Preservation Management

Content

Content Further information Preservation management Basic concepts Technology and community watch Strategic planning Risk management Execution of preservation action Migration Emulation Process diagram: preservation planning

Further information

TIB Preservation Policy

Event metadata

Preservation management

The Rosetta digital preservation system has a preservation planning module. With this module, users can create preservation plans, verify such plans using test sets, and undertake preservation actions based on the test results. The module includes Format Library, mechanisms for risk analysis and evaluation and for carrying out preservation actions. Conversion tools can be integrated as plug-ins. The type and deployment of conversion tools are specifically selected by TIB for each preservation plan, enabling the process to be adapted individually to its objectives, tasks and designated communities.

Basic concepts

The TIB Preservation Policy contains general principles on digital preservation, based on which the integrity, authenticity, interpretability and completeness of archived objects shall be ensured. TIB uses the concepts of bitstream and content preservation to preserve objects. In addition, content preservation is subdivided into the possible strategies of migration and emulation.

Technology and community watch

TIB continuously monitors the latest developments and best practice. The library does this by participating in networks, by enabling employees to attend relevant symposia, by sharing information and ideas with partner organisations, and by observing the digital preservation specialist community and the respective format and content communities. In addition, TIB plays an active part in further developing best practice by producing its own publications (for example: Bähr, Lindlar, Rechert: Functional Access to Electronic Media Collections Using Emulation-As-a-Service (2014)) and by participating in research projects (examples include: EU project "DURAARK – Durable Architectural Knowledge" (2013-2016) and the Web platform for editing, publishing and digitally preserving regional science research data (LaZAR)).

Strategic planning

Preservation management at TIB is based on the following activities and documents:

- Formal quality control by the librarians
- Preliminary analyses during the pre-ingest process and after ingest
- Format validation and format identification
- Extraction of technical metadata in the validation stack
- Significant properties
- Information in Format Library and
- The Preservation Policy

Risk management

In Format Library, 1-n risk factors can be defined for each file format specifically for each institution. Risks may be properties (technical, administrative or process metadata) or the results of analysis undertaken using tools. An example of a concrete risk are file formats that are not linked to any of the playback programs listed in Format Library, which is an indication of impending obsolescence. Another concrete risk is the production software used to create an object where it is known that the production software has converted the format specification incorrectly. Risk analysis is performed as a regular automated task, or is initiated manually. Based on the risk analysis, the institution decides whether preservation action is required. Preservation action may also be taken without a systems-based risk analysis having been conducted beforehand.

Execution of preservation action

The concepts of migration and emulation are used as preservation actions for currently archived objects so as to preserve their interpretability. In the process, TIB is guided by the high-level requirements of preservation planning, as described in the Planets project using Plato.

Emulation is used as a preservation strategy for CD and USB images; to this end, TIB has implemented a solution for mass imaging of CDs and USB sticks.

Migration

If preservation action has to be taken due to file format obsolescence or where a risk to existing file formats becomes known, the process described in the process diagram "Preservation planning" is taken:

- 1. Based on the risk report or the preservation watch process, a risk is detected and the institution sees a need for action.
- 2. A set is formed out of the objects concerned.
- 3. A preservation plan is drawn up, containing the following information:

- a. A name and the reason why the preservation plan is being drawn up. Where needed, a document can be inserted.
- b. Evaluation criteria are selected. Examples include selected significant properties from Format Library or alternative evaluation criteria such as costs or the position of content elements within the document.
- 4. A test set is formed.
- 5. An appropriate conversion tool is selected, and the conversion process is initiated for the test set.
- 6. The output is checked based on the evaluation criteria.
- If the output does not satisfactorily meet the defined evaluation criteria, the process is repeated using a new conversion tool. The preservation plan is only signed off once the output meets the evaluation criteria.
- 8. Only signed-off plans can be applied to complete data sets.

If objects are modified, the objects concerned are transferred from permanent archival storage to operational storage. In operational storage, the object is migrated and transferred to re-ingest as a new version of the AIP.

Emulation

An emulation framework (Emulation as a Service (EaaS)) can be integrated into Rosetta as a viewer; the parameters for the emulation environment can be saved in the technical metadata. When retrieved, the relevant object is transferred to the emulation framework, and loaded there.

Process diagram: preservation planning

Click to enlarge.

