



File Transfer Ingest Requirements

Thank you for choosing to archive your materials with the community governed CLOCKSS Archive. Your content will be ingested into the archive via "file transfer".

CONTENT DELIVERY

Required General

- Uniquely named deliveries, including full text and supplemental materials
- Clear, consistent file naming convention that correlates specific metadata to full text content
- Full-text content in a standard readable form (ex PDF)
- Notification to the delivery alias when a delivery has occurred
- A listing of files delivered with updates clearly identified
- Delivery of clearly identified "duplicative content" only for corrections and updates
- Delivery of non-proprietary formats (such as tar and zip) is required

Required Metadata

- CLOCKSS requires XML metadata in a standard schema to identify delivered content.
- JATS is used to identify journal article content. For general information about the schema, go to: <http://jats.nlm.nih.gov/index.html>
 - Sample files can be found at: <http://jats.nlm.nih.gov/publishing/tag-library/1.1d1/n-cq42.html>
 - See (1) below for a sample
 - CLOCKSS only uses the front matter portion of the XML, which provides metadata about the content.
- For books, CLOCKSS accepts ONIX2 and ONIX3, both long and short form tags.
 - For specific information about what is allowed in the schema, go to: <http://www.editeur.org/93/Release-3.0-Downloads/>
 - Sample files can be found at: http://www.editeur.org/files/ONIX%203/ONIX_Books_3.0_sample_2.zip
- Metadata to be supplied: DOI, publisher, publication date, ISSN or ISBN, and as appropriate: publication title, article/book title, series title, volume, issue.
- DOIs for journal articles and book titles are required

Desired General

- Content as individual PDFs without redundant versions such as manuscript, epub, mobi
- rsync is the preferred over FTP as the data delivery method
- Deliveries in tar or zip bundles should be completely self-contained. Multiple archives should not need to be expanded to map metadata to articles
- Updates to previously delivered content would include just the modified items
- Semi annual list of individual items sent for preservation - e.g. all the ISSNs and DOIs

SERVER STAGING SETUP

- If the publisher has a staging server, please tell us: the location and password; the update schedule; the content removal schedule
- If the publisher would like to send content to the CLOCKSS staging server, we will send you the location and password; please tell us your update schedule.

(1) JATS sample: PJID Example Journal Title Optional Subtitle 1111-1111 Publisher of Journal 10.0000/fo Article Title Goes Here Writer Sally Q Scribbles John P 13 4 2002 324 7342 880 882 PJID Example Journal Title Optional Subtitle 1111-1111 Publisher of Journal 10.0000/fo Article Title Goes Here Writer Sally Q Scribbles John P 13 4 2002 324 7342 880 882

NOTE: CLOCKSS does not routinely accept "published before print" articles. Please do not include this content in your preservation feed.