

Time is a thief of memory

Alicia Wise^{a,b,*}

^a*Executive Director, CLOCKSS, USA*

^b*CLOCKSS, UK*

Abstract. This paper is based upon a presentation during the session ‘The Version of Record under Attack! The Dark Side of the Scholarly Publishing Universe’ at the 17th APE Conference. It discusses the complex challenges of long-term digital preservation and how CLOCKSS is collaborating with publishers, libraries, and other stakeholders to preserve digital scholarship over the coming centuries.

Keywords: Digital preservation, long-term preservation, CLOCKSS, journal preservation, Open Access, book preservation

1. Introduction

There are many threats to the scholarly record that are in focus at present, including concerns about fraudulent papers and paper mills. There is another. This is the threat of time, if no long-term digital preservation arrangements are in place for published content. This means much more than a remote back-up copy. Long-term preservation requires active management to ensure that content remains healthy and vigilance in the face of changing technology, censorship, hacking, and more. It takes a community to safeguard the scholarly record. It’s too big a job for any single organisation, and too horrific for our species if done badly.

2. CLOCKSS

This paper is written from the perspective of CLOCKSS [1], a collaboration between world leading academic publishers and research libraries to provide a sustainable dark archive for the scholarly record. Initiated as a project by Stanford University 23 years ago, and incorporated as an independent not-for-profit 13 years ago, CLOCKSS is entrusted with more than 46 million journal articles, hundreds of thousands of books, protocols, software, and essential metadata including CrossRef DOIs.

The CLOCKSS Board is composed of an equal number of librarians and publishers, with two co-chairs. Duncan Campbell of Wiley is the current publishing co-chair. Other publishing representatives on the Board include professionals from the American Medical Association, the American Physiological Society, Elsevier, Emerald, IOPP, Oxford University Press, the Society for Industrial and Applied Mathematics, Springer Nature, Taylor & Francis, and Wolters Kluwer.

*E-mail: awise@clockss.org. ORCID: <https://orcid.org/0000-0002-7898-3428>.

3. Publishers preserve... but not everything soon enough

Sadly, failure to preserve at all (or until it is too late) is a key challenge. This challenge is greatest for less formally published scholarly communications and supplementary materials, but it remains firmly in scope for formal published content and established publishers too. Of 2.8 million ISSNs issued to date, only 68,960 are preserved in long-term digital preservation services according to the KEEPERS registry, and fewer than 20,000 of these are preserved in 3 or more of such services which is considered best practice [2].

Two studies show that hundreds of Open Access journals have disappeared entirely from the web in the last 20 years, and that more than 7,000 titles registered with the Directory of Open Access Journals (DOAJ) have no preservation policy or archive in place [3,4].

4. Project JASPER and the preservation of Open Access journals

To address this challenge, CLOCKSS is partnering with the DOAJ, the ISSN International Centre's KEEPERS registry, the Internet Archive, and the Public Knowledge Project on the JASPER project [5]. The purpose of project JASPER is to ensure more Open Access journals are preserved for the long-term. Many Open Access publishers are attentive to the need for long-term preservation of their content. Content published by Open Access publishing organizations such as Frontiers, MDPI, Peer J, and Science Open are preserved with the CLOCKSS archive, for example. However, there are many practical challenges for smaller publishers and project JASPER is focussed on the very long tail of small journal publishers and particularly those who publish under the diamond Open Access model [6].

Larger publishers can, and should, self-organize to preserve their publications with accredited archives [7]. This step is sometimes overlooked for Open Access titles which may not have a contractual commitment to preserve required by library customers for titles published under the subscription business model.

Preprints and accepted author manuscripts are very rarely formally preserved in accredited archives, and so the essential function of long-term digital preservation is an extremely important feature of the version of record.

5. Book preservation

Electronic books are also in need of long-term preservation, and in many ways the challenge for books is greater than for journals. Unlike the journal world, there is no single registry of all ISBNs that have been minted, nor any international infrastructure to record whether and where each book is preserved. This challenge continues to increase in complexity as book formats continue to evolve and books are published in increasingly interactive ways, meaning that multiple versions of a book need to be preserved in order for the content to be safe and usable in the long-term.

6. Conclusions

What practical step can publishers take to ensure the publications they are entrusted with by authors are preserved? They can ensure that their digital publications are preserved digitally in at least three accredited archives. The more copies, the safer the scholarship; the less correlated those copies are, the safer the

scholarship; the more reliable each copy, the safer the scholarship; and the faster failures are detected and repaired, the safer the scholarship.

References

- [1] <https://clockss.org/>.
- [2] The NDSA Levels of Digital Preservation: An Explanation and Uses. <https://ndsa.org/publications/levels-of-digital-preservation/>.
- [3] M. Laakso, M. Matthias and N. Jahn, Open is not forever: A study of vanished open access journals, *Journal of the Association for Information Science and Technology* (2021). doi:10.1002/asi.24460.
- [4] J. Bosman et al., OA Diamond Journals Study. <https://zenodo.org/record/4558704>.
- [5] <https://doaj.org/preservation/>.
- [6] J. Bosman, J.E. Frantsvåg, B. Kramer, P.-C. Langlais and V. Proudman, OA Diamond Journals Study. Part 1: Findings, *Zenodo* (2021). doi:10.5281/zenodo.4558704.
- [7] <https://www.crl.edu/archiving-preservation/digital-archives/certification-assessment>.