

CLOCKSS Collection Development Policy

Contents

Introduction	2
Collection and Archive Scope	2
Appraisal process policy.....	2
Inclusion Policy for Publishers	Error! Bookmark not defined.
Completeness policy	3
Journals.....	3
Books	3
E-journal format policy	4
Content Format.....	4
Content Delivery.....	4
E-book format policy	4
Eligibility Criteria	5
ISBN Requirement.....	5
Digital Format Requirements	5
Metadata Standards	5
Scope of Eligible Content.....	5
Other digital resource ingestion policy.....	6
Scope and Eligibility.....	6
Updates, modifications and deletions	6
Research Integrity Policy.....	6
Accessibility policy	6
Triggered Content Accessibility.....	7
Commitment to Continuous Improvement	7
Diversity, Equity and Inclusion Policy	7

Introduction

CLOCKSS is committed to the long-term preservation of digital scholarly content, ensuring that academic and scientific works remain accessible and intact for future generations. This policy outlines the procedures and guidelines for preserving various types of digital content, maintaining accessibility, ensuring sustainability, and promoting diversity, equity, and inclusion in our operations.

Collection and Archive Scope

The materials selected for inclusion in the archive must meet the following criteria:

- Relevance to Scholarly Communities: CLOCKSS focuses on archiving content that is collected by university libraries and thereby recognized as important to culture, literature, research, and scholarly communication.
- CLOCKSS preserves materials from publishers who are recognized as reputable and ethical within the academic community.
- CLOCKSS seeks to preserve content from a wide range of sources and to maintain the diversity and richness of preserved scholarly content.
- Journals and eBooks provided by publishers under various business models, including open access and subscription-based models, are eligible for archiving.
- Digital content published by libraries is eligible for inclusion in the archive if it is sold to other libraries or made available through Open Access (OA) initiatives and platforms.
- These materials encompass a wide range of publication types, including but not limited to journals, eBooks, conference proceedings, and supplementary materials.
- Works in any language are eligible for inclusion.
- Publications in any academic subject area are eligible for inclusion.
- The archive accepts materials published in any country.
- As well as current publications, digitized archives and backfiles fall within scope of the archive.

▪ Appraisal process policy

CLOCKSS evaluates content for preservation based on several key factors to ensure that the material meets archival standards and remains accessible for the long term.

1. Commitment to Preservation Standards: Content is evaluated to ensure it meets CLOCKSS's standards for long-term preservation. CLOCKSS adheres to best

practices for digital preservation, ensuring the integrity and authenticity of the content.

2. **Technical Suitability:** CLOCKSS evaluates the format of the digital content to ensure it can be effectively preserved over time. It uses a format-agnostic approach but gives priority to widely recognized and accessible digital formats that are more likely to be viable in the future.
3. **Licensing and Legal Agreements:** Content must be deposited into CLOCKSS with clear rights for long-term preservation and a signed specific contract must accompany the content.
- 4.

Completeness policy

CLOCKSS aims to preserve a comprehensive and vital portion of the scholarly record, working in partnership with other trusted archives to ensure broad coverage. While we strive to capture full runs of e-journals, e-books, and datasets, our focus is on preserving key scholarly content and critical supplementary materials. We recognize that no single archive can preserve the entire scholarly landscape, so we work collaboratively with other institutions to ensure a robust and sustainable approach to long-term preservation.

To achieve this, we would ask for the following:

Journals

1. **Full Issues and Articles:**
 - Complete journal issues, including all articles, editorials, letters, and supplementary materials.
2. **Metadata:**
 - Detailed metadata for each title, including title, authors, publication date, volume, issue number, DOI, ISSN, and license (if OA).
 - Detailed metadata for each article, including title, authors, abstract, publication date, volume, issue number, and DOI.
3. **Supplementary Materials:**
 - Any supplementary files associated with articles, such as datasets, multimedia files, or additional documentation.
4. **Manifest Pages:**
 - Publisher manifest pages listing the journal issues and articles, ensuring easy verification of completeness.

Books

1. **Complete E-Books:**
2. Full text of the e-books in relevant formats (preferably EPUB 3 or PDF) including all chapters, forewords, introductions, indexes, and appendices.
3. **Metadata:**
 - Detailed metadata for each book, including title, authors, publication date, ISBNs, DOI, and license information (if OA).
4. **Supplementary Materials:**
 - Any supplementary files associated with the books, such as images, charts, tables, and multimedia content.

E-journal format policy

CLOCKSS is dedicated to the preservation of e-journals across diverse disciplines. Our e-journals ingestion policy stipulates direct submission by publishers. We ensure that ingested e-journals are complete, correctly formatted, and accompanied by all relevant metadata.

Content Format

CLOCKSS preserves a wide array of e-journal content, including but not limited to:

- Textual Content:
 -
- Page Images: High-quality page images in accessible PDF or other accepted formats that accurately represent the original printed versions of the articles.
- Multimedia and Supplementary Files: Any accompanying media files, datasets, or supplementary materials are preserved in their original digital formats, ensuring the complete scholarly work is maintained.

Content Delivery

To ensure efficient and comprehensive ingestion of e-journal content, CLOCKSS supports multiple delivery methods:

- FTP Transfers: Content providers can securely upload files directly to CLOCKSS via FTP, allowing for the reliable transfer of large volumes of data.
- CLOCKSS can harvest content through RSS XML feeds or any web harvesting tools, including referenced images, multimedia, and other associated files, ensuring that dynamic and regularly updated content is captured and preserved.
-

E-book format policy

The preservation of e-books is a key component of our service. Our e-books ingestion policy includes the acquisition of complete e-books, including all chapters, figures,

tables, and supplementary materials. We work closely with publishers to ensure that e-books are provided in a preservation-friendly format and include comprehensive metadata.

Eligibility Criteria

ISBN Requirement

- All books submitted for archiving must possess a valid International Standard Book Number (ISBN) to ensure they are uniquely identifiable and meet global publishing standards.

Digital Format Requirements

- Books must be available in one or more of the following digital formats: EPUB, HTML, PDF, or XML. EPUB 3 is our preferred format. These formats are preferred for their compatibility with various digital preservation and access technologies.

Metadata Standards

- **Required Metadata Formats:** Metadata for all books must be provided in either ONIX (Online Information Exchange) or MARC (Machine-Readable Cataloging) format, ensuring that the information is standardized and can be easily integrated into our archiving system. For FTP content, we accept JATS, BITS, and any other well-formed XML (as mentioned above). For web-harvested content, we can use HTML meta tags and .ris citation files to harvest the metadata.
- **Disciplinary Codes:** Metadata must include relevant disciplinary codes to facilitate categorization and retrieval of the books across various academic and research fields.

Commented [GU1]: For FTP content, we also accept JATS, BITS and any other well-formed XML (see comment above). For web harvest content, we can use html meta tags as well as .ris citation files to harvest the metadata.

Scope of Eligible Content

- **Academic Books:** We accept a wide range of academic books, including both non-fiction and fiction works that contribute to scholarly discourse.
- **Serial Books:** Books that are part of a series or multi-volume works are also eligible for inclusion.
- **Conference Proceedings:** Peer-reviewed conference proceedings in book format are welcome, as they represent valuable academic and professional knowledge.

Other digital resource ingestion policy

CLOCKSS recognizes the importance of preserving other digital resources such as datasets, multimedia files, and software, which accompany scholarly publications. Our collections policy ensures that these materials are ingested alongside primary content, with appropriate metadata to facilitate their future use and accessibility.

Scope and Eligibility

1. Dataset Requirements

- **Digital Format:** Datasets must be in standard, widely accepted digital formats that support long-term preservation and accessibility, such as CSV, JSON, or XML.
- **Metadata:** Each dataset must be accompanied by comprehensive metadata that adheres to open standards like Dublin Core or DataCite. The metadata should detail the dataset's connection to the related article or book (if any), as well as its content, structure, and context.

Updates, modifications and deletions

CLOCKSS maintains a rigorous policy for handling updates, modifications, and deletions to ensure the integrity of the archived content. We track and record all changes to ingested materials, including version updates and errata. Content is not deleted from the archive but changed versions can be added to it. Publishers *must* send us corrected and retracted versions of articles.

Research Integrity Policy

CLOCKSS is committed to upholding the highest standards of research integrity in all its preservation activities. Our Research Integrity Policy is designed to ensure that all preserved content is authentic, reliable, and free from misconduct, thereby maintaining the credibility and trustworthiness of the scholarly record.

Accessibility policy

CLOCKSS is dedicated to ensuring that preserved content is accessible to all users, including those with disabilities. We are committed to adhering to international accessibility standards, such as WCAG 2.1, and encourage publishers to achieve this goal. Through collaboration with publishers, CLOCKSS strives to make content universally accessible and compliant with these important standards.

Triggered Content Accessibility

- **Accessible Content on Trigger Events:** In the event that content is triggered for public access, CLOCKSS will make every effort to ensure that content triggered for public access is available and accessible to all users.

Commitment to Continuous Improvement

- **Ongoing Evaluation:** CLOCKSS regularly evaluates its systems and processes to ensure they evolve with accessibility standards. We are committed to continuous improvement, adapting to new accessibility requirements and technologies as they emerge.

Diversity, Equity and Inclusion Policy

CLOCKSS is dedicated to promoting diversity, equity, and inclusion (DEI) in all aspects of our work. Our DEI policy ensures that our preservation practices are inclusive, reflecting the diverse voices and perspectives within the scholarly community. We strive to provide equitable access to preserved content and actively seek to include works from underrepresented groups and regions.

Inclusive Preservation Practices: We strive to make our preservation practices inclusive, ensuring that all aspects of our work embrace and reflect the diversity of the scholarly community.

Equitable Access: We are dedicated to providing equitable access to preserved content, ensuring that all users benefit from our resources, regardless of their background or location.

Inclusion of Underrepresented Works: We actively seek to include works from underrepresented groups and regions, working to address disparities and enhance the representation of diverse perspectives within our archive.