Communication

Christine Mwongeli Mutuku*

Engaging a Data Revolution: Open Science Data Hubs and the New Role for Universities in Africa

Abstract: This paper presents a new ideology for engaging Africa in a data revolution. It explores the idea of creating open-science-data-hubs (OSDH) at flag-ship universities in Africa to preserve and share both internally and externally produced data. Although limited in the technical aspect, the objective here is to explore the pragmatism of how and why such an endeavor in Africa should be undertaken. This paper argues that the African university is uniquely placed to play this new role in today's technological world and discusses the characteristics and foundational pillars necessary to set up such a program. The arguments provided here challenge Africa to be smart and adopt clever solutions to their data generation, collection and access problems, by finding value and a new role in the intellectuals and institutions of higher learning and in the necessity to involve them in the generation, preservation and sharing of data and knowledge that can be used in the policy formulation process.

1 Introduction

Since independence, Africa's public policy development process has functioned with limited, if any, contribution from intellectuals and institutions of higher learning. This is because the intellectual was, and largely still is, deemed as a critic of the data-less elite process of formulating public policies and allocation of national resources. Although studies have documented the dichotomy between intellectuals and the government in as far as utilization of knowledge in the policy process is concerned (Neilson, 2001; Weiss, 1977) and the complexity of the much politicized and volatile agenda setting process which further limits knowledge adaptation (Kingdon, 1998), the situation in Africa seems particularly wanting. The politics of the entire policy process as well as the poor quality of data collected and held in silos by national statistics offices in various countries in Africa basically eliminates any chance of it being interrogated by intellectuals and adapted by government. Moreover, the limited research volume (Andoh, 2015) and poor quality of data available for possible application in the policy formulation process all point to Africa's need to engage a data revolution as stated by the United Nations Economic Commission for Africa – UNECA- (2015) as well as to their need to facilitate knowledge communication (Africa Union – AU-, 2011). The recent call, particularly by UNECA to engage the 'African Data Revolution' has inadvertently challenged institutions of higher learning to facilitate data generation, preservation, and consequentially evidence-based policy formulation in Africa.

The objective of this paper is therefore to propose and assess the possibility of creating open-science-data-hubs (OSDH) at Africa's flagship universities in each African state to collect and preserve internally and externally produced data in a central repository (hub) for the purpose of sharing with various stakeholders

*Corresponding author, **Christine Mwongeli Mutuku,** Univerity of Nairobi, Nairobi, Kenya, mwongelimutuku@uonbi.ac.ke

such as students, academicians, researchers, public institutions, urban planners, and other interested parties, which will allow interrogations and hopefully, subsequently improve the policy analysis process. Although there are many hurdles, and varying perspectives in the adaptation of knowledge generated by intellectuals by the policy formulators (Girard, Neilson, 2001; Weiss, 1977), having a data hub is still a good idea because it is better to have data, in case there is a need, than to have none when it is needed. Such a hub will not only improve data preservation and access, which can positively affect data generation and hopefully knowledge utilization, such a creation will also serve to revive and facilitate a positive data culture and elevate the intellectual to their deserving position of interrogating evidence.

To achieve this objective, this paper begins by first defining the term OSDH, as applied herein, followed by a review of the historical role of the university in Africa. A review of open data advocacy, followed by the problem statement, research question, and significance of a data hub as well as why the university is the most ideal host will also be provided. The methodology, foundational pillars and the expected limitations will also be discussed. It is worth noting that the discussion engaged here-in does not discuss the technological aspect of the OSDH process and exclusively focusses on assessing the viability of this idea in Africa. Hopefully this assessment will stimulate a healthy debate on the new role that African universities can and should assume to contribute towards Africa's data revolution and subsequently in the policy development arena.

2 Definitions

What is an open-science-data-hub? To better understand what is being advocated for, it is necessary to first define open data. The definition of open data was first coined by Open Knowledge International in 2005 and denotes three principles;

- First, open data embraces the principle of a**vailability and access** which means data is available in whole, to be accessed by anyone.
- Second is *re-use and redistribution of the data which is possible when data is provided in formats and ways that allow* **it's** re-use redistribution and intermixing with other datasets.
- The third principle is *universal participation which* means that no one is discriminated from using the data, willing consumers are able to use, re-use and redistribute.

(Open Knowledge International -OKI-, 2018).

According to OKI, the existence of these three principles would then allow the fourth principle which is **interoperability to occur. Interoperability** denotes the ability of diverse systems and organizations to work together (inter-operate) which allows for the intermixing of different datasets. This, according to OKI, consequentially allows for different components to work together and so build large and complex data systems. In essence, consumers of such a system will be able to run frequencies, correlation and other statistical analyses with the data as they please. It is these three principles of open access which are suggested here in as crucial to define the OSDH at the hosting university.

In a similar respect, the term 'science' in this paper is used to refer to both the skill that is applied to systematically generate reliable and valid data concerning the natural world through the application of defined steps and standards that can be replicated by others. This term is also used to refer to the data which is a product of such a rigorous activity as engaged by researchers and students within the institution. So then what is a hub? Synonyms of the term hub include repository, bank, center, pivot, core, heart or nucleus of something, in this case, data. Fieldman (2018) defines a data hub as a data integration approach where data is physically moved, harmonized and re-indexed into a new system. The idea of an open-science-data-hub being explored in this paper is that of creating (institutionalizing), an unhindered, center or repository of datasets that are collected, digitized, harmonized and indexed within a given university and which are subsequently archived for sharing purposes. This is the new role for the African university that this paper presents and discusses as a strategy to facilitate a data revolution in Africa.

So then what is a data revolution? A data revolution is "the process of embracing a wide range of data communities and diverse range of data sources, tools, and innovative technologies, to provide disaggregated data for decision-making, service delivery and citizen engagement (UNECA, p. 2)." For a sustainable data revolution to occur in Africa each country must be willing to be part of, and invest in, harvesting and collecting a diverse range of valid and accurate data that will be held in an open and accessible manner so as to allow for its analysis and subsequent utilization in various arenas more so in the policy process.

Other terms that will be used in this discussion and that need definition is a data community and data ecosystem. A data community is "a group of people who share a social, economic or professional interest across the entire data value chain – spanning production, management, dissemination, archiving and use (UNECA, 2015, p. 2)." A data ecosystem comprises of "multiple data communities, all types of data (old and new), institutions, laws and policy frameworks, and innovative technologies and tools, interacting to achieve the data revolution (UNECA, 2015, p. 2)."

Both the data hub being suggested here and the data revolution will need good management to evolve from being an institutional data hub to a national data hub. In other words, it will be necessary for the flagship university to invest in achieving international standards of data stewardship.

3 Data Stewardship

The Data Governance Institute (2017) defines data stewardship as the concern of taking care of data assets that do not belong to the stewards themselves and considers the entire data value chain. Therefore, the institution to host this program will have to consider both the FAIR principles and the Africa Data Consensus principles on data stewardship in order to operate the OSDH. The FAIR principles for research data stewardship are Findable, Accessible, Interoperable, and Reusable (Boeckhout, Zielhuis, and Bredenoord, 2018; Wilkinson et al., 2016; Springer Nature, 2018).

- The principle of findability stipulates that data should be identified, described and registered or indexed in a clear and unequivocal manner.
- The principle of accessibility specifies that datasets should be accessible through a clearly defined access procedure, ideally by automated means.
- The principle of interoperability requires that data and metadata are conceptualized, expressed and structured using common, published standards.
- Finally, the principle of reusability means that characteristics of the data, including their provenance, should be described in detail according to domain-relevant community standards, with clear and accessible conditions for use.

The Africa Data Consensus document provides the following principles as well: consider data a public good, share data, user friendly, disaggregate data to the lowest levels of administration, factual and valid , needs driven, provide data management and curation services, educate and train, and embrace technology (UNECA, 2015, p. 4). Apparently, these principles are already encompassed by the FAIR Principles discussed above. However, the FAIR principles miss two particular principles of data stewardship that need to be included in the OSDH project. These are;

- Ethics: the need to adhere to ethical standards in research design, data collection, curation, and sharing can never be over emphasized. As a data steward, the university will need to responsibly uphold and mitigate issues of privacy, justice and benevolence. This means respecting and protecting privacy, intellectual and property rights of human subjects who participate in various research projects.
- Inclusivity; the university will have to ensure that all facets of this program are age, gender and ethnic sensitive, so as to have valid and reliable data to meet the variety of data required by researchers.

Additionally, it is imperative that the university continuously engages in data collection, curation and sharing so as to have a well-stocked, structured and managed data hub.

Another definition worth providing is the term flagship. As applied herein flagship refers to the 'star' or 'leading' university within each African state. The term could also be applied to mean the institution that is selected to host the program. This term is not necessarily applied in the plural, for it could be plural in one country and singular in another. For the purpose of understanding, the following section provides a summary of the university in Africa.

4 The Historical Role of the University in Africa

The university, as an institution of higher learning in Africa, has existed for a long time. Africa is home to the world's oldest existing, and still operational universities; the University of Karueein, founded in 859 AD in Fez, Morocco and the Al-Azar University in Egypt built in 970 AD (Guinness World Records, 2018; QS Top Universities, 2018). Although the existence of institutions of higher learning in Africa is well documented, however, such institutions, both during the pre-colonial and colonial period, were few and not widely distributed, more so in sub-Saharan Africa (Woldegiorgis & Doevenspeck, 2013; Zeleka, 2006). Although the colonial period saw the expansion of basic formal education in Africa, the development of higher education was quite limited. Essentially, the colonial government was more concerned with training clerks and typists and giving them basic education sufficient only to perform these duties and assimilate them into the new culture of their masters (Andoh, 2017; Woldegiorgis & Doevenspeck, 2013). Thus, minimum efforts and resources were geared to establishing higher education in Africa. As a result, African students seeking advanced education had to go abroad, to the country of their colonial masters, to advance their education. Upon independence this apparent vacuum in higher education provided the necessary inspiration for newly independent African governments to not only expand basic education so as to ensure that more of their citizens could access education, but to build institutions of higher learning as well. Since then, the university in Africa has been perceived as key in the generation of a highly skilled labor force which was, and still is, deemed as essential to facilitate the Africanization and development of the African state (Friesenhahn, 2014; Woldegiorgis & Doevenspeck, 2013; Zeleka, 2006). Although expansion of higher education in Africa was evident between the 1960s and 1980s, graduate level education was not equally expanded and much of this deficiency still remains. According to Friesenhahn (2014) only a minority of the estimated 1,500 public and private universities across Africa today offer graduate programs.

Over the years higher education in Africa has faced numerous challenges. The first challenge was due to the military coups experienced in a significant number of African countries between 1960s and early 1990s. However, both the military and civilian governments in Africa engaged in tactics that interfered with the administration of universities by appointing their political affiliates to positions of authority (Andoh, 2017). Another blow to higher education was wielded by the Structural Adjustment Programs (SAPs) that were imposed on African countries in the 1980s and 1990s by the International Monetary Fund (IMF) and the World Bank to instigate economic reforms. The austerity measures imposed through these SAPs led to severe government cutbacks in social expenditures which higher education relied on to survive. With limited funds and political intrusion in the management of universities, Africa's higher education existed in a state of crisis. This crisis was expressed in declining state funding, low research output, falling instructional standards, poorly equipped libraries and laboratories, shrinking wages and faculty morale (Andoh, 2017; Woldegiorgis & Doevenspeck, 2013; Zeleka, 2006).

In the past decade or so, African governments seem increasingly aware of these challenges and are exuding a willingness to reform, renew and revitalize higher education. According to Zeleka (2006) Africa has a reform agenda for higher education which is centered on five broad issues geared at enabling the continent to compete in a technology driven globalized world. The first issue on the agenda seeks to systematically examine the philosophical foundations of African universities. The second issue, labeled management, seeks to understand how African universities are dealing with the challenges of quality control, funding and governance. The third issue focuses on pedagogical and paradigmatic issues which touch on the dynamics and societal relevance of knowledge production in academic institutions. The fourth

issue focuses on the relations between universities, the state, civil society, and industry. The fifth issue on the agenda, and which is most relevant in this discussion, addresses issues of globalization and the impact of the new information and communication technologies on higher education in Africa (Ramphele, 2004; Zeleka, 2006). This agenda can be translated as a challenge to African universities to engage the data revolution.

5 Open Data Advocacy in Africa

Fortunately, Africans are progressively realizing the need for a data revolution and the benefit such a transformation could afford in terms of knowledge availability which has the potential to transform the policy analysis arena and subsequently contribute to development susta inability in their states. This recognition has led to the advocacy for data openness by private and public organizations on the continent. One such organization is the Research Data Alliance –RDA- (2016) and the Data Intensive Research Initiative of South Africa (DIRISA) which are focused on refining the sharing and re-use of data in Africa. The Africa Open Data Conference (2017) is another example of efforts to create awareness on data accessibility. This portal offers a convening space for tech industry experts and small businesses, entrepreneurs and private and public organizations to virtually connect and share advances in open data, share lessons and form new collaborations. Another platform is the Open Africa Forum by Code for Africa (2018) which aims to be the largest independent repository of open data on the African continent. As stated in their website, this forum is not a government portal but a grassroots initiative to offer data resources for ordinary citizens, civil society organizations and civic activists, the media, and government agencies (Code for Africa, 2018). The Code for Africa has branches in various African countries including Ghana, Kenya, Malawi, South Africa and Zimbabwe and these too seek to create online open data hubs to empower citizens make informed decisions. What Code for Africa does is solicit data from any contributor and anyone can upload data on this forum. However, this raises concerns regarding the validity of the data available as well as the credibility of the contributor. Moreover, datasets in their webpage are few and far from comprehensive. Effort has also focused on creating awareness through the organization of conferences related to the subject of open data. Corinium Connected Thinking (2018) organized the DataCon Africa conference in 2016 to convene the South African data analytics community for discussions on the future of data analytics. The United Nations Global Pulse (2018) organized the Data Science Africa 2018 conference in Kenya sometime this year.

Although these efforts seek to make data more accessible in Africa, besides Code for Africa, others seem focused on making journal articles and publications accessible and on creating awareness on the value of open access. Apparently, there is no study that assesses the new role that the African university could and should assume; that of creating OSDHs so as to engage Africa's data revolution as articulated in the Africa Data Consensus document (UNECA, 2015). Moreover, no discussion exists mentioning the need to populate such OSDHs with datasets that can be statistically analyzed. These two gaps are what this discussion is about; creating data hubs to host datasets which can be accessed, used and reused for statistical and content analytical purposes by different stakeholders including academic bodies, government planners, researchers, internal and even external agencies and researchers.

6 Problem Statement

The problem being interrogated here is the inherent wastage of both traditional format data and big data, which is being produced by the digitization of institutional, public and private service delivery.

In this paper, the use of the term traditional data is a reference to data that is collected using quantitative and or qualitative strategies, as well as data stored in fixed formats or files. This data is majorly collected by both graduate students while conducting their research, as well as by intellectuals at institutions of higher learning engaged in research for scholarship. On the other hand, the use of the term big data is to mean massive quantities of both structured and unstructured data that cannot be effectively processed

with traditional database and software techniques and therefore needs special analysis tools (Laney, 2001; Oracle integrated Cloud, 2019; Techopedia Inc., 2019). Given that this data (which exists in the form of photographs, texts, video/audio, networks, web log files, web searches and pages, emails, transactional applications like ATMs, and social media sites) is generated by and collected from sensors and devices that are used daily at home, in the work place and for commercial activities such as smart phones, tablets, security cameras, machines and many others (IBM, 2018; Marr, 2018) big data therefore has three peculiar characteristics referred to as the three 'V's; volume, velocity and variety (Laney, 2001).

The problem being interrogated here, which is the inherent wastage of both traditional format and big data, extends to both internally and externally generated data. The research output achieved by intellectuals in various African institutions of higher learning is simply wasted because there are no hubs to host this data. Ideally, such data wastage can be reduced or even solved by setting up an OSDH to collect, curate, preserve and make data accessible for interrogation by interested parties. This will result in the embracing of a "diverse range of data sources, tools, and innovative technologies, to provide disaggregated data for decision-making, service delivery and citizen engagement (UNECA, 2015, p. 2)." In other words, this will engage a data revolution in Africa.

7 Research Question and Objective

The basic research question this paper seeks to answer is: What role should the university in Africa play so as to address the problem of data wastage and engage a data revolution in Africa? The attempt to answer this question reveals the main objective of this proposal which is to assess the viability of institutionalizing both traditional and Big Data by creating an OSDH in flagship universities within each African state. Such a hub will serve as a repository for data sets collected by staff and students as well as by domestic institutions and provide access to that data in an open way to encourage national data sharing. The goal, and thus expected outcome, of this paper is to stimulate a continental discussion on the utility and role of universities in the creation of data hubs to facilitate a data revolution in Africa. The ultimate objective is to contribute to the redefinition of the role of the African university in this digital age. It is undeniable that academic institutions are at the heart of solving the continents data issues and consequentially public policy and development challenges.

7.1 Significance of a Data Hub

So why should African governments and universities consider creating an OSDH? The answer is simple; because of the associated benefits of doing so. These benefits include;
– Provide a trustworthy data repository: An OSDH at a flagship university is bound to be publicly viewed as being a trustworthy, reliable and good hub where one can go to search for credible data.
– Standardization of instruction: An OSDH will avail and promote the use of scientific data for teaching students the practical aspects of data analysis thus promote learning and skill development.
– Amplify researcher visibility: Having an OSDH will duly induce intellectual exchange, promoting the formation of epistemic communities and serve as a platform for research networks and partnerships (Matheka, Nderitu, Mutonga, Otiti and Siegel, 2014; Center for Qualitative and Multi-Method Inquiry-CQMI- Syracuse, 2017). This will therefore result in researcher publicity.
– Increase data visibility and impact: Once the OSDH is functional, data will be easier to find, access, and use and even to cite (Swan & Chan, 2009; CQMI Syracuse, 2017). This is the expected impact of the OSDH, where local and regional collaborations between intellectuals will thrive given the ability to share data and participate in solving everyday societal issues (Matheka et al., 2014).
– Transparency of research process: An OSDH will contribute to making the process and products of social research at the university more transparent by increasing the openness and access to data

which will consequently facilitates the replication, reproduction, and assessment of empirically based analysis (CQMI Syracuse, 2017).

– Standardize data generation requirements: An OSDH will promote a culture of good research skills to deliver quality and factual data for collection and preservation.

– Standardize data handling requirements: Data collection, preservation, harmonization and indexing activities require the standardization of procedures and services. Thus, having an OSDH will help staff meet data-management requirements that have been established by international institutions such as the National Science Foundation and other reputable organizations.

– Indigenization of the African development agenda: The governing of such a hub by the university will ensure that donor priorities in the collection of data do not override national needs and priorities.

– Manage the university's research capacity: Having a data hub will enable administrators to view and analyze performance of each discipline in terms of research output and hence manage the institutions research capacity (Swan & Chan, 2009).

– Status and opportunity: the university can use the existence of a data hub as a strategic marketing tool whose aim is to share research output with stakeholders (Swan & Chan, 2009).

– Facilitate effective public policy process: A university data hub will contribute towards improving both local and national public policy decision making as the open access will allow data communities to access and interrogate information for improved problem definition and policy formulation.

– Reduce redundancies in research output by identifying over researched topics and those that need to be researched on thus maximizing research and knowledge creation.

– Reduce costs by identifying open sources that are free to host such a program and which the university can easily acquire to institute this program.

8 Why the University?

Why is the university in Africa better placed to host such a program? So far, since the call by the Africa Data Consensus that "governments should identify a body authorized to provide credentials to data communities providing open data, based on established criteria for quality, reliability, timeliness and relevance to statistical information needs (UNECA, 2015, p. 5)" none has thus far been identified. It is only rational for this 'hub' to be set up somewhere. This paper proposes the university in Africa as the ideal host.

There are six reasons that make the university in Africa the ideal institution to host such an ambitious program. These include; role performance, home of intellectuals, enabling environment, increasing research output, a reducing digital divide, ability to consolidate government data.

First is the role the university has historically played in training Africa's human resources. Throughout its existence the African university has continued to disseminate its most vital role; that of teaching and imparting knowledge to the masses. Besides teaching, institutions of higher learning have struggled to perform the other two roles; research and community engagement (Andoh, 2017). Through research and analysis of national issues, African universities and intellectuals have always attempted to contribute to the development of their countries and should continue to do so in this technological age by boldly and courageously embarking on this project. Second, the university in Africa is 'home' of the intellectual and of various other skills (librarians and ICT experts) that are significant in the data value chain. This reality necessitates that efforts throughout the data value chain be focused at institutions of higher learning. Therefore, engaging the scholars, students, faculty members, librarians and all talented people 'residing' in academic institutions to participate in the OSDH initiative is necessary.

Third, there is an enabling environment in Africa for any institution to engage in such an effort. The African Union (AU) Charter, Article 6 (3) on public service and administration recommends African governments establish effective communication systems to inform and enhance access to information by the public (AU, 2011). Essentially this Charter created the necessary environment to initiate a data revolution in Africa. This charter, together with the open-data wave, is inducing the necessary political

support for institutions to start such a project. This political desire to ride the digital wave should be applied to emphasize the crucial role and significance of the African intellectual and institutions of higher learning in the realm of research and data collection and analysis. Moreover, the flagship university in Africa has the basic infrastructure like venues, computers and skills/labor to more comfortably embark on this project.

Fourth is the increased research output in Africa: Although research output by Africans has historically been low, studies indicate that research output in Africa has tremendously increased (Andoh, 2017; Duermeijer, Amir and Schoombee, 2018). Compared to other regions, Africa has by far the strongest growing scientific production. Between 2012 and 2016, research output increased by 38.6 percent and the number of authors in Africa has increased by 43 percent over that period (Duermeijer et al.). This increased research is mostly achieved by intellectuals connected to universities and hence the data being produced can be collected and preserved for future use and reuse by interested parties.

Fifth is a reducing digital divide in Africa and universities are at the forefront of this. The administration of the OSDH or any other database requires computer science and information technology hardware's and skills. In Africa, the existences of such skills and or services to the masses are typically based in public universities and government offices – which are not accessible to the public. This distinguishes the university in Africa as the ideal host of such a program. Moreover, Africa is getting increasingly digitalized as more people get to own computers, smart phones and also have internet access. According to GSMA Intelligence (2017), over the last two years, smart phone connections have doubled in sub-Saharan Africa to nearly 200 million. According to Internet World Stats (2018) internet usage in Africa has expanded by 9942% between the years 2000 and 2017. These numbers clearly indicate that Africa is becoming more and more techno savvy and can, and should thus, ride the data wave of development which oftentimes is directly tied to research, the availability of data, statistics and information (TL First Group, 2017). This means that engaging universities will be a wise and rational move to make.

Six, the African university, being a public institution, and given their geopolitical location and status, can facilitate the adoption and creation of "an inclusive data ecosystem involving government, private sector, academia, civil society, local communities and development partners that tackles the informational aspects of development decision-making in a coordinated way" (UNECA, 2015, p. 4). Thus, the university in Africa is also uniquely placed to consolidate data currently existing in government silos. Given that data in Africa's public institutions exists in silos (storage towers), also referred to as data warehouses, creating a hub at a public university would hopefully solve the most critical issue challenging its use and allow for the integration of the data which will then allow more precise insights into problems and their alternatives. Additionally, being a public institution and given its role in education, it will be more likely to get government funding due to its connection and service to the masses. These reasons, which are by no means exhaustive, distinguish the university as the ideal host for such a project.

9 Methodology

This section operationalizes the program as follows; understanding the OSDH; inputs and limitations.

9.1 Understanding the Open Science Data Hub (OSDH)

This section discusses the kinds of data, and its sources, that will be collected and preserved in this hub; the intended design, the viable policy alternatives as well as the inputs necessary for the university to start this program.

So what kinds of data will such a hub collect? Most African universities and public institutions are producers and consumers of mostly traditional data. This hub will therefore collect both the traditional (qualitative and quantitative data) and big data formats. Although the traditional format data is more prevalent in African institutions, big data is becoming increasingly relevant given the technological advancements and the digitalization of services. Potential sources of data for the OSDH are threefold;

internal, domestic and external. Internal sources are those within the particular university and will initially form the primary and most significant data contribution for the OSDH. To start this program, data normally generated by faculties within the humanities and social sciences should be collected. This is because the humanities and social sciences have diverse disciplines (data communities) which are mostly concerned with human issues, opinions, behaviors and challenges. Besides, the university administration can also provide big data that is being generated by their institution's digitalization of services such as in bookstores, admissions and library services. However, there are ethical concerns surrounding the utility of big data. In particular, the issue of privacy arises when data is collected and used without the consent of contributors because this has the potential to jeopardize their privacy, security, and rights (Marr, 2018). Therefore, the challenge for governments across the world, and particularly Africa, is to have proper legislation in place to ensure that the utility of big data, and all other data, protects contributors and consumers, both individuals and institutions, from unethical practices.

The domestic sources suggested herein refer to both public institutions (other public universities, national statistics office, departments/ ministries within government) and the private sector within each state. National statistics offices will most likely have demographic data which is also vital. External sources refer to foreign organizations, non-governmental organizations and other agencies that could also contribute data for this cause. Both the domestic and external sources will initially be the secondary sources, if at all they should be considered.

9.2 Intended OSDH Design

As earlier defined, an OSDH is the creation (institutionalization) of an unhindered, center or repository of datasets that are collected, digitized, harmonized and indexed within a given university and which are subsequently archived for sharing purposes. This section discusses the applicable data hub models which provide replicable examples.

Data hub models; The intended design of the OSDH is one which resembles the Inter-university Consortium for Political and Social Research (ICPSR) data dissemination model at the University of Michigan, and the Qualitative Data Repository (QDR) at Syracuse University in New York. The ICPSR model of quantitative data collection, archiving and sharing can be said to be the modern-day pioneer of such an effort in the United States of America (USA). Since its inception in 1962, ICPSR has continually encouraged and facilitated research and instruction in the social sciences and related areas by acquiring, developing, archiving, and disseminating data and documentation relevant to a wide spectrum of disciplines, and by conducting related instructional programs (ICPSR University of Michigan, 2018). Today, more than 40 disciplines are being supported by their data which has over 9200 studies and over 72,700 data sets (ICPSR University of Michigan, 2018). On the other hand, the QDR model provides a qualitative data institutionalization model. This model is a dedicated archive for storing and sharing digital data (and accompanying documentation) generated or collected through qualitative and multi-method research in the social sciences (CQMI Syracuse, 2017).

Both these models provide search tools to facilitate the discovery of data to interested parties and also serves as a portal to material beyond its own holdings, with links to USA and international archives. Additionally, both provide leadership and training in, and work to develop and publicize common standards and practices for managing, archiving, sharing, reusing, and citing qualitative data. Both these data hubs started small and have expanded exponentially. There are plenty of other such data hubs but these two were selected because they can provide a model for institutionalizing the traditional data format (quantitative and qualitative) that are the bulk of data generated in Africa. Additionally, these two models offer the most relevant examples for the African university to emulate since they too are universities, doing what we too should and can do.

9.3 Policy Alternatives

So what are the rational data-hub options available to the African university? Rational choice theory is an approach used by social scientists to understand how humans are likely to behave when confronted with a decision to make. According to rationalist, each alternative should be objectively analyzed and selected per the costs and benefits it affords (Dye, 2013; Green, 2002). In this paper, three alternatives emerge as the most rational and viable for universities in Africa to adopt, each with its own consequences (costs and benefits):

1. Alternative 1: Have departmental/disciplinary data hubs. This requires that traditional data, generated by different schools within the social sciences, be collected, and preserved in school/department-based repositories. Thus each school or department, e.g. economics, health, chemistry, political science, will have their own data hub and will be responsible for extending access to other disciplines or schools.
   The positive consequence of such a creation is that initial strategies to curb data wastage at the disciplinary level are engaged, which would facilitate a more centralized data hub later on.
   The negative consequences are that data access and sharing with 'outsiders' beyond the discipline is limited. Additionally, control of the pace of institutionalization of the data hub will also be weak as departments may delay initiating the program.
   Viability: This option is considerable as far as the initial stages are concerned.

2. Alternative 2: Have one central data hub populated by both kinds of data from internal sources. This means having all the internally generated traditional and big data be deposited within a central hub.
   Positive consequences are that a foundation for the collection and preservation of both kinds of data gets started. Additionally, such a hub will have greater data variety for access and will call for more collaboration and coordination with the administration. Moreover, this is a more centralized data hub which will be more preferable to the university administration because it is cohesive.
   A negative consequence is that coordination and management is trickier as it is necessary to ensure that harmonization indexing is well done. Various professional levels and skills will also be necessary.
   Viability: This would mean institutionalization of all kinds of data at the university has started. This could be a scary starting point. Coordination of multiple skills, and the large amount of data to be handled, may pose a management challenge. However, it is a viable immediate goal.

3. Alternative 3: Have one central data hub populated by both kinds of data from internal and domestic sources. This means having all the internally generated traditional data (both quantitative and qualitative) and domestic data be deposited within a central hub.
   A positive consequence is that this is a more inclusive and diverse collection effort that will require engaging other universities and willing domestic data generators.
   Negative consequence: Too much data may be collected, and this may be overwhelming as a starting point because it will negatively affect harmonization and preservation processes.
   Viability: Although this means institutionalization of data at the university is a serious endeavor, this alternative is potentially overwhelming as a starting point.

These 3 alternatives can be viewed as either separate choices or as steps towards achieving the ideal data institutionalization level. In considering potential resource constraints (foundational pillars – discussed below) it is advisable to initiate a slow, progressive program, starting with either Alternative 1 or 2 and working towards Alternatives 3. However Alternative 3 should not be viewed as the last stage but instead as a path to better more diverse data-hub processes. The ultimate goal of this discussion, and desired alternative (outcome), is to have institutional data hubs in flagship universities across Africa, which would eventually transition into national data hubs. A national data hub will be able to 'hub' all kinds of data from at least the internal and domestic sources and thus sustain the Africa data revolution. Hopefully, such an achievement will also transform the data valuation culture and thus drive Africa's development agenda.

9.4 Inputs

To get this program started, there are several contributions, efforts, sacrifices and costs that both the institution and the state government will have to incur. These are discussed as essential characteristics of the flagship university and the foundational pillars as well as the outcomes.

9.4.1 Essential Characteristics of Flagship Universities in Africa

As earlier defined, the term flagship is used herein to refer to the 'star' or 'leading' university within each African state. This term is not necessarily applied in the plural, for it could be plural in one country and singular in another. To engage this program each African government needs to critically assess which one (in case of several) of the flagship universities possesses most of the necessary attributes that will be discussed here. Once the identification (biding) and selection process for a host is done, then the operationalization of the program can be undertaken and the foundational pillars established. Although the state can assess and select which institution qualifies to host this program, the selection process being advocated for here is to have the universities engage in a competitive process of bidding to identify the most qualified, willing and equipped institution to host the data hub. Either way there are about eight characteristics that the African flagship universities should have in order to host the OSDH project. These are; prestige and status, a public institution, philosophical foundations, organizational culture, program diversity (disciplines), graduate program, location and infrastructure.

First, the university selected needs to be prestigious and have status in the nation as the epitome of higher learning. Most of the flagship universities will likely be among the first public institutions of higher education to be set up within their states. Such prestige and status have most likely nurtured positive public regard which will be crucial when legitimizing the idea to the people. Although this criterion may be discriminatory against universities that are young, older and more historical universities may fail because of other factors too. However, younger universities, with the drive and ambition to grow, may be more open to innovation and take up the initiative to bid for this opportunity. All in all, selling the idea to the public for financial support is what is crucial here. Second, the university selected should be a public institution. It is essential that the flagship university be public because, any responsible government would only be comfortable channeling public funds to a public institution for such a program. A national university in Africa means that students from various backgrounds and locations within the state can be admitted so long as they achieve the qualifications required. This heterogeneity of students and staff means a variety of interest in national socioeconomic issues, hence research/ data diversity. Moreover, a public institution will likely be home to a diverse set of skills that are necessary to launch and sustain this program.

Third, is the philosophical foundations of the university which is (and should be) reflected in the mission, vision and declared core function of the institution. The institutional philosophy needs to be reviewed as it is crucial to create the hub in a university that is research oriented. Evaluators need to ask; are the stated values in line with what is needed to engage in this program? How does it compare to the philosophies of other universities within the country? Fourth is the organizational culture which refers to a system of shared assumptions, values, and beliefs, which governs how people behave in organizations (McLaughlin, 2018). An academic institution with a positive work ethic, teaching track, research and technology, will most likely be successful in hosting such a data hub. However, as earlier stated a particular university may have the location and history but fail in the philosophical and organizational culture attributes.

Fifth is the diversity of programs at the university. As recommended by the Africa Data Consensus document, existence of data communities at the university level is an indication that a variety of data will be generated, collected and curated for archiving and sharing (UNECA, 2015). If the university has only one department or school, or has one discipline and its structure is very simple, then chances are that the research output is equally narrow. A university with diverse disciplines and programs can more easily

provide the expertise needed to operationalize and implement this program. Thus, everyday issues, as well as gender related ones, will be captured in the research and data collection championed by staff members and students in the various disciplines. Sixth, having a thriving graduate program is equally important. Thus, the institution selected should also have a well-established graduate program. Across many universities in Africa, traditional data (both quantitative and qualitative data) is continuously collected by graduate students engaged in both mono-method and mixed research methodologies, to fulfil their degree requirements. It is reasonable therefore that such data contribute towards an OSDH. Essentially, the existence of a sound graduate program indicates a university's research potential and skill pool needed to initiate this program.

Seventh, the university selected should be located within the city limits of a large heterogeneous town, preferably the capital city so as to cater for consumers who may not have personal technological connections to access the data online. However, if there is more than one 'star' university and the need is to implement the program in one of these, then the choice institution should be centrally located in the state, or where population is dense and heterogeneous. This will allow interested persons easy access to the data. Also, cities have the essential urban amenities such as electricity; infrastructure and technology to jump start this program. Eight, the university to host this program also needs to have the necessary infrastructure, in the name of computers, venue, electricity and internet connection, not to mention adequate bandwidth, for this to work. Better yet, such an institution of higher learning would most likely have an ethics review committee to help maintain research and data handling standards.

Fortunately, Africa is home to quite a number of good universities many of which will easily meet the criteria stated above. However, the bigger test is how well equipped and prepared the university is at setting up the foundational pillars needed to engage this program. The foundational pillars discussed can be viewed as necessities or as potential constraints and limitation on this program.

9.4.2 Foundational Pillars (Constraints)

The foundational pillars discussed herein could also be constraints. The Africa Data Consensus document identified six foundational pillars to frame the plan of action in engaging the data revolution in Africa. These pillars are; securing political commitment, building the evidence base, embedding the data revolution in African countries, financing and sustainability; building capacities and skills, and building partnerships and synergies (UNECA, 2015, p. 10). These same pillars are discussed herein as they are similarly key in establishing a well-stocked institutional OSDH at the university.

In regards to African universities, the foundational pillar of securing political commitment simply emphasizes the importance of governmental and legislative support in the data value chain. To initiate, as well as ensure that policy instruments and essential funds are easier to get, the African university needs to engage in 'advocacy-related actions' to initiate and secure political commitments at the highest levels of government (UNECA, 2015, p. 10). This is because ideas such as this one only become polices when and if there is political will and support. The second foundational pillar is building the evidence base. This means that each university needs to seek for "the body of knowledge needed to establish the state of readiness" of the institution to operationalize and successfully implement this program (UNECA, 2015, p. 10). It is essential for the university administration to engage the various schools and disciplines within the humanities and social science college(s) and work with them to establish baseline data. This means that the institution must consider what it has, what its deficient in and thus what is needed. This may include sending some people to train and or visit the institutions that have undertaken such programs, like the University of Michigan and/or to Syracuse University in New York, to understand the kind of data processing required and to be able to operationalize and initiate the OSDH project.

The third foundational pillar of embedding the data revolution within the country relates "to key structures and processes of dialogue and consensus building needed to provide the foundation for building national and sub-national data ecosystems that will support sustainable development, with the active participation of all stakeholders and data communities" (UNECA, 2015, p. 10). Basically, this pillar

necessitates the selected university to engage in marketing strategies to ensure the data institutionalization process takes off. Marketing and awareness strategies may need to be targeted towards the faculty members, data communities, students and all relevant stakeholders including the public. The fourth foundational pillar is to consider availability of financing and thus the sustainability of the initiative. Clearly, creating a data hub at a university in Africa is quite ambitious and requires initial investment. Money will be needed to implement this program; from paying for a venue, to hiring and training of staff to run the program, to buying the computers and servers. Besides, the program will require the use of modern-day technological tools (which will keep evolving) and maintenance of skills both of which are costly. These financial constraints can be mitigated through lobbying for sufficient funds from the government, seeking for private funds as well as engaging in international collaborations for support and funding purposes.

The fifth foundational pillar is building capacities and skills. Many universities in Africa will likely lack sufficient skills to start this program. What this pillar requires the university to do is "support the strengthening of country stakeholders' capacities and skills in generating and using data and technologies needed to realize the data revolution (UNECA, 2015, p. 10). The university will thus need to identify staff with the required expertise to start this program and keep identifying talent as time goes on and the program keeps developing. Importantly, the university must purpose to train more staff on research design, data entry and analysis and on all the necessary data handling activities to ensure this program is a success. Additionally, the university will have to be more clear and concise in their training of students in research and information communication and technology (ICT) to empower them to contribute in data generation or in the data chain.

The sixth foundational pillar is building partnerships and synergies. This requires the university to "build and strengthen partnerships with other existing data revolution-related initiatives and processes – at national, continental and global levels (UNECA, 2015, p. 10)." Given its stature in academia, it is natural for many public, private and international agencies to be interested, and to actively seek to be affiliated with a flagship university. Building partnerships with data communities, both private and public, is crucial for the success of this program as these partners will be data generators, contributors or consumers. The most important partnership is with various government institutions and ministries that are already in the business of collecting data. Other worthy collaborations (with international data generators) would be with international institutions such as the World Bank, The Statistics Department of the African Development Bank, The African Union, and the Pan African Institute for Statistics, United States Agency for International Development (USAID) and the United Nations. These collaborations would assist in various ways such as contributing data, infrastructure, expertise or even funds.

9.4.3 Outputs and Outcome

What is the expected output of these inputs? In regard to OSDH, the expected output is a policy to institute this program which will see the identification of a venue and set in motion a system of collecting, curating, preserving and sharing data in an open system. The desired outcome is findable, accessible, interoperable and reusable data that will enable the university's academic body, and other interested users, to enjoy the benefits that an OSDH has to offer.

9.5 Limitations

Although having the right combination of characteristics as well as establishing the foundational pillars is crucial, the institution should also be ready to address the following limitations as they could be detrimental to the fruition of this idea. These include; the cultural environment, political constraints,

capacity constraints, ethical concerns, issues of rigidity and time.

Cultural environment: The most pertinent limitation to this program is the cultural environment in which it is suggested and will exist. As earlier stated, Africa's culture of valuing data is weak thus this idea may sound unconventional to some politically placed people. Worse is the attitude of political figures towards data and knowledge contributions from intellectuals and think tanks (Weis, 1977; Nielson, 2001). Such an attitude, which is quite common in Africa, consequently limits the utility of research data in policy formulation (Yayboke, 2017). Moreover, Africa's administrative capacity has been seriously challenged by ethnic heterogeneity, weak organizational codes of ethics, rampant disregard and contempt of the law and corruption (Transparency International, 2015). Those required to somehow support the implementation of this program may be corrupt and ask for kickbacks thus strangling the program. Also, corruption, in the form of nepotism and tribalism, may be exercised or experienced in the hiring of staff and will needlessly affect the quality of work produced. Such a practice can be tamed by a recruitment and hiring system that is independent from the university and or closed such as the online application format.

Political constraints may exist in the form of lack of sufficient political commitment, complexity of the legislations necessary and bureaucracy. Given that the African intellectual and institutions of higher learning have hardly been appreciated in the policy arena, lack of the necessary political commitment may be experienced especially in regard to funds. Additionally, in countries that are resource constrained, leaders may fail to provide the necessary and sufficient commitments to such a program thus affecting output of the necessary policy instrument (legislation and funding) necessary to initiate this program. Lack or delay in delivering on the policy instruments could have adverse effects ranging from lack of funds and infrastructure, to staff members not committing to saving their work to contribute data for the hub, to even a delay in sensitizing staff on how and why they need to embrace this program (Otando, 2011). To counter this, it is crucial to engage all the significant stakeholders, more so the faculty and staff, in policy analysis earlier on. This will curb reluctance and lethargy towards the implementation of the program.

Complexity of the legislations is another political constraint. As Yayboke (2017) contends finding the appropriate, context specific policy and regulatory foundation for such a program to occur will be a big challenge given the weak government institutions that are typical of Africa. Clearly, finding or creating the right mix of laws, support, partnerships and awareness as well as capacity are crucial for this program to succeed. Also, various instruments are necessary to successfully implement this program and these need to be well coordinated and in synch to function and facilitate the program's success. The bureaucratic complexity and rigid nature of Africa's policy arena provides another political challenge. Even with the necessary political support, the limited capabilities of the African bureaucracy may complicate and delay the approval and release of the necessary instruments. These political constraints can create a positive or hostile environment to start this program. To counter this, African governments need to be informed and challenged to establish effective communication systems to inform and enhance access to information by the public (AU, 2011; Girard, 2012).

Capacity constraints: Although reducing, the digital divide is still prevalent in African countries, though at varying degrees, which adversely affects the data chain. These limitations are evidenced by lack of skills and infrastructure at both the individual and organizational levels. At the individual level, intellectuals may lack the technological skills to fully participate in data generation or utility. Apparently, many researchers in public institutions across Africa are yet to understand and utilize e-communication processes in research, including e-publishing and open-access initiatives (Mellado, 2018; Otando, 2011). At the organizational level, the African university will mostly likely be deficient in the necessary technology and skill to decipher the volume of big data they confront so as to mine and analyze any relevant data. This means that a new skill category that will be crucial in this program and which African universities are likely deficient in is the technical expertise, big data analysts and managers of such a hub. Deficiency of such skills may derail the initiation of the project in some countries. It is therefore necessary for governments to facilitate capacity building, collaborations and knowledge exchange by providing the infrastructure such as computers, and digital repositories as well as facilitating skill generation (Otando, 2011; UNECA, 2015).

Ethical concerns: The reluctance of authors to openly share their data has been related to ethical issues such as privacy, security, discrimination and property ownership rights (Marr, 2018; Parker, 2015).

In today's business world collectors of digital data can use it in unethical manner without the consent of the contributors, which is a violation of the contributor's rights. In relation to domestic data, the flagship university may experience widespread reluctance of government officials to share data that is held by their institutions. Such reluctance, contends Yayboke (2017), is typical of organizations whose data quality and accuracy is questionable. Another ethical concern may arise when data is manipulated for political reasons. Why would this happen? Well, because data has powerful political persuasion potential and so leaders wanting to subvert their obligation and public expectation may demand the data be manipulated for political purposes (Yayboke, 2017). Such cases may pose a challenge in harnessing the data revolution for sustainable development. Therefore, mature ethical behavior, work ethic and proper legislation and laws that protect both the contributor and consumer in the necessary and highest extend possible, need to be formulated and enforced to curb future challenges to data sharing (Fonseca, 2015).

Issues of rigidity: For the longest time, the university in Africa has played one major role; teaching. Requiring institutions that have been marginally participating in research production to start generating and collecting data in an open way is bound to receive some resistance. According to Cheal (2015) the agility of public universities has been compromised by the way they are organized which is for flat-bottomed stability. Given such a structure, many African universities may lack the interest to slowly or quickly alter their activities for a changing market. Essentially this means that convincing staff to value and contribute their data, to be shared in an open unrestricted format, will be difficult. Hopefully the technological revolution, open-data- wave and the necessity for data informed public policies will somehow force the African university to become agile. Agility calls for responsiveness to the evolving need of the society and of seeking new ways of understanding and resolving these (Cheal, 2015). Time is another resource that can limit the program. The institutionalization of this project is bound to take a long time. Predictably, from the passage of the policy to its implementation, a few years will have elapsed. This presents a unique opportunity for slow implementation or for resistance to brew. Either way, time must be carefully managed to tame skepticism and opposition to this program.

A careful analysis of the foundational pillars and strategic planning approach to these challenges and limitations will enable the African university to apply a realistic and informed strategy to initiate such an ambitious, time appropriate, program such as this one. Such determination, boldness and focused strategy will overcome any derailment and facilitate the achievement of desired outcome/goal of having an open-science data-hub at the African university. Such an outcome will mean that there is a defined space, with computers and servers, collecting, curating, archiving and sharing data for interrogation by interested stakeholders.

10 Conclusion

Back in the 1960s, the University of Michigan courageously and boldly embarked on a data sharing initiative at a time when such a move was unheard of and technology and human skills necessary for such an endeavor were quite limited. Nonetheless, they believed in the program and today, it is the standard and epitome of data generation, collection, sharing and dissemination in the academic world. In the same token of faith and commitment, the university in Africa can undertake and commit to set the foundational pillars and acquire the needed resources and ready themselves to boldly deal with any challenges and embark on creating an OSDH at their institution. The benefits of such a hub to the students, academicians, researchers and to the nation and continent at large are bound to endure for a long time. The flagship university(ies), in each African state, should be naturally compelled and supported by their governments to ride this wave of data (both digital data and open access) that is sweeping the continent. Those that will boldly embrace the challenge, making the necessary adjustments and investments, will reign high in the academic relevance chart for having pioneered and contributed knowledge necessary for informed policies and subsequent development in their countries.

References

Africa Open Data Conference (2017). About. Retrieved from http://africaopendata.net/index.php/about-aodc/

African Union (2011). The African Union Charter on values and principles of public service and administration. Addis Ababa, Ethiopia.

Andoh, Harris. (2017, Sept. 15). Research: The lost mission of African Universities. University World News, Issue No: 474.

Boeckhout, Martin., Zielhuis, Gerhard A., Bredenoord, Annelien L.(2018, May 17). The FAIR guiding principles for data stewardship: Fair enough? European Journal of Human Genetics. Retrieved from https://www.nature.com/articles/s41431-018-0160-0

Center for Qualitative and Multi-Method Inquiry, Maxwell school of Citizenship and Public Affairs, Syracuse University. (2017). Qualitative Data Repository. Retrieved from https://qdr.syr.edu/about

Cheal, Catheryn. (2015, Apr. 9). Five central characteristics of agile universities. The EvoLLLution.

Code for Africa. (2018). Open Africa. Retrieved from https://africaopendata.org/

Corinium Connected Thinking. (2018). DataCon Africa: Driving business value through data driven decisions. Retrieved from https://coriniumintelligence.com/dataconafrica/

Data Governance Institute, The (2017). Data governance and data stewardship. Retrieved from http://www.datagovernance.com/adg_data_governance_governance_and_stewardship/

Duermeijer, Charon., Amir, Mohamed and Schoombee, Lucia. (2018, Mar. 22). Africa generates less than 1% of the world's research: Data analytics can change that. Elsevier Connect.

Dye, Thomas R. (2013). Understanding public policy. (14th Ed.). Upper Saddle River, NJ: Pearson Education.

Fieldman, Damon (2018). Data lakes, data hubs, federations? Which one is best? Retrieved from https://www.marklogic.com/blog/data-lakes-data-hubs-federation-one-best/

Fonseca, Tomas Lopez da. (2015). Sharing is caring. Labfolder. Retrieved from https://www.labfolder.com/blog/data-sharing/

Friesenhahn, Irene (2014, Jun. 25). Making higher education work for Africa: Facts and figures. Retrieved from https://www.scidev.net/global/education/feature/higher-education-africa-facts-figures.html

Girard, Bruce (Ed.). 2012. New mechanisms for linking research and policy: What is impact 2.0? In Bruce Girard and Estela Acosta y Lara. Impact 2.0: New mechanisms for linking research and policy. Pp. 1 -20. Retrieved from https://www.apc.org/sites/default/files/impact2point0-final.pdf

Green, Steven. (2002). Rational choice theory: An overview. Retrieved from https://business.baylor.edu/steve_green/green1.doc

GSMA Intellignece. (2017). The mobile economy in sub-Saharan Africa. Retrieved from https://www.gsmaintelligence.com/

Guiness World Records Ltd. (2018). Oldest higher learning institution: Oldest University. Retrieved from http://www.guinness-worldrecords.com/world-records/oldest-university

IBM (2018). Big data analytics. Retrieved from https://www.ibm.com/analytics/hadoop/big-data-analytics

Internet World Stats. (2018). Internet user's statistics for Africa: Africa Internet Usage, 2018 Population Stats and Facebook Subscribers. Retrieved from https://www.internetworldstats.com/stats1.htm

Inter-university Consortium for Political and Social Research, University of Michigan. (2018). Understanding ICPSR: An orientation and tour of ICPSR data services and educational resources. Retrieved from https://www.slideshare.net/ICPSR/understanding-icpsr-an-orientation-and-tours-of-icpsr-data-services-and-educational-resources

Kingdon, John W. (2011). Agendas, alternatives and public policies. 2nd Ed., Updated. Glenview, IL: Pearson Education Inc.

Laney, D. (2001). 3D data management: Controlling data volume, velocity, and variety. Application Delivery Strategies. META Group. Retrieved from http://blogs.gartner.com/doug-laney/files/2012/01/ad949-3D-Data-Management-Controlling-Data-Volume-Velocity-and-Variety.pdf

Marr, B. (2018b). 3 massive big data problems everyone should know about. Retrieved from https://www.forbes.com/sites/bernardmarr/2017/06/15/3-massive-big-data-problems-everyone-should-know-about/#fbe36dc61862

Matheka, Duncan M., Nderitu, J., Mutonga, D., Otiti, M. I., Siegel, K., Demaio, A. R. (2014, Apr. 19). Open access: Academic publishing and its implications for knowledge equity in Kenya. Globalization and Health, 10, 26

McLaughlin, John. (2018). What is organizational culture? Definitions and characteristics. Retrieved from https://study.com/academy/lesson/what-is-organizational-culture-definition-characteristics.html.

Mellado, B. (2018). The big data challenge and how Africa can benefit. Retrieved from http://theconversation.com/the-big-data-challenge-and-how-africa-can-benefit-50664

Neilson, Stephanie. 2001. IDRC-supported research and its influence on public policy; Knowledge utilization and public policy processes: A literature review. Ottawa: International Development Research Centre. Retrieved from https://web.archive.org/web/20050225034426/http://www.idrc.ca:80/evaluation/litreview_e.html.

Open Knowledge International. (2018). Open data handbook: Why open data? Retrieved from http://opendatahandbook.org/guide/en/why-open-data/

Oracle integrated Cloud. (2019). What is big data? The definition of big data. Retrieved from https://www.oracle.com/big-data/guide/what-is-big-data.html

Otando, Rosemary M. (2011). ETD2011 paper: Building institutional repositories in KLISC member institutions in Kenya: Current status and emerging challenges. Proceedings of the 14th International Symposium on Electronic Theses and Dissertations. Cape Town, South Africa, 13-17 September 2011.

Parker, Michael. (2015, Sept 18). Exploring the ethical imperative for data sharing. The National Academies of Sciences, Engineering, and Medicine. Washington (DC): National Academies Press (US).

QS Top Universities. (2018). 10 of the oldest universities in the world. Retrieved from https://www.topuniversities.com/blog/10-oldest-universities-world

Ramphele, Mamphela. (2004). The university as an actor in development. Journal of Higher Education in Africa, 2, 1, pp. 15-33.

Research Data Alliance. (2016). Improving data sharing and re-use and use in Africa. Retrieved from https://www.rd-alliance.org/improving-data-sharing-and-re-use-and-africa-2-october-2016-cape-town-south-africa

Springer Nature. (2018). Scientific data. Retrieved from https://www.nature.com/sdata/about/principles

Swan, Alma., & Chan, Leslie. (2009). Institutional advantages from open access. Open Access Scholarly Information Sourcebook.

Techopedia Inc. (2019). Big data: Definition – what does big data mean? Retrieved from https://www.techopedia.com/definition/27745/big-data

TL First Group. (2017) Delivering tomorrow, today. Retrieved from http://www.tlfirst.com/challenges-of-access-to-data-in-africa-a-two-way-conversation/

Transparency International. (2015). People and corruption: Africa survey 2015. Global Corruption Barometer. Retrieved from file:///C:/Users/Dr.%20Mutuku/Downloads/2015_GCB_SubSaharanAfrica_EN.pdf

United Nations Economic Community for Africa. (2015, Mar 29). Africa data consensus: Addis Ababa Ethiopia.

United Nations Global Pulse. (2018). Data science Africa 2018. Retrieved from https://www.unglobalpulse.org/events/data-science-africa-2018

Weiss, Carol. (1977). Research for policy's sake: The enlightenment function of social science research. Policy Analysis, 3, 4: 531-545.

Wilkinson, Mark D., Dumontier, Michel et al., (2016). The FAIR guiding principles for scientific data management and stewardship. Sci Data; 3, 160018. Retrieved from https://www.nature.com/articles/sdata201618

Woldegiorgis, Emnet Tadesse & Doevenspeck, Martin. (2013). The changing role of higher education in Africa: A historical reflection. Higher Education Studies, 3, 6. Retrieved from https://files.eric.ed.gov/fulltext/EJ1079222.pdf

Yayboke, Erol. (2017). Harnessing the data revolution to achieve the sustainable development goals. Center for Strategic and International Studies. Rowman & Littlefield. Boulder Colorado.

Zeleza, Paul T. (2006, Aug. 30). Beyond afropessimissm: Historical accounting of African universities. Retrieved from https://www.pambazuka.org/governance/beyond-afropessimism-historical-accounting-african-universities