**SEMINAR PAPER**

on

“The role of Digital Curation in enhancing scholarly research visibility”

By

Kenneth Ohis IRENOA

PG/PHD/18/87905

**Department of Library and Information Science**

**UNIVERSITY OF NIGERIA, NSUKKA**

**October, 2019**

**Abstract**

*The marked upshift in digital content due to the impact of technologies are well documented. Adapting to the new information seeking patterns and expectations require that information professionals keep abreast of the changes and leverage on the technologies to keep the age-old role as memory of the society. This paper attempts to give a ‘bird’s eye’ view of the emerging area in archival and library and information science professions and how it can improve research visibility of the intellectual efforts out of Nigeria and Africa in general. It attempts to define and describe digital curation, an emerging field of theory and practice that embraces digital preservation, data curation, and management of information assets over their lifecycle. It dissects key issues and debates in the area while arguing that digital curation is a vital strategy for enhancing research visibility. The paper recommends proper packaging and publicity for research; funding agencies should impose data sharing and managing funded projects; and the need to design and implement data management and curation programs to help improve dataset fidelity and preservation for the future.*

**Keywords**: Data Curation, Digital Curation, Research Visibility

**Introduction**

Web and Internet technologies have revolutionized the nature, manner and way we conduct research today. While the internet has made information more accessible and with ease, much of the intellectual products out of Africa still remain in the fringes. So many tools have become available to help authors during research activities. Research activity in the 21st century to a very large extent now depend on digital sources of information (Tammaro, Ross, and Cararosa, n.d.). The increase in digital content is largely attributed to the impact of libraries and research institutions who host repositories and publish many of the intellectual works of academics and researchers within their institutions. The digital sources are available in a variety of formats ranging from publications, images, video, audio, databases, email, websites, and so on

As visibility of research has become more and more important to academics, higher institutions of learning, and the international community, the online audience has also become more significant for the researcher. Means of creating publicity with many researchers and scholars resorting not only by experimenting with platforms for the dissemination of their work but also with the format their audience receives and experiences. Achieving this requires a special effort at making sure that the digital content enables data discovery and retrieval, maintain high quality, with value added, and made available for re-use over time; this includes authentication, archiving, management, preservation, retrieval, and representation (Noonan and Chute, 2014), this is referred to as data curation

**Digital Curation**

Digital data and technologies have fast become an integral aspect of 21st century life. Accessing information in a variety of formats, taking digital photos, shopping online, accessing digital entertainment or online government services, and even electronic socialising and communications: all are commonplace activities for people today. It is understandable therefore, that public organisations - particularly libraries, archives and educational establishments – face increasing demands for digital services from users who routinely and unthinkingly use or depend upon digital information in other walks of life (Pennock, 2007). Digital curation is fast becoming recognised as the most viable way to meet this challenge and keep digital resources authentic and re-usable for future users. Digital curation is a recent (and we believe, useful) phrase to designate a set of opportunities for cross- institutional and cross-disciplinary engagement that have been evolving—with many valuable advances and contributions by archivists—for decades (Noonan and Chute, 2014).

Digital curation bridges research, practice, and training across nations, disciplines, institutions, repositories, and data formats (Gold, 2010; Ray, 2009). Digital curation involves maintaining, preserving and adding value to digital research data throughout its lifecycle. The active management of research data reduces threats to their long-term research value, thereby reducing the chances of digital obsolescence. Curated data in trusted digital repositories increased the reach of such work due to sharing among the wider spectrum of society. Digital curation helps in reducing duplication of effort in research data creation, and enhancing the long-term value of existing data by making it available for further high-quality research (Knight, 2017). Data Curation is a means of managing data that makes it more useful for users engaging in data discovery and analysis.

Digital curation is a specialization that encompasses both theory and applied knowledge. Digital curation and data preservation are ongoing processes, requiring considerable thought and the investment of adequate time and resources (Pennock, 2007).



The life cycle approach according to (Pennock, 2007) is necessary because:

* Digital materials are fragile and susceptible to change from technological advances throughout their life cycle, i.e. from creation onwards;
* Activities (or lack of) at each stage in the life cycle directly influence our ability to manage and preserve digital materials in subsequent stages;
* Reliable re-use of digital materials is only possible if materials are curated in such a way that their authenticity and integrity are retained.

The Digital Curation Centre (DCC, 2019) summarises that the digital curation lifecycle comprises of the following steps:

* **Conceptualise**: conceive and plan the creation of digital objects, including data capture methods and storage options.
* **Create**: produce digital objects and assign administrative, descriptive, structural and technical archival metadata.
* **Access and use**: ensure that designated users can easily access digital objects on a day-to-day basis. Some digital objects may be publicly available, whilst others may be password protected.
* **Appraise and select**: evaluate digital objects and select those requiring long-term curation and preservation. Adhere to documented guidance, policies and legal requirements.
* **Dispose**: rid systems of digital objects not selected for long-term curation and preservation. Documented guidance, policies and legal requirements may require the secure destruction of these objects.
* **Ingest**: transfer digital objects to an archive, trusted digital repository, data centre or similar, again adhering to documented guidance, policies and legal requirements.
* **Preservation action**: undertake actions to ensure the long-term preservation and retention of the authoritative nature of digital objects.
* **Reappraise**: return digital objects that fail validation procedures for further appraisal and reselection.
* **Store**: keep the data in a secure manner as outlined by relevant standards.
* **Access and reuse**: ensure that data are accessible to designated users for first time use and reuse. Some material may be publicly available, whilst other data may be password protected.
* **Transform**: create new digital objects from the original, for example, by migration into a different form.

**Digital curation tools**

Digital curation is not done in isolation, it requires tools and softwares which help in implementing and adding value to research data. These tools are meant for presenting collections in a graphically appealing layout that can easily be shared with others with possibility to include notes which explain context, offer opinions and solicit questions. These collections can be great resources for discovering and keeping up with information. And they’re great tools for students gathering resources for research projects (Cool Tools for School, 2019).

The Digital Curation Centre (DCC) (Digital Curation Centre, 2019) lists some of the tools it uses:

* **DMPonline** (https://dmponline.dcc.ac.uk/) - helps you to develop data management plans that meet research council and funding body mandates.
* **Collaborative Assessment of Research Data Infrastructure and Objectives (CARDIO**) - helps you to assess your data management support and infrastructure and to collaboratively plan for improvement.
* **Data Asset Framework (DAF)** - helps you to identify researchers' current data management activity, their data holdings and their data management requirements.
* **Digital Repository Audit Method Based on Risk Assessment (DRAMBORA)** - helps you to define and address the risks threatening your digital repository content and infrastructure.
* **Curation Costs Exchange** - helps you better understand what you are spending on digital curation and how your spending compares with your peers. This tool was developed through our work on the Collaboration to Clarify the Costs of Curation (4C) project.

**Research Visibility**

Research visibility is a function of the number of citations an article or publication accumulates over a period. Citation is a measure that shows the number of times an article or publication has been used by other articles/publications (Ebrahim, Salehi, Embi, Tanha, and Gholizadeh, 2014). Citations are applied to measure the importance of information contained in an article (Fooladi et al., 2013). “The more often a paper becomes cited, the greater its influence on the field” is a basic assumption of citation analysis (Martínez, Herrera, Lopez-Gijon, and Herrera-Viedma, 2013; Garfield and Merton, 1979).

This has made many universities and research institutions to encourage their researchers to publish high quality papers which can receive high citations and will reach the widest possible audience (Ale-Ebrahim et al., 2013). Which in the other sense will be affected negatively if the paper does not have online presence, because citations will be limited to the availability of the published article on the web (Lawrence, 2001). Literature has shown a marked increase in visibility by making research outputs available through open access repositories, with a resulting wider access and higher citation impact (Antelman, 2004; Hardy, Oppenheim, Brody, and Hitchcock, 2005; Amancio. Oliveira, and Da Fontura, 2012; Ertürk and Şengül, 2012; Dalton, 2013). A paper has greater chance of becoming highly cited whenever has more visibility (Egghe, Guns, and Rousseau, 2013)

Visibility can ultimately be translated to increased opportunity for attracting citations (Mirjana, Milos, Vladimir, and Jovana, 2013). Rewarding the time and effort put in by researchers during writing up their research for publication. However, publishing a paper in the journal which has a high impact factor is not guaranteed by analyzing citation rate. The World Wide Web has become an outstanding tool for the collection and dissemination of scholarly Information (Más-Bleda and Aguillo, 2013). The majority of freely available articles will be found on author personal website rather than in a repository or in an open-access journal (Antelman, 2004). Alternatively, institutional repositories make articles visible and increase the chances for use by other scholars and exchange ideas among similar disciplines (Ngah, 2010). Eysenbach (2006) posed the challenge of visibility directly resulting to more citations, increased chances of the article getting cited sooner.

Lack of visibility caused that some senior scholars in some African universities may not have a significant citation impact (Rotich and Musakali, 2013). Writing an article for online distribution needs to cover some techniques from writing a search engine-friendly title and abstract to maximize visibility once it is published (Norman, 2012). Depositing the paper in the institutional repository is another way of increasing paper’s visibility. Commonly cited benefits of using an institutional repository are to increase the visibility and citation impact of the institution’s scholarship (Tate, 2010). By making a few adjustments in how and where to publish and present the research findings, researchers can become more productive, better known for their research, and better connected with the members of your professional community (Pfirman, Balsam, Bell, Laird, and Culligan 2007).

**The role of Digital Curation in enhancing scholarly research visibility**

Scholarly research has seen a new paradigm characterized by the massive scale of data creation and accumulation, as well as scientific discovery based on intensive data (Hey, Tansley, and Tolle, 2009; Jahnke, Asher, and Keralis, 2012). With the mind-boggling amount of content on the Internet, content that varies enormously in quality, there’s huge value in relying on experts to select the best content for a topic. And with that need, many new tools have been developed to make it easy for anyone to select, collect and share their own collections of digital resources (Tammaro et al., n.d.).

Academic libraries and librarians have been identified as curatorial liaisons on campus in the data curation movement due to their long-standing history, credentials and commitments (Fox, 2013; Heidorn, 2011; Lyon, 2012; Schubert, Shorish, Frankel, and Giles, 2013). As a result, several metadata standards for data management and curation have been developed to manage massive large- scale data sets (Ogier, Hall, Bailey, and Stovall, 2014; Weber, Palmer, and Chao, 2012).

Tyler Walters and Katherine Skinner cited in: (Noonan and Chute, 2014) observed in New Roles for New Times: Digital Curation for Preservation that “Historically, librarians and archivists have been looked to as the custodians of physical cultural heritage. This focus on content curation—including selection, management, and preservation—is a defining component of research libraries’ ongoing work. A new role within this arena is emerging—that of digital curation (Walters and Skinner, 2011, Cited In: (Noonan and Chute, 2014). Over the past decades, researchers and others in higher education have become more aware of the need for curation of research data to enhance the general accessibility and visibility of research.

1. Digital curation outlines carefully planned processing actions for all data deposited in order to best arrange, transform, and prepare the data according to established procedures.
2. Curation allows for the application of the appropriate descriptive metadata to digital content and enhances author submitted metadata to best facilitate easy discovery on the Internet, databases, repositories, and so on, making visibility more achievable and database and search engine optimized (SEO).
3. Facilitate access through discovery, dissemination, retrieval, and download functionality. Discovery is a key way of enhancing access to a digital content. Digital curation helps improve the digital content to ensure the content can stand the test of time and future access.

**Challenges**

Digital curation is a growing area of the information professions. It is not without its challenges, some of which are highlighted below.

1. Scalability and complexity of the systems are always big challenges we face as technology evolves. Technology has shown no signs of slowing down in terms of discovery, innovation and updates/upgrades. This means continuous altering of how they operate and requirements to keep them running. Meeting up with this constant change is a challenge for digital curators, librarians, archivists, and data handlers.
2. Stewardship challenges in the digital environment, is another important puzzle that needs to be addressed. The need to clearly define who handles what, how and when in the continuously changing digital environment.
3. The effects of pervasive and intensive personalization in various kinds of information delivery and social media systems. Privacy issues, data breaches, and a whole litany of issues created by social media and emerging technologies are also another daunting area of concern.
4. Problem of extremely aggressive technological obsolescence and what that means to creators and their willingness to engage with the affordances of technology. Planning Ahead for Software, Hardware, and Operating Systems Obsolescence Software developers update and improve their applications. Hardware slows down as it ages and manufacturers build faster machines. Operating systems also evolve and adapt to changes in computing environments. (Smithsonian Institution Archives, 2019).

**Conclusion**

The challenges of visibility for research and scholarship is real, and it is high time that universities, government agencies, and research policy makers need to do more for the future. Concerted efforts must be made to increase the impact of our intellectual efforts to the world. Data management policies and Digital curation programmes need to be given attention in order to create a presence in the academic space and contribute to the body of knowledge with our research.

**Recommendations**

To increase research visibility of the intellectual products from the country, the paper recommends intentional efforts be put in place to properly package and publicize innovative and ground-breaking research. Also, that funding agencies such as tertiary education fund (Tetfund), NITDA and others should impose requirements for data sharing and management plan for funded projects. The paper also recommends that as a response to the challenges for research and scholarship, more institutions and academic libraries in Nigeria should start designing and implementing data management and curation programs to help improve dataset fidelity and the preservation of research products for future use.

**Reference**

Antelman, K. (2004). Do open-access articles have a greater research impact? College & research libraries, 65, 372-382.

DCC. (2019). DCC Curation Lifecycle Model. Retrieved 2 September 2019, from http://www.dcc.ac.uk/resources/curation-lifecycle-model

Digital Curation Centre. (2019). Tools and applications. Retrieved 2 September 2019, from http://www.dcc.ac.uk/resources/tools-and-applications

Ebrahim, N. A., Salehi, H., Embi, M. A., Tanha, F. H., & Gholizadeh, H. (2014). Visibility and Citation Impact. *International Education Studies*, *7*(4), 120–125. https://doi.org/10.5539/ies.v7n4p120

Egghe, L., Guns, R., & Rousseau, R. (2013). Measuring co-authors’ contribution to an article’s visibility. *Scientometrics*, 95, 55-67. http://dx.doi.org/10.1007/s11192-012-0832-4

Ertürk, K., & Şengül, G. (2012). Self-Archiving in Atılım University. In S. Kurbanoğlu, U. Al, P. Erdoğan, Y. Tonta, & N. Uçak (Eds.), E-Science and Information Management. Berlin, Heidelberg: Springer.

Fooladi, M., Salehi, H., Yunus, M. M., Farhadi, M., Aghaei-Chadegani, A., Farhadi, H., & Ale-Ebrahim, N. (2013). Do Criticisms Overcome the Praises of Journal Impact Factor? *Asian Social Science*, 9, 176-182. http://dx.doi.org/10.5539/ass.v9n5p176

Fox, R. (2013). The art and science of data curation. *OCLC Systems & Services*, 29(4), 195-199.

Garfield, E., & Merton, R. K. (1979). Citation indexing: Its theory and application in science, technology, and humanities. Wiley New York.

Hardy, R., Oppenheim, C., Brody, T., & Hitchcock, S. (2005). Open Access Citation Information.

Heidorn, P. (2011). The Emerging Role of Libraries in Data Curation and E-science. *Journal of Library Administration*, 51(7/8), 662-672. doi:10.1080/01930826.2011.601269

Hey, T., Tansley, S., & Tolle, K. (2009). The fourth paradigm: Data-intensive scientific discovery.

Knight, M. (2017, November 30). Data Curation 101: The What, Why, and How. Retrieved from Dataversity: https://www.dataversity.net/data-curation-101/#

Lawrence, S. (2001). Free online availability substantially increases a paper’s impact. *Nature*, 411, 521. http://dx.doi.org/10.1038/35079151

Lee, C.A. & Tibbo, H. (2011). Where’s the archivist in digital curation? Exploring the possibilities through a matrix of knowledge and skills. *Archivaria*, 72, 123-168.

Lyon, L. (2012). The informatics transform: Re-engineering libraries for the data decade. *International Journal of Digital Curation*, 7(1), 126-138.

Martínez, M. A., Herrera, M., López-Gijón, J., & Herrera-Viedma, E. (2013). H-Classics: Characterizing the concept of citation classics through H-index. *Scientometrics*, 1-13.

Más-Bleda, A., & Aguillo, I. (2013). Can a personal website be useful as an information source to assess individual scientists? The case of European highly cited researchers. *Scientometrics*, 1-17.

Ngah, Z. A. (2010). Open Access repositories and journals for visibility: Implications for Malaysian libraries. *Malaysian Journal of Library & Information Science*, 15.

Noonan, D., & Chute, T. (2014). Data Curation and the University Archives. The American Archivist, 77(1), 201–240.

Norman, E. R. (2012). Maximizing Journal Article Citation Online: Readers, Robots, and Research Visibility. *Politics & Policy*, 40, 1-12. http://dx.doi.org/10.1111/j.1747-1346.2011.00342.x

Ogier, A., Hall, M., Bailey, A., & Stovall, C. (2014). Data management inside the library: Assessing electronic resources data using the data asset framework methodology. *Journal of Electronic Resources Librarianship*, 26 (2), 101-113. doi:10.1080/1941126X.2014.910406

Pennock, M. (2007). Digital Curation: A Life-Cycle Approach to Managing and Preserving Usable Digital Information. *Library and Archives*, *1*(January), 1–3. Retrieved from http://www.ukoln.ac.uk/ukoln/staff/m.pennock/publications/docs/lib-arch\_curation.pdf

Pfirman, S., Balsam, P., Bell, R., Laird, J., & Culligan, P. (2007). Maximizing Productivity and Recognition, Part 1: Publication, Citation, and Impact.

Popoola, S. O. (2008). The Use of Information Sources and Services and Its Effect on the Research Output of Social Scientists in Nigerian Universities. *Library Philosophy and Practice (e-journal)*.Paper 183. <http://digitalcommons.unl.edu/libphilprac/183>.

Rotich, D. C., & Musakali, J. J. (2013). Publish or Perish: Remaining Academically Relevant and Visible In the Global Academic Scene through Scholarly Publishing. In Ocholla, N. E. D. N. (Ed.), Conference and Programme Chairs, 2013 South Africa. Department of Information Studies-University of Zululand, 58.

Schubert, C., Shorish, Y., Frankel, P., & Giles K. (2013). The evolution of research data: Strategies for curation and data management. *Library Hi Tech News*, 30(6), 1-6.

Tammaro, A. M., Ross, S., & Cararosa, V. (n.d.). Research Data Curator: the competencies gap. Parma.

Tate, J. (2010). An evaluation of SOAS Research Online, the Institutional Repository of the School of Oriental and African Studies-SOAS Research Online. CITY UNIVERSITY.

Weber, N. M., Palmer, C. L., & Chao, T. C. (2012). Current trends and future directions in data curation research and education. *Journal of Web Librarianship*, 6(4), 305-320.