

(1) Introduction

Who are we?

(2) *Digital Preservation and Preservation Planning:*

What is a Preservation Plan and why do we need it?

(3) *Preservation Planning:*

How do we build a Preservation Plan? How does Plato help?

(4) *Exercise (& Coffee Break):*

Which objectives should we meet for preserving scanned images?

(5) Decision Criteria in Digital Preservation:

How do we specify and measure them? What are good criteria?

(6) Lessons learned:

What are common misconceptions?

1. Is all this really necessary?
2. What are the costs and benefits of planning?
3. What are the prerequisites of planning?
4. Who is supposed to do planning?
5. What is the scope of one plan?

Is all this really necessary?

- Challenges when evaluating preservation actions
 - Quality varies across tools
 - Properties vary across content
 - Usage varies across communities
 - Requirements vary across scenarios
 - Risk tolerance varies across collections
 - Preferences and constraints vary across organisations
 - Cost structures and compatibility varies across environments
 - Constraints, priorities and requirements shift constantly
 - Evaluation is complex
- Trust requires evidence
 - Trust has to be evaluated in a realistic context
 - Controlled experimentation, repeatable documentation, and scenario-specific requirements assessment

- Primary cost factors
 - Maturity of organisational framework:
Constraints, goals, drivers and responsibilities
 - Degree of familiarity with the planning approach
 - Technical complexity of the content to be preserved
 - Technical proficiency of the staff assigned to do planning
- Learning curve
 - First intent generally effort-intensive
 - Subsequent activities significantly easier and faster
- Return on Investment
 - Hard to quantify
 - ... but shouldn't we rather ask: What are the costs of NOT planning?
 - This is quite easy to quantify

What are the prerequisites of planning?

- A clear and concise documentation of the organisation itself
 - Constraints
 - Drivers
 - Goals
 - Responsibilities
 - Infrastructure and technical capabilities
 - Cost structures
- Context must be known and explicitly defined
- This is a Critical Success Factor

Who is supposed to do planning?

- A full understanding of the planning *role* has yet to be formed
- Combination of expertise and skills required
 - Understanding of business goals to achieve
 - In-depth knowledge of technical intricacies
 - Not all planning activities need and should be carried out by the same person or role in an organisation
- Preservation Planning needs to take place on an operational level
 - This should include an escalation path

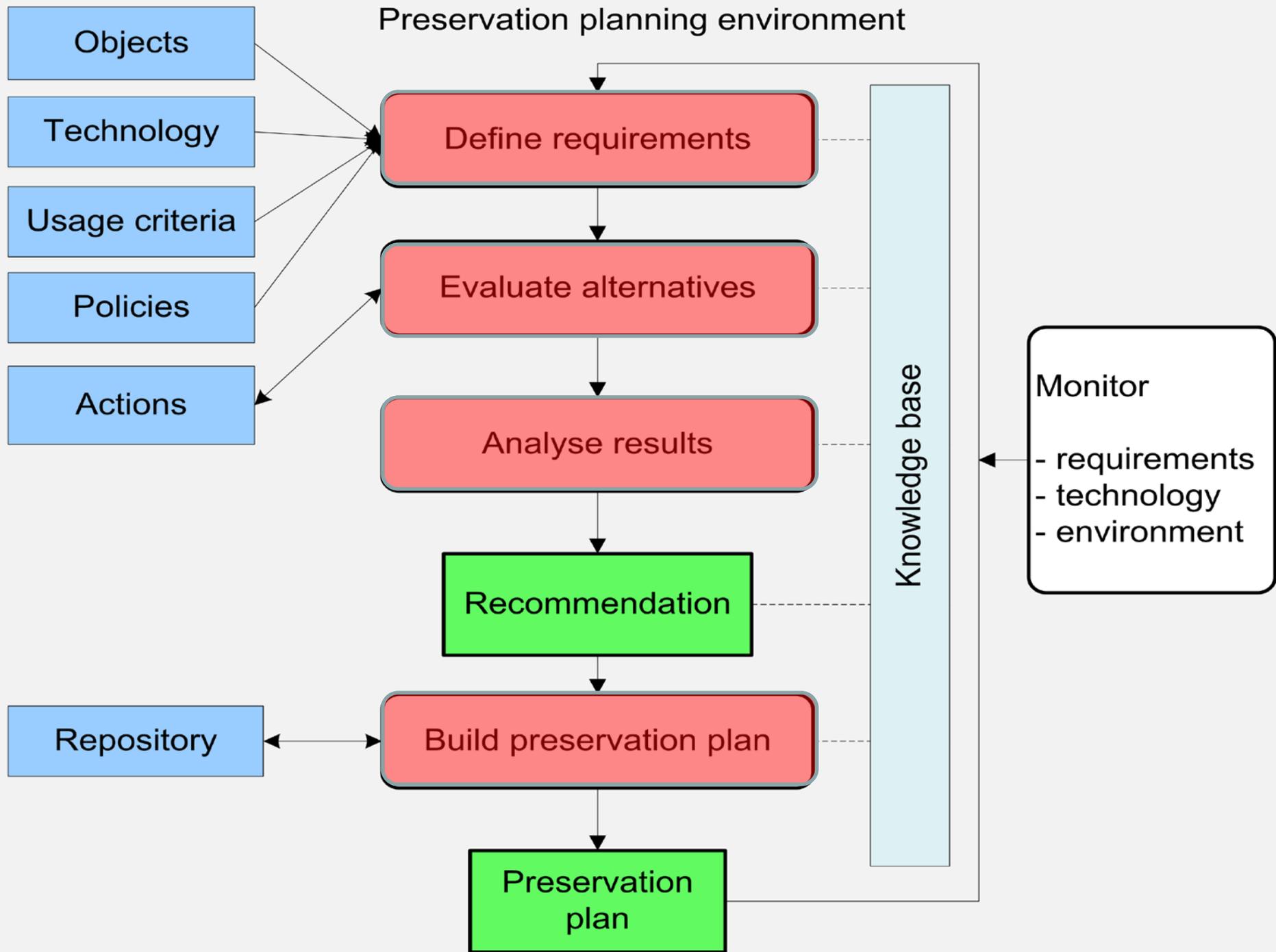
What is the scope of *one* plan?

- ‘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).’ [IJDL 2009]
- 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.
- See www.ifs.tuwien.ac.at/dp/plato for full definition

What is the scope of *one* plan?

- A plan specifies the treatment for one *collection*
- A *collection* is the maximum set of objects
 - to which the same requirements set applies and
 - which can be covered with one operational preservation action
- For example...
 - the set of images that can be normalized using a certain conversion workflow (which includes Quality Assurance!)
 - the set of applications that may be deployed in a certain emulation platform with good results
- During planning, assumptions may turn out false
 - Collection treatment may need to be split

Preservation planning environment



1. Requirements definition
2. Measurement specification
3. Measurement and assessment
4. Weighting requirements
5. The method, the tools, and the services

What is a requirement?

- Requirements definition is the most critical part
 - Ill-defined or incomplete requirements may lead to wrong actions
- Common mistakes...
 - Too abstract scales
 - Too subjective scales
 - Insufficient semantics definition
 - Unclear specification of the evaluation procedure
 - Mix between solution space and problem space

What is being measured?

- Measurement specification needs to be precise
 - What is the difference between these?
 - *Text should be searchable*
 - *Text needs to be fully preserved*
 - *Text font? Text encoding? Text characters? White space?*
 - Scale needs to reflect the objective in question
 - ❑ Image width measured in pixels: positive number?
 - ✓ Image width, measured in pixel, is unchanged: yes/no
- It's easy to miss something...
 - Falsify criteria sets by imagining things that could go wrong

What is 'acceptable'?

- Measurement and Assessment should provide clear distinction between objective facts and subjective assessment
 - Fonts should be preserved: Yes/Acceptable/No ?
 - ✓ Fonts should be preserved: Ordinal
 - Identical
 - Replacement with font family
 - Replacement with standard font
 - Loss of font information
 - ...
- Later changes in the environment or the organization may require a reassessment of facts
 - Only possible if facts are separated from their assessment

What is important?

- Weighting requirements
 - Assigns relative importance factors on all level of the tree
 - Low level changes in relative importance have little influence
 - Criteria often have a total weight of 1-5%
- Weighting vs. utility function
 - Key effects of criteria with low weight: Acceptance or rejection
 - Output range of utility function may include 0.0
 - Utility function is much more critical on the level of criteria

- Method is very generally applicable
 - From computer games to scanned images
 - From databases to born-digital art
 - From private photographs to national heritage institutions
 - Entirely independent of the kind of action (migration, emulation, virtualization.... Even bitstream planning)
- Tool support varies
 - Degree of automation strongly dependent on content and preservation actions
 - But: Manual evaluation is always possible as fallback!
- Integrated services
 - Action services may or may not work on specific content
 - Failure of a service simply means that the service is not suitable
 - Planning and thorough evaluation is important

- Alignment of technology and business
 - Continuum between business and technology
 - User requirements vs. IT operations
 - Technology obsolescence vs. technological opportunities
 - Conflicts between ends and means
 - Conflicts between strategy and tactics
- Organizational capabilities and processes
 - Enterprise Architecture
- Relationships of DP processes and capabilities....
 - IT Governance
 - Governance, Risk and Compliance

A DP decision space

Strategy

Tactic

Business

Technology



