

IPRES 2012

Preservation Health Check: introduction to the pilot

Workshop 4
PREMIS implementation fair 2012

Titia van der Werf

Senior Program Officer
OCLC



What is the Preservation Health Check Pilot?

Joint initiative:

- Open Planets Foundation

A community hub for digital preservation whose main goal is to jointly manage and improve tools and research outcomes for practical use.

- OCLC Research

A community resource for shared R&D that addresses challenges facing libraries and archives in a rapidly changing information technology environment.

What is the Preservation Health Check Pilot?

Aim of the Pilot

To help digital repository/archive managers answer the question:

“What does my preservation metadata tell me about the risks to which my digital collections are exposed?”

What is the Preservation Health Check Pilot?

Problem statement

Digital preservation is an ongoing investment

- taking preventive actions (virus checks, tracking file integrity, etc.)
- creating and managing preservation metadata
- carrying out preservation actions (media refreshment, format conversions, etc.)
- using tools, that need to be managed as well

What is the Preservation Health Check Pilot?

We take all these measures and precautions to avoid risks

... BUT ...

How do we know that we are taking the right measures?

How effective are our standards, tools and practices in addressing digital preservation risks?

What is the Preservation Health Check Pilot?

There are quite a few risk assessment instruments (checklists, frameworks and even standards) for digital preservation

... BUT ...

how many repositories actually perform risk assessments and how regularly?

What is the Preservation Health Check Pilot?

Proposal

Repository managers should be able to perform regular preservation risk assessments (health checks) as part of their preservation management task.

- risk assessment should be made easy (automatically generated reports)
- risk assessment should be based on objective data, generated by the repository (preservation metadata)

What is the Preservation Health Check Pilot?

Proposal

1. Collect the preservation metadata generated by repositories
2. Analyse the metadata collected
3. Report back on the risks to which the objects in the repository are exposed

What is the Preservation Health Check Pilot?

Pilot goal

The pilot will investigate the potential of preservation metadata for assessing digital preservation risks

What is the Preservation Health Check Pilot?

Pilot scope

1. Identify 3 pilot sites with 10-15 years of experience running a digital repository/archive
2. The preservation metadata should be PREMIS or easily mappable to PREMIS
3. The preservation metadata should at least include technical metadata and audit trails of preservation actions (events)

What is the Preservation Health Check Pilot?

Methodology

- Research methodology will be based on mapping the SPOT risk model with PREMIS;
- Apply the mapping to real metadata
- Check and correct assumptions and validate findings with the pilot sites and the digital preservation community

What is the Preservation Health Check Pilot?

Roles

- OPF: community engagement/pilot sites; feed-back to findings via OPF-hackathons; dissemination.
- OCLC: research methodology and activity; dissemination.
- Pilot sites: provide preservation metadata and feed-back to interim findings.

What is the Preservation Health Check Pilot?

Anticipated results

- return meaningful feed-back to the pilot sites on:
 - what their metadata reveal about risks
 - how they could improve their metadata
- recommendations on the usefulness and value of PREMIS for risk assessment
- recommendations for the improvement of preservation metadata quality

What is the Preservation Health Check Pilot?

Desired impact

- The pilot will demonstrate the value of producing and maintaining preservation metadata
- The pilot will demonstrate the importance of regular health checks to meet audit and certification requirements
- The pilot outcomes will provide useful feed-back for the further development of standards, best practices and tools

What are we going to do today?

Workshop agenda

- Standards, Risk analysis, Common sense and Evidence - Bram
- Expectations of a candidate pilot site: Bibliothèque Nationale de France (BNF) - Sébastien
- Presenting the SPOT Model - Priscilla
- Mapping Preservation Metadata to Risk Factors - Brian
- Plenary discussion
- Wrap-up - Titia

Titia van der Werf
titia.vanderwerf@oclc.org

Q&A