Preservation planning with Planets

Hannes Kulovits

Vienna University of Technology
http://www.ifs.tuwien.ac.at/~kulovits

Sofia, September 2009
Outline and schedule

• Recap

• Plato walk-through

• Scenario

• Exercise 1: Select sample objects
• Exercise 2: Create an objective tree
Preservation plan in context
Definition of a Preservation Plan

• ‘A preservation plan defines a series of preservation actions to be taken by a responsible institution to address an identified risk for a given set of digital objects or records (called collection).’

• The Preservation Plan takes into account the preservation policies, legal obligations, organisational and technical constraints, user requirements and preservation goals.
• It also describes the preservation context, the evaluated alternative preservation strategies and the resulting decision for one strategy, including the rationale of the decision.
Evaluating preservation strategies

- Variety of solutions and tools exist
- Each strategy has unique strengths and weaknesses
- Requirements vary across settings
- Decision on which solution to adopt is complex
- Documentation and accountability is essential

- Preservation planning assists in decision making
- Evaluating preservation strategies on representative samples according to specific requirements and criteria
Preservation Planning Workflow
Preservation Planning in Plato

- Web based planning tool implementing the Planets preservation planning workflow
- Publicly available
- Automation of the planning process
  - Integration of registries and services for
    - File format identification
    - Preservation action (migration, emulation...)
    - Characterisation and comparison
- Knowledge base to support planning
Welcome to Plato, the Planets Preservation Planning Tool

Introduction

What is Plato?

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. Different strategies such as migration and emulation have been proposed; however, the decision for a specific tool e.g. for format migration or an emulator is very complex. The process of evaluating potential solutions against specific requirements and building a plan for preserving a given set of objects is called preservation planning. So far, it is a mainly manual, sometimes ad-hoc process with little or no tool support. The planning tool Plato is a decision support tool that implements a solid preservation planning process and integrates services for content characterisation, preservation action and automatic object comparison in a service-oriented architecture to provide maximum support for preservation planning endeavours.

This software is licensed under the CC-GNU LGPL version 2.1 or later. The source code can be downloaded from our project repository.

Click here to enter Plato.
Scenario

- National library
- Scanned yearbooks archive
- GIF images
- The purpose of this plan is to find a strategy on how to preserve this collection for the future, i.e. choose a tool to handle our collection with.
- The tool must be compatible with our existing hardware and software infrastructure, to install it within our server and network environment.
- The files haven't been touched for several years now and no detailed description exists. However, we have to ensure their accessibility for the next years.
- Re-scanning is not an option because of costs and some pages from the original newspapers do not exist anymore.
Exercises

• Exercise 1: Basic questions, Collection
  • Describe your collection, your objects
  • Describe the designated community, organisation...
    – Document that shortly to have a common basis
    – Select representative sample objects from the collection on the USB stick

• Exercise 2: Requirements definition
  – Define the assigned branch of the tree
  – Assign measurable units
  – Set high-level importance factors
<table>
<thead>
<tr>
<th>Group</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Focus</td>
<td>Object+Process characteristics</td>
<td>Object+Process characteristics</td>
<td>Technical+Costs</td>
<td>Technical+Costs</td>
</tr>
<tr>
<td>Room</td>
<td>Arena Hall</td>
<td>Serdica Hall</td>
<td>Sofia Hall</td>
<td>Forum Hall</td>
</tr>
<tr>
<td></td>
<td>Maurice van den Dobbelsteenn</td>
<td>Clive Billenness</td>
<td>Amir Bernstein</td>
<td>Sara van Bussel</td>
</tr>
<tr>
<td></td>
<td>Nataliya Angelova</td>
<td>Kim Sommerville</td>
<td>Marija Milhova</td>
<td>Serap Kurbanoglu</td>
</tr>
<tr>
<td></td>
<td>Zoran Bjin</td>
<td>Niels Hoppe</td>
<td>Brian Hole</td>
<td>Joyce Rambo</td>
</tr>
<tr>
<td></td>
<td>Luga Behije</td>
<td>Frank Houtman</td>
<td>Temenuga Kalcheva</td>
<td>Stefan Stefanov</td>
</tr>
<tr>
<td></td>
<td>Dina Crnec</td>
<td>Mile Jovanov</td>
<td>Nicolaie Constantinescu</td>
<td>Dumitru Tkacenko</td>
</tr>
<tr>
<td></td>
<td>Bruno Dobric</td>
<td>Vanja Jovišić</td>
<td>Eleni Novakovska</td>
<td>Yasar Tonta</td>
</tr>
<tr>
<td></td>
<td>Gabrijela Gavran</td>
<td>Rudolf Müller</td>
<td>Krunoslav Rendulic</td>
<td>Alen Vodopijevce</td>
</tr>
<tr>
<td></td>
<td>Maurizio Gentilini</td>
<td>Tarvo Kärberg</td>
<td>Johann-Baptist Schrall</td>
<td>Kaloyan Zdravkov</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Bojan Marinkovic</td>
<td>Tanya Stoyanova</td>
<td></td>
</tr>
</tbody>
</table>
Schedule

10:30-10:45 Coffee break
10:45-11:15 Define objective tree
11:30-11:45 Presentation and discussion
11:45 Walk-through in Plato
12:00 Lunch break
Questions?

kulovits@ifs.tuwien.ac.at

www.ifs.tuwien.ac.at/dp/plato
www.planets-project.eu