

# **IDRC proposal : Surveying Data Sharing Practices in African Research**

**Louise Bezuidenhout, Anelda van der Walt and Thomas Mboa Nkoudou**

## **Research problem and justification**

The nature of science is changing from a closed system to an open and sharing one. It affects virtually all components of doing science and research, and shifts in particular the focus from “publishing as fast as possible to sharing knowledge as early as possible” (von Schomberg & Hankins 2019: 19).

### **Open Data and data sharing**

Open Science is defined as research and development (R&D) that is “collaborative, transparent and reproducible and whose outputs are publicly available” (European Commission 2018: 4). It represents a new approach to research and development that encourages cooperation and sharing through the use of digital technologies and new collaborative tools.

The Open Science movement reflects a systemic change to the way science and research has been carried out for the last fifty years: shifting from the standard practices of publishing research results in scientific publications towards sharing and using all available knowledge at an earlier stage in the research process (European Commission 2012). Open Science is more than open access to publications or data; it includes many aspects and stages of research processes. It is commonly recognized that the Open Science movement rests on three key pillars: open access to scientific information, open data and open engagement with society (including firms) (OECD 2015).

Open Data forms a subset of Open Science activities. Open Data promotes the idea of research data being shared for scrutiny and re-analysis without unnecessary barriers. Open Data activities promote the notion that “making data accessible is not enough” (Science International 2015: 4). They advocate for “intelligent openness” (Royal Society 2012: 15) by which data are “discoverable, accessible, intelligible, accessible, and usable (Science International 2015: 4). In recent years, an increasing number of funders, publishers and institutions have made commitments to fostering data sharing practices within the research community. The IDRC, for instance, issued an Open Data Statement of Principles last year which outlines commitment to support the researchers they fund in their activities to manage and where possible, publish their research data (English: <https://www.idrc.ca/en/open-data-statement-principles>).

Commitments to Open Data are global. In 2019 the African Union recognised the importance of Open Science (open access and open data), declaring: “Open Science could be a game changer for achieving the Sustainable Development Goals, particularly in Africa, least developed countries, landlocked developing countries, and small island developing states (SIDS), as it has the potential to significantly increase scientific discovery and facilitate adoption of well-adapted technologies for enhancing economic competitiveness, supporting sustainable development and alleviating poverty” (AU 2019: 3).

Nonetheless, sharing data is both a practical challenge and a change of traditional attitudes to data ownership. The Open Data movement combines the “use of new digital tools, a specific set of values, and/or practices of collaboration and sharing” (Levin & Leonelli 2016: 4). Activities thus range from education and the promotion of cultures of sharing, to research data management and standard setting. These activities include not only researchers, but also scientific publishers, research institutions, national governments and funders.

Two key challenges face the current Open Data movement. The first is the understanding that there can be no “one size fits all” when it comes to research data management and “intelligent openness”. Responsible data practices may vary considerably based on the discipline, types of data and community requirements. Second, there remain considerable technical and social challenges to sharing data. These range from infrastructural challenges to scientists’ concerns about losing control of their data and being “scooped”. For researchers in Africa, these challenges are often compounded by resource-constrained research settings, historical academic legacies and a paucity of established data sharing cultures in situ.

#### Surveying research data sharing practices in different research contexts

The growing support for the Open Data movement urgently requires evidence on current data sharing practices and challenges within research communities. This has led to the dissemination of a range of surveys examining trends in data sharing practices. One example of this background work was a panel hosted at the Research Data Alliance conference in Gaborone, Botswana in 2018. The focus of the panel was understanding what quantitative research (in the form of national and global surveys) reveals about how, where and why researchers share their research data.

The findings of these surveys have considerable implications for both policy and practice. Funding agencies, research institutions and national governments can make use of this data to stimulate national dialogues to advance open research data, to develop policy and monitoring strategies. As Open Science moves into legislation (such as Plan S), this data will also become a means of benchmarking practices across countries to foster collaboration and cooperation.

Nonetheless, as there is “no one size fits all” when it comes to data management practices, there is also considerable heterogeneity in the framing and design of existing surveys. This heterogeneity reflects the difficulty of encapsulating the diversity of data sharing challenges, including:

- the ways that researchers generate, curate, store, share and reuse data
- research cultures and attitudes to sharing
- interpretations of terms such as “sharing”, “data” and “openness”
- digital and research infrastructure provision supporting data sharing activities

The variety of existing surveys of data sharing practices draw attention to two key issues addressed by this project. First, that the findings of surveys identifying attitudes to, and challenges of, data sharing cannot automatically be extrapolated to other national research contexts, disciplines or research communities. Second, that there is a great need for standardization in the design of survey questions to facilitate inter-national/disciplinary comparisons.

As yet, there have been very few surveys of data sharing practices on the African continent. Despite increased funding for research on the continent, there is little empirical evidence of current data sharing practices. A recent scoping study by the IDRC looked at the geographic coverage of the 9 largest global open data / open science surveys. These 9 surveys reported a total of 24,652 responses worldwide but only 642 responses were from the African continent (or 2.6 % of responses). None of these 9 surveys could provide any meaningful country-level picture in Africa. The surveys that focused specifically on data sharing practices amongst African researchers, such as Bezuidenhout and Chakauya 2018 made use of research networks (the NEPAD Southern African Network for the Biosciences) meaning that respondents were spread over 14 countries. The results were therefore not suited for in-depth, in-country analyses.

The following are examples of open data surveys

Bauer, B. F. (2015). *Researchers and their data. Results of an Austria survey – Report 2015. Version 1.2.* doi:10.5281/zenodo.34005

Berghmans, S. C. (2017). *Open Data: the researcher perspective - survey and case studies v1* . Mendeley Data. Retrieved from <http://dx.doi.org/10.17632/bwrnfb4bvh.1>

Bezuidenhout, L. &. (2017). *INASP survey final.pdf (Version 3)* . . figshare. doi: <https://doi.org/10.6084/m9.figshare.4818043.v3>

Nature Research. (2018). *State of Open Data 2018*. figshare. doi:<https://doi.org/10.6084/m9.figshare.7234985.v1>

Wiley. (2017). *Wiley Open Science Researcher Survey 2016*. figshare.  
doi:10.6084/m9.figshare.4748332

Wiley. (2016). *Wiley Data Sharing Survey*. figshare. Retrieved from  
<https://doi.org/10.6084/m9.figshare.3468368.v2>

Moreover, there is a paucity of evidence surrounding the impact of low-resourced research environments on data sharing practices. This impact was recently recognized by the AU, who stated: “But some challenges, such as access to internet across the continent, political governance of Open Science and the standardization of services and platforms, to name a few, still remain for Open Science in Africa. As the movement grows globally, it is important to ensure that all countries in all regions have a voice on Open Science” (AU 2019: 3).

### **Overview of project contribution**

This project has two key objectives that directly address the issues raised above. These are:

1. to survey data sharing practices amongst African researchers in 8 countries to provide in-depth evidence on data sharing practices on the continent
2. to enhance comparability between surveys by:
  - a. conducting a review of existing data sharing surveys and formulating interoperable questions
  - b. developing a re-usable survey and surveying resources (website, analysis template, report template) to encourage further surveying – both in Africa and beyond
  - c. to maximize openness in all stages of the project to further promote Open Science

### **Surveying data sharing practices amongst African researchers**

In order to effectively support research capacity building on the African continent, it is vital that academic publishers, research funding agencies and government agencies better understand researchers’ data sharing practices, challenges and perceptions. This data will be generated via the survey proposed in this project. The aim of the survey is to better understand how researchers in universities and other public research organizations manage, share and access research data. This will contribute towards the AU vision of Open Science contributing to the implementation of the Science, Technology and Innovation Strategy for Africa (STISA 2024). At present, this kind of data has not been collected at scale in Africa so this initiative would represent a novel contribution. The results of this survey will be useful for a variety of purposes, such as developing new support services, benchmarking current practices, and informing national consultations.

### Enhancing survey comparability

The design of the survey questions will be informed by existing surveys, to facilitate increased interoperability between quantitative datasets on data sharing practices. A database of questions in use by the different surveys will be grouped according to categories and made available as a resource for future research. The final survey tool, dataset, analysis code, report template and any additional project resources will be made openly available to foster re-use of the survey and enhance comparability between future data sharing surveying. This approach will underpin the commitment of this project to Open Science, and set a positive example for future quantitative research in the area.

In addition to the sharing of project resources, the findings will be published in a consolidated report. This report will be promoted to a variety of stakeholders, including the Research Data Alliance, funding councils and African libraries/research institutions. The data would also be circulated amongst those participating in the science Granting Council Initiative to inform country-level reports. It is anticipated that the data will assist these reports in identifying actions to raise awareness of data sharing challenges and opportunities in-country. The questions relating to barriers and valued services, for example, could facilitate a discussion with key stakeholder groups on how to support researchers and research organizations better manage and make best use of publicly funded research data.

Finally, the data will create a baseline profile that could be monitored over time to identify changes in the open data landscape. Disseminating the findings creates numerous opportunities for agencies involved in building research capacity on the African continent. In particular, this data will be of particular significance to agencies looking to develop Open Science, Open Data and data sharing policies. In light of the recent activities of the African Open Science Platform to promote data sharing policies and infrastructural integration (see <http://africanopenscience.org.za/>), this survey is extremely timely.

### Objectives

The general objective of this project is to create a re-usable, interoperable survey that examines data sharing practices and to disseminate it amongst African research communities. Data sharing practices include curation, archiving, dissemination and re-use strategies. This general objective will be supported by a number of specific objectives that will include:

- Develop a survey template with questions that are interoperable with other international data sharing surveys
- developing a re-usable survey and surveying resources (website, analysis template, report template) to encourage further surveying – both in Africa and beyond
- Generate a dataset of data sharing practices amongst research communities in 8 African countries
- Disseminate the research findings to stakeholders via an open dataset, open access publications, a project report, presentations and feedback workshops
- Share the research findings with international academic networks, such as the Research Data Alliance

## **Methodology**

### **Conceptual framework**

This survey takes a broad socio-technical view on data sharing. It anticipates data sharing practices to be influenced by aspects of the respondent's research environment, including social aspects, infrastructural availability, technical support and training, institutional support and national regulations. This position is in line with existing studies on data sharing practices in Africa that foreground the role of socio-technical systems in determining data sharing practices. The survey will be designed so as to gain information about these issues using Elinor Ostrom's Institutional development and Analysis framework as a guide. This will allow the information of data sharing practices to be contextualized and to provide a richer narrative on data sharing in Africa.

### **User participation**

The survey will be designed by the project team with reference to the existing surveys identified by the RDA working group on surveying data practices. The survey will be translated into French and Portuguese using professional translation services and a survey of relevant literature for appropriate terminology. It is important to offer the survey in a variety of languages, to ensure that the concepts and questions being asked are accessible to the respondents, some of which will have low levels of English literacy.

The three versions will be piloted in-country amongst research communities. Feedback from these piloting exercises will be used to refine the survey prior to dissemination. The national

funding councils, academic networks, and research and teaching institutions will be invited to assist with the dissemination of the survey.

### Data collection

Once the survey has been piloted it will be loaded onto an online survey site (JISC Online Surveys). The landing page for the survey will contain a link to the project website, a full project description, details of data management and a consent statement. The latter will highlight that participation in the survey is entirely voluntary, that respondents have the right to withdraw at any time, and that their responses are anonymous. This landing page will be available in English, French, and Portuguese.

While disseminating the survey electronically has challenges associated with it (penetrance and uptake), it is the best model available during the Corona Virus 2019 outbreak, where travel and social contact are severely limited.

### Sampling

The aim of this survey is to gather information on data management and sharing practices amongst academics generating research data in African institutions. The survey will be distributed to 8 countries that are selected to be geographically, financially and linguistically representative of the continent. These countries are drawn from the list of countries currently participating in the IDRC Science Granting Councils Initiative. They are Uganda and Ethiopia from East Africa; Burkina Faso and Senegal from West Africa and Botswana, Malawi, Mozambique and Zimbabwe from Southern Africa.

Within each of these countries the survey will be open to any researcher who meets the inclusion criteria, which are:

- Currently employed by an institute of higher education or a research facility
- Hold a bachelor degree
- Engaged in data-generating activities, or have been involved in the last 3 years
- Over 18 years of age

### Dissemination of survey

The survey will be disseminated through a number of different channels to ensure the maximal penetrance within the research community. These will include reaching out to national funding councils, academies of science, the African Association of Universities, NRENs, academic societies, and professional associations to request that they distribute the survey request. Other dissemination pathways will include social media, such as Twitter and Facebook. Reminders will be sent out each week for a month. The responses from each country will be monitored, and targeted requests may be sent to ensure adequate disciplinary and gender representation. After a month, the survey will be closed and the data analysed.

### Data analysis

Survey data will be analysed using open and freely available tools such as R. The scripts will access the published data set (see below under Data Management for details) to facilitate reproducibility and re-usability. Raw survey data will be cleaned in a freely available open tool such as OpenRefine or R. Scripts for cleaning data will be shared as described under the Data Management section.

Quantitative data will be analysed with the tidyverse package from R - currently a gold standard globally for doing data science. Tidyverse include packages for data visualisation (ggplot2), data manipulation (dplyr), and various functions for data analysis and calculation of descriptive statistics and advanced statistics.

Descriptive statistics will be calculated for all continuous variables and counts and proportions for categorical variables. Relevant modeling techniques will be decided on after development of the survey questions as it will depend on the final format of the survey instrument and will be informed by the total number of responders and quality of responses. Statistical analysis will be run within the R package.

Qualitative data will be analysed with the RQDA package, a free R package that seamlessly integrates with R to allow for statistical analysis on coding and extension of data manipulation and analysis through complementary R packages mentioned above. After development of the survey instrument an analysis plan will be drafted [see Table 1 in <https://files.eric.ed.gov/fulltext/ED567753.pdf> for an example]. The analysis plan will be reviewed and updated prior to the data analysis phase.



## **Gender considerations**

Ensuring adequate gender representation amongst respondents is a priority. Networks supporting women in research (such as the South African Women in Science and Engineering) may be contacted directly to assist with dissemination should there be an initial low response. Every effort will be made to ensure that at least 30% of respondents will be female.

## **Ethical considerations**

This survey will undergo ethics review by the School of Anthropology and Museum Ethnography ethics committee (University of Oxford, UK) prior to piloting. Via the existing Science Granting Council Initiative, the IDRC will provide a contact at each of the National Research Council. These contacts will be asked to review and comment on your survey and to endorse its dissemination. This will ensure that the correct procedures have been followed in each country. The survey participants will have their anonymity protected throughout the project, and no personal identifiers such as names, institutional names or physical locations will be collected.

As mentioned above, the landing page for the survey will contain a full project description, details of data management and a consent statement. The latter will highlight that participation in the survey is entirely voluntary, that respondents have the right to withdraw at any time, and that their responses are anonymous. This landing page will also have a full description of the data management plans for the dataset (see below) enabling participants to make informed decisions regarding their participation.

## **Training**

The survey will be developed and disseminated by the project team. Similarly, the analysis will be conducted by the same individuals. However, this project will contribute to training in two key ways. First, a detailed description of the survey development and dissemination process, as well as the survey template will be made available via FigShare or Zenodo. This will enable its future reuse by interested parties. This resource will be publicized to funding councils and research organizations across Africa.

The second way in which this project will contribute to training is via the Research Data Alliance. There is an existing working group looking at surveys of open data practices, and the experiences of this project will contribute to their future activities. A detailed presentation of the project roll-out (and the challenges faced) will be prepared for this group.

The data will also be published in an open repository and can be used subsequently in hackathons or open data days to identify challenges with data sharing and come up with strategies to address those. The data along with the published analysis scripts can be used to train researchers to perform data analysis on other survey data sets that may be relevant to their research projects.

### **Organizational matters**

This project will interact with a range of national and international organizations during survey dissemination. In order to safeguard the transparency of this project, these organizations will receive two feedback messages. The first will be issued after the survey closes to inform them of the national level of response and to thank them for their support. The second will be issued after the data analysis and contain a summary of the findings. This message will also contain links to the full project report, dataset, survey and any relevant presentations/papers.

### **Collaboration with Canadian institutions**

This project will work directly with the IDRC. No other Canadian institutions will be involved.

### **Data management plan**

A shared project folder will be created by LB in Sharepoint, the University of Oxford secure cloud storage facility. This folder will contain meeting minutes, datasets, and related documentation and scripts will be kept for the duration of the project before publication. LB

will have full ownership rights and will still have access to the folder upon completion of the project. Only project members will have access to the data stored in the folder. Long term archiving of the raw data and related documentation and scripts will be done in accordance with the relevant policies of Oxford University.

#### *Pilot data*

Data from the pilot phase of the project will be exported from the online survey tool database in comma-separated value (CSV) format. The data will be stored in the shared project folder. Cleaned datasets will be saved in the shared folder with readme files and associated scripts. If the pilot data is of sufficient quality, the project team may decide to publish the pilot data on an openly accessible data repository.

#### *Project data*

Project data will be extracted from the online survey tool database in CSV format. The data will be stored in the shared project folder. Cleaned datasets will be saved in the shared folder with readme files and associated scripts. De-identified datasets will be published in an online open data repository under an open license.

#### *Project scripts*

Scripts will be developed under version control and shared via a commonly used open source development platform such as Github. Development will happen in the open to increase scrutiny for bugs and allow the community to participate in data analysis.

## **Project schedule**

<b>Dates</b>	<b>Activities</b>	<b>Deliverables</b>
April - June 2020	Analysis of existing surveys of data sharing practices and development of survey	Survey questions
April - June 2020	Landscaping study of current data sharing regulations and legislations in Africa	Landscape report on data sharing regulations in Africa
April - June 2020	Collation of contact details for dissemination stakeholders	Dissemination strategy
June 2020	Preparation and submission of ethics review	Ethics review from SAME, University of Oxford
June 2020	Engagement with national ethics committees (or similar) in the survey countries and submission of applications (where appropriate)	Ethics clearance from survey countries
June 2020	Translation of survey materials and preparation of online survey platform	Online survey tool
June 2020	Preparation of project website	Project website
July 2020	Piloting of survey in-country and amendments to survey where necessary	Finalized survey tool
August - November 2020	Dissemination of survey (2 countries per month) and collection of data	Survey data
November 2020	Presentation of preliminary data at RDA Plenary in Costa Rica	Presentation, feedback from discussion
December 2020 - January 2021	Continued analysis of data and preparation of feedback documents and articles	Analysed dataset
January 2021	Preparation of project resources for deposition in open repositories. Final project wrap up	Feedback, Open project resources

## **Results and dissemination**

### **Feedback of project findings**

In order to maximize the dissemination of the project findings, the results will be disseminated in a number of different ways.

#### *Project feedback report and in-country workshops*

A short summary of the project findings will be prepared as a feedback report for dissemination to pertinent stakeholders, such as funding organizations, in-country research institutions, governmental departments relating to tertiary education. This dissemination will be accompanied by an in-country workshop in each of the participating countries. This will be organized in conjunction with the National Research Council and the format may vary according to the preferences of the stakeholders. This feedback report will also be available on the project website for download. The findings of the project will be linked to the curated datasets, code and methodologies (see below).

#### *Academic papers*

At least one academic paper will be prepared detailing the findings of the survey. This will be submitted to a relevant journal, such as Data Science Journal. Another paper may be prepared for a data journal that includes the dataset and a detailed description of the project methodology. These papers will be published in Open Access journals. The University of Oxford will cover the Article Processing Charges associated with Open Access publication.

#### *Conference presentations*

The results of the survey will be presented at a range of academic conferences. These will include the 2020 Research Data Alliance plenary in Costa Rica and the Society for the Social Studies of Science annual meeting. The findings will also be presented at a range of meetings determined by the IDRC in line with their capacity-building work in African funding councils.

#### *Public articles*

The findings of the project will be discussed in articles prepared for public consumption in outlets such as The Conversation and SciDevNet. These articles will also have links to the project resources (data, code, reports and articles).

*Longevity of survey and methodology*

Dataset, survey and analysis code will be curated as described above in the data management plan.

## **References**

African Union (2019). *Open Science for Africa*. Addis Ababa: African Union

Bauer, B. F. (2015). *Researchers and their data. Results of an Austria survey – Report 2015*.

Version 1.2. Zenodo: doi:10.5281/zenodo.34005

Berghmans, S. C. (2017). *Open Data: the researcher perspective - survey and case studies v1*

Mendeley Data: .doi.org/10.17632/bwrnfb4bv1.1

Bezuidenhout, L., Chakauya, E. (2017). *INASP survey final.pdf (Version 3)*. Figshare:

doi.org/10.6084/m9.figshare.4818043.v3

Bezuidenhout, L., Chakauya, E. (2018). Looking Beyond the Binaries: Hidden Concerns of Sharing Research Data by Low/Middle-Income Country Researchers. *Global Bioethics*: 29(1): 39 - 54.

European Commission (2018). *Open Science Policy Platform Recommendations*. Brussels:

European Commission.

European Commission (2012). *Commission Recommendation of 17.7.2012 on access to and preservation of scientific information*. Brussels: European Commission.

Levin, N., Leonelli, S. (2016). How Does One “Open” Science? Questions of Value in Biological Research. *Science, Technology and Human Values*: 42(2): 280 - 305.

Nature Research. (2018). *State of Open Data 2018*. Figshare.

doi:<https://doi.org/10.6084/m9.figshare.7234985.v1>

Science International (2015). *Open Data in a Big Data World*. Paris: International Council for Science (ICSU), International Social Science Council (ISSC), The World Academy of Sciences (TWAS), InterAcademy Partnership (IAP)

OECD (2015), *Making Open Science a Reality* (OECD Science, Technology and Industry Policy Papers, No. 25). OECD Publishing: Paris.

Royal Society (2012). *Science as an Open Enterprise*. London: Royal Society.

Von Schomberg, R., Hankins, J. (2019). *International Handbook on Responsible Innovation - a Global Resource*. Cheltenham: Edward Elgar.

Wiley . (2017). *Wiley Open Science Researcher Survey 2016*. Figshare:

doi:10.6084/m9.figshare.4748332

Wiley. (2016). *Wiley Data Sharing Survey*. figshare. Retrieved from

Figshare: <https://doi.org/10.6084/m9.figshare.3468368.v2>

### **Institution and personnel**

Dr Louise Bezuidenhout is a research fellow at the Institute for Science, Innovation and Society at the University of Oxford. She has extensive experience conducting research on data sharing practices in Africa. She has led both qualitative and quantitative research projects on Open Data in sub-Saharan Africa. In her work she has also collaborated with key African data sharing networks such as the African Open Science Platform, the Southern African Network for the Biosciences and the Data Intensive Research Initiative of South Africa. She also works closely with the Research Data Alliance and CODATA.

Ms Anelda van der Walt is the director and a digital scholarship consultant at Talarify, a South African-based consulting company. She has extensive experience working with a large network of African researchers in the area of digital scholarship, open science, reproducible research, and computational literacy. She has worked with CODATA and RDA in setting up the first CODATA-RDA Research Data Science Summer School in Trieste in 2016 and has been collaborating with various researchers within the RDA and CODATA network. She has experience with rolling out online surveys, data analysis, and data visualisation. Most of her past work is available under open licenses.

Dr Thomas Mboa is an expert in Open Science practices in Africa. Dr Mboa currently directs the African Institute for Open Science and Hardware in Cameroon, and runs an Open Science workspace called Mboa Labs. He has extensive experience conducting research in Francophone African countries and will lead the surveying in the Francophone and Lusophone countries.



## Curriculum Vitae: Louise Martha Bezuidenhout

### Qualifications

**2009 – 2013.** PhD (sociology). University of Exeter, UK. Pass without corrections. Funding: Wellcome Trust.

**2007 – 2008.** Erasmus Mundus MA in Bioethics: KU Leuven, Belgium; Radboud University Nijmegen, Netherlands; University of Padova, Italy. Pass summa cum laude. Funding: Erasmus Mundus Foundation.

**2003 – 2007.** PhD (cardiothoracic surgery): University of Cape Town, RSA. Pass with minor corrections. Funding: Medtronic.

**2003 – 2007.** BA (English and creative writing): University of South Africa. First class pass.

**2002.** BSc (Med) Honours (cell biology): University of Cape Town. First class pass. Funding: merit scholarship.

**1999 – 2001.** BSc (molecular and cell biology): University of Cape Town. Upper second class pass. Funding: merit scholarship.

**1997.** Matriculation: Pretoria High School for Girls, South Africa. 6 distinctions.

### Research Employment

**2017 – current.** Research fellow: University of Oxford, UK. Institute for Science, Innovation and Society

**2016 – 2017.** Research fellow: University of Notre Dame. Center for Theology, Science and Human Flourishing

**2014 – 2016.** Research fellow: University of Exeter, UK. Egenis Centre for the Study of the Life Sciences.

**2013 – 2013.** Contract researcher: Academy of Sciences of South Africa, RSA

**2008 – 2009.** Postdoctoral Researcher: University of Edinburgh, UK. Cardiovascular Sciences

### Teaching and Other Employment

**2019 – current.** Departmental lecturer: University of Oxford. Department of Continuing Education.

**2019.** General manager. TREND in Africa NGO: University of Sussex.

**2018.** Temporary teaching fellow: University College London. Department of Science and Technology Studies. Course coordinator: University College London. *Interdisciplinary Research Methods* for 1<sup>st</sup> year Bachelor of Arts, Science and Culture.

**2016 - present.** Course coordinator: CODATA-RDA Data Science Summer School. *Responsible Data Science* for Data Science Summer School

**2014 – 2015.** Part-time lecturer: University of the Witwatersrand, RSA. Steve Biko Centre for Bioethics. Course coordinator: University of the Witwatersrand (RSA). *Advanced Research Methods* for MA in Biomedical Ethics and Health Law.

**2013.** Visiting teaching fellow: Katholieke Universiteit Leuven, Belgium. *Dual-use and Data Security* for Erasmus Mundus Masters in Bioethics

### **Visiting fellowships and current honorary affiliations**

**2017 – 2019:** Member of common room, Kellogg College, University of Oxford

**2017 – 2019:** Affiliated member of Center for Theology, Science and Human Flourishing, University of Notre Dame, USA.

**2016 – 2019:** Honorary lecturer at the Steve Biko Centre for Bioethics, University of the Witwatersrand, South Africa

**2016 – 2019:** Honorary fellow at the Department of Anthropology, University of Durham, UK

**2016 – 2017:** Honorary fellow at the Egenis Centre for the Study of the Life Sciences, University of Exeter, UK

**2016 – present:** African Research Excellence Fund College of Experts

**2016:** Visiting fellow: University of Durham, UK. Department of Anthropology

**2014.** Visiting fellow: University of Oxford, UK. Helix Centre for Health, Law and Emerging Technologies

**2013.** Visiting scholar: Katholieke Universiteit Leuven, Belgium. Centre for Biomedical Ethics and Law

**2013.** Visiting Research Fellow: University of the Witwatersrand, RSA. WITS Centre for Ethics

### **Current service activities**

**2018 – present:** Lead representative for Social Sciences – University of Oxford Research Staff Society

**2018 – present:** Working group member – University of Oxford working group on health and well-being (OxRSS representative)

**2017 - present:** Committee member – University of Oxford Research Staff Society and departmental representative (anthropology)

**2017 - present:** Staff representative – Institute for Science, Innovation and Society steering panel

**2017 – present:** Associate editor for ORBIT Journal - Online Journal for Responsible Research and Innovation in ICT

**2016 – present:** Section editor for Data Science Journal

## Curriculum vitae: Thomas Hervé Mboa Nkoudou

ORCID ID: <https://orcid.org/0000-0001-9678-7765>

email: [thomasmboa@gmail.com](mailto:thomasmboa@gmail.com)

### Positions

- 2020 - .... : PostDoc – Queen Elizabeth II Scholar – University of Ottawa ;  
2020 - ... : DOAJ - Ambassador for West and Central Africa ;  
2020 - ... : Member of the Advisory Board for the BMBF (German Ministry of Education) project « Analyse the state of research and innovation in the field of digitization in Africa » ;  
2018 - ...: Researcher at Open AIR ([Open African Innovation Research](#)), Université d'Ottawa, Canada;  
2015 – 2020 : PhD student in Information and Communication, Université Laval, Canada  
2017 - ...Founder of the [Mboalab](#), Yaoundé, Cameroon.

### Scientific activities

#### Books' chapters

- “Epistemic alienation in African Scholarly communications: Open access as a Pharmakon” in Reassembling Scholarly Communications: Histories, Infrastructures, and Global Politics of Open Access, edited by Martin Ève. 2020, MIT Press ([In press](#)).
- “Towards African and Haitian universities in service to sustainable local development: the contribution of fair open science” in Contextualising Openness, edited by Leslie Chan. 2019, University of Ottawa Press (in press). ISBN: [9780776626666](#).
- “Les injustices cognitives en Afrique subsaharienne: réflexions sur les causes et les moyens de lutte.” in Justice cognitive, libre accès et savoirs locaux, edited by Florence Piron. 2016, ISBN : [978-2-924661-13-0](#)

#### Peer-reviewed Papers

- “Benefits and the hidden face of the maker movement: Thoughts on its appropriation in African context”. LIIMC EM REVISTA 2017, DOI: <http://dx.doi.org/10.18617/liinc.v13i1.3774>
- “Le libre accès vu d’Afrique francophone subsaharienne”, Revue française des sciences de l’information et de la communication. 2017. DOI: [10.4000/rfsic.3292](#)
- “Stratégies de valorisation des savoirs locaux africains : questions et enjeux liés à l’usage du numérique au Cameroun”, éthique publique. 2015, DOI: [10.4000/ethiquepublique.2343](#).

#### Scientific Reports

“Global open science hardware roadmap : making open science hardware ubiquitous by 2025”.  
(2018). By Mboa Nkoudou, T H, Bild, D, Grey, F, Toivonen, H, Serrano, J, Maestre, J P, Paz,  
B. Edited by Murillo, L F R, Molloy, J, Dosemagen, S,

### **Invited Professor**

“Hacker methodologies and Makers Space”. Program : Media Tactics for Development Advocacy  
and Social Change. 2019, University of Toronto at Scarborough.

"A Decolonized Approach to Scholarly Communication: Foundations, Challenges, and Perspectives  
in Practice and Research". 2019, FORCE11 Scholarly Communications Institute.  
<https://www.force11.org/fsci/2019/course-abstracts#PM-B4>

### **Scientific achievements with the Mboalab**

Developing a locally-manufacturable typhoid diagnostic in Cameroon: new project for 2020  
Local manufacturing of enzyme, 2019. Funded by Open Bioeconomy Lab, Cambridge UK.  
DIY Incubator, Funded by the BFOSH (Building the Free Open Science and Hardware) program of  
the Mozilla Foundation.

### **Networks**

Africa Open Science and Hardware summit (AfricaOSH). Since 2017 | Founding member and Co-  
organizer, [www.africaosh.com](http://www.africaosh.com)

Gathering for Open Science Hardware (GOSH). Since 2016 | Member

Open and Collaborative Science in Development, Network (OCSDnet). Since 2015 | Member,  
[www.ocsdnet.org](http://www.ocsdnet.org)

Projet Science Ouverte en Haiti et en Afrique Francophone (APSOHA)

### **Prizes and Awareness**

Shuttleworth Foundation Flash Grant

Expert in Digitalisation of Higher Education by DAAD

Session Chair and Reviewer in Open Science, UNESCO TECH4DEV EPFL (École polytechnique  
fédérale de Lausanne)

“Étudiant au Parcours Inspirant de l'Université Laval » Banque Desjardins

# Anelda Van der Walt

EMAIL [anelda@talarify.co.za](mailto:anelda@talarify.co.za)  
ADDRESS South Africa  
PHONE +27 (0)72 838-6969  
WEBSITE <http://www.talarify.co.za>  
ORCID 0000-0003-4245-8119  
LinkedIn <https://www.linkedin.com/in/aneldavanderwalt/>

---

## OBJECTIVE

I am driven by knowledge creation and sharing with the aim of improving lives. This includes:

- the use of technology and modern analytical skills;
- the development of human networks (communities of practice) through which knowledge can be shared; and
- an understanding of and compassion for pressing problems within communities that can be solved through knowledge creation and/or sharing.

---

## EXPERIENCE

### **Founder, Consultant & Trainer**

Talarify (PTY) LTD  
South Africa  
Dec 2014 - Currently

*Talarify exists because of the dire need for agile, contextualised, and customisable digital & computational training, knowledge sharing, and community development.*

Through Talarify we have:

- designed and implemented various capacity development initiatives around digital and computational literacy in higher education and research settings;
- supported the development of communities of practice in these areas;
- grown awareness around the importance of Open Science, Open Source Software, Open Data, Open Educational Resources, and Research Data Management practices in research; and
- developed competency in Data Analysis and Visualisation.

**Capacity Building:** <http://www.talarify.co.za/Capacity-Building/>  
**Presentations & Talks:** <http://www.talarify.co.za/Presentations/>

### **eResearch Analyst**

University of Cape Town  
South Africa  
Mar 2014 - Dec 2014

- Appointed as researcher within an academic IT support department;
- supported the establishment of the first eResearch Centre in Africa;
- provided in-house support to the IT director and core eResearch team around 21st century research needs in terms of technology and training; and
- assisted with development of infrastructure and support, as well as training for data intensive research at UCT.

### **Senior Bioinformatics Analyst**

Stellenbosch University  
South Africa  
Sep 2011 - Feb 2014

- Established the Central Analytics Facility's Bioinformatics Core within the Next Generation Sequencing Facility;
- project design, data QC, data analysis and interpretation;
- grant writing and proposal development;
- research article and report writing;
- management of high performance computer infrastructure; and
- capacity development in next generation sequencing data analysis and the use of technology in research.

## EDUCATION

### **MSc Bioinformatics (cum laude)**

University of the Western Cape  
South Africa  
2003 – 2004

### **BSc Hons Genetics (cum laude)**

Stellenbosch University  
South Africa  
2002

### **BSc Genetics, Biochemistry with Psychology**

University of Pretoria  
South Africa  
1998 – 2000

---

## SKILLS

- Project Design, Development & Management
  - Community Development & Communities of Practice
  - Remote/International Collaboration (Virtual Meeting Platforms, Collaborative Writing, Slack, etc)
  - Written & Oral Communication
  - Project Management Tools – Trello, Asana
  - Data Cleaning, Analysis & Visualisation – OpenRefine, R, Python, Jupyter, Rstudio, SQL
  - Linux Shell
  - Research computing infrastructure – High Performance Computing, cloud
  - Github, git, Github Pages, HTML, Markdown
  - Linux, Windows, & Mac
  - Office Suite & Open Source Alternatives
  - Familiarity with re-using open source/open data resources, and interpreting licensing
  - Leadership, teaching and mentorship
  - Attention to details
  - Certified Scrum Master (expired 2017)
- 

## RECENT PROJECTS

### **SADiLaR – National Capacity Development Programme**

Conceptualised a Digital Humanities Champions and capacity development programme for public universities in South Africa. Contributed to the development of project proposal and budget. Lead conversations with researchers to get buy-in. Collaborated on grant proposals with Programming Historian and other European colleagues to develop lesson materials.

**Aug 2019 – Currently**

### **Reproducibility for Everyone**

Worked with the Reproducibility for Everyone (a global group of volunteers from research and industry) to enable the project to grow and be more accessible for new members to contribute, collaborate, and benefit. Developing new contribution guidelines, website, Figshare project site and more.

**May 2019 – July 2019**

<https://www.repro4everyone.org>

### **Rural Campuses Connection Project Capacity Development Initiative (RCCP II)**

Developed and implemented a national capacity development initiative for public universities' staff and students to acquire digital and computational skills, grow communities of practice, discuss topics including Open Science, Reproducible Research, Social Media, and more. RCCP II was funded by DHET, managed by Universities South Africa, implemented by TENET and Talarify.

**Jan 2018 – March 2019**

<https://tenet-rccpii.github.io/rccpii-2018/>

## **North-West University (NWU) eResearch Initiative**

Developed an institutional eResearch initiative in collaboration with the NWU IT Director. Collaborated within IT and across the institution with Libraries, Research Office, Postgraduate Offices, and Research Faculties. Supported training initiatives, community building, awareness creation, funding applications.

**Jul 2015 – Dec 2017**

<https://www.nwu.ac.za/eresearch>

---

## **OTHER CERTIFICATIONS & QUALIFICATIONS**

### **Short Course: Women in Leadership**

GetSmarter: University of Stellenbosch  
Business School  
South Africa  
*Completed 2018*  
<http://tiny.cc/USB-Women>

### **Short Course: Project Management**

GetSmarter: University of Cape Town  
  
South Africa  
*Completed 2013*  
<http://tiny.cc/UCT-ProjectManagement>

### **Software & Data Carpentry Instructor**

The Carpentries  
The Carpentries teaches foundational coding and data science skills to researchers worldwide.  
*Qualified in 2015*  
<http://tiny.cc/Carpentries-Instructor>

### **The Carpentries Instructor Trainer**

The Carpentries  
Carpentries Instructor Trainers receive training in evidence-based methods that underpin successful pedagogy.  
*Qualified in 2016*  
<http://tiny.cc/Carpentries-Trainer>

### **Scientific Computing & Python for Data Science**

WorldQuant University/The Data Incubator  
Comprehensive introduction to scientific computing, Python, and the related tools data scientists use.  
*Completed March 2019*  
<http://tiny.cc/WQU-DataScience>

### **Machine Learning**

Coursera – Stanford University  
Learn to apply advanced machine learning algorithms to real world problems, selecting the right algorithm for the job, debugging, and improving learning algorithms' performance.  
*Completed May 2019*  
<http://tiny.cc/Coursera-ML>

### **Data Science Ethics**

Coursera – Michigan University  
Defining ethics in the big data era. Covers informed consent, data ownership, privacy, anonymity, data validity, algorithmic fairness, societal consequences, and proposes a code of ethics.  
*Completed May 2019*  
<http://tiny.cc/Coursera-Ethics>

### **Deep Learning Specialisation**

Coursera – deeplearning.ai  
A 5 module course:  
– Neural Networks & Deep Learning  
– Improving Deep Neural Networks  
– Structuring ML Projects  
– Convolutional Neural Networks  
– Sequence Models  
*Completed May 2019*  
<http://tiny.cc/Coursera-DeepLearning>