Journal Submission Guidelines

Portico does not impose strict requirements for publisher content submissions; however, we do have some baseline needs for publishers to meet so that we can successfully ingest content into the archive.

• Portico requires both metadata and rendition files (typically PDF) for all content

• PDF renditions must be accompanied by metadata files (preference is XML)

• There must be a clear way to identify which metadata files belong to which article rendition files (through file references in the metadata, naming convention, packaging structure, etc.)

• The minimum metadata fields that we expect for each article supplied are:

Journal title
Journal P-ISSN (if available)
Journal E-ISSN
Article ID (DOI if available; if you have DOI plus internal article IDs, we would like to receive all of them, if possible)
PDF file name
Article title
Authors
Publication date
Volume
Issue*
Supplement (if relevant)
First page
Last page
Copyright statement

*Not required for article-level publishing.

Please contact us if you have any questions or concerns regarding the minimum metadata fields.

• All images and other files referenced in the metadata must be included in the content deliveries to Portico.

• We prefer to receive all supplemental materials related to content, whether or not it is referenced in the metadata (if not referenced, we need to know your naming convention so that we can relate the files to the correct article).

• While we can be flexible regarding the packaging of content, we do require content to be delivered consistently. We realize that publishers switch platforms or change formats from time to time, and work with our publishers to ensure that modifications due to planned changes continue to meet our needs.

Note: Portico is happy to use CrossRef or PubMed metadata for publishers who cannot otherwise export metadata from their publishing platform.

Portico's preferred packaging is:

For issue level deliveries
• One Zip file per issue
• Where images or other related files are being supplied with content: one directory per article

For article level deliveries
• One Zip file per article, OR
• One Zip per delivery, with one directory per article

Portico's preferred naming convention
• For issue level deliveries:
journalID_Vol_Iss_Year
• For article level deliveries:
journalID_Vol_Year_articleID

Overview of Portico's Ingest Process
Portico analyzes each publisher's content, develops a preservation plan for it, and creates tools that ensure your content is properly ingested and preserved. Tools typically developed include:

Profile (required)
• Understands the packaging of and expected file types for your content

Transform (required)
• Creates an NLM XML version of your metadata

Loader (required)
• Allows us to batch and load your content into Portico's Content Preparation System

Delayer tool (optional)
• Creates individual article XML files out of single issue-level XML file

To illustrate the packaging and naming conventions, an issue-level delivery for content with full-text markup would look like this:

```
abc_12_3_2012.zip
    12345/
        12345.xml
        12345.pdf
        fg01.jpg
        fg02.jpg
    23456/
        23456.xml
        23456.pdf
    . . . .
```

Our process to ingest publisher-supplied content into the archive includes a step that determines the file format validity of every file being preserved. Each file is matched to its format specification to decide what preservation level will be assigned to the file. All files in the Portico archive are assigned one of four format validities:

Well Formed and Valid
• The file completely matches its format specification

Well Formed and Not Valid
• The file matches its format specification, but has an ill-formed value in the metadata (or, in the case of XML, if the file parses but does not conform to its DTD or schema)

Not Well Formed
• The file does not match its format specification

Not Determined
• The file format does not yet have a validation module in JHOVE (as the community develops validation modules for these other formats, we will check these files for format validity)