

# **Beyond Scholarly Communications: The Role of Open Access in Facilitating Digital Preservation**

Laughton, P.A.<sup>1</sup> , Alemneh, D.G.<sup>2</sup>

[Paul.Laughton@up.ac.za](mailto:Paul.Laughton@up.ac.za) & [Daniel.Alemneh@unt.edu](mailto:Daniel.Alemneh@unt.edu)

University of Pretoria, South Africa & University of North Texas, USA

## **1 Introduction**

Ensuring long-term access to digital information resources is one of the key challenges facing cultural heritage institutions today. According to the UNESCO's Charter for the Preservation of Digital Heritage, resources of human knowledge or expression, whether cultural, educational, scientific and administrative, or embracing technical, legal, medical and other kinds of information, are increasingly created digitally, or converted into digital form from existing analogue resources (UNESCO, 2021). This trend has created a need for sustainable digital preservation and is the driving force behind much development and research. Much of the efforts to sustain preservation are rooted in standardisation and proliferation of standards. This conference paper will look broadly into the contributions that Open Access (OA) has made towards facilitating the preservation of data.

## **2 Statement of the research problem**

The open access movement has seen much progress in the last two decades and is considered to be a global trend that is gaining momentum (Zhang, Wei, Huang & Sivertsen, 2022). The shift in the way citizens access information through the increasing volume of open access content comes with the need for changes in the way publishers distribute, collate, fund and preserve these publications. The recent pandemic further highlighted the potential roles of open access in terms of addressing some of the inequities in the established scholarly communication system through which scholarly contents are created, disseminated, and preserved for future use. Many commentators (Alemneh, 2022, ...among others) argue that despite the progress, the case on the ground may not be where it needs to be and more studies are required to gauge the extent to which what more is required to enhance the levelling the scholarly communications, so as to promote scientific progress through inclusivity. In particular, articulating the role of open access in facilitating digital preservation can only be determined through a comprehensive assessment in the context of different communities. This paper seeks to review the current status and identify the impact that open access has had on digital preservation in the context of cultural heritage institutions' resources.

### **3 Methodology**

This conference paper is a conceptual paper and the above identified research problem was addressed through a literature review. A number of resources were consulted to understand how the related topics were viewed in academic literature. Upon the completion of the reviewed literature the findings were summarised and discussed under the relevant themes.

### **4 Best practices in digital preservation**

The open access movement has continually evolved, starting as a scholarly communication issue and expanding to include issues integral to the social justice movement, opportunity, inclusion, equity and rights. In fact, the world learned the critical role of open access to research during COVID-19 pandemic. Thanks to the swift actions by governments around the world, in 2020, paywalls were lifted for coronavirus articles, enabling scientists and doctors to access the latest research without delay, making a significant contribution to the development of the fastest vaccine for COVID-19, in our history. Why would not such openness continue or remain as a default to solving such challenges to humanity at large?

In addition to the usual discussions on how open access publishing is transforming scholarly communications in general, this paper pushes the boundaries further to examine various aspects to advance the role of OA in digital lifecycle management on a global level. In particular, the efforts of the United States and the European Union initiatives highlighted as models for others to learn and make concerted efforts to align with UNESCO's recent recommendation on Open Science, bringing their respective nations to equal footing with governments across the world to establish strong open access policies to promote their national innovation agendas. Such global legislation and policy will go a long way in helping craft and promote the role of open access in ensuring long-term access and to broaden the potential of the ecosystem by levelling the playing field for all players (UNESCO, 2021).

#### **4.1 *European initiatives:***

The European Commission is committed to supporting beneficiaries in supporting open access publishing as the main mode of publishing research in the context of open science. As early as 2012, the European Commission published a recommendation on access to and preservation of scientific information encouraging all EU countries to put publicly-funded research results in the public domain in order to strengthen science and the knowledge-based economy (European Commission, 2012).

Building on the pro-open access environment, in March 2021, the European Commission launched Open Research Europe, a publishing platform for scientific papers accessible to everyone. This platform presents the results of research funded

by Horizon Europe, the EU research and innovation programme for 2021-2027, and its predecessor, Horizon 2020. This platform makes it easy for beneficiaries to comply with the open access terms of their funding and offers researchers a publishing venue to share their results and insights rapidly and facilitate open, constructive research discussion (Kosmopoulos, 2022).

#### **4.2 North American initiatives:**

The White House Office of Science & Technology Policy (OSTP) issued an updated memorandum in August 2022, "*Ensuring Free, Immediate, and Equitable Access to Federally Funded Research*." The memo calls on federal agencies to make "*articles resulting from all U.S. federally funded research freely available and publicly accessible by default in agency-designated repositories without any embargo or delay after publication.*" The policy updates eliminate the previous 12-month embargo period and make articles (and the underlying data needed to validate results) openly available in machine readable formats (White House, 2022).

More importantly, such policy guidance makes taxpayer-funded research available immediately, at no cost to the public. This in turn increases transparency, improves scientific research integrity, and promotes equity in the publishing of and access to publicly funded research. This policy guidance is a major milestone in the collective effort that the cultural heritage community have been working toward since the beginning of the open access movement in the past two decades. One such community is SPARC (the Scholarly Publishing and Academic Resources Coalition), an international alliance of academic and research libraries which promotes open access to scholarship (SPARC 2022).

#### **4.3 Other partnership initiatives:**

Many regions of the world recognize the fact that partnerships and networking have become essential arrangements for cultural heritage and/or related institutions to leverage their resources for greater impact in general. There are several examples of such partnerships among institutions, regions, and nations working together to achieve greater outcomes, and these include the African Research Universities Alliance (ARUA). The ARUA is intended to develop local research excellence through collaboration to help find solutions to the problems of Africa. It is set to become a pan-African network for bringing research and academic excellence to the fore throughout the continent by developing strong and viable research universities.

While these regional developments have led to capacity building and tangible outcomes to increase openness and access to scholarly work, there have been long-term collaborative efforts to ensure the reality of OA, developing standards and frameworks that make open access a reality. After all, scholarly communication is essentially the system through which scholarly contents are created, disseminated,

and preserved for future use. In the following sections we will discuss the standards and initiatives on which open access is based.

## **5 The synergy of open access, open source, open standards, and open archives**

Along with the Open Science and OA movements came some significant developments that spearhead change, better integration of content and the opening of access to content. Not all of these developments are considered as part of the OA movement, but are related and have a role to play in making the OA movement a reality today. In this section Open Standards, Open Source and the Open Archives Initiative will be examined to identify their contribution and role in the OA movement and development of standards to support data preservation.

Open access, open source software, and open standards each individually and collectively offer a number of benefits to libraries, archives and museums. When they are combined they can offer greater formalisation to ensure access. Open source and open standards ensure patrons with easier access to open access materials and other resources. Due to the normal delay in the adoption of open standards, there are thousands of open access titles available and without open standards, making it very difficult to find what one is looking for. Open standards are key to making it possible to create interoperable systems to access the content in open access publications (Corrado, 2005).

Open access, open source and open standards have been playing a greater role recently when we consider the concept of borderless libraries (greater access and integration of digital libraries). Along with these developments came the Open Archives Initiative, which led to further developments in the open standards and open source communities. Together these developments contribute to a number of benefits including lower cost, greater access and more opportunity for long term preservation of scholarly works (Baruah, 2010).

### **5.1 Open Standards**

An open standard is a standard that is freely available for adoption, implementation and updates. Some popular examples of open standards are XML, SQL and HTML. "Open Standards" are standards made available to the public and are developed/updated (or approved) and maintained through a collaborative and consensus driven process. "Open Standards" are necessary to facilitate interoperability and data exchange among different products or services and are intended for ease of use and widespread adoption.

Other elements of Open Standards include, but are not limited to (ITU, 2005):

- **Collaborative process:** voluntary development (or approval) following a transparent consensus driven process, reasonably open to all interested parties.
- **Reasonably balanced:** ensuring the lack of bias and that the process is not dominated by a single interest group.
- **Due process:** protocol is used to ensure all comments are considered and responded to by interested parties.
- **Intellectual property rights (IPRs):** IPRs are essential to implement the standard to be licensed to all applicants on a global, non-discriminatory basis, either for free and under other reasonable terms and conditions.
- **Negotiations:** negotiations where necessary are handled by parties concerned and are performed outside the Standard Developing Organisations (SDO).
- **Quality and level of detail:** sufficient to allow for the development of a variety of competing implementations of interoperable products or services.
- **Publicly available:** easily available for implementation and use, at a reasonable price. Publication of the text of a standard by others is permitted only with the prior approval of the SDO.
- **On-going support:** maintained and supported over a long period of time.

Organisations share open standards because it allows them to bring significant value to both themselves and their customers. Standards are often jointly managed by a foundation of stakeholders. Often there are typically rules about what kind of adjustments or updates users can make, to ensure that the standard(s) maintains interoperability and quality.

Along with the abovementioned elements of open standards, Walli, Endsley, Bryant & Ruff (*n.d.*), outline six criteria an open standard must satisfy:

1. **Availability:** open standards are available for all to read and implement.
2. **Maximise end-user choice:** open standards create a fair, competitive market for implementations of the standard, straying away from locking in their customers.
3. **No royalty:** open standards are free for all to implement, with no royalty or fee.
4. **No discrimination:** open standards and the organizations that administer them do not give unequal favor to one implementor over another for any reason other than the technical standards compliance. Certification organisations must provide a path for low and zero-cost implementations to be validated.

5. **Extension or subset:** implementations of open standards may be extended, or offered in subset form.
6. **Predatory practices:** open standards may employ licence terms that protect against subversion of the standard by embrace-and-extend tactics.

Open standards create a community and conversation or the exchanging of ideas that is necessary for broadstream adoption. The open source movement is another important and related movement that works towards greater openness.

## **5.2 Open-Source Software**

Open-source software commonly referred to as open source, is software that includes source code (allowing for changes and customisation) and is usually available at no charge. There are a number of requirements other than the availability of source code that a software program must meet before it is considered open source, these include: the software must be free to redistribute; customisation must be allowed and the licence is not allowed discriminate against any persons or discriminate against any fields of endeavour. Software that is licensed under an open source licence typically allows for a community of developers from around the world to optimise the software by providing enhancements and fixes to any bugs (Corrado, 2005). Open source software may be beneficial to libraries, museums and archives in lowering initial and ongoing costs, eliminating vendor lock-in and allowing for increased customisation and flexibility (Baruah, 2010).

When one looks at two concepts of open source and open standards the two phrases use the word "open" so differently: For open source, open means that the source code must be made available with every copy of an executable application and every is allowed to modify and distribute the source code freely. In open standards, open means that the standards process is open to participation from the public and interested parties and that the completed standards are made available to all. When working with open standards, many standards organisations distribute copies of their standards for free and the right to implement a standard is typically also free (Siegel & Soley, 2008).

On closer examination, these two approaches are seen to be complementary:

- Open standards need implementations to provide: i) confirmation of their suitability; ii) a market presence; and iii) feedback from implementers and users
- Open source development projects need guidance and direction regarding their interfaces for interoperability and portability

Open Standards and open source software have paved the way for a number of initiatives to increase access and interoperability which are key to open access. One of these initiatives is the Open Archives Initiative.

### **5.3 Open Archives Initiative**

The result of the Santa Fe Convention, was an agreement among the parties to subscribe to a common standard for interoperability based on transfer of metadata from repositories using a minimal protocol and utilising accessible technology to achieve this (van de Sompel & Lagoze, 2000). Initial Technical Efforts of the Santa Fe Convention paved the way for future efforts by defining the foundational principles of what soon was named the Open Archives Initiative (OAI) (Suleman & Fox, 2001).

The Open Archives Initiative (OAI) represents an attempt to develop a framework for a “universal e-print archive” (digital content) that establishes interoperability standards to support the searching and retrieval of digital papers from all disciplines. At the most basic level interoperability is defined as “the capacity of a user to treat multiple digital library collections as one,” and is considered a key to unlocking the digital library challenge (Luce, 2001).

The increasing development of digital content provides an opportunity to reconsider many aspects of the current research communication process and the roles each of us play. Creating an ideal opportunity to experiment and rethink the assumptions that underlie our systems. Ginsparg (1994) believes “*we should take advantage of this opportunity to map out the ideal research communication medium of the future. It is crucial that the researchers, who play a privileged role in this as both providers and consumers of the information, not only be heard but be given the strongest voice. In particular, we need to dislodge definitively the curiously prevalent notion that the future electronic medium will strictly duplicate, inadequacy for inadequacy, the current print medium.*” This gives libraries an ideal scenario to rethink their vision and roles. Once this has been done, it is up to leadership to provide an environment that encourages creativity and safe space to think outside of the box in this transition phase (Luce, 2001).

With all the developments in open standards, open source and the Open Archive Initiative, significant work has gone into the refinement and improvement of elements such as metadata. Below is a discussion of the contribution that metadata plays towards the preservation of data.

## **6 Role of metadata in preservation**

Metadata is used to capture the characteristics and features of an object (both digital and analogue) as well as a collection. The concept of metadata came about with the need to organise and manage information. Metadata can be used to arrange, describe and enhance access to objects. The creation of well defined, accurate metadata is key to the access and preservation of digital library material, including cultural heritage collections and scholarly publications in digital repositories (Xie & Matusiak, 2016).

To better understand the role of metadata in preservation, it is important to acknowledge the role metadata plays in the maintenance and migration of digital content to new formats to accommodate evolving technologies and varied user needs. To better ensure the survival of digital objects, preservation metadata needs to exist independently from the systems which were used to create them. Without preservation metadata, digital objects will be lost. Unlike the physical analogue items like books, if the metadata gets damaged (like the spine on a book that comes loose and can be repaired easily), it can lead to digital objects becoming corrupted or obsolete, often impossible to repair. Along with the use of metadata for discovering and use of objects, metadata is essential integration, organisation and management of digital objects.

The development and adoption of open standards was a substantial step towards progress in digital preservation. Conceptual frameworks and metadata standards which incorporate these open standards provide a theoretical baseline for developing sustainable preservation systems and services. Two standards from the early millennia that have been recognized as particularly influential are the OAIS reference model (2001) and PREMIS metadata standard (2005) (Xie & Matusiak, 2016).

### **6.1 Complexities of metadata creation and management**

The creation of preservation metadata by multiple users or institutions can complicate issues of metadata ownership, access and responsibility. Fieldhouse (2012) notes that “the creation of metadata is shifting from collections managers to suppliers and publishers”. McDonough (2010) identifies the “benefits of multiple partners collaborating to improve metadata records around an object with preservation metadata as a key in cross-institutional communication”. While Corral (2012) notes that “the creation and management of preservation metadata represents the intersection between libraries, IT management and archival practice”.

### **6.2 Linked open data**

New standards and development of systems across the Library Archives Museums and publishing communities are focused on defining techniques, standards, systems, and services that keep up with the evolution changing information sources and needs of LAM communities. The LAM community’s information requirements are grounded both in a need for physical (analogue) and digital information objects and also in the



establishment of connections among these different objects. While within the LAM profession some are focused on exploring how information systems and structures support these needs and enable cross-community and cross-repository access and aggregation. Along with these identified changing needs the development of metadata systems and standards are required to achieve the desired functionality of linked open data. This has caused experts to rethink and revisit common questions of metadata exchange, migration, interoperability, scale, sustainability, and value (Mitchell, 2014).

## **7 Role of open access in facilitating the preservation activities in the context of cultural heritage institutions**

The drive to make content open and accessible to all is a real answer to humanitarian issues of inequality in society. Open access and the spirit of open access is breaking down traditional barriers and challenging the status quo. The developments around open access have been gradual, and even the broader acceptance of the movement has been gaining momentum. With so many challenges to overcome, new solutions are required to overcome the challenges. Developments around open standards, open source, open archives and open data are creating solutions to the myriad of challenges, and as more development and effort goes into these initiatives the more resolutions and answers we have for these challenges.

The way forward is to reduce the current limitations through refinement and development of open initiatives. The overall goal and objective should be to make it easier to publish open content and ensure that content remains open. The diverse and ever expanding digital information resources in the Web, has evolved without much regard to resources management and organisation issues. As noted by Segaletsho (2021) the proliferation of digital resources in cultural heritage institutions has also increased the dire need for digital preservation for continued access.

Digital curation activities related to managing digital contents throughout their lifecycle. It is essentially ensuring that digital contents are properly appraised, selected, and securely stored, while appropriately maintaining logical and physical integrity and authenticity. According to Jain, Mnjama & Oladokun (2021) digital curation is generally the process of strategically establishing and developing electronic repositories and adding value for long term access. The multifaceted challenges of digital preservation often require a mixture of approaches. Open access contents support the digital curation process. By removing any barriers, open access facilitates implementation of appropriate digital preservation strategies.

## 8 Conclusion

Advancing knowledge requires not only enhancing our capacity to generate more knowledge, but also cultivating our ability to comprehend, communicate, and ensure long-term access to the vast quantities of digital contents that we continue to generate. The recent COVID-19 pandemic, the spread of fake news on social networks, among other issues, have shown us the need for mechanisms to provide easy, immediate and continued access to accurate and scientific information that is controlled by the right stakeholder/community. These communities are well established and are growing in the OA community. This paper discussed the need to ensure long-term access to digital resources by examining the potential role of open access scholarly publishing in general, particularly for the cultural heritage contents.

Digital Preservation is essentially a comprehensive set of managed activities that are necessary to provide continued access to digital objects. Continuous efforts through communities such as open standards, open source, open archives and open data have made considerable inroads to making OA a sustainable initiative. Considering the fact that it is a shared community responsibility, open access plays a huge role in facilitating digital preservation activities.

## References

Alemneh, D. G. (Ed.). (2022). Handbook of Research on the Global View of Open Access and Scholarly Communications. IGI Global. <https://doi.org/10.4018/978-1-7998-9805-4>

Baruah, B.G. (2010). Role of open access, open standards and open sources in libraries : a study.

Corrado, E. M. (2005). The Importance of Open Access, Open Source, and Open Standards for Libraries. *Issues in Science and Technology Librarianship*.

Corrall, S. (2012). Fieldhouse, M.; Marshall, A. (eds.). Collection Development in the Digital Age. Facet Publishing. pp. 3–25. ISBN 978-1856047463.

Eby, J. (2021). Open Access and the 17 UN Sustainable Goals. eContent Pro International: <https://www.econtentpro.com/blog/open-access-17-un-sustainable-goals/199>

European Commission (2012) Commission Recommendation on Access to and Preservation of Scientific Information. *Official Journal of the European Union*, 2012/417/EU. Available online: <https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32012H0417&rid=1>

Fieldhouse, M. (2012). *The Process of Collection Management*. In Fieldhouse, M.; Marshall, A. (eds.). Collection Development in the Digital Age. Facet Publishing. pp. 27–43. ISBN 978-1856047463.

Ginsparg, P. (1994). First steps towards electronic research communication. *Computers in Physics*, 8(4).

International Telecommunications (ITU). (2005). *Definition of "Open Standards"*. Available online: <https://www.itu.int/en/ITU-T/ipr/Pages/open.aspx> (Accessed 19 September 2022).

Jain, P., Mnjama, N., & Oladokun, O. (Eds.). (2021). *Open Access Implications for Sustainable Social, Political, and Economic Development*. IGI Global. <https://doi.org/10.4018/978-1-7998-5018-2>

Kosmopoulos, C. (2022). From Open Access Publishing to Open Science: An Overview of the Last Developments in Europe and in France. In D. Alemneh (Eds.), *Handbook of Research on the Global View of Open Access and Scholarly Communications* (pp. 1-22). IGI Global. <https://doi.org/10.4018/978-1-7998-9805-4.ch001>

Luce, R. E. (2001). The Open Archives Initiative: Interoperable, Interdisciplinary Author Self-Archiving Comes of Age. *Serials Librarian*, Apr2001, Vol. 40 Issue 1/2, p173-182, 10p; DOI: 10.1300/J123v40n01\_15

Mitchell, E.T. (2014). *Library Linked Data: Research and Adoption*. American Library Association,

McDonough, J.P. (2010). Packaging Video Games for Long-Term Preservation. *Journal of the American Society for Information Science and Technology*. 62 (1): 171–184. doi:10.1002/asi.21412

Open Research Europe: <https://open-research-europe.ec.europa.eu>

Segaetsho, T. (2021). The Concoct of Digital Preservation in Open Access: A Case of the University of Botswana Research, Innovation, and Scholarship Archive. In P. Jain, N. Mnjama, & O. Oladokun (Eds.), *Open Access Implications for Sustainable Social, Political, and Economic Development* (pp. 118-138). IGI Global. <https://doi.org/10.4018/978-1-7998-5018-2.ch007>

Siegel, J., & Soley, R. M. 2008. Open Source and Open Standards: Working Together for Effective Software Development and Distribution. *Open Source Business Resource*, (November 2008). <http://timreview.ca/article/207>

SPARC. <https://sparcopen.org/>

Suleman, H., & Fox, E. (2001) The Open Archives Initiative, *Journal of Library Administration*, 35:1-2, 125-145, DOI: 10.1300/J111v35n01\_08

UNT Libraries (2017). UNT Libraries' Digital Preservation Policy Framework. <https://library.unt.edu/policies/digital-preservation-framework/>

Van de Sompel, H., & Lagoze, C. (2000) The Santa Fe Convention of the Open Archives Initiative. *D-Lib Magazine*, 6, No. 2.

Walli, S., Endsley, R., Bryant, D., & Ruff, N. (n.d.). *What are open standards?* Available online: <https://opensource.com/resources/what-are-open-standards> (Accessed 19 September 2022).

White House (2022). OSTP Issues Guidance to Make Federally Funded Research Freely Available Without Delay. Press Release, August 25, 2022. Available online: <https://www.whitehouse.gov/ostp/news-updates/2022/08/25/ostp-issues-guidance-to-make-federally-funded-research-freely-available-without-delay/>

Xie, I., & Matusiak, K. (2016). *Discover Digital Libraries: Theory and Practice*. Elsevier, St Louis.