

Long term preservation – from prototype to tools in operation

Ronnie Kolehmainen, Stefan Andersson, Hanna Blomquist,
Uwe Klosa and Anna-Karin Anderson

Electronic Publishing Centre, Uppsala University

Abstract

The project *Long term preservation – from prototype to tools in operation* is carried out by the Electronic Publishing Centre at Uppsala University Library in cooperation with a number of national and international partners.

The aim of the project is to facilitate the management and processing of digitally created material at both local and national levels. An infrastructure that makes it possible to guarantee access to documents in a longer time perspective has been created building on the earlier SVEP-project.

This infrastructure consists of persistent identifier services (previously developed in the SVEP project) and services based on standardised methods and tools for transfer of data between local producers and the local and/or national archives.

Within *this* project the following has been developed:

1. A format and data environment register which includes information on file formats that can be used by the packaging tool.
2. A packaging tool for packaging and transfer of objects based on METS or DIDL and for checking and reception of objects at the local/national archives. This overall tool is assisted by some individual subtools:
 - the DIDL/METS tools are designed to read XML input streams and augment the contents of referenced resources and to extract individual resources.
 - the metadata mapping tool which facilitates a semantic mapping between metadata formats provided by data providers (repositories) and an internal vocabulary used for definition of specific metadata elements (definition list). This mapping is used as a base for transformation to a harmonised metadata format (schema). It also provides a base for a control mechanism and feedback to the data provider (repository) on the quality of the delivered metadata.
 - the contract tool which purpose tool is to create and maintain a specification which helps to make sure that the data and digital objects delivered by data providers (repository) satisfies the requirements specified by the archive. This tool is integrated with the packing workflow and provides both a control mechanism and feedback to the data provider (repository) on the quality of the delivered data (objects).

This infrastructure and the tools has been successfully tested on a smaller scale and is about to move up to full scale production.