa, while a specialised data portal dedicated to the Ebola crisis was captured from Sierra Leone. The dates for the creation of the data portals span from 2011 to the current year (*2015). The result indicates an appreciable effort is being made in the creation of data portals however more countries need to take the giant stride to data provisioning in the open formats. More data sets also need to be populated onto the portals. Movement towards attaining the 5-star status of open data portals is in progress but a lot needs to be done towards attaining this. There is also a need to complement the efforts of national governments and this could be done by regional/state, cities, universities and independent organisations.*

Keywords: Data Portal, Open Data, Visual Data, Data Formats, 5 Star of Open Data

Introduction

Different types of data are produced and collected by many public organisations in the course of performing their tasks. These organisations, such as government, education and development cooperation, are in a transition from closed, formal organisations towards open and networked models of organisation (Broek et al., 2012). The basis for the open and networked nature of organisations is more enshrined in the use of information and communication technologies which is a veritable tool in the preparation, packaging and distribution of data and information alike.

Making public sector data open helps citizens to better understand what a government or does and the level of performance achieved. Citizens can therefore hold their government accountable for not meeting up with stated goals and misconducts. As a considerable amount of these government data are increasingly becoming more easily accessible and can be used in conjunction with information from other sources, achieving this is feasible. This position is more

strengthened in that open data initiatives are springing up across the world, including Africa. These initiatives are government, academic and not for profit organisation driven. Apart from the transparency and accountability perspective to open data platforms, they can also help generate insights into how to improve government's performance. Ubaldi (2013) opined that increased data transparency provides the basis for public participation and collaboration in the creation of innovative, value-added services. Additionally, data openness is eventually expected to improve the decision making of both governments and individuals. Finally, open government data (OGD) is also seen as an important source of economic growth, new forms of entrepreneurships and social innovation.

As initiatives towards open data are springing up in government, academic and private domains across the continent, there is a need to take a cursory look at the present status of the data portals so as to identify their strong points as well as weak points to serve as a basis for standardized deployment. Ubaldi (2013) is of the opinion that, the public is expected to be able to use government data to make better decision and improve the quality of their lives, for example making specific databases easily accessible, such as through mobile apps, to better inform their choices; while governments are expected to be able to more easily access a wider range of datasets to foster evidence-based decision making. Achieving this requires data to be made available through the open data initiatives, come with the right licensing regime, be in the right format, standard and built on the right technology.

While it can be argued that there are about six dimensions to an open data initiative, this study takes a critical look at the technical dimension as it relates to the development and management of data portals. Other dimensions could include political, legal, organisational, social and economic. In evaluating the technical perspective, issues such as *five stars* of open data, deployment technology, licensing types, data formats/presentation, applications and search functions were investigated. This study therefore aims to provide an evaluative view of the data portals at present and compare them with set standards. This could form a basis for improvement and provide more insight into how such initiatives could be addressed.

Literature Review

The Open Data Handbook (2010) defines Open Data as data that can be freely used, reused and redistributed by anyone- subject only, at most, to the requirement to attribute and share-alike. It refers to data that is available as well as accessible in a convenient and modifiable form, reusable and redistributable and should enable universal participation, for example, everyone must be able to use, reuse and redistribute – there should be no discrimination against fields of endeavour or against persons or groups. On the other hand, the Open Data White Paper of HM_Government (2012), published by the UK government defines open data as data that meets the criteria of being accessible at no more than the cost of reproduction, without limitations based on user identity or intent, in a digital, machine readable format for interoperation with other data and free of restriction on use or redistribution in its licensing conditions (HM_Government, 2012).

Recently, open data has begun to gather considerable momentum. This is evidenced by the provisioning of wide array of government data that are significant not only because of the quantity and centrality but also because most government data are public by law (OKF, 2012) and that entails access and reuse of public information (Naser and Concha, 2012). In order to become more transparent and to work closer with citizens and companies, public administrations worldwide are starting up open data portals (ODP), which are repositories providing structured

access to the opened up data. This is stimulated by the idea that open government data reuse can open up economic opportunities, can promote transparency and accountability or can support the reform of public services (Davies, 2010).

The adoption and creation of open data portals is widespread in advanced democracies. However, countries and independent organisations in Africa, are also making considerable efforts to not only open up data, but to also provide a central repository for access, which is a welcome development. The attention of governments and organisations to open up data is not only stimulated by the strategies of the front runners as demonstrated by the US government, but also by the development of technologies which enable the creation of new services based on the open data (Huijboom and Van den Broek, 2011). Various technologies thus exist that can be used to evaluate and analyse individual elements of these portals, but there is the absence of a single scale of measure that can be used to gauge the overall performance of these portals.

Data portals are built on technologies and platforms that support the provisioning of the data in a prescribed format. CKAN (Comprehensive Knowledge Archive Network) is the world's leading open-source data platform (Almon, 2014). It is a software solution that makes data accessible by providing tools to streamline publishing, sharing, finding and using data (CKAN). CKAN has been adopted by various levels of Open Data Portals (ODPs), the Edo State Open data portal, publicdata.eu, data.gov.uk, data.gouv.fr, among others run on CKAN. CKAN requires technically-savvy people to implement and maintain their project solution (Open Data Monitor 2014). Socrata provides a commercial platform to streamline easy data publishing, management, analysis and reusing, by allowing export of data in many formats such as comma separated values (CSV), JavaScript object notation (JSON), xls, xml, portable document format (PDF) as well as Resource Description Framework (RDF) (Wilson and Cockburn, 2014) empowering users with the ability to customise the data set metadata according to individual requirements (Open Data Monitor 2014).

The Kenya, Chicago, Bristol and New York City government open data portals are hosted by Socrata. Junar is a cloud-based open data platform with integrated features of data collection, enrichment and analysis. Junar allows the publisher to choose what data to collect and how to present them. The platform encourages social conversations between open data administrators and end users in order to help publishers to understand what data the end users want and find valuable. The Bahia Blancc City open data portal is hosted by Junar. DKAN is a Drupal-based open data platform with a full suite of cataloguing, publishing and visualisation features.

Compared with CKAN, DKAN is seamlessly integrated with Drupal content management system, thus it can be easily deployed with Drupal and customised using different Drupal themes. DKAN provides user analytics and data users can upload, tag, search, and group data sets via a web front-end or APIs. In addition, they can also collaborate, comment, and share information via social network integration. The current deployment of DKAN instances include the Morroco open data portal, Sierra Leone's Ebola data jam, Nigeria's Visual Data among others. The Open Government platform is a set of open source tools that allow any user to "promote government transparency and greater citizen engagement by making more government data, documents, tools and processes publicly available". The listed platforms are not exhaustive as developers can also develop platforms from the scratch for implementation on the various sites and data portals.

In 2006, Tim Berners-Lee proposed a 5-star model to describe different characteristics of open data, and its usefulness for people wishing to reuse it, it is being used globally as a model

for accessing data readiness for re-use. In essence, this 5-star model is a re-user's Maslow pyramid wherein the first star reflects its basic needs and the fifth star its finest hour (EPSI, 2010). The 5-Star model was described in terms of the data availability on the web, machine-readability, non-proprietary formats, RDF Standards, and linked RDF standards, with each star representing each of these stated properties (Berners-Lee, 2006).

A 1-star rating as provided by Berners-Lee (2006) refers to the availability of the data on the web, readable by the human eye, usually represented in portable document format (PDF) which is arguably the most widely used file format for representing documents in a portable and universally deliverable manner (King, 2013). The ability to capture the exact appearance of output from nearly any computer application in a form that can be subsequently viewed or printed on nearly any computing device has made it invaluable for the presentation of content for which the author wishes to have total control of the presentation. However, data in the first star categories cannot be easily reused because it is in a closed document format (EPSI, 2010).

When the data is not only available on the web, but also available in a structured, machine-readable format, then it is deemed to have a 2-star rating. The re-user can thus access these files in the spreadsheet (xls) format, this provides a step up to the single star rating. A 3-star rating is achieved when the user is able to access to the files in established Comma Separated Version (CSV) files. CSV files are very useful formats because they are compact and thus suitable to transfer large sets of data with the same structure. However, the format is so cumbersome that data is often useless without documentation since it can be almost impossible to guess the significance of the different columns. It is therefore particularly important for the comma-separated formats that documentations of the individual fields are accurate (Open Data Handbook, 2010). A data set having four-star rating means that the data is now in the web, as opposed to being on the web, through a URI, Universal Resource Identifier, which allows for bookmarking and linking. A five-star rating means that the data is not only in the web but also linked to other data, fully exploiting its network effect and available in (RDF). RDF is a W3C-recommended format that makes it possible to represent data in a form that makes it easier to combine data from multiple sources. RDF encourages the use of URLs as identifiers, which provide a convenient way to directly interconnect existing open data initiatives on the web.

Facts cannot be copyrighted, but that does not mean that data and databases are exempt from legal discussions and licensing requirements, even if the intention is to share the data openly (Watters, 2011). This therefore justifies the introduction of licensing within the context of open data. Open data licenses facilitate the use and potential reuse of data, and provide the benefits of enhanced organisational efficiency and cost saving, leading to greater interoperability of data as well as increased user awareness of the license terms, enabling better compliance (Korn and Oppeinheim, 2011).

Licensing is very important for the community to trust that the data will not be closed off, however, finding a suitable open access license for data can be tricky, in part because intellectual property in data is treated fairly differently in different jurisdictions (Rochkind, 2008). Creative Commons, CC Zero, Open Data Commons and The Open Government License are some of the most recognised licenses. The Creative Commons open data license is fast becoming one of the most used and recognised standard licenses for providing access to data and other resources (Korn and Oppenheim, 2011). They permit the free of charge copying, reuse, distribution and, in some cases, the modification of the initial creator's creative work, without having to obtain permission every time from the rights holder. This is found to be useful within open data context.

CC Zero (CC0) is a tool also created by Creative Commons to facilitate the release of content, data, datasets and databases into the public domain, for example by the copyright owner waiving all its rights, including the database right and the right to be identified as the creator. The open government licence facilitates the reuse of government and other public sector information. As with creative commons licences, the open government licence is available in a machine-readable form as well as a "human-readable" form. Unlike the creative commons license, it (open government licence) defaults to the governing legislation to which the licensor has their place of business. It does not permit copying of logos, registered trademarks and other IP such as patents, and includes specific non-endorsement clauses (Korn and Oppeinheim, 2011).

The functionality provided by the data portals provides another criterion for assessing them. A fully functional data portal can be largely described as one that is easily accessible, searchable, provides datasets that are downloadable and in no one specific export format, provides applications for information provided on the portals such as the use of infographics as demonstrated by Ghana open data initiative, Kenya open data, Visual Data Nigeria, and also provides a medium for analysis.

The usefulness of the information provided to the users can be described as a function of its accessibility, as only the data that is available and accessible can be used for the required purpose. It is important, therefore, that information is presented in an accessible way, in a range of languages and formats that can be easily used and understood by the intended audience (NHS, 2010). Data analysis as defined by the Northern Illinois University is the process of systematically applying statistical and /or logical techniques to describe and illustrate, condense and recap, and evaluate data.

According to Shamoo and Resnik (2003) various analytic procedures "provide a way of drawing inductive inferences from data and distinguishing the signal (the phenomenon of interest) from the noise (statistical fluctuations) present in the data". The use of infographics and visuals present a way of data analysis. Infographics are increasingly becoming a common useful tool in education, entertainment, storytelling, broadcasting, and more. Madsen (2014) defined Infographics as graphic visual representations of information, data or knowledge intended to present complex information quickly and clearly. The provision of such functionality on a portal increases the level of accessibility such a portal provides as studies have shown that 90% of the information we remember is based on visual impact (Costill, 2013).

A typical data portal holds a large database of information, searching is by flipping through the various pages available which can prove strenuous and time-consuming, therefore, a functional portal will not only make the data available, but provide means of searching through the database by employing keywords and filters and other variables. The acquired data should also be downloadable by users in formats that can facilitate reusability. Functionality is thus enhanced by providing a wide array of formats in which the desired dataset can be exported into.
Open data initiatives around the world are characterised by a great extension of the number of data sets made available for access by public administrations, constituencies, businesses and other actors, such as journalists, international institutions and academics, to mention a few. These datasets usually rely on selection criteria, based on a technology-driven perspective, rather than on a focus on the potential public or social value of the data to be published (Viscusi, Castelli and Batini, 2014). The aforementioned technologies and criteria are the premises upon which the analysis in this paper is based.

Methodology

This survey seeks to evaluate data portals across African countries with open data initiatives. The countries selected for this assessment were arrived at based on their efforts in moving forward towards requiring proactive disclosure of government data as part of their right to information (RTI) laws as presented by Open Data Barometer (2015). Through a comprehensive search of the World Wide Web (www) data portals in these countries were identified. Other initiatives were also identified outside the countries presented in the Open Data Barometer report. Therefore, this survey conducted from 16th to 19th February, 2015 reflects data portals in the entire Africa continent as at the time of this research.
Specifically, the evaluations of the various data portals from the different countries are based on the following parameters:

Technology Used for implementation:
The technology used for building the open data site/applications including those present on the portals such as Ckan, Dkan, Junar, Socrata, prognoz, opensoft etc.

- **Open data 5-star Evaluation**

A 5-star evaluation (Berners-Lee, 2006) based on the availability of data on the portal, making the available data structured, the use of non- proprietary formats, the use of URIs to denote things, and linking of the data to other data to provide context.

- **Format**

The formats in which data is published on the data portal such as csv, xls, json, xml etc. and the adherence to the use of non-propriety formats.

- **Major datasets**

The major Datasets that are frequently published on the data portals.

- **Open data licensing**

The open data licensing agreement used to share and modify datasets on the data portal if available.

- **Functionality**

Features on the data portal that enable users to perform specific tasks while using the portal.

Results

The varieties of open data portals surveyed include initiatives from national governments, regional/state government, universities, city, independent organisations as well as specialised portals. About 60% of the portals under investigation are national initiatives. Nigeria is identified with the only regional/state initiative and the only two independent organisation data portals. South Africa however, is credited with the two university related initiatives as well as the only city based data portal recorded. Sierra Leone provided a specialised data portal dedicated to the current Ebola crisis. The dates for the creation of the data portals span from 2011 to 2015 when the study was done.

Technology used for Implementation

Data portals can be implemented and managed using various technologies including web application tools such as Ckan, Dkan, Lunar, Socrata, JavaScript, Qu, OpenSoft and more. These technologies are important in the day to day management, maintenance of the data portals and regular updates of datasets contained in them so as to make such up to date and allow for easy access to current data and meaningful information by users. Within the open data portal

implementation sphere of African countries, the result shows a combination of proprietary and open source platforms. Evidence also shows that some portals are designed from the scratch without leveraging the power of existing data portal technologies
Presented in Figure 1 are the technologies used in the development of the data portals and the proportion of their deployment across the investigated platforms.

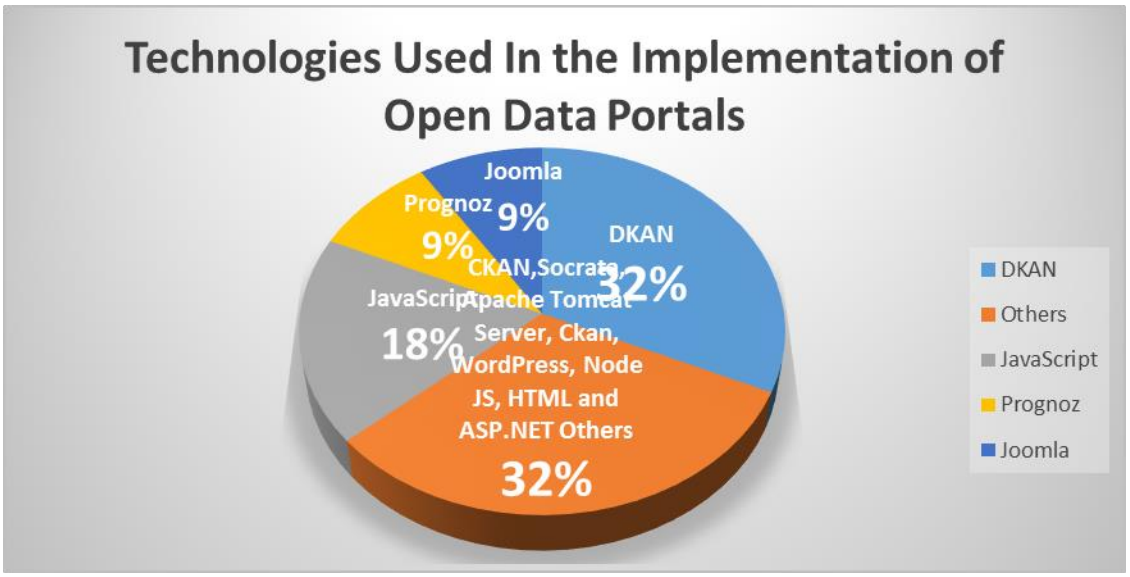

Figure 1: Data Portals Implementation Technology

A total of twenty-two data portals from seventeen different countries were assessed. Seven of the data portals representing thirty-two percent of the total number of data portals were implemented using Drupal tool (Dkan). Four of the data portals representing Eighteen percent of the total number of data portals were implemented using JavaScript. Two of the data portals, representing nine percent of the total number of data portals were implemented using Prognoz. Another two of the data portals representing nine) percent of the total number of data portals were implemented using Joomla. The remaining seven data portals representing 4.5 percent each were implemented using different technologies such as; Socrata, Apache Tomcat Server, Ckan, WordPress, Node JS, HTML and ASP.NET. While Ckan seems to be a popular platform for the creation of data portals, its complexity and difficulty to adapt might have contributed to its very low patronage.

Open Data 5 Star Evaluation

The data portals under investigation, were evaluated using the open data 5-star evaluation. This is based on the 5-star deployment schemes for open data as suggested by Berners-Lee (2006). A careful presentation of the attributes relating to each star level has been discussed in the introduction. Each data portal was rated based on the presence of a higher level open data star feature, this implies that if a data portal has any of its data in non-proprietary formats such as csv but not in excel, it will be given a 3-star rating, this applies to all other star level ratings.

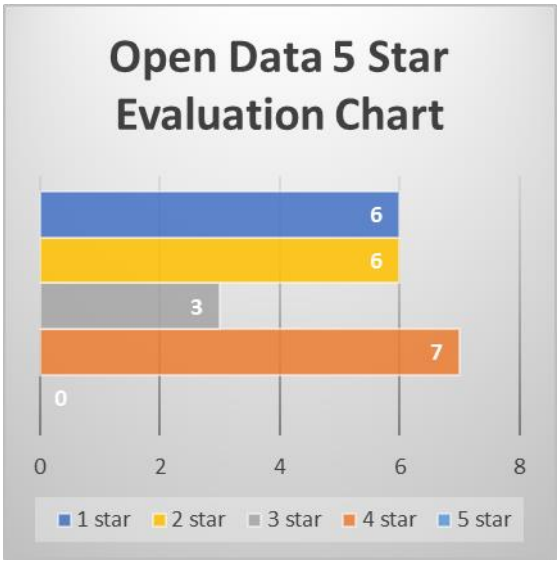

Figure 2: 5 Star Evaluation of the Data Portals

None of the data portals under investigation was able to attain a 5-star rating as a result of the inability to link the data provided on their portal to other related data to provide context (Open linked Data). This could be as a result of the concept of open linked data just beginning to be topical within the open data paradigm. However, seven open data portals managed to score a rating of 4 stars. This list includes two of Nigeria's platforms open data Edo state and Visual data Nigeria. The Ghana open data portal, Morocco's open data portal, Senegal Ouvert data portal, Kenya Open data portal and Burkina Faso open data initiative portal also scoring a 4- star rating. It was also discovered that the three of these seven data portals (Visual data Nigeria, Ghana open data portal and Morocco's open data portal) used the latest version of Dkan in implementing their open data portals which went a long way in helping them achieve the 4-star rating. Edo state open data portal was implemented using the Ckan technology, though complex, this also had a significant impact on its achieving the 4-star rating. However, Burkina Faso open data portal was implemented using python and JavaScript technology.

Data Format

Datasets can be represented in a variety of formats such as Excel, DOCX, PPTX, CSV, JSON, RDFs, Pdf, and more. However, the standard formats for open data are non-proprietary formats such as CSV, JSON and RDFs. These formats enable the users to further better reuse the datasets without having to purchase software to process them. These formats in all make the process of reuse a painless one and have the tendency to promote the use of such data sets. Formats like the RDF promote the concept of open linked data, which is the apex of the 5-star rating and its use is entirely absent in this survey. This is another pointer to the inability of the portals assessed to attain the 5-star rating. The use of non-proprietary data formats such as JSON and CSV which is machine readable and do not need proprietary software to process turned out to be low. Fourteen (14) of the 22 data portals which accounts for more than half the total number of data portals did not adhere to the use of non-proprietary and machine readable data formats at all. The result of this being that the flow of the available data to a wider pool of audience who would love to reuse them could be limited which boils down to the issue of

accessibility. Users could have to purchase programs or software that would be needed to process the data in proprietary format such as excel, pdf or word before it can be reused by them.

Major Datasets

Datasets within a data portal span a variety of categories such as education, finance, agriculture, transportation, population, health, weather and others. The categories of data sets on a portal could serve as a pointer to its versatility and ability to draw users to it. Similarly, availability of data from different spheres of life could really indicate the extent to which a government is opening up to the public. It is to this extent that this survey takes a critical look at the major data sets available on the data portals for this survey**.** Datasets on education, administration and finance were found to be most prevalent among the data portals**.** Education and administrative datasets were more prominent on nine (9) of the data portals.

Finance datasets were also prevalent in seven (7) data portals**.** Five of the 22 data portals focused solely on one dataset. They are BudgIT from Nigeria (Finance), OpenUCT from South Africa (Education), Centre for Higher Education Transformation from South Africa (Education), Egypt's Government Services Portal from Egypt (administration) and Ebola data jam from Sierra Leone (Health/Ebola). It is also interesting to note that all four data portals from Nigeria had financial datasets on their portals which shows that open data initiatives in the country are demanding the right to information on government financial data. This may go a long way in tackling corruption in the country.

The diversity, though not comprehensive, in categories of datasets within the data portals across Africa shows that open data can be useful in various aspects of life and not just limited to a particular aspect. Areas such as agriculture, science and technology, and weather could definitely be improved.

Open data licensing

For datasets/data published on the data portal to be reusable by users, it has to be published under an open data license. The most recognised and commonly used data licenses are; Creative Common (CC) Licenses, CC zero (CC0), Open Data Commons and the Open Government License. It is on this premise that this survey assesses the selected data portals based on the open data licences under which the data is published. The use of open data licenses by the data portals turned out to be low as fourteen (14) of the twenty-two data portals published datasets without the use of open data licenses.

Open data commons open database license (ODbL) was used in publishing datasets by four (4) data portals. Open data commons was used in publishing datasets by two data portals and Creative Commons Attribution was used by one data portal in publishing its datasets. The use of open data licenses should be highly encouraged as this has the propensity to increase data reuse. This could come in the form of other users being able to republish the content or data on their own website and derive new content or Information from them.

Functionality

Users should be able to perform a variety of functions while accessing the data portals. Data portal functionality refers to specific tasks users can perform on the datasets available while using the data portal. This survey assesses the selected data portals based on accessibility, downloadable datasets, exportable datasets, dataset analysis, dataset visualisation and search function. These functions are explained below.

Accessibility: access to the datasets within the data portal by users without any form of restriction.
Download: datasets should contain download links to enable user download datasets from the data portal
Export: datasets available for download should also be exportable into several other formats to enhance reusability.
Analysis: elements of datasets within the data portal should be broken down into simpler parts to enhance user understanding.
Visualisation: There should be a graphical view or representation of what a particular dataset represents.
Search: Datasets should be accessible using the search form instead of having to navigate through the pages of the data portal.

The result of the survey shows that all the data portals are accessible. Seven of the data portals; Morocco open data portal, Senegal ouvert, Burkina Faso initiative, Visual Data Nigeria, National Bureau of Statistics Nigeria, Open Data Edo State and the Kenya open data portals provided all the required functionality. It is worthy to note that Egypt's and Tanzanian data portals might even fall short of the definition of open data portals in the real sense of it as these portals only give access to view data sets available on them but such cannot be downloaded or exported which violates the definition of open data as provided by the Open Data Handbook.

Conclusion

Opening up data in the African continent is witnessing an upsurge. Within a four-year period (2011-2015), 22 data portals have sprung up from 17 countries. The present uptake is still very low (about a third) having in mind that there are 54 countries on the continent. Significant efforts will be required from national governments in ensuring the development and sustenance of national initiatives towards open data. Such efforts can also be complemented by other units like state/regional governments including local governments, cities, independent bodies and universities. In future implementations however, there is a need to pay more attention to the right data formats which have a significant bearing on how re-users are able to key into the enormous data being made available for their various personal use. Employing the right approach is tantamount to achieving the desired success. Therefore, it is expedient that technologies that are capable of fast tracking the deployment of these platforms are adopted.

References

Almon, G. 2014. Introduction to World's Leading Open Data Platform: CKAN. Retrieved 19th February 2015 from http://www.tech4relief.com/2014/05/22/introduction-worlds-eading-open-data-platform-ckan/

Berners-Lee, T. 2006. 5 Star Deployment Scheme for Open Data. Retrieved 16th February, 2015 from http://www.w3.org/DesignIssues/LinkedData.html

Broek, T., Rijken, M., and Oort, S. (2012). Towards Open Development Data: A Review of Open Development Data From A NGO Perspective. TNO whitepaper, publication number P10098. Retrieved 18th February 2015 from http://openforchange.info/sites/pelle.drupalgardens.com/files/TNO%20whitepaper%20%20Towards%20open%20development%20data.pdf

Costill, A. 2013. 6 Benefits of Using Infographics. The Search Engine Journal (SEJ). Retrieved 19th February 2015 from http://www.searchenginejournal.com/6-benefits-using-infographics/70917/

Davies, T. 2010. Open Data, Democracy and Public Sector Reform. Retrieved 17th February 2015 from http://practicalparticipation.co.uk/odi/report/wp-content/uploads/2010/08/How-is-open-government-data-being-used-in-practice.pdf

EPSI 2010. The Five Stars of Open Data. The EPSI platform, 2010. Retrieved 16th February 2015 from http://www.epsiplatform.eu/sites/default/files/The%205%20stars%20of%20Open%20Data_MdV_PR2.pdf

HM_Government 2012. Open Data White Paper Unleashing the Potential. Retrieved 19th February 2015 from https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/78946/CM8353_acc.pdf

Huijboom, N. and Van den Broek, T. 2011. Open Data: An International Comparison of Strategies. *European Journal of e Practice*. Nº 12 March/April 2011. Pp 4-16. Retrieved 13th March 2015 from https://joinup.ec.europa.eu/sites/default/files/76/a7/05/ePractice%20Journal-%20Vol.%2012-March_April%202011.pdf

King, C.J. 2013. Role of Pdf and Open Data. Retrieved 17th February 2015 from www.w3.org/2013/04/odw/odw13_submission_52.pdf

Korn, N. and Oppenheim, C. 2011. Licensing Open Data: A Practical Guide. Joint Information Systems Committee (JISC). Retrieved 18th February 2015 from http://www.discovery.ac.uk/files/pdf/Licensing_Open_Data_A_Practical_Guide.pdf

Madsen, L.B. 2014. Data-driven Healthcare: How Analytics and BI are Transforming the Industry. Hoboken, (NJ) John Wiley and Sons Inc.

Naser, A. and Concha, G. 2012. "Datos Abiertos: Un Nuevo Desafío Para Los Gobiernos De La Región," Naciones Unidas. Retrieved 18th February 2015 from http://www.eclac.org/publicaciones/xml/7/46167/DatosAbiertos_17_04_2012.pdf.

NHS 2010. Accessible Information Policy. Retrieved 19th February 2015 from www.wdchcp.org.uk/EasySiteWeb/GatewayLink.aspx?alId=71777

OKF 2012. Open Government Data Handbook. Retrieved 30th April, 2015, from Open Knowledge Foundation: http://opendatahandbook.org/pdf/OpenDataHandbook.pdf

Open Data Barometer (2015). Open Data Barometer Global Report. 2nd Edition. Retrieved 23rd

May 2015 from http://opendatabarometer.org/assets/downloads/Open%20Data%20Barometer%20-%20Global%20Report%20-%202nd%20Edition%20-%20PRINT.pdf

Open Data Handbook 2010. What Is Open Data? Retrieved 17th February 2015 from http://opendatahandbook.org/en/what-is-open-data/

Open Data Monitor (2014). Monitoring, Analysis and Visualisation of Open Data Catalogues, Hubs and Repositories. Retrieved 17th February 2015 from: http://project.opendatamonitor.eu

Open Knowledge Foundation (2012). "Manual de los Datos Abiertos," Portal de Datos Abiertos del Gobierno de Misiones, 2012. Retrieved 18th February 2015 from http://www.datos.misiones.gov.ar/directorio/documentos/Manual_de_Datos_Abiertos.pdf .

Rochkind, J. 2008. Creative Commons Is Not Appropriate For Data. Retrieved 19th February 2015 from https://bibwild.wordpress.com/2008/11/24/creative-commons-is-not-appropriate-for-data/

Shamoo, A.E. and Resnik, B.R. 2003. Responsible Conduct of Research. Oxford University Press.

Ubaldi, B. 2013. Open Government Data: Towards Empirical Analysis of Open Government Data Initiatives, OECD Working Papers on Public Governance, No. 22, OECD Publishing. http://dx.doi.org/10.1787/5k46bj4f03s7-en

Viscusi, G., Castelli, M. and Batini, C. 2014. Assessing Social Value in Open Data Initiatives: A Framework. *Future Internet* 2014, No. 6, 498-517

Watters, A. 2011. Choosing The Right License for Open Data. Retrieved 19th February 2015 from http://radar.oreilly.com/2011/06/openstreetmap-creative-commons-open-database-license.html

Wilson, P. and Cockburn, M. 2014. Bristol's New Socrata Open Data Platform. Bristol future. Retrieved 18th February 2015 from https://bristolresearchnetwork.files.wordpress.com/2014/07/bristol-research-network-socrata-open-data-18-july-2014.pdf

Olayiwola Bello is a researcher with interest in information and decision systems among other areas of information technologies. A winner of Social Entrepreneurship Initiative for West Africa (2014), foundation member and project lead of Visual Data Nigeria. A beneficiary of research grants in Nigeria and Internationally, Olayiwola is currently with the Department of Information and Communication Science, University of Ilorin, Nigeria.

Victor Akinwande is currently a graduate student at Carnegie Mellon University in Rwanda, and is always interested in ways to apply Data Science to tackle social and business challenges. An Idea to solve issues in Energy, Education, or Agriculture in Africa is a great way to start a conversation with him. A winner of Social Entrepreneurship Initiative for West Africa (2014) and foundation member of Visual Data Nigeria.

Jolayemi Oluwatoyosi is a Bsc. Information and Communication Science degree holder from the University of Ilorin, Nigeria. She is currently on her national youth service. An enthusiastic researcher, hardworking and dedicated IT optimist. She is also a member of the Visual Data Nigeria family.

Ibrahim Ahmed is a Computer Science graduate of University of Ilorin Nigeria. He is a full stack software developer and technology enthusiast who is passionate about software and is dedicated to building technological solutions to solve both human and societal challenges. Ahmed is an active member of the Visual Data family.