

CeBIT 2008, March 4-9, Hannover, Germany



Preservation and Long-Term Access through Networked Services

Christoph Becker

Vienna University of Technology

www.ifs.tuwien.ac.at/~becker

Agenda

- Introduction to Planets
 - Who are we?
 - What are we doing?
 - Why are we doing it?
- The Planets architecture and components
- Preservation Planning in Planets
- Progress and next steps

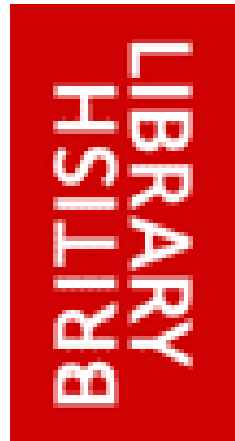


The Planets project

- ❑ 4-year research and technology development project co-funded by the European Union
- ❑ Address 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

STATS BIBLIOTEKET

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

A

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



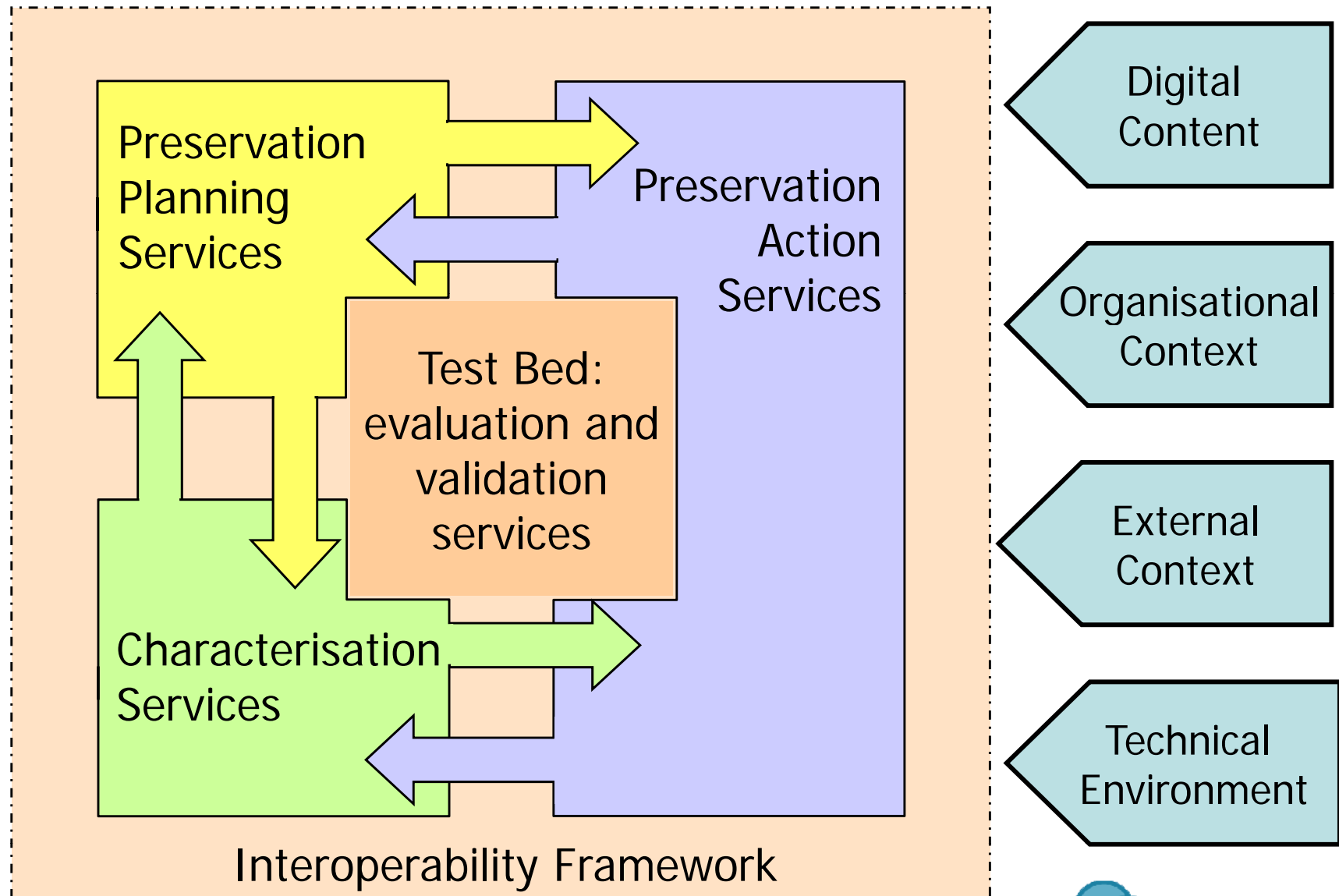
Aims and objectives

- ❑ **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



Planets Architecture



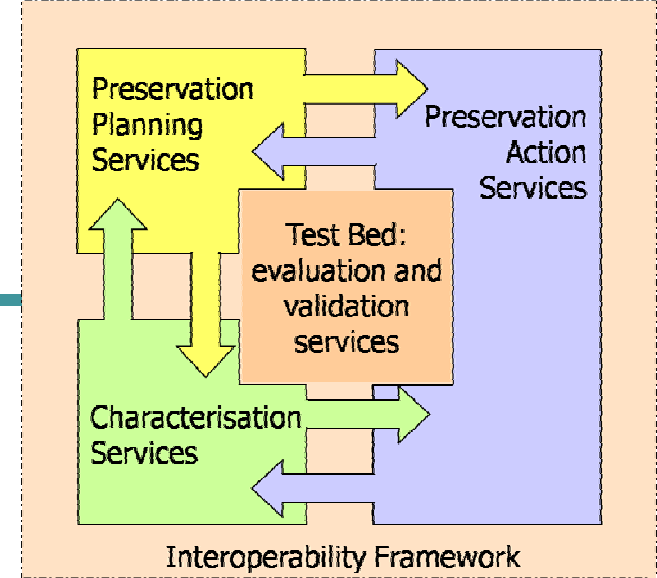
Preservation Action

- Transform content
 - Pluggable infrastructure for third-party migration tools

- Transform environment
 - Dioscuri:
Modular emulation of the full hardware/software environment

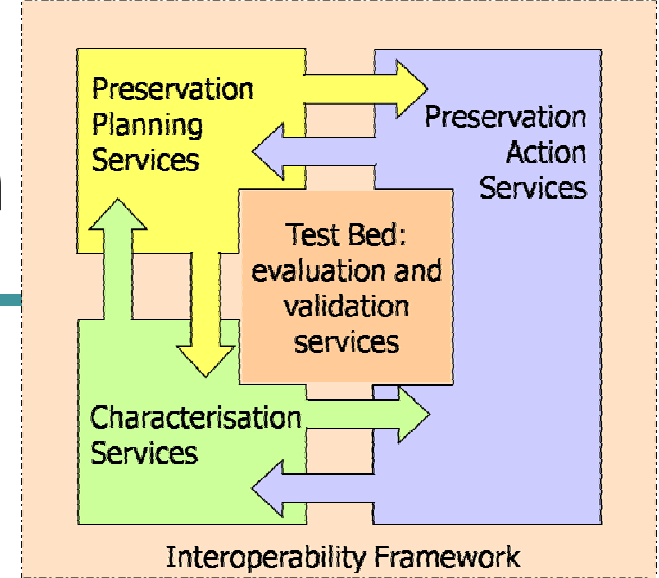
 - Universal Virtual Computer (UVC):
provides a layered durable approach to emulation

- Preservation Action Tools registry
- XML language for describing preservation action tools



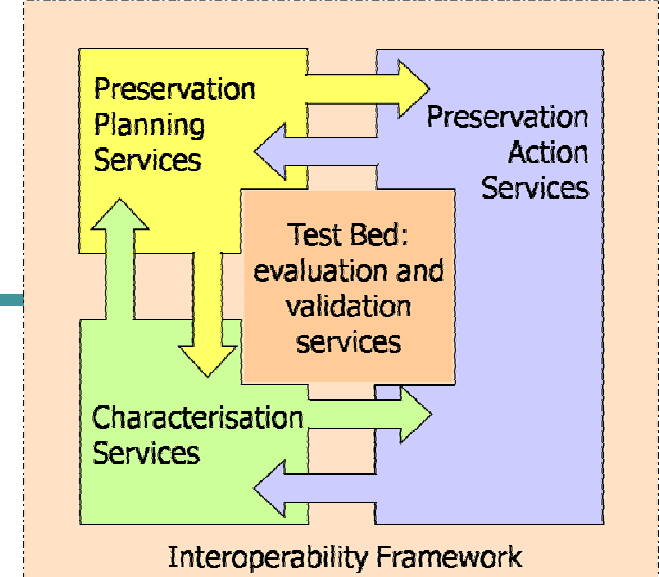
Preservation Characterisation

- 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



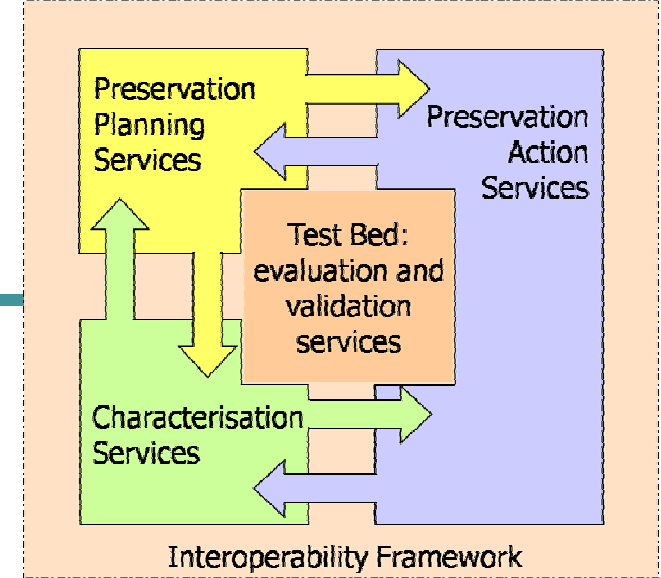
Infrastructure and Testbed

- ❑ 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



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



Decision support for preservation planning

- Systematic procedure for evaluating preservation strategies and building preservation plans
 - By conducting experiments on sample content

- Case studies
 - Electronic documents, interactive art, web archives...
 - Identify essential characteristics of objects and requirements for preservation strategies
 - Evaluate strategies and build plans

- Develop a decision support software
 - Plato – Planning Tool
 - Web application supporting the workflow



Scenario: Changes in user community

- ❑ Repository of electronic publications
 - ❑ Policy: 90% of users can access all published reports
 - ❑ Usage profile: 98% of users can not view dvi files
 - ❑ Content profile: 5% of published reports in dvi format
-
- ❑ Mission: Build and execute a plan for preserving access to these documents for the designated user community



Scenario (2)

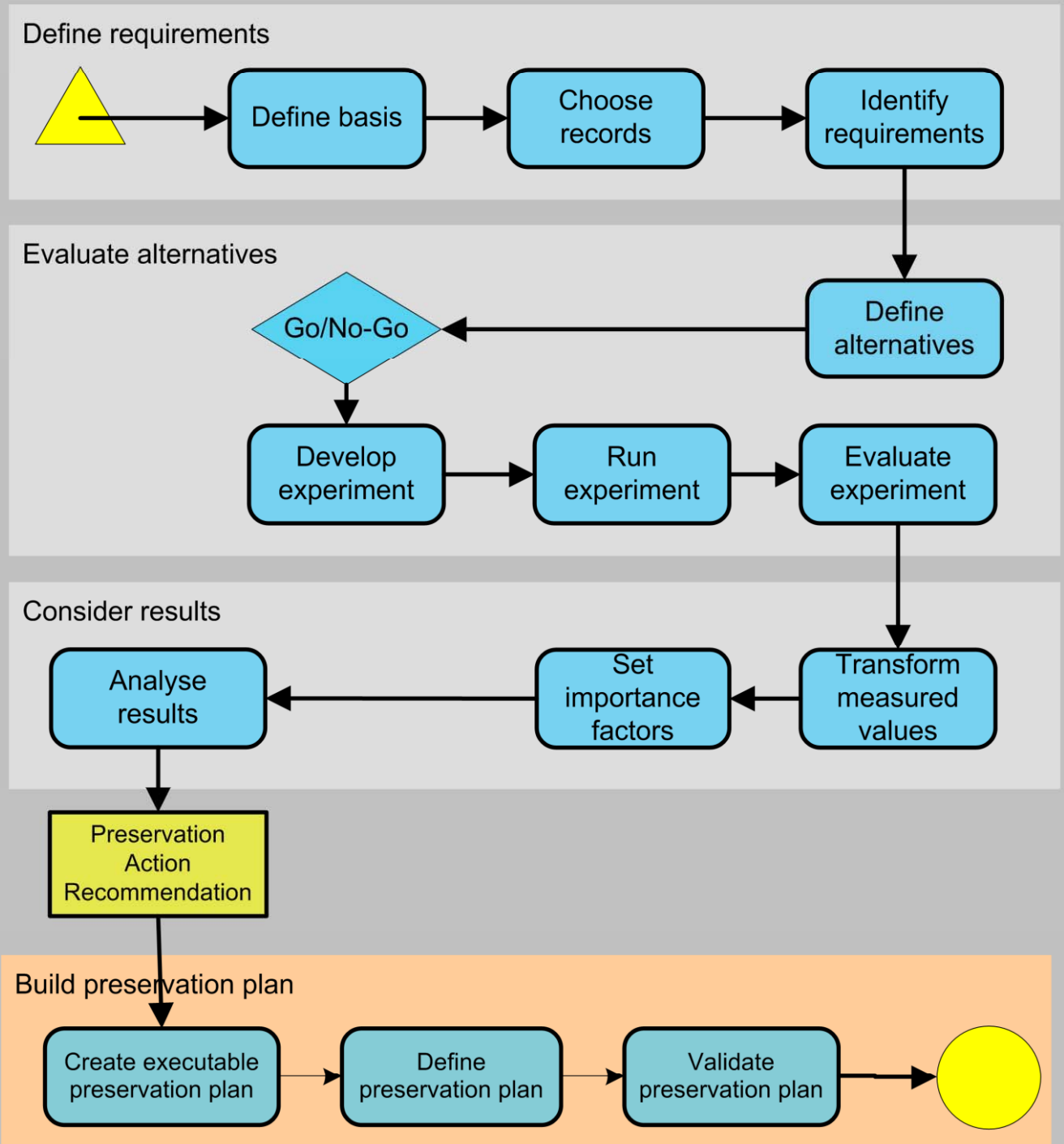
- Create a preservation plan with Plato
 - Define requirements
 - Identify possible actions using registries
 - Convert to PDF with Tool A
 - Convert to TIFF with Tool B
 - Provide users with a viewer plug-in
 - Evaluate actions on sample content
 - Build a preservation plan
 - Convert content (using data registry)
 - QA results (using comparison services)
 - Ingest results into repository (using adaptor)



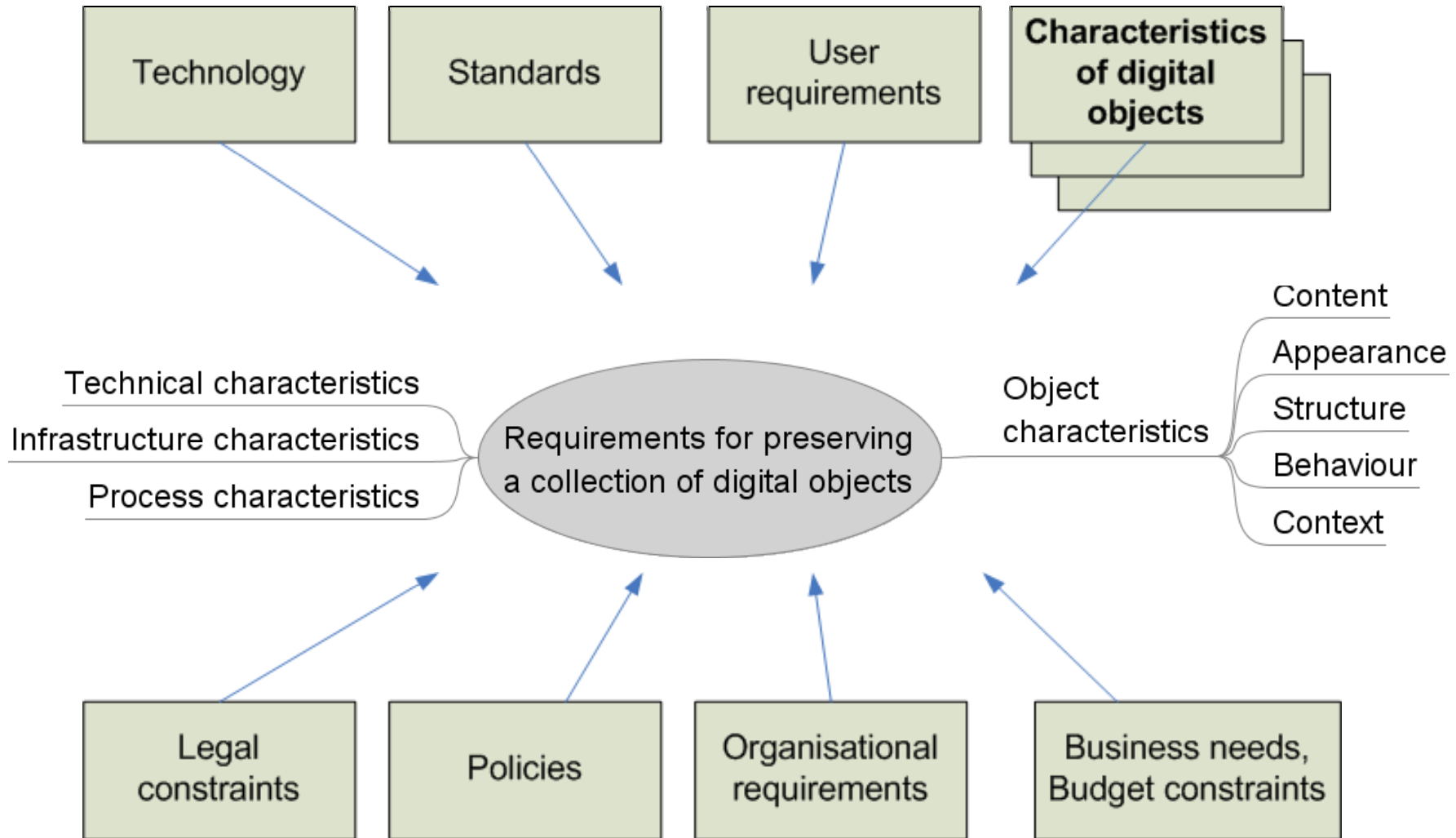
Workflow

1. Define requirements
2. Evaluate potential actions
3. Analyse the results
4. Build a preservation plan

Preservation Planning workflow

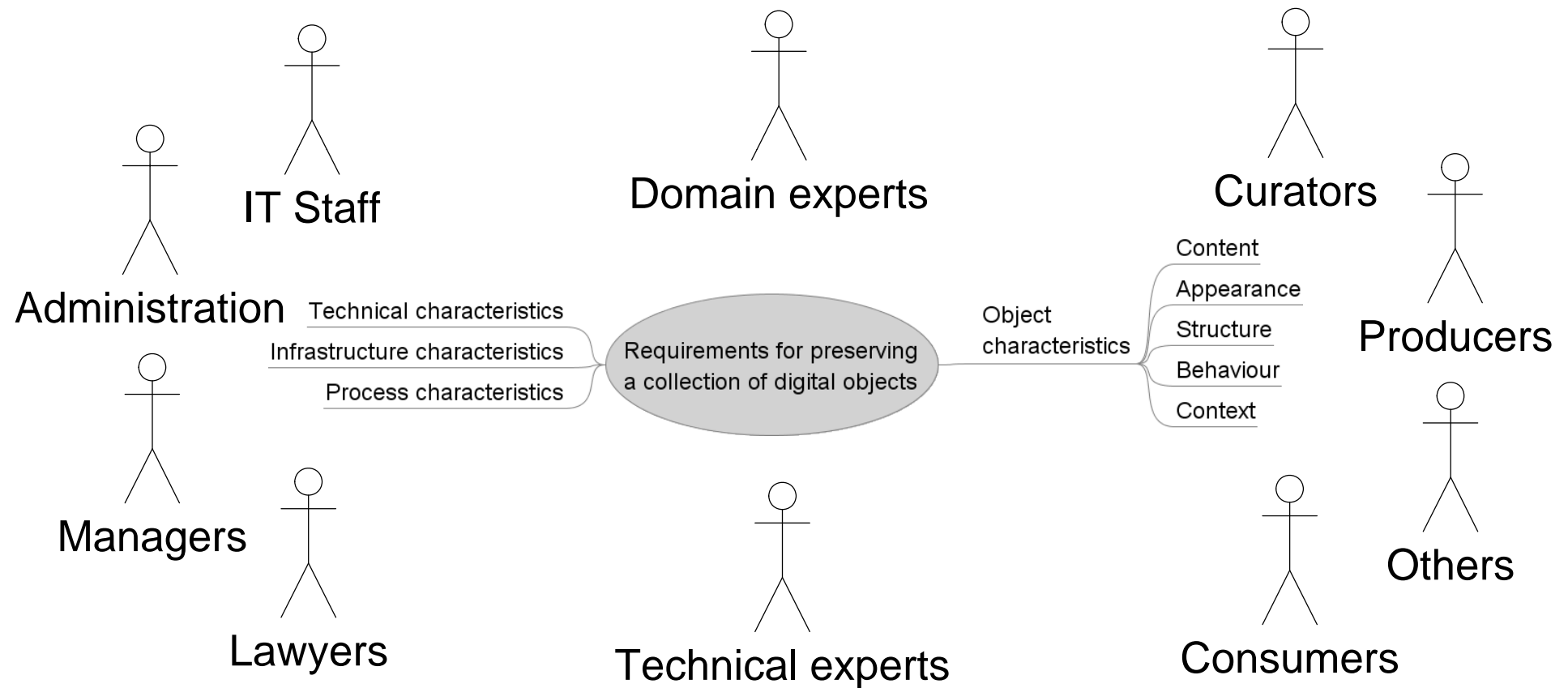


Influence Factors

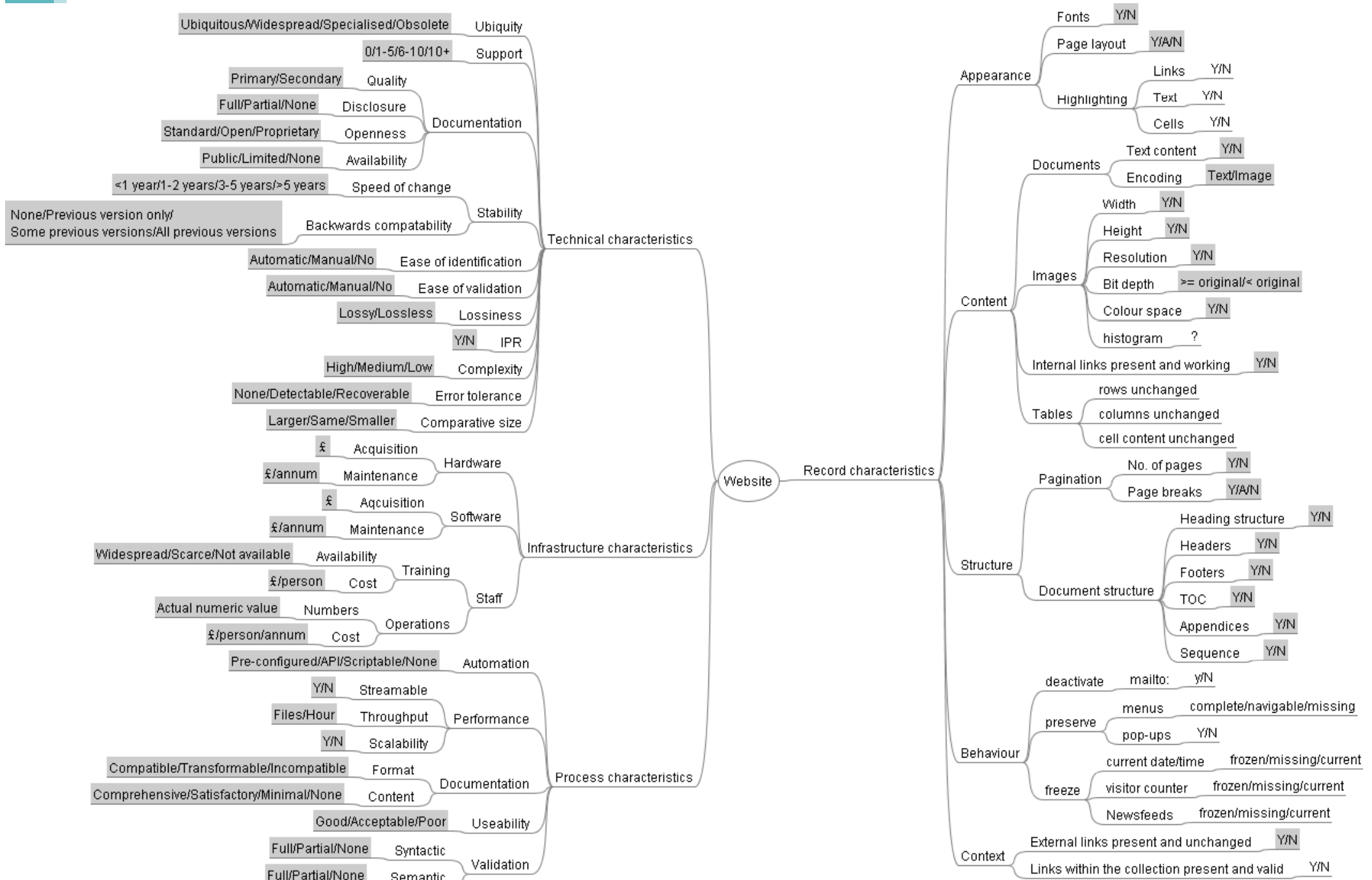


Stakeholders

- Input from a wide range of persons, depending on the institutional context and the collection



The Objective Tree



The Objective Tree



Identify Requirements

[Objective Tree](#)

[Descriptive Information](#)

[How can I define the objective tree?](#)

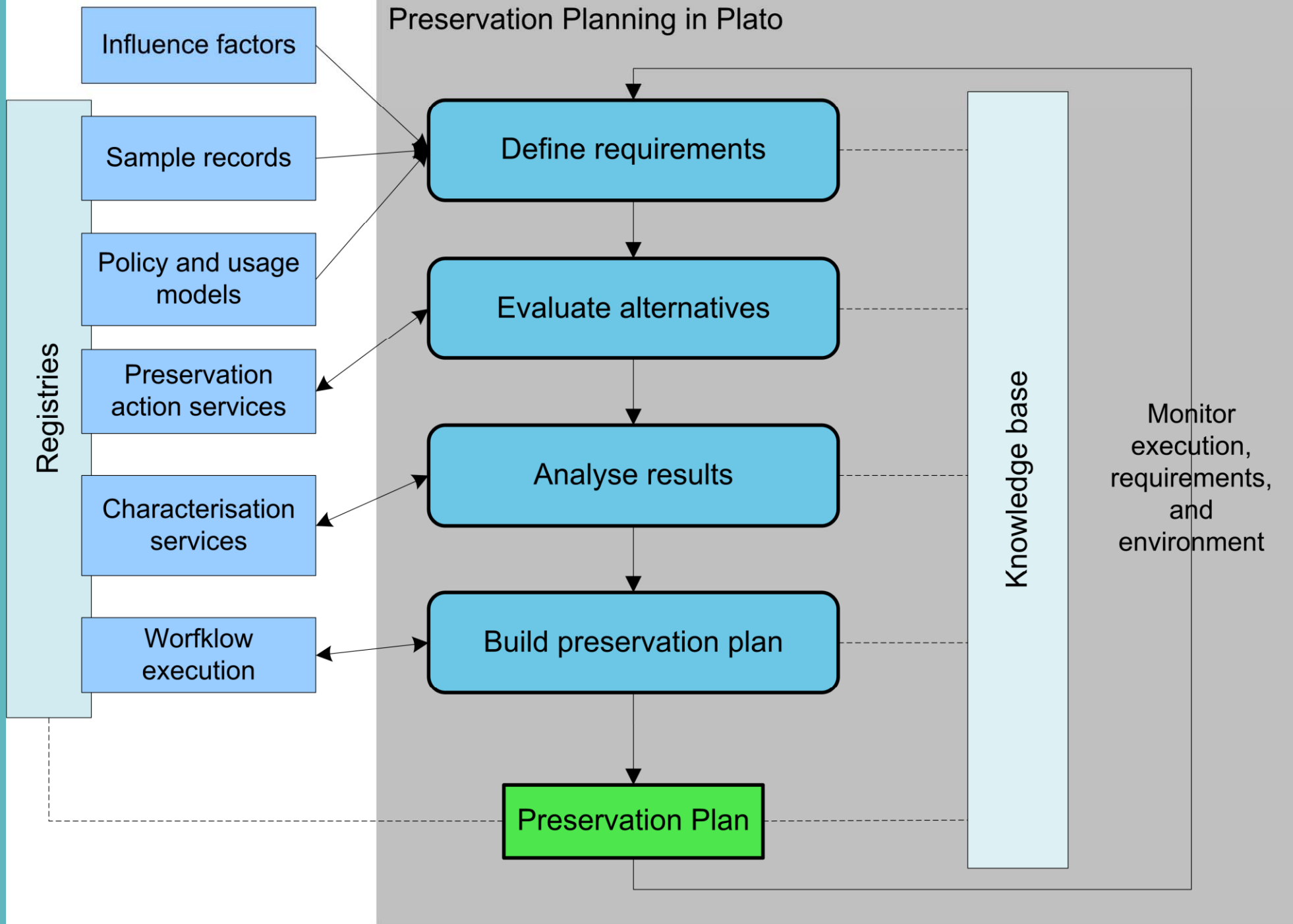
[+] Objective Tree

[Expand All](#) | [Collapse All](#)

[Website](#)

Focus	Node	+	+	-	Single	Scale	Restriction	Unit
	Website	+	+					
X	Record characteristics	+	+					
X	Appearance	+	+					
X	Content	+	+					
X	Structure	+	+					
X	Behaviour	+	+					
X	deactivate	+	+					
X	mailto:				<input type="checkbox"/>	Boolean	Yes/No	
X	preserve	+	+					
X	menus				<input type="checkbox"/>	Ordinal	complete/navigable/missing	
X	pop-ups				<input type="checkbox"/>	Boolean	Yes/No	
X	freeze	+	+					
X	current date/time				<input type="checkbox"/>	Ordinal	frozen/missing/current	
X	visitor counter				<input type="checkbox"/>	Ordinal	frozen/missing/current	
X	Newsfeeds				<input type="checkbox"/>	Ordinal	frozen/missing/current	
X	Context	+	+					
X	External links present and und				<input type="checkbox"/>	Boolean	Yes/No	
X	Links within the collection pres				<input type="checkbox"/>	Boolean	Yes/No	
X	Technical characteristics	+	+					
X	Ubiquity				<input type="checkbox"/>	Ordinal	Ubiquitous/Widespread/Specia	
X	Support				<input type="checkbox"/>	Ordinal	0/1-5/6-10/10+	
X	Documentation	+	+					

Preservation Planning in Plato





Analyse Results

Aggregation method:

Select	Alternative
<input checked="" type="checkbox"/>	PDF/A ToolA
<input checked="" type="checkbox"/>	PDF/A ToolB

Show

[Expand All](#) | [Collapse All](#)

Minimalist root node

Focus	Name	Result
	▼ Minimalist root node	PDF/A ToolA: 2,98 PDF/A ToolB: 3,19
X	▶ Image properties	PDF/A ToolA: 0,70 PDF/A ToolB: 0,80
X	▼ Karma	PDF/A ToolA: 0,40 PDF/A ToolB: 0,00
X	▼ Filesize (in Relation to Original)	PDF/A ToolA: 0,78 PDF/A ToolB: 0,99
X	▼ A Single-Leaf	PDF/A ToolA: 0,40 PDF/A ToolB: 0,80
X	▼ IntRange 0-10	PDF/A ToolA: 0,70 PDF/A ToolB: 0,60

Recommendation

Recommendation:

Reasoning:

Save recommendation

Generate final report

What's next in Planets? (1/2)

- ❑ Preservation Planning tool (Plato)
 - Version 1.1 released just now
 - Integration of registries and services
 - 2nd version in the second half of 2008
- ❑ Integrated preservation planning services
 - risk assessment
 - automated collection profiling
 - technology watch
- ❑ A description language for preservation action tools
- ❑ Planets-compliant migration tools for digital objects
- ❑ Emulation tools for specific environments



What's next in Planets? (2/2)

- Characterisation tools
 - extract significant properties from digital objects
 - compare different objects
- A characterisation description and extraction language (XCL)
- Characterisation and preservation action registries
- A controlled environment for the empirical assessment of services (Planets Testbed)
- Planets Interoperability Framework
 - as downloadable “click-and-install” software package



Dissemination and Take-up programme

- ❑ Workshops and training events
- ❑ Course in Vilnius, Lithuania, October 2007
- ❑ Scientific publications
- ❑ Newsletter and web page

Planning the future with Planets: A Preservation Planning tutorial

- Austrian Computer Society, Vienna, April 14-15
- <http://www.planets-project.eu/events/pp-vienna-2008/>



Conclusion

- ❑ 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>



Thank you very much for your attention.

www.planets-project.eu

becker@ifs.tuwien.ac.at
www.ifs.tuwien.ac.at/~becker

