



Portico Content Type Action Plan: E-Book Content

1. Content Type Introduction

- 1.1. E-book content is comprised of books or monographs in digital format.

2. Policy Statement

- 2.1. The fundamental archival unit for e-book content is a book.
- 2.2. E-book content as received from the content provider is batched and processed into Submission Information Packages (SIP) by the Portico ConPrep System.
 - Each content provider's content stream package structure is analyzed by Portico to determine the directory structure and file naming conventions used by the publisher to associate the components which comprise a single e-book.
 - The results of this analysis are described in the Turnover Document associated with the content stream. The results are also expressed more formally in the XML profile which informs the ConPrep system as it processes publisher content and assembles the components of each archival unit (book).
 - Analysis sometimes reveals superfluous 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.
 - With the exception of these excluded files, all assets provided by the publisher are validated and characterized with the JHOVE tool, associated with the appropriate e-book, and are associated with a Portico METS file containing technical and descriptive metadata, in the archival unit.
- 2.3. Portico may migrate content provider supplied XML files from their "native" XML format to the Portico archival e-book XML format. This decision will be made on a case-by-case basis (unlike e-journals, there are strictly enforced e-book standards and it is unlikely Portico will migrate content in those formats as part of the ingest process). Should Portico migrate supplied XML, Portico will preserve the original e-book format artifacts along with the transformed content.
- 2.4. Portico will determine, on a licensor-by-licensor basis, whether or not to preserve the supplied delivery package.
- 2.5. With the exception of any "excluded files," all assets provided by the publisher are validated and characterized with the JHOVE tool, associated with the appropriate book, and are associated with a Portico metadata file containing technical and descriptive metadata, in the archival unit.
- 2.6. Should Portico make the determination that any formats of the assets comprising the e-book content are at risk, these components will be migrated in conformance with the Format Action Plan for that format.
- 2.7. All e-book content will be associated with a business artifact comprising the contract authorizing Portico to store and manage those assets.
- 2.8. Portico will maintain descriptive and technical metadata associated with the technical artifact in the Archival Information Package (AIP) containing the e-book.

3. Implementation Examples

3.1. Portico performs two migrations on e-book content that looks like e-journal content. Portico transforms the content provider’s mark-up to a Portico version of the NLM e-book XML and transforms the content provider’s packaging into the Portico content model. The Turn Over Document associated with each stream of content from a specific content provider is considered the Format Action Plan for these two initial migrations. The migration of the packaging and the migration of the publisher-provided XML and SGML files in the publisher’s “native” XML or SGML format to the Portico NLM XML format are performed for ease of collection management and the delivery needs demanded by the trigger event model associated with this content-type and the fact that the publisher has not used a standard e-book format. The publisher-provided XML/SGML file is preserved in the archive along with the other assets comprising the article.

4. Document History

- 4.1. Approved by: Amy Kirchhoff
- 4.2. Last Review Date: 8/5/2009
- 4.3. Reviewed by: Stephanie Orphan, Sheila Morrissey
- 4.4. Change history:

Version	Date	Change	Author
0.1	6/04/2009	Drafted document	Sheila Morrissey
0.2	06/25/2009	Edit and format in standard template	Sheila Morrissey
0.3	07/07/2009	Add info re: contract	Sheila Morrissey
0.4	07/27/2009	Add Stephanie’s copy-edit suggestions	Sheila Morrissey
1.0*	7/29/2009	Made it look more like e-journals and added in the bit about 2 migrations.	Amy Kirchhoff
1.1*	8/5/2009	Copy edited and added review by line.	Amy Kirchhoff

* An approved version of this document.