

Potential opportunities and risks of sharing agricultural research data in Tanzania

Anajoyce Samuel Katabalwa¹, Jo Bates² and Pamela Abbott³

Abstract

Purpose: The purpose of this paper was to examine the potential opportunities and risks of sharing agricultural research data in Tanzania identified in the existing research literature.

Design/methodology/approach: The study involved a review of the literature on research data sharing practices.

Findings: The findings indicate that, research data sharing has significant positive benefits among researchers such as increased high research impact; enhancing international community collaboration among researchers with same interests; improving scientific transparency and accuracy of data (Rappert and Bezuidenhout, 2016); increasing research output whereby a single dataset can be used to generate more than one article by different authors; and many more. The risks hampering data sharing practices include researchers' fears that data will be scooped, poached or misused (Onyancha, 2016); unreliable electric power; lack of funding to support research data sharing activities; absence of institutional governmental support for data management; perceived lack of benefits for data sharing (Leonelli, Rappert and Bezuidenhout, 2018); and others. However, in Tanzania research data sharing is relatively new, thus, there are no governmental agency(ies) mandating or encouraging research data sharing; therefore, there is no research data management; no research open data repositories and no research data sharing policy at any agricultural institution in Tanzania. The study recommends that agricultural researchers should be encouraged to share their data, research data policy and data repositories should also be established to support data sharing practices in Tanzania.

Originality and usefulness: From the available literature, this has been the first time that an effort has been made to examine the potential opportunities and risks of sharing agricultural research data in Tanzania. The study could be used by agricultural institutions and other institutions to assess the researchers' needs in supporting research data sharing. Also, it can be used by the government and institutions to see the need of establishing open data repositories and open data policies to support research data sharing.

Keywords

Research data sharing, open data repositories, research data policy, agricultural institutions, Tanzania.

Introduction

The development of Information and Communication Technology (ICT) has brought tremendous changes in research. ICT supports Open Science in which data dissemination have triggered the development of Open Access and Open Data (Bezuidenhout *et al.*, 2017). Open access is defined as: 'free availability on public internet, permitting any user to read, download, copy, distribute, print, search, or link to full-texts of these articles, crawl them for indexing, pass them as data to software, or use them for any other lawful purpose, without financial, legal, or technical barriers other than those inseparable from gaining access to the Internet itself' (Budapest Open Access Initiative, 2002).

Open Data refers to 'data that can be freely used, re-used and redistributed by anyone, subject only, at most, to the requirements to attributed and sharealike' (Open Data Handbook, 2013). Onyancha (2016) defines research data as data from instruments including 'telescope or raw data from a mass spectrometer to digital maps or full-text documents such as those in the creation of critical editions' and consists observational data, experimental and computational data.

Various initiatives have been established in supporting open data. Initiatives such as Elixir, OpenAIRE and European Open Science Cloud supports data sharing and re-use (Leonelli *et al.*, 2018). Moreover, the Guardian's 'Free our Data' campaign forced the UK government to release 8000 open datasets (Randalls, 2017). In Africa, the African Open Science Platform seeks to foster openness in African science (Leonelli *et al.*, 2018). In Tanzania, some related initiatives started early in the 2000s. The government's open data portal was launched in 2014 (Twaweza, 2017). The portal hosts data on health, education, water services and Tanzania National Bureau Statistics and efforts are underway to include Agriculture and Transport datasets (URT, 2017). Furthermore, Tanzania hosted an Africa Open Data Conference where open data sharing was emphasized (URT, 2015).

Methodology

The study involved a review of the literature on research data sharing practices. Specifically, a thematic review was conducted using some selected papers on data sharing practices. The researchers consulted databases such as Google scholar and Google to select the papers used in this study. The keywords used were 'open data', 'open data AND Africa', 'open data AND Tanzania', 'research data sharing', 'research data AND Africa', and 'research data AND Tanzania'.

Research environments in Tanzania

Tanzania is among low-resource countries which (Leonelli *et al.*, 2018) comment that they are challenged with 'low-resourced environment with intermittent access to a broadband connection and inexpensive outdated instruments' affecting excellence of research conducted and research goals choices. Researchers' (faculty members) working environments at Tanzania's agricultural institutions are equivalent to other developing countries. Researchers work from offices having desktop computers connected with Internet. However, not all researchers have computers; forcing others to buy their laptops and software using their own money; many offices have old computers, challenged with unreliable bandwidth and frequent power outage as reported by Bezuidenhout *et al.* (2017) and Rappert and Bezuidenhout (2016). Confirming Bezuidenhout's *et al.* (2017) study concerning institutions' lacking proxy servers; very few institutions have a proxy server supporting access to information resources outside the IP range of the institution, enabling researchers to continue working from home.

Agricultural institutions have field-based laboratories and research-based laboratories that are relatively in good conditions. Although, institutions are doing laboratories maintenance, buying new equipment whenever funds are available; they are lacking good, modern equipment, chemical reagents and other important tools. Therefore, sometimes researchers are forced to send compounds to other laboratories within Tanzania or abroad for activity testing. Rappert and Bezuidenhout (2016) argue that sending compounds abroad entails costs and delays in paper publishing.

High teaching and supervision workloads (Bezuidenhout *et al.*, 2017; Rappert and Bezuidenhout, 2016) is challenging researchers at Tanzania's agricultural institutions; limiting time for research activity. Researchers have big numbers of undergraduate students to teach (ranging from 100 to 600 students – especially for shared courses) and postgraduate students to supervise, yet same staff might have three to four courses per semester.

Although research publications are used for promotions, researchers are not given time and support to conduct research and publish. Like the researchers in Kenya (Rappert and Bezuidenhout, 2016) some researchers use their own money to pay for data collection and/or for publication fee in some journals to acquire publications to support their promotions. Moreover, each academic year researchers are evaluated in terms of research effectiveness using various criteria whereby aside from number of publications and students supervised, funds attracted by a researcher is also applied. Very few can engage in projects that attract funds at institutions. Thus, many research activities are connected to graduate students as reported by Bezuidenhout *et al.* (2017) and Rappert and Bezuidenhout (2016).

Researchers acquire research project funds either from COSTECH (a governmental agency) or international funders from abroad. To acquire funds, researchers must write a proposal and win the specific project call. Finding time for funds application and establishing collaboration is taxing (Rappert and Bezuidenhout, 2016).

Tanzania's institutions have libraries with both printed and electronic resources on subscription and open access basis. However, due to insufficient funds, only few information resources are through subscriptions. Many libraries rely on open access and resources supported by Research4Life and EIFL (Electronic Information for Libraries). Contrasting Raju's (2018) study regarding humanities and social sciences backgrounds of librarians; in Tanzania, there are varied backgrounds including Library and Information Science, Social sciences, Information studies, ICTs, and others. Nevertheless, librarians' skills are lacking to cope with changing technology (use of IT, artificial intelligence) and ever-changing needs of users.

The question is: What are the potential opportunities and risks of sharing agricultural research data in Tanzania? In order to address this question, some other sub-questions were asked as follow: Does data have any value? Are there social and political dynamics influencing research data sharing in Tanzania? Are there data sharing policies and open data repositories at agricultural institutions in Tanzania? Is there Research Data Management at agricultural institutions in Tanzania? What are the benefits of research data sharing? What are the risks of sharing agricultural research data?

Potential opportunities for sharing agricultural research data

Agricultural sector in Tanzania

Agriculture is a mainstream activity and key contributor to overall growth and development in Tanzania. Agriculture provides about 66.9% of employment, 29% of GDP, 30% of exports and 65% of inputs to the industrial sector (CPF, 2017). As a mainstream activity of many people in Sub-Sahara Africa (SSA), agriculture takes a centre stage in data sharing (Onyanha, 2016). Data sharing can contribute to the agricultural sector's improvement. The government predicts that agriculture will lead to growth and transformation of the economy (CPF, 2017). This is an opportunity for agricultural researchers to share data to contribute to economic growth.

Big data, data as a commodity

Many organisations depend on data to understand things and make correct decisions. The meteorological organisations, for example, depend heavily on data to explain the changes in climate, why it is changing and what should be done to respond (Bates, 2017a). In developed countries like the US, big companies like Bayer and Monsanto control the food system whereby data collection involves 'use of sensors, ranging from pieces attached to farm machinery to satellites' (Davidson, 2018). In Tanzania, researchers and farmers access data from the Ministry of Agriculture, and/or through the government's open data portal. This is an opportunity for researchers and other companies to invest in digital agriculture.

Data has been regarded as a precious commodity and 'new oil' fueling Fourth Industrial revolution (Bates, 2017a). According to Bates (2017a), financial markets depend on data and key actors in markets are businesses and government; 'weather derivatives are types of climate risk product traded' whereby volumes of weather data required by climate risk industry are treated as a commodity traded by the national meteorological agency. Likewise, environment data in the US and UK are traded as market goods rather than public goods (Randalls, 2017). In Tanzania, the Tanzania Meteorological Agency (TMA) is a government agency responsible for meteorological services, climate services and warnings, weather services and advisories information (URT, 2020). Not all data are considered public goods (Randalls, 2017); some data collected by TMA are treated as market goods including data requested by various industries. TMA confirms that one of its responsibilities is to 'collect fees for data, products and services

rendered' by the agency (TMA, 2020). Therefore, there is a commercial value of data and services. Bates (2017a) and Randalls (2017) questioned whether these practices are good for society, which identifies that there are some possible risks that need to be considered in sharing such research data.

Social and political dynamics influence on research data sharing

Social and political forces influence research data sharing. One force comes from journal publishers' requirement for authors to submit data as they submit papers for publishing, and the other comes from funding agencies and professional societies (Bates, 2017b; Onyanacha, 2016; Rappert and Bezuidenhout, 2016). In Tanzania, researchers are mandated only to submit their data when publishing in journals demanding data submission or to international funders of different projects who require data submission. Unlike in the US and South Africa where the US National Science Foundation and National Research Foundation (NRF) Scheme require fund beneficiaries to submit their research products (datasets, software, non-traditional research outputs, final peer-review publications) (Onyanacha, 2016); in Tanzania, there is no governmental agency mandating agricultural researchers for the same. Researchers receiving funds through COSTECH must submit reports and publications published out the project but only encouraged to share data, it is not a requirement. Thus, Tanzania contributes fewer (0.1%) data records in the Data Citation Index (Onyanacha, 2016). This is an opportunity for Tanzania to encourage researchers to share their research data and for the Tanzanian government to think of establishing a national mandate to support research data sharing.

Data sharing policies and open data repositories

Various regulatory efforts including national funding councils and institutions develop policies to encourage or mandate research data sharing (Bates, 2017b). These regulatory frameworks 'come to shape the collection, use and accessibility of data' (Randalls, 2017). Also, governments develop policies supporting data sharing. Change of governmental regulations and establishment of the UK government data policy made weather data collected by the meteorological organization, available free for re-use by anyone (Bates, 2017a). Like in many African institutions (Leonelli *et al.*, 2018); in Tanzania, there is no data sharing policy (Mushi *et al.*, 2020; Shao and Saxena, 2019), only a draft for open data policy has been developed (URT, 2017). Nevertheless, few academic institutions have started the process of establishing the policies which will support research data sharing.

Despite having institutional repository policies supporting publications sharing in few of Tanzania's academic institutions, there are no institutional open data policies and hence no institutional research open data repositories at various academic institutions including agricultural academic institutions. It would have been important to have open data policies stipulating guidelines regarding data sharing (Bezuidenhout *et al.*, 2017). Having open data policies will provide an opportunity of establishing research open data repositories at agricultural academic institutions.

Libraries have a very important role in establishing and implementing research open data repositories in Tanzania. Mushi *et al.* (2020) argue that libraries should be involved from the early stage especially in creating a needs assessment of the academic research community, participation in creation of research data management (RDM) policies, creating awareness, training researchers, and establishing open data repositories.

Research Data Management

Many African institutions have no formal research data management (RDM) (Bezuidenhout *et al.*, 2017; Leonelli *et al.*, 2018). Researchers require data management and curation skills, mentorship skills and knowledge in data engagement activities for enhanced research community sharing norms (Bezuidenhout *et al.*, 2017). Research data sharing is relatively new in Tanzania, few academic institutions are at the initial stage of establishing RDM policies and support, both are important and training for researchers is needed.

Benefits of research data sharing

Various scholars outlined the benefits regarding research data sharing. These include: increase in high research impact as many prominent scientific researches refer to data than articles because data attract more citations than articles; enhance international collaboration and community among researchers with same interests; improve scientific transparency and accuracy of data hence fostering scientific integrity (Onyanacha, 2016; Rappert and Bezuidenhout, 2016); increases research output whereby a single dataset can be used to generate more than one article by different authors; accelerate socio-economic development in SSA; increase accessibility and availability of research findings (Onyanacha, 2016); increase researchers' visibility; fostering identification of research questions and direction; help in attaining intellectual credit and peer recognition through standard citation; and improves data management capacities through greater resourcing (Rappert and Bezuidenhout, 2016). The perceived benefits of research data sharing among agricultural researchers in Tanzania require investigation.

Risk of sharing agricultural research data

Scholars have mentioned several risks hampering research data sharing among researchers. These include: researchers' fears that data will be scooped, poached or misused (Bates, 2017b; Bezuidenhout *et al.*, 2017; Onyanacha, 2016; Rappert and Bezuidenhout, 2016); unreliable power; insecure/absence of ICT infrastructure (Bezuidenhout *et al.*, 2017; Leonelli *et al.*, 2018; Rappert and Bezuidenhout, 2016); lack of skills on using repositories (Onyanacha, 2016; Rappert and Bezuidenhout, 2016; Bates, 2017b); lack of fund to support research data sharing activities (Leonelli, Rappert and Bezuidenhout, 2018; Onyanacha, 2016); absence of institutional governmental support for data management (Bezuidenhout *et al.*, 2017; Leonelli *et al.*, 2018); inadequate time to deposit data in repositories (Bezuidenhout *et al.*, 2017; Onyanacha, 2016); perceived lack of personal benefits (Leonelli *et al.*, 2018; Rappert and Bezuidenhout, 2016); confidentiality issues (Bates, 2017b; Rappert and Bezuidenhout, 2016); lack of ICT support and curation solutions; low bandwidth conditions; lack of agency to actively counter misappropriation of data; lack of confidence in researchers' data compared to data produced in developed countries; lack of training in data engagement activities (Bezuidenhout *et al.*, 2017); technological obsolescence; absence of research data management policies; unequal power relation in the setting of standards for what counts as good science worldwide (Leonelli *et al.*, 2018); and fear of losing to those who can publish easily (Rappert and Bezuidenhout, 2016). The perceived risks of data sharing by agricultural researchers in Tanzania should be investigated.

Conclusion

Research data sharing is very valuable to various institutions including agricultural institutions in Tanzania. Research data sharing is being necessitated by different forces. However, in Tanzania research data sharing is relatively new, thus, no governmental agency mandating or encouraging research data sharing, therefore, there is no research data management; no research open data repositories and no research data sharing policies at agricultural institutions in Tanzania. Moreover, the perceived benefits and risks of research data sharing among agricultural researchers in Tanzania is not known. This situation calls for further research to fill this gap.

Table 1: Summary of benefits and risks of data sharing

Benefits of data sharing	Scholars
Increase in high research impact as many prominent researches refer to data than articles	Onyanacha, 2016; Rappert and Bezuidenhout, 2016
Enhance international collaboration and community among researchers with same interests	Onyanacha, 2016; Rappert and Bezuidenhout, 2016

Improve scientific transparency and accuracy of data hence fostering scientific integrity	Onyancha, 2016; Rappert and Bezuidenhout, 2016
Increases research output whereby a single dataset can be used to generate more than one article by different authors	Onyancha, 2016
Accelerate socio-economic development in SSA	Onyancha, 2016
Increase accessibility and availability of research findings	Onyancha, 2016
Increase researchers' visibility; fostering identification of research questions and direction	Rappert and Bezuidenhout, 2016
Help in attaining intellectual credit and peer recognition through standard citation	Rappert and Bezuidenhout, 2016
Improves data management capacities through greater resourcing	Rappert and Bezuidenhout, 2016
Risks of data sharing	
Researchers' fears that data will be scooped, poached or misused	Bates, 2017b; Bezuidenhout <i>et al.</i> , 2017; Onyancha, 2016; Rappert and Bezuidenhout, 2016
Unreliable power; insecure/absence of ICT infrastructure	Bezuidenhout <i>et al.</i> , 2017; Leonelli <i>et al.</i> , 2018; Rappert and Bezuidenhout, 2016
Lack of skills on using repositories	Onyancha, 2016; Rappert and Bezuidenhout, 2016; Bates, 2017b
Lack of fund to support research data sharing activities	Leonelli, Rappert and Bezuidenhout, 2018; Onyancha, 2016
Absence of institutional governmental support for data management	Bezuidenhout <i>et al.</i> , 2017; Leonelli <i>et al.</i> , 2018
Inadequate time to deposit data in repositories	Bezuidenhout <i>et al.</i> , 2017; Onyancha, 2016
Perceived lack of personal benefits	Leonelli <i>et al.</i> , 2018; Rappert and Bezuidenhout, 2016
Confidentiality issues	Bates, 2017b; Rappert and Bezuidenhout, 2016
Lack of ICT support and curation solutions	Bezuidenhout <i>et al.</i> , 2017
Low bandwidth conditions	Bezuidenhout <i>et al.</i> , 2017
Lack of agency to actively counter misappropriation of data	Bezuidenhout <i>et al.</i> , 2017

Lack of confidence in researchers' data compared to data produced in developed countries	Bezuidenhout et al., 2017
Lack of training in data engagement activities	Bezuidenhout et al., 2017
Technological obsolescence	Leonelli et al., 2018
Absence of research data management policies	Leonelli et al., 2018
Unequal power relation in the setting of standards for what counts as good science worldwide	Leonelli et al., 2018
Fear of losing to those who can publish easily	Rappert and Bezuidenhout, 2016

References

- Bates, J. (2017a) 'Big data, open data and the climate risk market', in Brevini, B. and Murdock, G. (eds) *Carbon Capitalism and Communication*. Palgrave Macmillan, Cham, pp. 83–93. doi: https://doi.org/10.1007/978-3-319-57876-7_7.
- Bates, J. (2017b) 'The politics of data friction', *Journal of Documentation*, 74(2), pp. 412–429. doi: <https://doi.org/10.1108/JD-05-2017-0080>.
- Bezuidenhout, L. M.; Leonelli, S.; Kelly, A. H; and Rappert, B. (2017) 'Beyond the digital divide : Towards a situated approach to open data', *Science and Public policy*, 44(4), pp. 464–475. doi: <https://doi.org/10.1093/scipol/scw036>.
- Budapest Open Access Initiative (2002) *Budapest Open Access Initiative | Read the Budapest Open Access Initiative*. Available at: <https://www.budapestopenaccessinitiative.org/read> (Accessed: 23 April 2020).
- CPF (2017) *Country Programming profile for United Republic Tanzania*. Available at: <http://www.fao.org/3/a-bt133e.pdf>. (Accessed: 23 April 2020)
- Davidson, J. (2018) *Bayer, Monsanto and Big Data: Who will control our food system in the era of digital agriculture and mega-mergers?* Available at: https://medium.com/@foe_us/bayer-monsanto-and-big-data-who-will-control-our-food-system-in-the-era-of-digital-agriculture-aae80d991e4d (Accessed: 22 April 2020).
- Leonelli, S., Rappert, B. and Bezuidenhout, L. (2018) 'Introduction : Open Data and Africa', *Data Science Journal*, 17(5), pp. 1–3.
- Mushi, G. E., Pienaar, H. and Deventer, M. Van (2020) 'Identifying and Implementing Relevant Research Data Management Services for the Library at the University of Dodoma , Tanzania', *Data Science Journal*, 19(1), pp. 1–9.
- Onyancha, O. B. (2016) 'Open Research Data in Sub-Saharan Africa : A Bibliometric Study Using the Data Citation Index', *Publishing Research Quarterly*, 32, pp. 227–246. doi: <https://doi.org/10.1007/s12109-016-9463-6>.
- Open Data Handbook (2013) *What is Open Data?* Available at: <http://opendatahandbook.org/guide/en/what-is-open-data/> (Accessed: 23 April 2020).
- Raju, R. (2018) 'From “ life support ” to collaborative partnership: A local/global view of academic libraries in South Africa', *College & Research Libraries News*, pp. 30–33.
- Randalls, S. (2017) 'Commercializing environmental data: seeing like a market', in. Routledge.
- Rappert, B. and Bezuidenhout, L. (2016) 'Data sharing in low-resourced research environments',

- Prometheus*, 34(3-4), pp. 207–224. doi: <https://doi.org/10.1080/08109028.2017.1325142>.
- Shao, D. D. and Saxena, S. (2019) 'Barriers to Open Government Data (OGD) initiative in Tanzania : Stakeholders ' perspectives', *Growth and Change*, 50, pp. 470–485. doi: <https://doi.org/10.1111/grow.12282>.
- TMA (2020) *Majukumu ya Mamlaka ya Hali ya Hewa Tanzania*. Available at: <http://www.meteo.go.tz/pages/functions-of-tma> (Accessed: 21 April 2020).
- Twaweza (2017) *ICTs and National Agricultural Research Systems - Tanzania*. Available at: <https://www.twaweza.org/go/monitoring-series>. (Accessed: 23 April 2020)
- URT (2015) 'Africa Open Data Conference: Developing Africa throguh Open Data.' Available at: http://africaopendata.net/wp-content/uploads/2017/06/AODC_Report-FINAL_withLogos.pdf (Accessed: 18 April 2020).
- URT (2017) *Africa Open Data Conference: Developing Africa through Open Data*. Dar es Salaam. doi: [https://doi.org/10.1016/0144-2449\(89\)90018-3](https://doi.org/10.1016/0144-2449(89)90018-3).
- URT (2020) *Tanzania Meteorological Agency (TMA)*. Available at: <https://www.kilimo.go.tz/index.php/en/stakeholders/view/tanzania-meteorological-agency-tma>. (Accessed: 24 April 2020)

Endnotes

¹ Anajoyce Samuel is a Lecturer at the Sokoine University of Agriculture. She can be contacted through joykatabalwa@sua.ac.tz

² Dr. Jo Bates is a senior Lecturer at the University of Sheffield. She can be contacted through jo.bates@sheffield.ac.uk

³ Dr. Pamela Abbott is a senior Lecturer at the University of Sheffield. She can be contacted through p.y.abbott@sheffield.ac.uk