



Portico Content Type Action Plan: Supplied Files for E-Book Content

1. Content Type Introduction

1.1. E-book supplied files comprise the digital packages supplied by the e-book content producers to Portico or harvested by Portico (for example, by web crawling) for processing into Submission Information Packages by the Portico ConPrep system for ultimate ingest into the Portico archive. Portico preserves the entirety of the supplied packages in this content-type for all content. This will allow us to “redo” the ingest process for the content, should we discover we made a mistake. This content type was introduced simultaneously with the E-Book content type.

2. Policy Statement

2.1. As there is no standard package format for e-book content, the files comprising each publisher’s content stream differ. Publishers might provide one or more container (ZIP or TAR) files, or they might provide a subset of a hierarchical file system (one or more directories, containing files and possibly other sub-directories).

- Analysis sometimes reveals extraneous files that are included in the publisher package, but which do not comprise components of the books (e.g. thumbs.db files). These files are not ingested into the Portico archive.
- As there is some risk of an erroneous characterization of these files as extraneous, Portico packages all supplied e-book content (as a ZIP file for each supplied batch) and retains all the original publisher-supplied files.

2.2. Portico may determine at some future date that it is no longer necessary to retain the publisher-supplied files, and may, at its discretion, remove these assets from the archive.

2.3. Portico will maintain technical metadata associated with the supplied files in the Archival Information Package containing the e-book.

2.4. All supplied files for e-book content will be associated with a business artifact comprising the contract authorizing Portico to store and manage those assets.

3. Implementation Examples

3.1. The supplied files artifact is processed by the Portico Conprep system, including processing by the JHOVE tool for validation and technical metadata extraction. The Portico-created container and its associated metadata file are then ingested into the archive

4. Document History

4.1. Approved by: Kate Wittenberg

4.2. Last Review Date: 8/7/2023

4.3. Reviewed by: Amy Kirchhoff, Karen Hanson, Kate Wittenberg, Sandra Parr

4.4. Change history:

Version	Date	Change	Author
1.0*	3/27/2016	Created document, based on e-journal supplied files doc.	Amy Kirchhoff



1.1*	8/7/2023	Reference web harvesting as well as content delivered by producers. Minor copy edits.	Karen Hanson
------	----------	---------------------------------------------------------------------------------------	--------------

* An approved version of this document.