eArchiving Conformance Report for the Generic Services Project: The Science and Engineering Library

Document Control Information

Settings	Value
Document Title:	eArchiving Conformance Report for the Generic Services Project: The Science and Engineering Library
Project Title:	eArchiving
Evaluators:	Karin Bredenberg, Carl Wilson, Sven Schlarb.
Project Owner (PO):	Árpád Welker, Fulgencio Sanmartín, DG CNECT
Doc. Version:	1.0
Sensitivity:	Limited
Date:	15/08/2023

Document Approver(s) and Reviewer(s):

NOTE: All Approvers are required. Records of each approver must be maintained. All Reviewers in the list are considered required unless explicitly listed as Optional.

Name	Role	Action	Date
Árpád Welker/Fulgencio Sanmartín	Project Owner	<Approve / Review>	

Conformance Result: The Science and Engineering Library is 100% E-ARK/eArchiving conformant.

Conformance Submission details:

A selection of 3 Information Packages (IPs) per type (SIP, AIP, DIP) were requested by the E-ARK/eArchiving Consortium, i.e. a total of 9 information packages (container packages in *.zip, *.tar, or *.tar.gz format) plus system log files (*.log or *.txt format) related to the creation of information packages.

9 IPs were supplied by The Science and Engineering Library together with system log files:

- 3 SIPs;
- 3 AIPs;
- 3 DIPs;

All 9 IPs validated correctly so The Science and Engineering Library is 100% E-ARK/eArchiving conformant.

Observations

- All of the packages submitted had package and representation level METS documents.
- There was no real difference between the SIPs, AIPs and DIPs, with no evidence of value added during the package lifecycle.
- Packages 2 and 3 contain excessive file format documentation and rendering software packed as documentation. This would be better normalised to a collection level. Currently the documentation is the largest part of the package.
- Some elements from the E-ARK SIP have been used but not all mandatory elements are present, for example the pointer to the correct profile when using E-ARK SIP.

Improvement Opportunities

- While the ERMS metadata files may be technically valid, most of the fields have "not defined" values, which provides no context for users.
- All of the packages contain empty metadata folders, two contain empty representation folders. Some of these are even recorded in the METS documents. This is unnecessary and just adds to the package complexity.
- For packages 2 and 3 the "Dublin Core" metadata file only adheres to a proprietary standard, for linking ATOM and Archivematica records. These cannot be validated and are not documented. Dublin core XML would be more useful.
- Packages 2 and 3 contain book page JPEG images packaged in a zip file. There is no useful page level metadata, including structural information needed to present the original item.
- The CITS identification/classification is misleading for all packages, these use the package type declaration of Datasets for all packages. None of them are structured datasets. More appropriate descriptors are available.
- The AIP and DIP for package 1 are identical to the SIP in structure and content, except:
  - The representation was renamed from "representation-1" to "rep01" for reasons that are unclear and undocumented.
  - The AIP has a new, empty representation folder "rep01-preservation", again the reason for this is unclear.
  - There is no evidence of any additional metadata generated by the ingest or dissemination processes, including preservation metadata.