

Preservation Planning: Planning for optimal long-term digital preservation solutions

Andreas Rauber

Vienna University of Technology,
Austria

June 24, 2008

Abstract. The rapid technological changes in today's information landscape have considerably turned the preservation of digital information into a pressing challenge. Without appropriate measures, digital objects will be inaccessible in a very short time.

A lot of different strategies, i.e. preservation actions, have been proposed to tackle this challenge. However, which strategy to choose, and subsequently which tools to select to implement it, poses significant challenges. The creation of a concrete plan for preserving an institution's collection of digital objects requires the evaluation of possible preservation solutions against clearly defined and measurable criteria. Preservation planning aids in this decision making process to find the best preservation strategy considering the institution's requirements, the planning context and possible actions applicable to the objects contained in the repository. Performed manually, even evaluating a rather small number of possible solutions against requirements takes a good deal of time. Plato, a web based, interactive software tool, supports and partly automates this process.

This tutorial will start with an introduction to the challenges of digital preservation, highlighting the core strategies available to address them. After introducing OAIS, the ISO standard for an Open Archival Information System, the tutorial will focus specifically on the aspect of creating a sound preservation plan by following a well-defined workflow incorporating documented requirements specification and empirical evaluation of tools. In a series of hands-on exercises, participants will be developing a preservation plan for a specific object collection.

Level: This tutorial is at an introductory level. Some basic understanding of the challenges of digital obsolescence and digital preservation is helpful, but not required.

Expected number of participants: Approximately 35

Contact information: Andreas Rauber
Vienna University of Technology
Institute of Software Technology and Interactive Systems

Favoritenstrasse 9-11/188
A-1040 Wien

E-mail: rauber@ifs.tuwien.ac.at
Phone: +43 1 58801 188{26,63}
Homepage: <http://www.ifs.tuwien.ac.at/dp>

Requirements: An internet connection is required to access the application for demonstrations. During group work, participants would need to connect to the net with one laptop per group, preferably via WLAN. If Internet connectivity constitutes a problem, the exercises can be made off-line for the whole planning process, with the exception of the actual migration and emulation steps performed via webservices. These would be called manually on a local server.

1 Outline

The fast changes of technologies in today's information landscape have considerably shortened the lifespan of digital objects. Digital preservation has become a pressing challenge. A lot of different strategies, i.e. preservation actions, have been proposed to tackle this challenge: migration and emulation are the most prominent ones. However, even the decision to either go for migration or emulation is quite hard, and it gets more complex when it comes to pick a certain migration tool for instance. The process of evaluating potential solutions against specific requirements and building a plan for preserving a given set of objects is called preservation planning which presently is a mainly manual process with little tool support.

This tutorial will introduce the major challenges that digital preservation activities face. It will then discuss the advantages and disadvantages of the various strategies that exist to tackle these problems. The OAIS reference model, an ISO standard describing the functional model as well as the information model of an archival information system will be presented, before focussing on a specific task in digital preservation, namely planning a sound strategy for preserving one's digital assets. To this end a solid workflow that leads to a profound, well-documented decision that one can be held accountable for, will be discussed. Additionally, we will present Plato, a software tool support that is supporting this workflow, automating the core steps in this endeavour.

In this tutorial attendees will then create a preservation plan on the basis of a representative scenario and receive an accountable and informed recommendation for a particular preservation action. This will be supported by Plato, a decision support tool that implements the preservation planning workflow and further integrates web-services for content characterisation, preservation action and automatic object comparison to provide maximum support for preservation planning endeavours.

2 Target Audience Characterisation

The target audience for this tutorial comprises people who are aware of or actively concerned with long-term preservation issues. People intending to ensure the accessibility of their private photo collection for their offspring belong to this group as well as librarians and archivists with a mandate to long-time preserve the institution's digital objects, or researchers caring for the preservation of research data.

3 Learning Objectives

Attendees will learn how to build a preservation plan from scratch by undergoing all steps required to arrive at a well-informed recommendation for a solution to adopt. The main objectives will be to recognise the importance of becoming clear about the institution's inert requirements, evaluating possible preservation strategies and analysing them.

4 Required Equipment

The creation of the preservation plan will be demonstrated using Plato, a web application running on a server at the Vienna University of Technology. Therefore, a standard DSL internet connection is needed which would, for ease of usage, preferably be made available via WLAN. Alternatively, the break-out exercises can be performed using flip-chart and post-it notes, with experiments being demonstrated using locally installed migration tools.

5 Biographies of the presenter

Andreas Rauber is Associate Professor at the Department of Software Technology and Interactive Systems at the Vienna University of Technology. He is actively involved in several research projects in the field of Digital Libraries, focusing on the organization and exploration of large information spaces, as well as Web archiving and digital preservation. His research interests cover the broad scope of digital libraries, including specifically text and music information retrieval and organization, information visualization, as well as data analysis and neural computation. He is involved in numerous initiatives in the area of digital preservation, such as DPE - Digital Preservation Europe; Planets - Preservation and Long-term Access Networked Services; nestor - Network of expertise in Digital long-term preservation. He has been lecturing extensively on this subject at different universities, as part of the DELOS and nestor summerschools on digital preservation, as well as during a range of training events on digital preservation.