

Publishing trends on research data management in Sub-Saharan Africa: A bibliometrics analysis

Tom Kwanya¹

Abstract

Research data management (RDM) is the all-encompassing term used to describe the processes and activities related to the creation, storage, security, preservation, retrieval, reuse and sharing of research data. As is often the case, researchers from Sub-Saharan Africa are lagging behind their counterparts in developed countries in embracing best practices in research data management. One of the factors to which this slow pace of adoption of research data management could be attributed, is inadequate research on the subject. This paper analyses the authorship, volume, visibility and quality of publications on research data management in Sub-Saharan Africa. The analysis was done using bibliometrics. The units of analysis were publications on research data management from, and on, Sub-Saharan Africa which are currently indexed in Google Scholar. This index was chosen because it is free and is reputed for its liberal selection criteria which does not favour, or discriminate, any discipline or geographic region. Data from Google Scholar was retrieved using Harzing's "Publish or Perish" software and analysed using Nees Jan van Eck's VOSviewer software. The findings of the study revealed that authorship collaboration, visibility, quality, and quantity of scholarly publications on research data management in Sub-Saharan Africa is low when compared to developed countries like the United States of America, the United Kingdom, Canada and Australia.

Keywords

Research data management, RDM, bibliometrics, informetrics, Sub-Saharan Africa, publishing trends

1. Introduction

Schöpfel et al. (2018) assert that research data is multifaceted and dynamic. This makes it easier to describe than define. However, this paper adopts the definition of research data provided by Ray (2014) as any qualitative or quantitative evidence collected, observed, generated or created through, and for, scientific research. This data is typically used to interpret, describe or understand the phenomena under study. The data enables researchers to make conclusions which validate or generate knowledge about the phenomena being studied. Briney (2015) argues that research data consists of facts and statistics which are collected for analysis and reference in regard to a specific research project. According to Ng'eno and Mutula (2018) research data is a valuable and unique resource which is irreplaceable and costly to reproduce. Therefore, research data should be managed adequately to preserve its value while enhancing its use and usability.

According to Ng'eno and Mutula (2018), RDM is the all-encompassing term applied to refer to the processes used in data creation, storage, security, preservation, retrieval, re-use and sharing. Adika and Kwanya (2020, p. 447) argue that "RDM also encompasses activities and processes aimed at enhancing the preservation, security, visibility and ethical use of research data. These activities and processes are related to the creation, organisation, structuring, naming, backing up, storage, conservation, and sharing of research data as well as all actions that guarantee the security of research data". Whyte and Tedds (2011) opine that the purpose of RDM is to facilitate effective verification of research data and enable new research to be anchored on existing research.

Effective RDM practices yield several benefits. For instance, the effective management of research data provides a sound and reliable foundation upon which to anchor future research and thereby advancing scholarship. Thus, effective RDM ensures the continued existence of valid data upon which current and future research can be founded. Effective RDM also enables reuse of research data generated, thereby saving costs. RDM also saves the researchers' time since loss of data or duplication of efforts by recreating existing data are avoided. Briney (2015) avers that access and use of existing research data enables researchers to use existing research to generate new knowledge in a cost-effective manner.

The place of research data management in the modern research ecosystem is increasingly prominent. Indeed, many research funding agencies and publishers have instituted comprehensive RDM requirements. Patterton et al. (2018) explain that researchers who apply for funding from the National Research Foundation (NRF) in South Africa are required to demonstrate how they plan to disseminate the data generated by their proposed research projects to other researchers. One recommended strategy of ensuring wide access to research data is archiving it in publicly accessible repositories. Similar trends are being applied by funding agencies in the United States of America, Australia, Canada, and the United Kingdom (Kahn et al., 2014). It has also been observed that reputable journals such as Nature require authors to avail data anchoring their manuscripts for validation (Nature, 2014). Koopman (2015) asserts that research data is often required by peer reviewers to enable them to verify the findings and prevent scholarly fraud.

Due to the growing significance attached to research data, effective RDM literacy is now an essential skill amongst researchers. According to Adika and Kwanza (2020, p. 461), RDM literacy encompasses the capacity of researchers to effectively “plan for, search, find, organise, store, secure, share research data competently”. The authors recommend a training for researchers on effective RDM. Besides basic data management, such training, they suggest, may also include how to use digital platforms for sharing research such as institutional repositories, use of credible databases of research material, as well as citation and reference management for scholarly purposes.

2. Literature review

According to Koopman (2015, p. 1), data is “the currency of academic research”. Denny et al. (2015, p. 294) assert that researchers describe data as “the lifeblood of their work”. This is due to the fact that research data is intricately intertwined with research performance, and excellence, which influence research networks and support. On their part, Chawinga and Zinn (2019) opine that scholarly advancement is propelled by research data. The growing acceptance of the significance of data in research is fanning the increased production of data from research projects. Several scholars also opine that the rapid growth of research data is catalysed by the ubiquity of digital technologies, devices and networks which make creation, processing and preservation of research data easier (Ajibaje & Mutula, 2020; Asher, 2012; Kuo & Kusiak, 2019; Kwanza et al., 2014; Neubert & Trischler, 2021). Consequently, research data is increasingly being produced in vast volumes and diverse formats (Kibe et al., 2020). The existing abundance of research data holds great potential for the advancement of science and innovation (Denny et al., 2015). However, Kibe et al. (2020) explain that the value of the abundant data cannot be unlocked without effective RDM in terms of access and analytics. According to Koopman (2015), a large portion of the existing research data is imperceptible and does not make any meaningful contribution to scientific research and development. Therefore, many scholars conclude that access to research data should be enhanced as a means of optimising its potential (Adika & Kwanza, 2020; Kuo & Kusiak, 2019; Lucas-Dominguez et al., 2021; Vlahou et al., 2021).

The visibility of existing research data can be enhanced by executing a number of strategies. One of these strategies is to sensitise researchers to understand that by sharing their research data, their research becomes more visible and thereby attracts more citations (Van Noorden, 2014). However, Chawinga and Zinn (2019) argue that publicising the benefits of sharing research data alone may not yield the desired results. Koopman (2015) opines that some researchers invest immense resources in generating, creating or accumulating research data. Thus, they hold the data sets as valuable resources which they would not easily share with anyone else. This data is so valuable to their career that they guard it jealously. Also, according to Denny *et al.* (2015), researchers are reluctant to share their data because of the fear that data may be misunderstood and applied in situations which jeopardise the integrity of the original research. Nonetheless, Anane-Sarpong *et al.* (2018) underscore the inevitability of research data sharing and urge researchers to make adequate preparations for it.

Several factors hinder research data sharing. These include “lack of time and data misappropriation at the individual level; inadequate data sharing training, absence of compensation and unfavourable internal policies at the institutional level; as well as weak policies, inadequate ethical and legal norms, lack of data infrastructure and interoperability issues at the international level” (Chiwanga & Zinn, 2019, p. 109). Denny *et al.* (2015) also explain that there is a feeling amongst some researchers, especially in developing economies, that research data sharing may in some cases lead to neo-colonialist behaviour of foreign researchers taking away valuable resources from their territories. Anane-Sarpong *et al.* (2018) explain that risks and fears of researchers in under-resourced regions, like Sub-Saharan Africa are less-reported. These include “risks faced by under-resourced scientists and institutions which are slower in translating data produced into new knowledge; absence of harmonised guidelines and structures to help address the risks and institute fairness in data-sharing rewards; and inadequate confidence in available protective safeguards including guidelines” (p. 404).

Despite all the challenges hindering effective sharing of research data, Patterton *et al.* (2018) argue that their *ubuntu*² spirit, researchers in Sub-Saharan Africa exhibit a general willingness to share data with other researchers. They suggest that this positive attitude can be harnessed to inspire researchers to disseminate research data within local research communities and thereafter, gaining the confidence to share it outside. Chawinga and Zinn (2019, p. 404) suggest that research data sharing can be enhanced by “recognising researchers who share data through data citations, acknowledgement and incentives; investing in infrastructure, conducting training and advocacy programmes; as well as formulating stringent and fair policies for data sharing”. Anane-Sarpong *et al.* (2018) also suggest that commodifying research so as to facilitate fees for shared data may motivate researchers in developing economies to share their data. The authors add that this arrangement may provide additional resources for research in developing economies.

According to Patterton *et al.* (2018), the behaviour of researchers is essentially similar globally. However, Ng’eno and Mutula (2018) conducted an analysis of diverse perspectives of RDM in the United Kingdom, Australia, Canada, United States of America, South Africa and Kenya and concluded that African researchers are lagging behind their contemporaries in the developed world in adopting research data management tenets. Pisani *et al.* (2016) identify the reasons why scholars in Africa lag behind the rest of the world in RDM. The reasons include fear to lose control of their shared research data; inadequate incentives for sharing research data; as well as lack of effective technological capacity and infrastructure relevant to RDM. Adika and Kwanya (2020) also argue that inadequate RDM literacy among researchers in Sub-Saharan Africa constrains their capacity to effectively share research data. In spite of these challenges, there are some efforts amongst researchers in Africa to share data. For instance, in South

Africa, Denny et al. (2015) found that RDM practices in the country were ad hoc and less formal. Nonetheless, the authors report that there were efforts to institutionalise and enforce RDM policies by scholarly and research funding organisations. Townsend (2021), with a focus on legal frameworks for sharing health research data in South Africa, holds a similar view and suggests new legal mechanisms to regulate and strengthen data sharing in and out of the country. The situation is more or less similar in Kenya where Ng'eno and Mutula (2018) acknowledged that initial efforts are being made but pointed out the need to strengthen institutional RDM capacities and invigorate resource mobilisation for research data sharing, preservation and reuse.

3. Rationale and methodology of study

The literature reviewed reveals that there is increased acceptance of data-driven research. Consequently, there is increased production of research data. This calls for effective RDM. As is often the case in other scholarly metrics, researchers from Sub-Saharan Africa are lagging behind their counterparts in developed countries in embracing best practices in managing research data. One of the factors to which this slow pace in research data management could be attributed, is inadequate research on the subject. Whereas several studies, as indicated in the literature review, have been conducted on diverse aspects of RDM in Sub-Saharan Africa, no study was found which has investigated the publication trends on RDM. The purpose of this paper is to analyse the quality, quantity, visibility and authorship of publications on research data management in Sub-Saharan Africa as a means of bridging the gap in literature on this subject.

Norton (2001) explains that bibliometrics is an approach in the measurement of information. Kwanya et al. (2021) explain that in this approach, metadata of publications, including author, publication date, publication channel, and citations are used to assess the quantity, quality and visibility of research output. Over the years, bibliometrics has been traditionally linked to quantitative measurement of scholarly materials as a means of quantifying research productivity and excellence. Wormell (2001) opines that bibliometrics has been widely applied to assess the production of research output. The advantages of using bibliometrics in research are numerous. However, its capacity to quantify research productivity and excellence has been acclaimed. It is also reputed to be an objective approach to examining knowledge exchange among scholars (Dayu, 2012). Nonetheless, Neuhaus and Damiel (2008) as well as Kwanya et al. (2021) acknowledge the demerits of bibliometrics. Paradoxically, one of the major demerits is linked to its quantitative focus which some scholars view as constricting qualitative perspectives to issues under research (Peng & Luo, 2021). There is also a view that bibliometrics may be prone to manipulations on issues such as citations (Kwanya et al., 2021). These demerits notwithstanding, bibliometrics offered the best mechanism for conducting this study since it can unravel issues of research which other approaches may fail to detect. Furthermore, it can also enable researchers to examine how knowledge is created and shared in scholarly communities.

Bibliometrics approaches were used to analyse publications on research data management from, and on, Sub-Saharan Africa which are currently indexed in Google Scholar. The index was chosen because it is free and is reputed to have liberal selection criteria which do not favour, or discriminate, any discipline or geographic region. The publications were identified using Harzing's "Publish or Perish" software. The search was conducted using two key phrases which were "research data management" and "Sub-Saharan Africa". A total of 184 publications were retrieved.

4. Findings and discussions

The findings are structured according to the key themes of the objectives of the study. These are quantity, quality, visibility, and authorship of research publications on RDM in Sub-Saharan Africa.

4.1 Quantity of research publications on RDM in Sub-Saharan Africa

The findings indicate that of the 184 publications retrieved, the latest were published in 2020 while the oldest was published in 1985. This implies that RDM has been a research issue in Sub-Saharan Africa for about 35 years. The publication trend over the years is as indicated in Figure 1. It is evident from the data that the number of publications on RDM grew exponentially from 2006. This indicates a growing interest in the subject over the period. The largest number of publications was put out between 2016 and 2020. The highest number of publications per year was 25, attained in 2018. This was followed by 24 publications in 2019 and 2017 as well as 23 in 2020. Only 12 publications were produced in 2016.

A quick search on Google Scholar using the Harzing's software yields more than 1,000 publications on RDM from the United States of America, Australia and the United Kingdom. For the United States of America, for instance, the first publication on RDM was registered in 1941 which is nearly fifty years before the first paper on the subject in Sub-Saharan Africa. Comparatively, in 2020 alone, 929 publications on RDM were produced in the United States of America. Similarly, a total of 369 and 330 publications on the subject were produced in the United Kingdom and Australia, respectively, in the same period. Therefore, it can be deduced from the findings above that the quantity of publications on RDM in Sub-Saharan Africa is low.

The publication trend revealed above generally follows the overall trends in the production of knowledge by Sub-Saharan Africa when compared to other economies. According to Siyanbola et al. (2016), the level of production and use of scientific knowledge by countries in Sub-Saharan Africa is low compared to the rest of the world. They add that the gap between Sub-Saharan Africa and other regions in knowledge production and use has persisted in spite of myriad strategies being executed to bridge it. According to Tijssen (2007), the impact of the research out of Africa is significantly below the world's average. It can, therefore, be concluded that the production of scientific publications on research data management in Sub-Saharan Africa, just like in other subject areas, is lower than that of the rest of the world. This can be attributed to many factors key of which are inadequate research funding and infrastructure.

Using VOSviewer, the keywords in the titles and abstracts of the retrieved papers were identified and visualised. The colour coding is used to distinguish the nodes (keywords) in the visual presentation. The connecting lines indicate linkages between the keywords. Therefore, connected keywords imply that they appeared together in either the titles (Figure 2) or abstracts (Figure 3) of the identified papers. From the figures, it is evident that the most prominent themes covered include research data management, university, research, data, data management, and South Africa. From this, it can be concluded that the publications cover a range of topics in research data management. The prominence of South Africa (as seen in Figure 3) implies that most of the publications are produced in or about South Africa. The possible higher production rate of publications on research data management by South Africa may be attributed to the fact that it is a bigger economy with greater advancement in education, science and technology than other countries in the region. Similarly, the presence of "university" as a key term in both the titles and abstracts indicates that most of the studies were focused on or conducted in universities. This is expected because universities are critical institutions in terms of research data production and use.

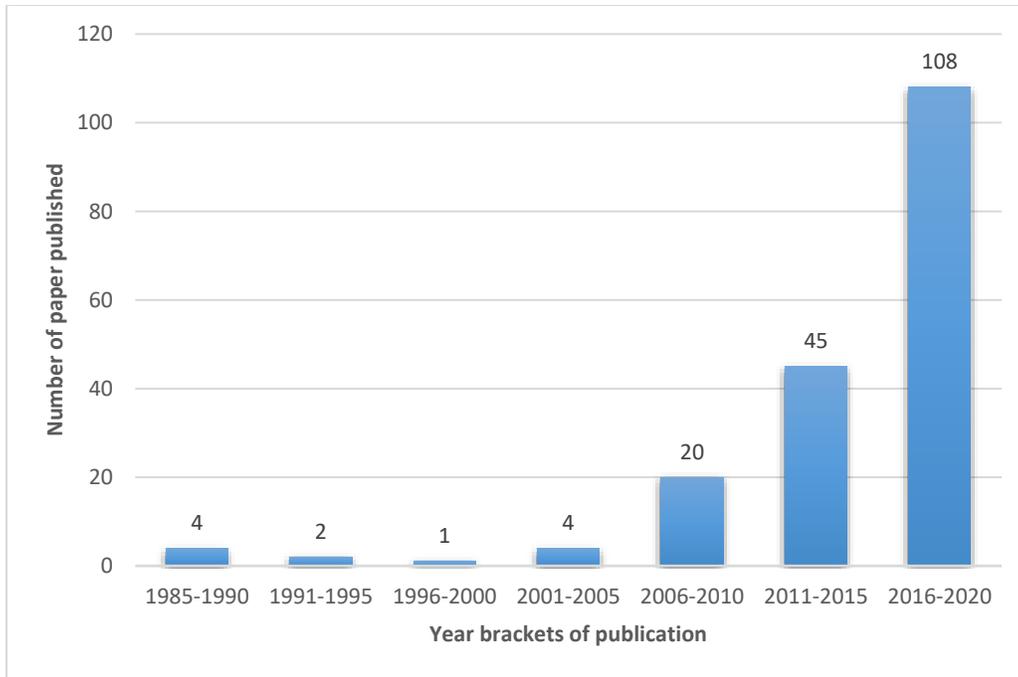


Figure 1: Publishing trends on RDM in Sub-Saharan Africa

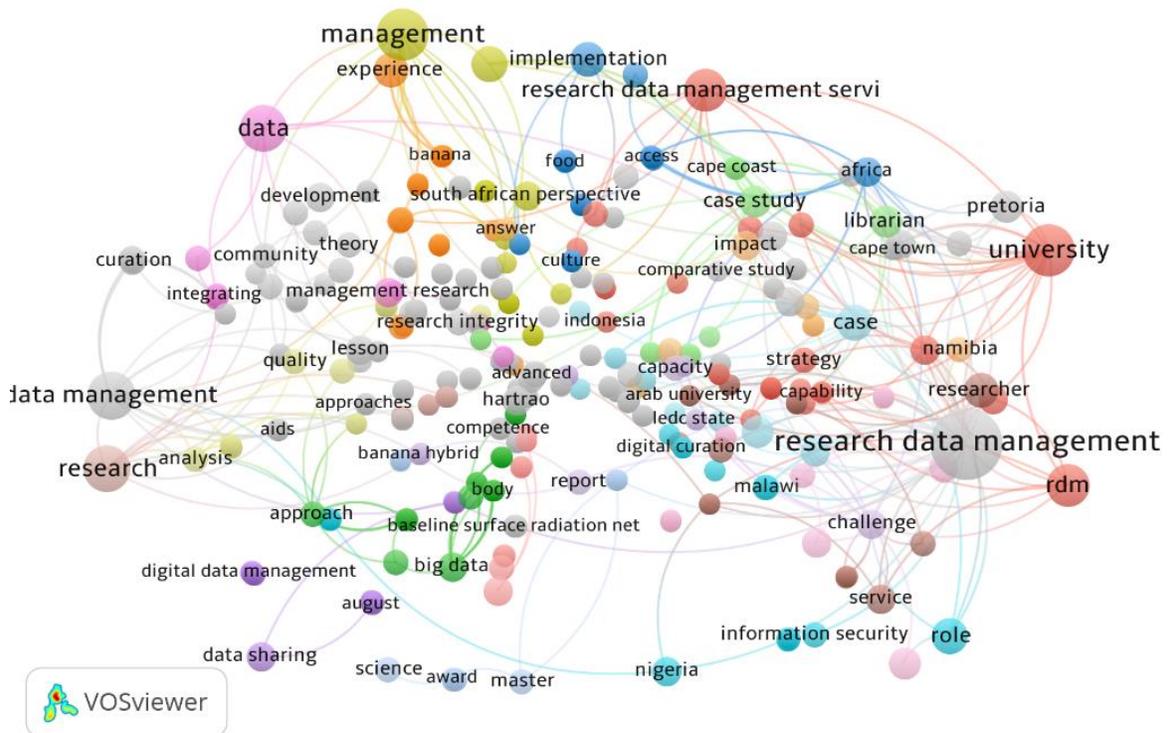


Figure 2: Keywords in titles of the publications

therefore a net “importer” of scientific products and does not influence research agenda on any thematic areas of research, including research data management.

4.3 Visibility of research publications on RDM in Sub-Saharan Africa

In the context of this paper, visibility refers to the extent to which researchers can identify, access and use a scholarly publication. According to Miguel et al. (2011), many factors determine the visibility of a research publication. They point out, however, that of these factors, the channel in which the research is published plays a pivotal role. Thus, Ale Ibrahim et al. (2014) opine that high impact channels of research publication expose research published therein more and gives them a greater possibility to be identified, accessed and cited. Similarly, Miguel et al. (2011) assert that since the full-text copies of research published through open access platforms are readily downloadable, such publications are more visible than those published in subscription-based channels. Therefore, impediments to their usability are minimal. Thus, they are likely to attract more use and citations than their counterparts published in subscription-based channels.

An analysis of the channels of publication of the works revealed that a large majority was published in subscription channels such as journals and books. Only 71 out of the 184 works were published in openly accessible channels such as institutional repositories and library web sites. Recognising the fact that many researchers in Sub-Saharan Africa, and other developing economies, have limited access to subscription-based publication channels, they are less likely to access or use the publications. This limitation on the visibility of the publications is one of the factors contributing to the low citation of the publications.

The low visibility of research products from Sub-Saharan Africa can also be attributed to the fact that most researchers do not promote their publications. In the age of social media and other open platforms, most researchers have embraced social networking sites to promote and enhance the reach of their publications. In Africa, however, Neylon et al. (2014) observed less involvement of researchers in social networking sites. They argue that by neglecting social media, researchers in Africa lose the opportunity to market their research output directly to their potential users.

4.4 Authorship of research publications on RDM in Sub-Saharan Africa

Most (118) of the works were published by more than one author. This demonstrates a high level of collaboration in terms of co-authorship of the publications. Available evidence argues that co-authored research is typically of a higher quality than singly-authored works (Hilmer & Hilmer, 2005; Hart, 2007; Andrade et al., 2010; Bidault & Hildebrand, 2014). According to Franceschet and Costantini (2010), it is not easy to avoid collaboration and co-authorship in the current scholarly communication landscape. Besides, Bidault and Hildebrand (2014) explain that co-authorship is a strategic instrument for mentoring junior academics by experienced researchers.

The authors who co-authored more than one publication were J van Wyk (6), H Pienaar (6), D Hoffmeister (4), M van Deventer (4), C Curdt (4), WD Chawinga (4), BK Avuglah (3), S Kralisch (3), F Zander (3), L Lotter (3), BV Cendon (2), D Nicholson (2), ER Chiware (2), FG Almeida (2), FJ abduldayan (2), G Coetzer (2), GL Coetzer (2); J Davidson (2), L Horton (2), L Jacobs (2), MB Macanda (2), MW Rammutloa (2), N Nhendodzasho (2), P Zibani (2), R Botha (2), S Zinn (3), SM Mutula (2), T Kramm (2), U Lang (2), and V van den Eynden (2).

Figure 4 shows the social networks created through co-authorship of the publications on RDM in Sub-Saharan Africa. It shows a number of loosely connected and less dense social networks around H Pienaar,

J van Wyk, C Curdt and F Zander. This finding indicates that the authors have collaborated less with authors with similar interests.

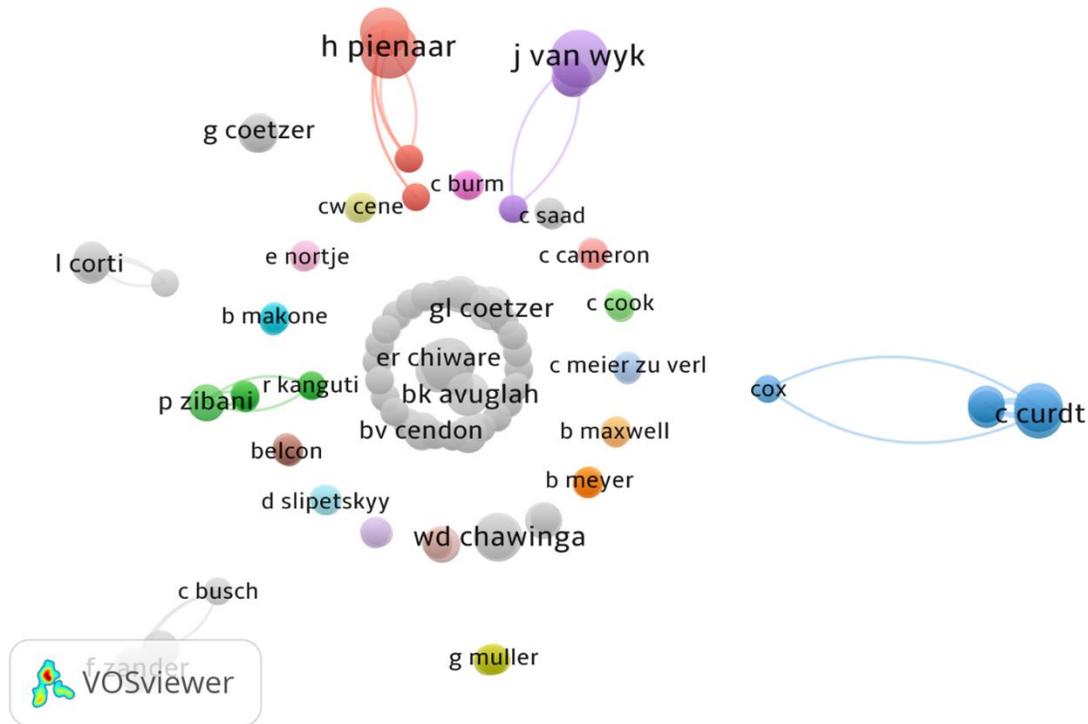


Figure 4: Social networks of co-authors

5. Conclusion

The value of research data to scientific innovation and development has grown in the recent past. Consequently, there is an over-production of research data. Researchers can benefit greatly by accessing and reusing existing data. The biggest hindrance to the effective use of research data is inadequate sharing of the data by their creators or collectors. These challenges can be overcome by scientists embracing effective research data management. Researchers from most regions have embraced the concept of research data management. However, Sub-Saharan Africa is lagging behind the rest of the world in research data management. Research on research data management can contribute greatly to the understanding and adoption of the practice. Evidence from this study reveals that the quantity, quality, visibility and scholarly collaboration on research data management in Sub-Saharan Africa is low. Therefore, Sub-Saharan perspectives on research data management are less voiced.

6. Recommendations

Based on the findings of the study, the author recommends as follows:

1. There is need to sensitise researchers in Sub-Saharan Africa about the concept of research data management as a means of stimulating them to adopt the concept. This can be done through appropriate training and publicity programmes.

2. Universities, research institutions and research funders in Sub-Saharan countries should develop policy guidelines which promote and perpetuate research data sharing and preservation. The policies should cover important elements of research data management.
3. Researchers are encouraged to collaborate with each other and create social networks which can be used to strengthen scholarly work. Increased collaboration will help the scholars to overcome some of the challenges such as lack of adequate resources. This would be achieved through pooling of resources.
4. Researchers are encouraged to publish their work in less restricted channels such as open access platforms and institutional repositories. This will likely increase the reach of the publications thereby enhancing their use and propagation.
5. Research institutions are encouraged to develop adequate infrastructure for research data management. The infrastructure may include ICT networks, laboratories and libraries. There can be no meaningful research without the requisite infrastructure.

Table 1: Citations analysis showing the articles with at least 10 citations

Cites	Authors	Title	Year
82	RJ Calantone, SK Vickery	Introduction to the special topic forum: using archival and secondary data sources in supply chain management research	2010
61	LM Delserone	At the watershed: preparing for research data management and stewardship at the University of Minnesota Libraries	2008
55	T Tuti, M Bitok, C Paton, B Makone	Innovating to enhance clinical data management using non-commercial and open source solutions across a multi-center network supporting inpatient pediatric care and research in Kenya	2016
52	DE Winickoff, K Saha, GD Graff	Opening stem cell research and development: A policy proposal for the management of data, intellectual property, and ethics	2009
48	MA Pirog	Data will drive innovation in public policy and management research in the next decade	2014
48	HH Tsai	Knowledge management vs. data mining: Research trend, forecast and citation approach	2013
40	E Chiware, Z Mathe	Academic libraries' role in research data management services: a South African perspective	2015
36	HH Tsai	Research trends analysis by comparing data mining and customer relationship management through bibliometric methodology	2011
32	M Kahn, R Higgs, J Davidson, S Jones	Research data management in South Africa: how we shape up	2014
32	QJ Groom, P Desmet, S Vanderhoeven, T Adriaens	The importance of open data for invasive alien species research, policy and management	2015
22	TA Reynolds, M Bisanzo, D Dworkis	Research priorities for data collection and management within global acute and emergency care systems	2013
19	C Willmes, D Kürner, G Bareth	Building research data management infrastructure using open source software	2014
18	M Van Deventer, H Pienaar	Research data management in a developing country: a personal journey	2015
18	F Cloete	Data analysis in qualitative public administration and management research	2007
16	AU Aydinoglu, G Dogan, Z Taskin	Research data management in Turkey: perceptions and practices	2017
14	AM Elsayed, El Saleh	Research data management and sharing among researchers in Arab universities: An exploratory study	2018
14	S Friedhoff, C Meier zu Verl, C Pietsch, C Meyer	Social research data: Documentation, management, and technical implementation within the SFB 882	2013
14	DS Bullock, M Boerngen, H Tao, B Maxwell	The Data-Intensive Farm Management Project: Changing Agronomic Research Through On-Farm Precision Experimentation	2019
13	FDS Choi	International data sources for empirical research in financial management	1988
12	O Maduka, G Akpan, S Maleghemi	Using Android and Open Data Kit Technology in Data Management for Research in Resource-Limited Settings in the Niger Delta Region of Nigeria: Cross	2017
11	ER Chiware, DA Becker	Research data management services in southern Africa: a readiness survey of academic and research libraries	2018
10	C Neylon	Building a culture of data sharing: policy design and implementation for research data management in development research	2017
10	F Huettmann	On the relevance and moral impediment of digital data management, data sharing, and public open access and open source code in (tropical) research: the Rio convention revisited towards mega science and best professional research practice	2015
10	RI Leihy, GA Duffy, E Nortje, SL Chown	High resolution temperature data for ecological research and management on the Southern Ocean Islands	2018

References

- Adika, F. O., & Kwanya, T. (2020). Research data management literacy amongst lecturers at Strathmore University, Kenya. *Library Management*, 41(6/7), 447-466.
- Ajibade, P., & Mutula, S. M. (2020). Big data research outputs in the library and information science: South African's contribution using bibliometric study of knowledge production. *African Journal of Library, Archives & Information Science*, 30(1), 49-60.
- Ale Ebrahim, N., Salehi, H., Embi, M. A., Habibi, F., Gholizadeh, H., & Motahar, S. M. (2014). Visibility and citation impact. *International Education Studies*, 7(4), 120-125.
- Anane-Sarpong, E., Wangmo, T., Ward, C.L., Sankoh, O., Tanner, M. & Elger, B.S., (2018). You cannot collect data using your own resources and put it on open access: Perspectives from Africa about public health data sharing. *Developing World Bioethics*, 18, 394–405.
- Andrade, H. B., de Los Reyes Lopez, E., & Martín, T. B. (2009). Dimensions of scientific collaboration and its contribution to the academic research groups' scientific quality. *Research Evaluation*, 18(4), 301-311.
- Bidault, F., & Hildebrand, T. (2014). The distribution of partnership returns: Evidence from co-authorships in economics journals. *Research Policy*, 43(6), 1002-1013.
- Briney, K. (2015). *Data management for researchers: Organize, maintain and share your data for research success*. Exeter: Pelagic Publishing.
- Chawinga, W.D. & Zinn, S. (2019). Global perspectives of research data sharing: A systematic literature review. *Library & Information Science Research*, 41, 109-122.
- Dayu, J. (2012) Bibliometric analysis tutorial. Retrieved October 18, 2020 from https://www.slideshare.net/dayu_jin/bibliometric-analysis-tutorial-by-dayu-jin
- Denny, S.G., Silaigwana, B., Wassenaar, D., Bull, S. & Parker, M. (2015). Developing ethical practices for public health research data sharing in South Africa: The views and experiences from a diverse sample of research stakeholders. *Journal of Empirical Research on Human Research Ethics*, 10, 290–301.
- Franceschet, M., & Costantini, A. (2010). The effect of scholar collaboration on impact and quality of academic papers. *Journal of Informetrics*, 4(4), 540-553.
- Hart, R. L. (2007). Collaboration and article quality in the literature of academic librarianship. *The journal of academic librarianship*, 33(2), 190-195.
- Hilmer, C. E., & Hilmer, M. J. (2005). How do journal quality, co-authorship, and author order affect agricultural economists' salaries?. *American Journal of Agricultural Economics*, 87(2), 509-523.
- Jahnke, L.M. & Asher, A. (2012). The problem of data: Data management and curation practices among university researchers. *The Problem of Data*, 3-31.
- Kahn, M., Higgs, R., Davidson, J. & Jones, S. (2014). Research data management in South Africa: how we shape up. *Australian Academic & Research Libraries*, 45, 296-308.
- Kibe, L., Kwanya, T. & Owano, A. (2020). Relationship between big data analytics and organisational performance of the Technical University of Kenya and Strathmore University in Kenya. *Global Knowledge Memory and Communication*. <https://doi.org/10.1108/GKMC-04-2019-0052>
- Koopman, M. M. (2015). *Data archiving, management initiatives and expertise in the Biological Sciences Department, University of Cape Town* (Master's thesis, University of Cape Town).
- Kuo, Y. H., & Kusiak, A. (2019). From data to big data in production research: the past and future trends. *International Journal of Production Research*, 57(15-16), 4828-4853.
- Kwanya, T., Kogos, A. C., Kibe, L. W., Ogolla, E. O., & Onsare, C. (2021). Cyber-bullying research in Kenya: a meta-analysis. *Global Knowledge, Memory and Communication*. <https://doi.org/10.1108/GKMC-08-2020-0124>

- Kwanya, T., Stilwell, C., & Underwood, P. (2014). Mainstreaming grey literature in research library collections in Kenya. *Libri*, 64(2), 134-143.
- Lucas-Dominguez, R., Alonso-Arroyo, A., Vidal-Infer, A., & Aleixandre-Benavent, R. (2021). The sharing of research data facing the COVID-19 pandemic. *Scientometrics*, 126(6), 4975-4990.
- Miguel, S., Chinchilla-Rodriguez, Z., & de Moya-Anegón, F. (2011). Open access and Scopus: A new approach to scientific visibility from the standpoint of access. *Journal of the American Society for information science and technology*, 62(6), 1130-1145.
- Nature. (2014). Scientific Data: data policies. Retrieved 24 May, 2020 from <http://www.nature.com/sdata/data-policies>
- Neubert, C., & Trischler, R. (2021). "Pocketing" research data? Ethnographic data production as material theorizing. *Journal of Contemporary Ethnography*, 50(1), 99-119.
- Neuhaus, C., & Daniel, H. D. (2008). Data sources for performing citation analysis: an overview. *Journal of documentation*, 64(2), 193-210.
- Neylon, C., Willmers, M. & King, T. 2014. Illustrating impact: applying altmetrics to southern African research. <https://open.uct.ac.za/handle/11427/2316>
- Ng'eno, E., & Mutula, S. (2018). Research Data Management (RDM) in agricultural research institutes: a literature review. *Inkanyiso: Journal of Humanities and Social Sciences*, 10(1), 28-50.
- Norton, M. J. (2001). *Introductory concepts in information science*. Information Today.
- Patterton, L., Bothma, T.J., & Van Deventer, M.J. (2018). From planning to practice: an action plan for the implementation of research data management services in resource-constrained institutions. *South African Journal of Libraries and Information Science*, 84(2), 14–26.
- Peng, X., & Luo, Z. (2021). A review of q-rung orthopair fuzzy information: bibliometrics and future directions. *Artificial Intelligence Review*, 1-70.
- Pisani, E., Aaby, P., Breugelmanns, J.G., Carr, D., Groves, T., Helinski, M. & Mboup, S. (2016). Beyond open data: realising the health benefits of sharing data. *Bmj* 355 I5295.
- Ray, J. M. (Ed.). (2014). *Research data management: Practical strategies for information professionals*. Purdue University Press.
- Schöpfel, J., Ferrant, C., André, F. & Fabre, R. (2018). Research data management in the French national research center (CNRS). *Data Technologies and Applications*, 52(2), 248-265.
- Siyabola, W., Adeyeye, A., Olaopa, O., & Hassan, O. (2016). Science, technology and innovation indicators in policy-making: the Nigerian experience. *Palgrave Communications*, 2(1), 1-9.
- Stewart, R. 2015. A theory of change for capacity building for the use of research evidence by decision makers in southern Africa. *Evidence & Policy*, 11(1): 547-57.
- Tijssen, R. J. W. 2007. Africa's contribution to the worldwide research literature: new analytical perspectives, trends, and performance indicators. *Scientometrics*, 71(2): 303-327.
- Townsend, B. (2021). The lawful sharing of health research data in South Africa and beyond. *Information & Communications Technology Law*, 1-18.
- Van Noorden, R. (2014). Confusion over open-data rules. *Nature*, 515, 478.
- Vlahou, A., Hallinan, D., Apweiler, R., Argiles, A., Beige, J., Benigni, A., ... & Vanholder, R. (2021). Data sharing under the general data protection regulation: Time to harmonize law and research ethics?. *Hypertension*, 77(4), 1029-1035.
- Whyte, A., & Tedds, J. (2011). *Making the Case for Research Data Management In DCC Briefing Papers*. Edinburgh: Digital Curation Centre.
- Wormell, I. (2001, September). Informetrics for informed decision making. In *A paper presented at Swedish-Lithuanian Seminar on Information Management Research Issues on* (pp. 21-22).

Endnotes

¹ Tom Kwanya is Professor of Knowledge Management in the Department of Information and Knowledge Management at the Technical University of Kenya. He can be reached on tkwanya@tukenya.ac.ke.

² African socialist philosophy of taking care of each other.