Association for Information Systems
AIS Electronic Library (AISeL)

SAIS 2019 Proceedings

Southern (SAIS)

Data and Information Literacy: Achieving Sustainable Development Goals in Africa

Shimelis Assefa
University of Denver, shimelis.assefa@du.edu

Dawit Demissie
University at Albany, State University of New York (SUNY), ddemissie@albany.edu

Follow this and additional works at: https://aisel.aisnet.org/sais2019

Data and Information Literacy: Achieving Sustainable Development Goals in Africa

Shimelis Assefa
The University of Denver
Shimelis.Assefa@du.edu

Dawit Demissie
The State University of New York at Albany
ddemissie@albany.edu

ABSTRACT

Libraries, especially academic and research libraries are increasingly taking on the task of supporting their community of users in data collection, analysis, management, and preservation – an area of charge that came to be known as research data management (Tenopir, Sandusky, Allard, & Birch, 2014). As Africa enters the post-2015 development agenda, i.e. Sustainable Development Goals, that has 17 goals and 169 targets (United Nations, 2015), the successful accomplishment and outcome of such goals and targets can only be known if appropriate data is collected, analyzed and evidence is presented. In view of these ambitious and daunting challenges, it is imperative to prepare librarians with skills and knowledge needed to manage and preserve research data. This paper aims to shed light on the requisite skills and knowledge needed by librarians to play active role in the SDG agenda. Important concepts, skills, tools, methods, and applications on the topics of data and information literacy will be investigated.

Keywords

Information literacy, sustainable development goals, Africa, academic and research libraries.

INTRODUCTION

In information literacy programs, libraries offered instructions on how to find, evaluate, and use information. Libraries have helped users to critically evaluate the veracity of information sources and the proper and ethical use of relevant information to create new knowledge. Especially in academic institutions, libraries work with scholars, faculty and students on regular basis to offer information literacy programs. Such activities and programs often take the form of new student orientation, first year seminar through graduate student workshops, and the topics range on search, authorship, ownership, and use of content (Association of College and Research Libraries. Working Group on Intersections of Scholarly Communication and Information Literacy, 2013).

Since its inception in the early 1970s (Zurkowski, 1974), information literacy has received significant focus and the amount of research and scholarly activities surrounding the topic grew exponentially in the last several decades. National and regional level societies and associations have developed standards and frameworks to guide information literacy programs. In the United States, the American Library Association (ALA) and Association for Educational Communications and Technology's landmark publication titled "*Information Power*," and the Association of College and Research Libraries (ACRL) publication "Information Literacy Competency," have both become the de facto standards for information literacy competencies from kindergarten to college, both across the US and in many other countries throughout the world (Majid, Chang, Hnin Nu Aye, May Win, & San Yu, 2015). In 2012, ACRL established a Task Force to revise and propose a new framework for Information Literacy Competency Standards for Higher Education (ACRL 2015).

In addition to information literacy sessions, recently libraries are confronted with new developments in research data and the management of large research datasets, especially in academic and research libraries. Because the effort to find, read, evaluate, and use data is different compared to bibliographic content, libraries have now started paying more attention to services and programs around data management. Data literacy is one of the efforts that libraries are trying to up the skills of their staff to support the needs of researchers and scientists. According to Mandinach and Gummer (2013) data literacy skills include knowing how to identify, collect, organize, analyze, summarize, and prioritize data, as well as developing hypotheses, identifying problems, interpreting the data, and determining, planning, implementing, as well as monitoring.

Taking the initiative to prepare librarians in Africa with skills and knowledge in data and information management could not have come at a more opportune moment. As Africa embarks with the new development goal of the United Nations, i.e., Sustainable Development Goals (SDG), it is very important that libraries play a key role in collecting, organizing, analyzing,

monitoring, and presenting data around the key goals and targets of the SDG (United Nations. General Assembly, 2015). As data-intensive research increases and researchers are required to comply with mandates from funding agencies, as journals publishers require authors to make available their datasets, and as data-driven decision making to inform policy and practice grow, the responsibilities for librarians have expanded to work with students, researchers, scientists, development professionals, and policy experts to support data management services.

The question, however, is what are the sets of skills and competencies required of librarians to become data literate? What is the current state of practice or models in terms of preparing current and future librarians in information and data literacy? What are the opportunities and challenges for libraries in Africa to provide research data services that are aligned to the SDG? What are the tools, technologies, and overall infrastructure to design data services in academic and research libraries? This paper is intended to offer a review of current practices and methods around data literacy, skills and knowledge components, tools, and infrastructure required to achieve these goals.

In other words, based on a systematic review of the literature on all aspects of data and information literacy – Research Data Management (RDM), data science, metadata, research data planning, data repository, data citation, open data planning, etc. – the purpose of this paper is to highlight trends and current practices around data management, and at the same time show what libraries in Africa can do to prepare for this area of service. Particular focus will be given to tie the discussion to the SDG and the numerous data sources that libraries in Africa can monitor, access, and archive to track progress around the 17 goals and 169 target areas of the SDG.

LITERATURE REVIEW

Data and Information Literacy

Data and information literacy, also known as DIL, goes beyond information literacy sessions that libraries have been doing for decades. In the case of information literacy, libraries are largely developing instructional programs and sessions to library users on how to find, evaluate, and use information. However, with data becoming one important asset in the research lifecycle, the information literacy program has entered a new frontier that is all new to libraries. For example, the data information literacy (DIL) project at Purdue University, working with faculty and graduate students, developed 12 data competencies that include - Data Processing and Analysis; Data Management and Organization; Data Preservation; Databases and Data Formats; Ethics and Attribution; Data Quality and Documentation; Data Curation and Reuse; Data Conversion and Interoperability; Data Visualization and Representation; Discovery and Acquisition; Metadata and Data Description; and finally Cultures of Practice (Carlson, Johnston, Westra, & Nichols, 2013).

Although there appears to be a slight distinction between information literacy and data literacy, with the former being more focused on published literature and the latter on data associated with the research itself, the boundaries between the two is fast disappearing (Perry et al., 2005). As a result, more studies are recommending the integration of data and information literacies to advance knowledge creation and support the scientific and research communities (Calzada Prado and Marzal, 2013). According to Sapp Nelson (2017), data literacy also referred to as data information literacy – a competency matrix was developed based on five existing works on the topic – i.e. (Carlson et al 2013; Qin & D'Ignazio 2010; Piorun et al 2012; Calzada Prado and Marzal Miguel 2013; Schneider 2013) – and although not all these competencies exist across the five works cited – the most comprehensive is given by Carlson et al (2013) and it includes – databases and data formats, discovery and acquisition of data, data management and organization, data conversion and interoperability, quality assurance, metadata, data curation and re-use, cultures of practice, data preservation, data analysis, data visualization, and ethics, including citation of data.

Four major universities – Purdue University, the University of Minnesota, the University of Oregon, and Cornell University – with funding from the Institute of Museum and Library Services (IMLS) have published a book titled "Data Information Literacy," and directory of case studies (http://www.datainfolit.org/). The title 'data information literacy' was adopted for the book with a deliberate attempt to tie these two emerging roles of librarians together by treating information literacy and data services as complementary rather than separate activities (Carlson and Johnston, 2015)

Data Infrastructure

There is not uniform agreement as to the nature and content of 'data literacy' education. Much of the discussion around data literacy programs center around the components of skills and knowledge that students or data service professionals need to acquire. Even here, the programs and educational efforts around 'data literacy,' are often known by different names such as –

research data management, e-science, science data literacy, data information literacy, research data literacy (e.g., Carlson et al., 2013; and Schneider, 2013) . Libraries are coping with this new and emerging area of data services through professional development and re-skilling efforts because very few libraries are able to hire specialized staff in data librarianship (Christensen-Dalsgaard et al., 2012; Ramirez, 2011). The field is relatively new for libraries and it is clear that the model and learning outcomes for data literacy programs have yet to be articulated (Koltay, 2017).

The focus of this section is to review exemplar educational efforts, as well as tools and software, repositories, policies and guidelines that combined amount to data infrastructure. It is not a realistic approach to try to list the numerous tools, tutorials, educational modules, repository systems that are out there in the public domain. Therefore, only select and highly used resources will be listed and annotated below under the following categories

Data Repositories

re3data.org. re3data.org is a global registry of research data repositories that covers research data repositories from different academic disciplines. It presents repositories for the permanent storage and access of data sets to researchers, funding bodies, publishers and scholarly institutions. re3data.org promotes a culture of sharing, increased access and better visibility of research data. The registry went live in autumn 2012 and is funded by the German Research Foundation (DFG). http://www.re3data.org/

Research Data Archive (RDA). RDA at the National Center for Atmospheric Research (NCAR) with its Global Change Master Directory Portal (*GCMD*) portal has datasets in the following topics – agriculture, atmosphere, biosphere, climate indications, land surface, oceans, solid earth, and many more. The NCAR RDA contains a large and diverse collection of meteorological and oceanographic observations, operational and reanalysis outputs, and remote sensing datasets to support atmospheric and geoscience research. The RDA contains greater than 600 dataset collections which support the varying needs of a diverse user community. https://rda.ucar.edu/

World Bank Open Data. Free and open access to global development data. The World Bank's Open Data initiative is intended to provide all users with access to World Bank data. The data catalog is a listing of available World Bank datasets, including databases, pre-formatted tables, reports, and other resources. Each of the listings includes a description of the data source and a direct link to that source. Where possible, the databases are linked directly to a selection screen to allow users to select the countries, indicators, and years they would like to search. http://data.worldbank.org/

Food and Agriculture Data. FAOSTAT provides free access to food and agriculture data for over 245 countries and territories and covers all FAO regional groupings from 1961 to the most recent year available. www.fao.org/faostat/en/

CGIAR Center for Spatial Information. The CGIAR-CSI GeoPortal is able to provide SRTM 90m Digital Elevation Data for the entire world. The SRTM digital elevation data, produced by NASA originally, is a major breakthrough in digital mapping of the world, and provides a major advance in the accessibility of high quality elevation data for large portions of the tropics and other areas of the developing world. http://srtm.csi.cgiar.org/

DATA MANAGEMENT TOOLS AND SOFTWARE; DATA MANAGEMENT PRINCIPLES, GUIDELINES

DataONE Software Tools Catalog. A comprehensive, searchable database of software tools that will facilitate your research, compiled and edited by members of the DataONE network. The database provides a brief description of a wide range of tools that are recommended for use by scientists and students, as well as additional information and links to further resources. https://www.dataone.org/software_tools_catalog

DMPTool. The DMPTool is a free service that helps researchers and institutions to create high-quality data management plans that meet funder requirements. https://dmp.cdlib.org/ or https://dmptool.org/about

DMPonline. A flexible web-based tool to assist users to create personalized plans according to their context or research funder. The tool also provides researchers examples of guidance and best practice via 'crowdsourced' links to Digital Curation Centre's (DCC) resources and external advice. https://dmponline.dcc.ac.uk/

The RSpace Platform**.** RSpace is a multi-faceted research tool in a single, secure system accessed conveniently through any modern web browser. RSpace is 'an electronic lab notebook that enables researchers to do the following. http://www.researchspace.com/enterprise/

- capture and organize data;
- use as collaboration software for groups to organize and share data;
- use as a management tool for Principal Investigator (PI) to observe and manage lab workflows; and
- use as a platform for institutions to capture, publish and archive data.

Globus. Globus is software-as-a-service for research data management, used at dozens of institutions and national facilities for moving and sharing big data. Globus provides easy-to-use services and tools for research data management, enabling researchers to access advanced capabilities using just a Web browser. Globus APIs can be leveraged to integrate Globus capabilities with third party portal and application. https://www.globus.org/

DataCite. DataCite is a leading global non-profit organization that provides persistent identifiers (DOIs) for research data. Our goal is to help the research community locate, identify, and cite research data with confidence. https://www.datacite.org/

EZID (easy-eye-dee). EZID makes it easy to create and manage long-term globally unique IDs for data and other sources. EZID is used to support the following activities. http://ezid.lib.purdue.edu/

- Create identifiers for anything: texts, data, bones, terms, etc.
- Manage your research objects more easily with shareable, unbreakable links
- Store citation information for the objects in a variety of formats
- Fit identifiers into your automated workflows with our standards-based API

THE STATE OF DATA EDUCATION IN AFRICA

In the extant literature data and information literacy is very rarely discussed in the context of African countries. When it exists, it is mainly about information literacy and not related to data literacy. Only in South Africa we have a considerable literature often in a form of conferences and workshop. For example, there was an attempt to develop a toolkit through a consultative workshop called - The Centurion Workshop and Information Literacy in the African context, held in Centurion, South Africa, on 15-17, February 2012. Another extensive program in information literacy is organized and offered by Information Training and Outreach Centre for Africa (ITOCA). ITOCA is a capacity building organization aimed at enhancing information and communications technology skills for librarians, information specialists, scientists, researchers and students in sub-Saharan Africa and so far it claims to have trained over 20k academic and research professionals in information literacy courses. (http://www.itoca.org/)

To determine the level of activities around research data services in Africa, the most comprehensive global repository system, re3data.org, is looked at by countries in Africa and very few data repository systems exist across the continent. These countries include – South Africa (5 repositories); Kenya (3 repositories), 1 repository in Cameroon, Ghana, Senegal, Ivory Coast, Tunisia, Egypt, and 2 repositories in Burkina Faso. Data portals related to the Sustainable Development Goals (SDGs) for each of the 54 countries in Africa is available (http://sdg.opendataforafrica.org/). Independent repository based on open data platform also called 'openAFRICA,' is also available (https://africaopendata.org/).

DISCUSSIONS, IMPLEMENTATION PLANS

This paper sets out the agenda to offer a conceptual framework for libraries and librarians in Africa to become data and information literate. The goal was how we prepare libraries and librarians in data information skills so they play active role to support the Sustainable Development Goals (SDG) agenda. Data and information literacy is constrained in this paper as the confluence of efforts to integrate data literacy and information literacy together. Although the literacy program was focused on information professionals in academic and research institutions, in order to support the SDG agenda in a meaningful manner, it is imperative there needs to be multi-stakeholders' partnership and collaboration that needs to happen. The academic or research library need to work, for example, with the health ministry, agricultural agencies, education departments, economic and planning offices, national level statistical offices, etc. to support data and information services.

As shown in this paper several national and international frameworks were developed to guide data and information literacy programs (for example, ALA, ACRL, SCONUL, CAUL, ANZIL, IFLA, and UNESCO). In addition, more and more scholars have published research on the same topic – data and information literacy skills for libraries and librarians (e.g., Calzada Prado & Marzal, 2013; Carlson, Johnston, Westra, & Nichols, 2013). Institutions have equally developed online educational modules and tutorials such as MANTRA (University of Edinburgh, 2017). From the reviewed materials, the takeaway message is that libraries and librarians should be prepared with skills and knowledge to understand the overall research data lifecycle – from creation, processing, analyzing, preservation, access, use, and reuse. In addition, libraries and librarians should also gain adequate knowledge of the tools, methods, and software required to support research data management such as – metadata, data organization, data cleaning, data management plans, data curation, data citation, and data repositories.

This study is also aimed at preparing libraries and librarians in Africa to play active role towards the accomplishment of the SDG through data services. When healthcare, education, agriculture, climate change, etc. (that are integral part of the SDG) supported on data services that libraries can play a role in, it is not that difficult to ascertain with confidence that Africa will be in a better position to fulfill the 17 goals and 169 targets of the SDG. In view of this it is very encouraging that there exist already a data portal for the SDG that is built for all 54 countries to track and monitor the SDG activities (http://sdg.opendataforafrica.org/). This will make the work of libraries in Africa easier. If academic and research libraries start coordinating the task of collecting, organizing, and managing data across all relevant sectors in their respective countries, the data portal will be one central outlet to deposit evidence and monitor where respective countries are in regards to the SDG goals and targets.

The SDG is organized under 5 P's –i.e. – People, Planet, Prosperity, Peace, and Partnership (United Nations. General Assembly, 2015). In the following section, attempt is made to map the 17 goals under these clusters and it is equally a good idea to organize data service activities by libraries around these major themes. As shown above, in addition to the sets of skills and knowledge that librarians need to acquire to support data services, there are extensive tools, software, repositories, guidelines, and principles that libraries in Africa can tap on to educate their professionals and start services around the SDG. The following list maps the major clusters (the 5 P's) around the SDG goals.

- People (poverty, hunger, agriculture, health, quality education, gender equality)
- Planet (climate change, sustainable consumption, sustainable cities and communities, natural resources, clean water, clean energy, life on land, life below water)
- Prosperity (decent work, economic, industry/innovation/infrastructure, social, technological progress)
- Peace (inclusive societies, free from violence, reduced inequalities)
- Partnership (global partnership, participation of all countries, all stakeholders, and all people)

In conclusion, libraries and librarians in Africa don't need to re-invent the wheel. A significant amount of work is already done in areas of competencies, skills, models, frameworks, and educational models around data and information literacy that libraries in Africa can easily adopt. Although there are educational activities largely supported by international organizations (such as Research4Life), and institutions in the west (such as the British Library for Development Studies and Cornell University) to train professionals in Africa around information practices, very little is done to offer data literacy in the African context. While there are plenty of opportunities that libraries in Africa can utilize to get started with data and information literacy educational programs with little startup costs, the challenges remain in terms of building the necessary data infrastructure. A careful introspection into this infrastructure issue also gives us hope when we especially consider the number of openly available repository and data workflow management tools that exist on the web today. The African Development Bank Group has already developed a data portal for all the 54 African countries and that, with proper coordination of efforts, be a starting place to launch data services in respective countries.

REFERENCES

Association of College and Research Libraries. Working Group on Intersections of Scholarly Communication and Information Literacy. (2013). *Intersections of Scholarly Communication and Information Literacy: Creating Strategic Collaborations for a Changing Academic Environment*. Chicago, IL: Association of College and Research Libraries. Available at: http://acrl.ala.org/intersections/

Association of College and Research Libraries. (2015). Framework for Information Literacy for Higher Education. Chicago, IL: Association of College and Research Libraries. Available at: http://www.ala.org/acrl/standards/informationliteracycompetency

Calzada Prado, J., & Marzal, M. Á. (2013). Incorporating Data Literacy into Information Literacy Programs: Core Competencies and Contents. *Libri: International Journal of Libraries & Information Services*, *63*(2), 123-134.

Carlson, J., & Johnston, L.R. (2015). *Data information literacy: Librarians, data, and the education of a new generation of researchers*. West Lafayette, IN: Purdue University Press.

Carlson, J., Johnston, L.R., Westra, B., & Nichols, M. (2013). Developing an approach for data management education: A report from the data information literacy project. *International Journal of Digital Curation, 8*(1). 204-217. doi:10.2218/ijdc.v8i1.254

Carlson, J., Nelson, M. S., Johnston, L. R., & Koshoffer, A. (2015). Developing data literacy programs. Bulletin of the Association for Information Science and Technology, 41(6). Retrieved from https://www.asist.org/files/bulletin/aug-15/carlson_etal.pdf

Chen, H., Chiang, R. H. L., & Storey, V. C. (2012). Business intelligence and analytics: From big data to big impact. *MIS Quarterly*, *36*(4), 1165-1188.

Christensen-Dalsgaard, B., et al. (2012). Ten recommendations for libraries to get started with research data management: Final report of the LIBER working group on E-Science/research Data management. Retrieved from http://libereurope.eu/wp-content/uploads/The%20research%20data%20group%202012%20v7%20final.pdf

Fagerberg, J., & Srholec, M. (2008). National innovation systems, capabilities and economic development. *Research Policy, 37*(9), 1417-1435.

Koltay, T. (2017). Data literacy for researchers and data librarians. *Journal of Librarianship and Information Science, 49*(1), 3-14.

Koltay, T. (2015). Data literacy: in search of a name and identity. *Journal of Documentation, 71*(2), 401-415.

MacMillan, D. (2014). Data sharing and discovery: What librarians need to know. *Journal of Academic Librarianship, 40*(5), 541-549.

Majid, S., Chang, Y., Hnin Nu Aye, M., May Win, K., & San Yu, W. (2015). Analyzing publishing trends in information literacy literature: A bibliometric study. *Malaysian Journal of Library & Information Science*, *20*(2), 51-66.

McKinsey & Company. (2011). Big data: The next frontier for innovation, competition, and productivity. McKinsey Global Institute, (June), 156.

Perry, G.J., Roderer, N.K., & Assar, S. (2005). A current perspective on medical informatics and health sciences librarianship. *Journal of the Medical Library Association 93*(2): 199–205.

Piorun, M., Kafel, D., Leger-Hornby, T., Najafi, S., Martin, E., Colombo, P., & LaPelle, N. (2012). Teaching Research Data Management: An undergraduate/ Graduate Curriculum. *Journal of eScience Librarianship,1*(1): e1003. http://dx.doi.org/10.7191/jeslib.2012.1003

Prado, J. C., & Marzal, M. Á. (2013). Incorporating data literacy into information literacy programs: Core competencies and contents. *Libri, 63*(2), 123-134.

Qin, J., & D'Ignazio, J. (2010). Lessons learned from a two year experience in science data literacy education. In: Proceedings of the 31st annual IATUL conference, 20–24 June 2010, 2. Available at: http://docs.lib.purdue.edu/iatul2010/conf/day2/5

Sapp Nelson, Megan R. 2017. "A Pilot Competency Matrix for Data Management Skills: A Step toward the Development of Systematic Data Information Literacy Programs." *Journal of eScience Librarianship* 6(1): e1096. https://doi.org/10.7191/jeslib.2017.1096

Schield, M. (2004). Information Literacy, Statistical Literacy and Data Literacy. *IASSIST Quarterly*, (6), 6-11.

Schneider, R. (2013). Research Data Literacy. *Communications in Computer and Information Science, 397*, 134-140.

Soehner, C., Steeves, C., & Ward, J. (2010). E-Science and data support services: A study of ARL member institutions. Washington, DC: Association of Research Libraries.

Tenopir, C., Sandusky, R. J., Allard, S., & Birch, B. (2014). Research data management services in academic research libraries and perceptions of librarians. *Library and Information Science Research*, *36*(2), 84-90.

United Nations. General Assembly (2015, October 21). Transforming our world: the 2030 Agenda for sustainable development. Retrieved from http://www.un.org/en/ga/search/view_doc.asp?symbol=A/RES/70/1&Lang=E

University of Edinburgh (2017). MANTRA: Research Data management Training. Retrieved from http://datalib.edina.ac.uk/mantra/

Zinn, S., Stilwell, C., & Hoskins, R. (2016). Information Literacy Education in the South African Classroom: Reflections from Teachers' Journals in the Western Cape Province. *Libri, 66*(1), 31-44.

Zurkowski, P.G. (1974). The Information Service Environment Relationships and Priorities. Washington, D.C.: National Commission on Libraries and Information Science, Related paper No. 5, 30p.