

Digital Preservation Planning

July 29 2008, London, UK



Digital Preservation Planning: Principles, Examples and the future with Planets

organized in cooperation with DPC

Andreas Rauber

Vienna University of Technology

<http://www.ifs.tuwien.ac.at/~andi>

Outline

- ❑ Introduction to Planets
 - Who are we?
 - What are we doing?
 - Why are we doing it?
- ❑ The Planets architecture and components
- ❑ A first glimpse at Planets Preservation Planning



The Planets project

- ❑ 4-year research and technology development project co-funded by the European Union
- ❑ Addresses core digital preservation challenges
- ❑ Started June 2006 with €15m budget
- ❑ Coordinated by the British Library
- ❑ 16 partners
 - national libraries and archives
 - leading technology companies
 - research universities
- ❑ Builds on strong digital archiving and preservation programmes



Planets partners



KB

Koninklijke Bibliotheek



STATSBIBLIOTEKET

Österreichische

Nationalbibliothek



- ❑ The British Library
- ❑ National Library, Netherlands
- ❑ Austrian National Library
- ❑ State and University Library, Denmark
- ❑ Royal Library, Denmark



DET KONGELIGE BIBLIOTEK

NATIONALBIBLIOTEK OG KØBENHAVNS UNIVERSITETSBIBLIOTEK



the national archives



Schweizerische Eidgenossenschaft

Confédération suisse

Confederazione Svizzera

Confederaziun svizra

Swiss Confederation

- ❑ National Archives, UK
- ❑ Swiss Federal Archives
- ❑ National Archives, Netherlands

nationaal archief



Planets partners



- ❑ Tessella Plc
- ❑ IBM Netherlands
- ❑ Microsoft Research
- ❑ Austrian Research Centers GmbH



rechenzentrum
universität freiburg



- ❑ Hatii at University of Glasgow
- ❑ University of Freiburg
- ❑ Vienna University of Technology
- ❑ University of Cologne



The Planets team



All Staff Meeting, February 2007



Aims and objectives

- ❑ **Increase Europe's ability to ensure long-term access to its cultural and scientific heritage**
 - Improve decision-making
 - Control costs through increased automation and scalable infrastructure
 - Ensure wide adoption across the user community
 - Establish a market place for preservation services and tools

- ❑ **Build practical solutions**
 - Integrate existing expertise, designs and tools
 - Deliver tools and services for operational environments

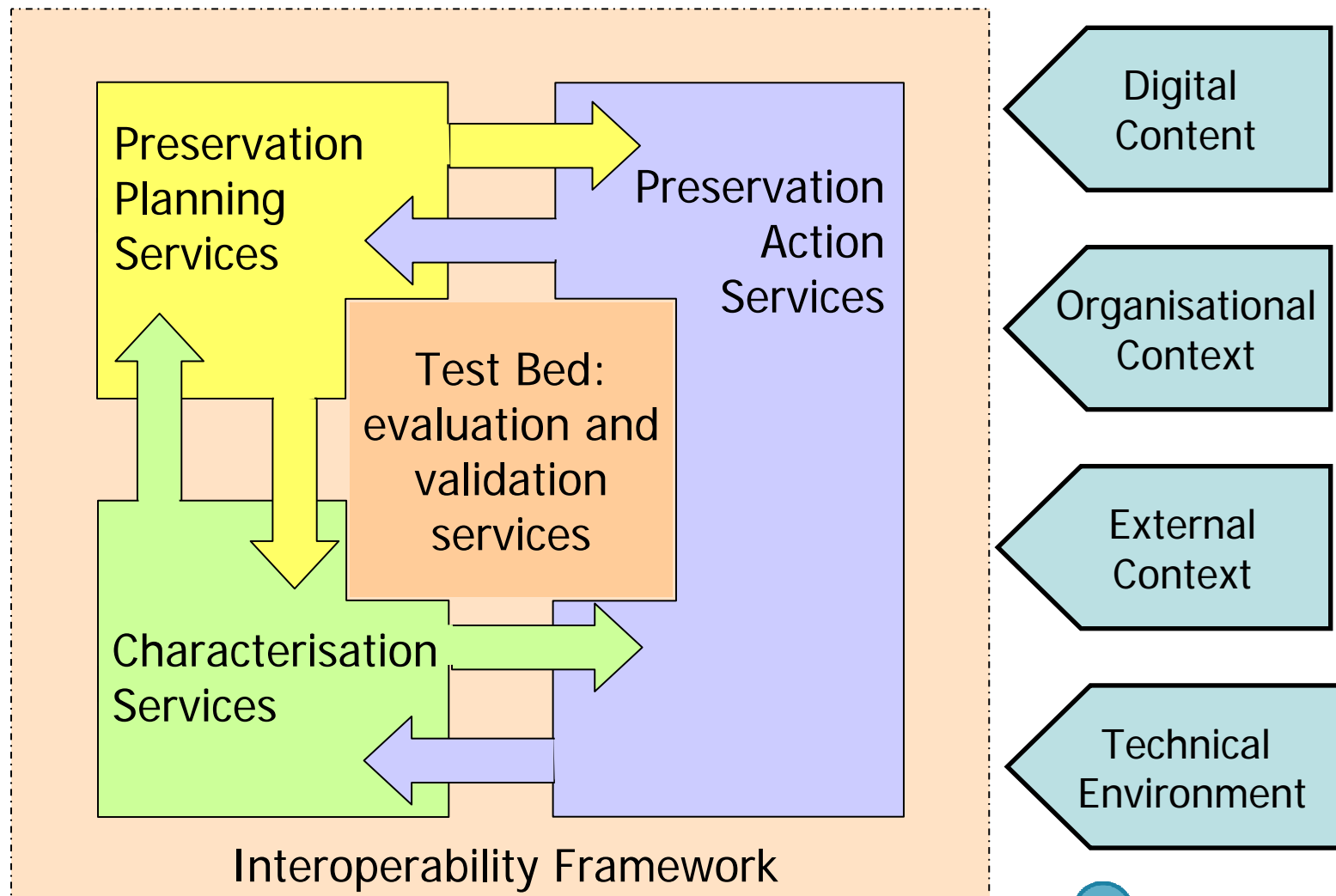


Outline

- ❑ Introduction to Planets
 - Who are we?
 - What are we doing?
 - Why are we doing it?
- ❑ The Planets architecture and components
- ❑ A first glimpse at Planets Preservation Planning

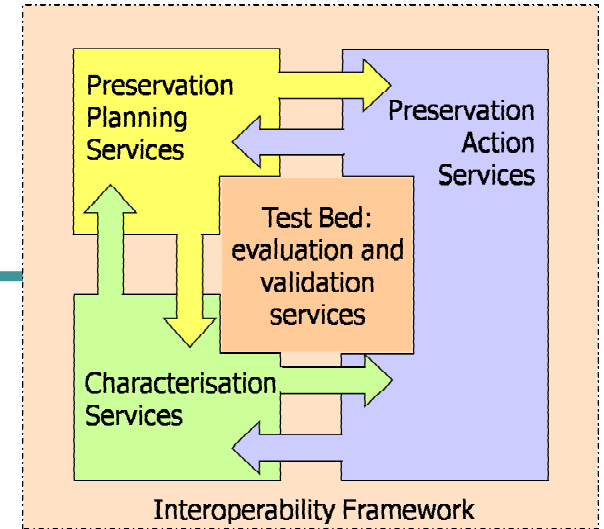


Planets Architecture



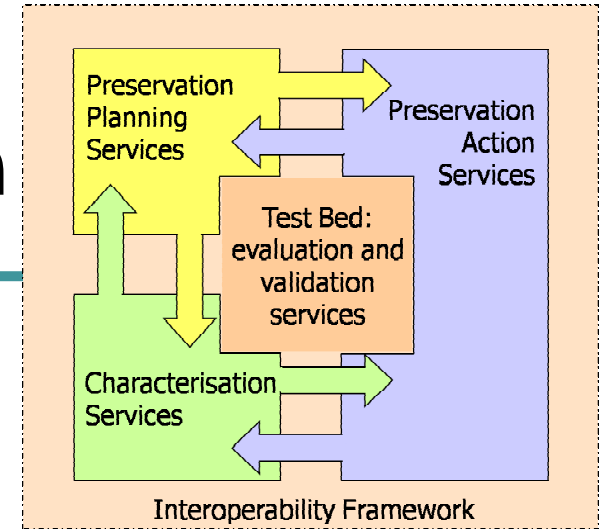
Preservation Action

- ❑ Transform content
 - Pluggable infrastructure for third-party migration tools
- ❑ Transform environment
 - Dioscuri:
Modular emulation of the full hardware/software environment
 - Universal Virtual Computer (UVC):
provides a layered durable approach to emulation
- ❑ Preservation Action Tools registry
- ❑ XML language for describing preservation action tools



Preservation Characterisation

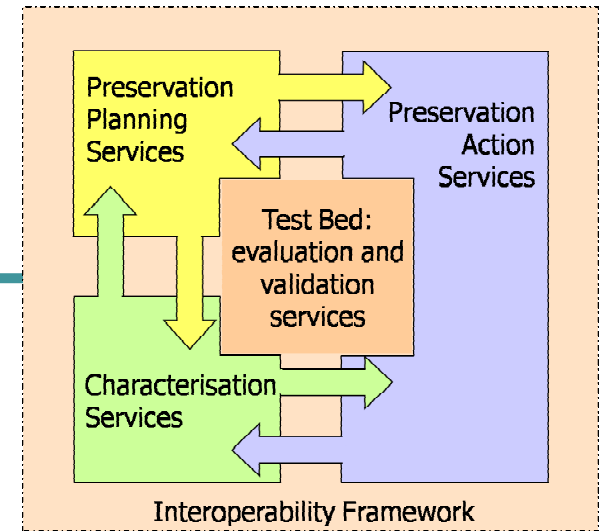
- ❑ Characterisation framework
 - Unifies tools for identifying file formats and extracting object properties
- ❑ Characterisation registry
 - Based on the file format registry PRONOM
- ❑ eXtensible Characterisation Languages (XCL)
 - Family of XML languages for characterising digital objects
- ❑ Comparator verifies effects of preservation actions



Infrastructure and Testbed

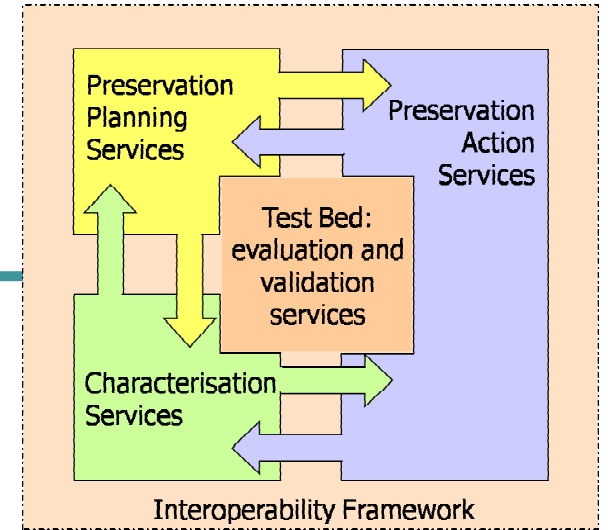
- ❑ Interoperability Framework provides common basis
 - JBoss Application Server
 - Logging, Security Services
 - Registry services
 - User management and Single-Sign-On

- ❑ Planets Testbed
 - Controlled environment for the execution of experiments
 - Accumulated experience base collected in registry



Preservation planning

- ❑ Collection profiling services
- ❑ Technology watch services
- ❑ Risk assessment of digital objects
- ❑ Preservation planning methodology
- ❑ Tool support: Plato, the Planning Tool



Summary

- ❑ Planets methods, tools, and services help organisations diagnose and treat problems with their digital objects
- ❑ High levels of automation and scalable components reduce costs and improve quality
- ❑ Empirical data enables improved decision making
- ❑ Find out more: <http://www.planets-project.eu>



Outline

- ❑ Introduction to Planets
 - Who are we?
 - What are we doing?
 - Why are we doing it?
- ❑ The Planets architecture and components
- ❑ A first glimpse at Planets Preservation Planning



Preservation Planning

Why Preservation Planning?

- Several preservation strategies developed
 - For each strategy: several tools available
 - For each tool: several parameter settings available
- How do you know which one is most suitable?
- What are the needs of your users? Now? In the future?
- Which aspects of an object do you want to preserve?
- What are the requirements?
- How to prove in 10, 20, 50, 100 years, that the decision was correct / acceptable at the time it was made?



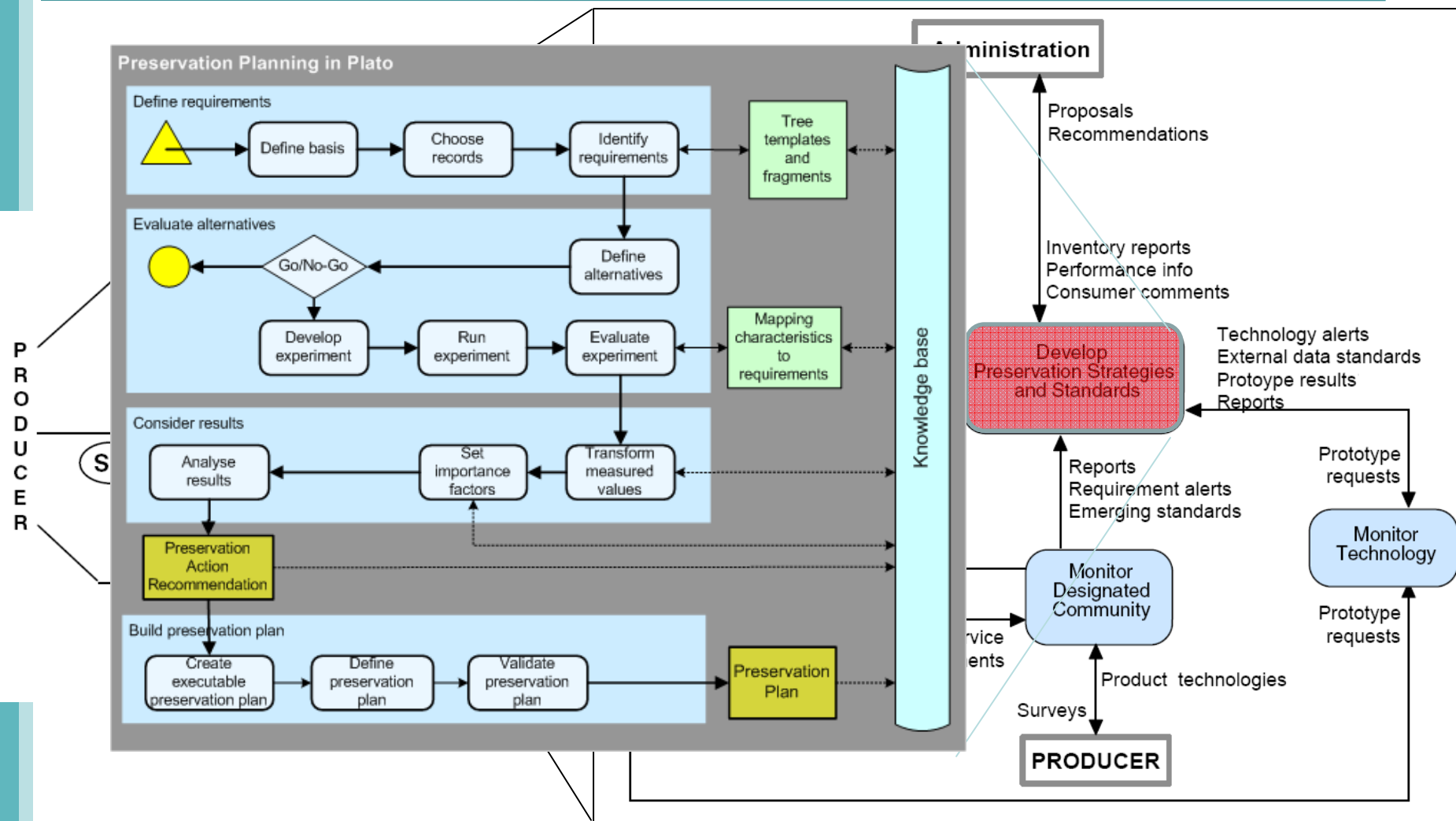
Preservation Planning

Preservation Planning Workflow

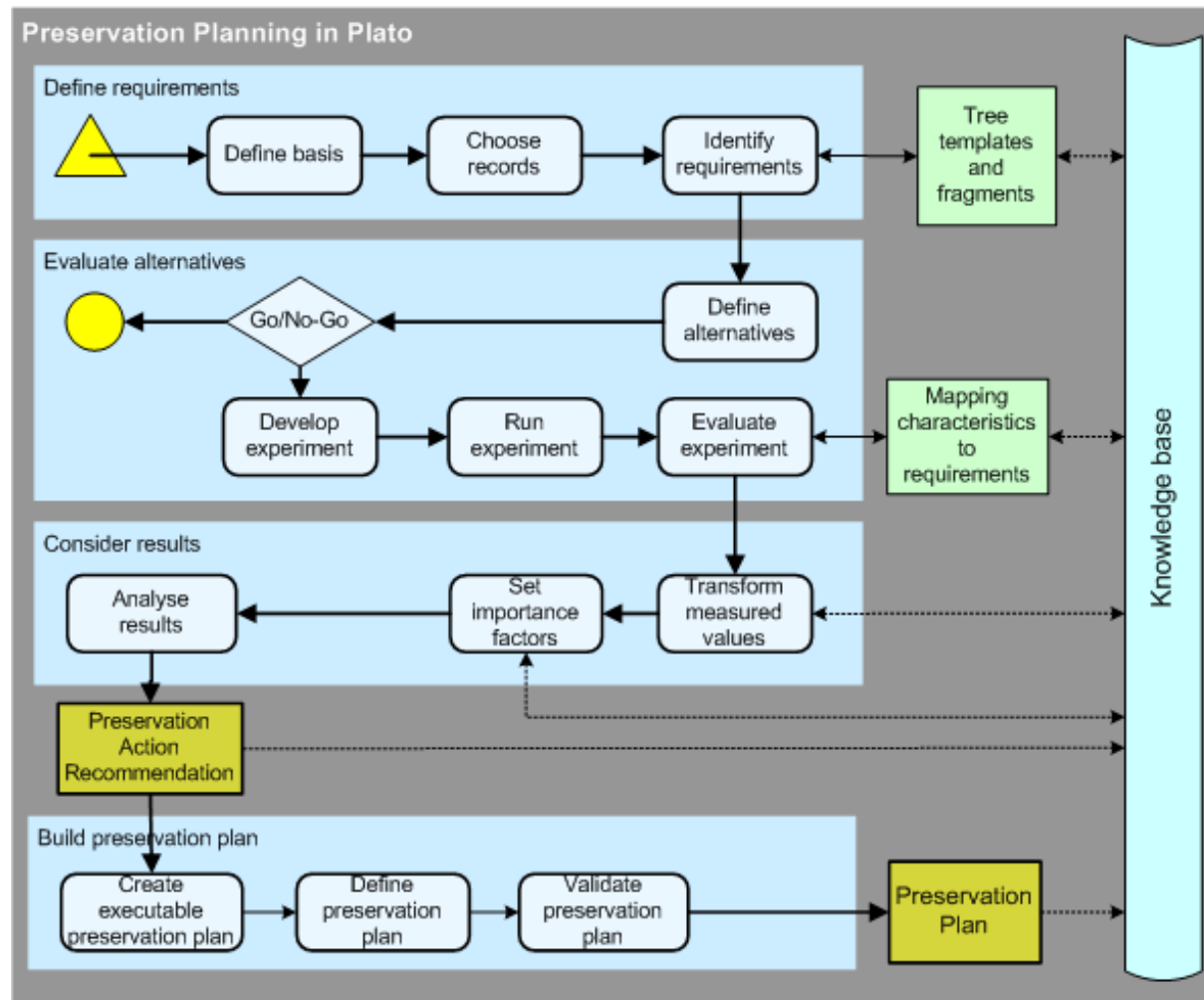
- Originally developed within the DELOS DP Cluster now refined and integrated within PLANETS
- Based on Utility Analysis
- Follows the OAIS model
- Consistent with requirements specified by OCLC/TRAC and Nestor criteria catalogue



Preservation Planning



Preservation Planning Workflow



Plato



PLANETS Preservation Planning Tool (*PLATO*)

PlanningTool > Home

Home
Load Project
New Project

Define Requirements

Define Basis
Define Sample Records
Identify Requirements

Evaluate Requirements

Define Alternatives
Go/No-Go
Develop Experiment
Run Experiment
Evaluate Experiment

Consider Results

Transform Measured
Set Importance Factors
Analyse Results
Sum
Multiplication
Sum of Priority
Austin Slight



PLANETS Preservation Planning Tool (*Plato*)

Project Define Requirements Evaluate Requirements Consider Results

Analyse Results

Aggregation method: Sum of Advantages

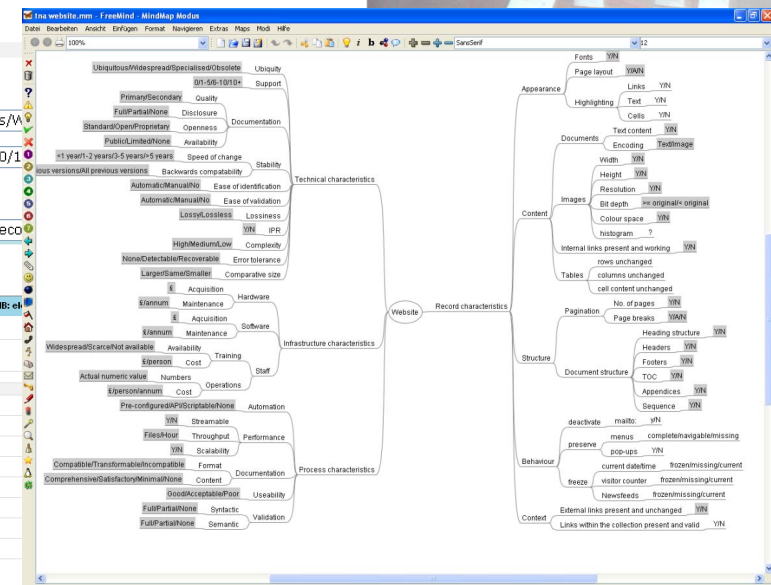
Select	Alternative
<input checked="" type="checkbox"/>	PDF-A
<input checked="" type="checkbox"/>	PDF-unchanged
<input checked="" type="checkbox"/>	TIFF
<input checked="" type="checkbox"/>	EPS
<input checked="" type="checkbox"/>	JPEG2000
<input checked="" type="checkbox"/>	RTF-acrobat
<input checked="" type="checkbox"/>	RTF-convertdoc
<input checked="" type="checkbox"/>	TXT

Expand All | Collapse All

ONB Master thesis

Focus	Name	Result
ONB Master thesis	PDF-A:	4,64
	PDF-unchanged:	5,61
	TIFF:	3,77
	EPS:	3,57
	JPEG2000:	3,53
	RTF-acrobat:	1,53
	RTF-convertdoc:	1,07
Object characteristic	TXT:	1,75
	PDF-A:	2,99
	PDF-unchanged:	2,88
	TIFF:	2,72
	EPS:	2,71
	JPEG2000:	2,67
	RTF-acrobat:	0,73
	RTF-convertdoc:	0,97
	TXT:	0,43
	PDF-A:	2,31
	PDF-unchanged:	2,31
	TIFF:	2,31
	EPS:	2,31
	JPEG2000:	2,31

Release 1.2 - Institute of Software Technology and Interactive Systems: « off-ice bears »



Quick Access: [Help]

© 2007 Institute of Software Technology and Interactive Systems



Preservation planning

- ❑ 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
 - ❑ Evaluation of strategies on representative sample content according to specific requirements



Thank you very much for your attention
and
Enjoy the Workshop!

www.planets-project.eu

<http://www.ifs.tuwien.ac.at/dp>

