Copim Open Archiving Criteria

Copim Open Archiving Criteria for Open Access Publications

Eight criteria are identified for the effective archiving of open access publications. Although developed primarily for books, we believe the criteria are appropriate for open access publications generally.

This page reproduces the text of Cole, G., Gatti, R., Higman, R., Stokes, P., Steiner, T., & Turpin, H. (2026). *Copim Open Archiving Criteria* (1.0). Zenodo. https://doi.org/10.5281/zenodo.19882469

Introduction

Open access publications carry specific characteristics that make appropriate archiving solutions both more urgent and more complex. Published under open licences, such as Creative Commons Licences, they may be freely downloaded, redistributed, and in many cases adapted without restriction beyond basic attribution to the original work (itself highlighting the importance of access to the original work). They often contain embedded or interactive content hosted by third parties, active hyperlinks to external material essential to understanding the work, and are more readily updated or revised than print equivalents. All these features, while central to the openness and scholarly value of such publications, raise important challenges for monitoring version control and ensuring the integrity of content, and create vulnerabilities that a robust archiving framework must directly address.

At the heart of the criteria is a commitment to openness not just in access, but also in process. An archiving solution for open access content should itself be transparent: openly accessible to readers without DRM restrictions, built on openly licensed metadata, and governed by verifiable, publicly stated processes for preservation, version control, and content integrity. Checksums, permanent URLs, and redundant geographic distribution are not technical niceties — they are the foundations of trust that allow future readers, researchers, and institutions to be confident they are engaging with authentic, unaltered scholarship.

These criteria have been developed as part of the COPIM and Open Book Futures projects, which focused primarily on the needs and priorities of smaller, often scholar-led, open access book publishers — organisations that typically operate with few staff, limited in-house expertise in preservation and metadata, modest annual output and necessarily rely on third-party

archiving solutions for their publications. (1) However we believe the criteria are quite general and are equally applicable to all open access scholarly publications - not just books - and to publishers of all sizes.

The intention is that these criteria can be used by publishers as a basis to assess their existing archiving processes and to design effective archiving solutions for open access content, and by funders and policy makers as a basis for identifying archiving requirements for open access content without necessarily identifying or mandating specific solutions.

The criteria recognise that open access archiving is a shared responsibility across multiple stakeholders — authors, publishers, repositories, and preservation organisations — and that no single solution will suffice. They are not intended as a barrier to participation, but as a common standard that raises the floor for archiving practice across the open access publishing ecosystem. They are designed to inform the development of shared solutions, guide the selection of existing services, and support the training and guidance that will help the whole community meet them. (2)

These eight criteria emerged from a workshop by the participants in the Open Book Futures Archiving work package (Higman et. al. 2025) and have been refined and validated through extensive community consultation with publishers, editors, solution providers, and digital preservation specialists via numerous discussions, conference presentations, workshops and an expert gathering convened by the Digital Preservation Coalition in October 2025. We are extremely grateful to all the participants who engaged in this process and for the feedback we received. The resulting framework distinguishes between what we believe an open archiving solution must provide and what it should provide, as set out in full below.

The Copim Open Archiving Criteria

The Copim Open Archiving Criteria

Must Haves

The following four conditions are considered essential requirements for any open archiving solution serving open access publications.

1. Openly Accessible Content

The original published content — in all its file formats — must be freely available to download by any reader with access to the internet, without DRM restrictions and under the same licence as the work itself. The archive must permit bulk downloads of content, enabling independent third-party preservation and ensuring that content can be rescued in the event of archive failure. A permanent URL must be available for each archived work, so that the location of the archived version can be recorded within the work's metadata and DOI record, creating a stable

and citable point of access for future readers.

2. Openly Accessible Metadata

Book metadata must be archived under a CC-0 (or equivalent) licence and made freely available for all readers to download alongside the work itself. Open metadata is not merely a convenience — it is a prerequisite for discoverability, interoperability, and the long-term ability of the scholarly community to find, verify, and build upon published research. Proprietary or restricted metadata undermines the openness that open access publishing is intended to deliver.

3. Openly Verifiable Archiving and Preservation Processes

For both content and metadata, the processes by which material is archived and preserved must be transparent and independently verifiable. This requires the publication of checksums to allow readers and institutions to verify content integrity; transparent version control so that the history of a work can be traced and understood; clearly stated processes for verifying that archived content matches the published version of record; and transparent preservation processes — including the maintenance of original files — so that the integrity of the archive can be assessed over time.

4. Adherence to Accepted Good Practice in Digital Archiving Operations

An archiving solution must demonstrate that it operates in accordance with established good practice for digital archives, such as membership of the ISSN Keepers Registry or compliance with a CRL TRAC audit, ISO:16363 or the Core Trust Seal. This includes reliability and documented strategies for ensuring the long-term continuation of the archiving service, as well as the maintenance of multiple geographically redundant copies of all archived content. The long-term viability of the archive — not merely its current operation — must be a central consideration in the design and governance of any solution.

Should Haves

The following four conditions are strongly recommended, reflecting the particular needs and characteristics of open access content, and the broader responsibilities of any archiving solution operating in this space.

5. Support for Retrieving and Archiving Additional Content Associated with the Work

Open access publications frequently depend upon material beyond the core text — supplementary datasets, media files, and web-based resources linked within the work itself. An archiving solution should actively support the retrieval and preservation of such additional materials, including supplementary content provided by publishers and authors to accompany the main text, and web pages referenced via URLs within the publication. Archiving linked web content is a critical mechanism for mitigating link rot, which poses a particular threat to the integrity of digitally native scholarship. (3)

6. Clearly Stated Policies Around Removal of Content

Because open access content is freely and publicly available, archiving platforms are more likely than restricted repositories to receive requests — or demands — for content to be taken down. An archiving solution should have clearly stated, publicly available policies for handling such requests, including tomestoning policies and processes for distinguishing between legitimate legal obligations and requests motivated by a desire to suppress lawfully published content. The question of retractions also requires explicit policy: where a publisher retracts a work, there may be a continuing scholarly and historical interest in preserving access to the retracted version alongside appropriate contextual information.

7. Independence from Private or Government-Controlled Entities

To guard against censorship and the concentration of control over the scholarly record, an archiving solution should be structured so as to minimise dependence on any single private or government-controlled entity. This is best achieved through the distribution of content across diverse legal jurisdictions, the use of diverse technologies and hosting environments, and the adoption of networked solutions involving multiple providers rather than a single point of failure. Redundancy of entity — not merely of copies — is a distinct and important dimension of resilience for openly published content.

8. Collation and Provision of Usage Statistics

Since open access content can be freely downloaded from archiving platforms without passing through a publisher's own systems, authors and publishers may otherwise have no visibility of how their work is being accessed and used. An archiving solution should therefore collect and make available usage statistics — including download counts and access data — to inform authors and publishers about the reach and impact of archived works. This supports both the practical needs of publishers and authors and the broader goal of demonstrating the value of open access scholarship.

---

1. An open archiving toolkit for small publishers is also available (link to Copim Compass)
2. A review of how several existing archiving solutions meet the various criteria is contained in an associated report (link) - based on publically available information.
3. An associated report on link-rot and procedures for archiving related content has also been prepared as part of the Open Books Futures project (link)

Revision #1
Created 2026-04-27 15:17:12 UTC by Rupert Gatti
Updated 2026-04-30 13:54:02 UTC by Rupert Gatti