

# **Appraisal, selection and preservation planning: concepts and issues in a digital environment**

**Hans Hofman**

Nationaal Archief Netherlands

**We Preserve Conference, Nice,**

29 October 2008

# Overview

- Appraisal and selection
  - what to keep? Risk analysis
  - archives and libraries
  - issues
- Preservation Planning
  - how to maintain digital material?
  - evaluate strategies and taking decisions
- Summary

This work is licenced under the Creative Commons Attribution-Non-Commercial-No Derivative Works 3.0 Netherlands License. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/3.0/nl/> or send a letter to Creative Commons, 171 Second Street, Suite 300, San Francisco, California 94105, USA.

# What to keep? What to create?

- Archives
  - long standing experience in appraising government records (documenting government activities)
  - shift in focus from appraising records to activities ('functional' appraisal): Australia, Netherlands, Canada
    - different flavours....
    - not exclusive
  - complementary approach: sampling (at random)
  - in digital era shift from what to destroy to what to keep
    - given the fragility of the digital material
    - (pro) active attitude required (starting with design/creation)
- Libraries
  - legal or voluntary deposit
  - web-archiving: try to keep it all (?) within geographical boundaries

- Appraisal is always biased and time-bound
  - not a one time but ongoing activity and related to all processes of managing information
  - criteria will change over time
  - risk analysis (context-sensitive)
- Why not keep it all?
  - Why would you? What is the value? We don't need to
  - Privacy issues
  - Maintenance/ costs (finding the right balance between effort/cost and result)
  - Semantic issues ('needle in a haystack')
- Appraisal at different layers
  - Starting from business activities to records to components/characteristics of the object

## Appraisal as part of all preservation activities

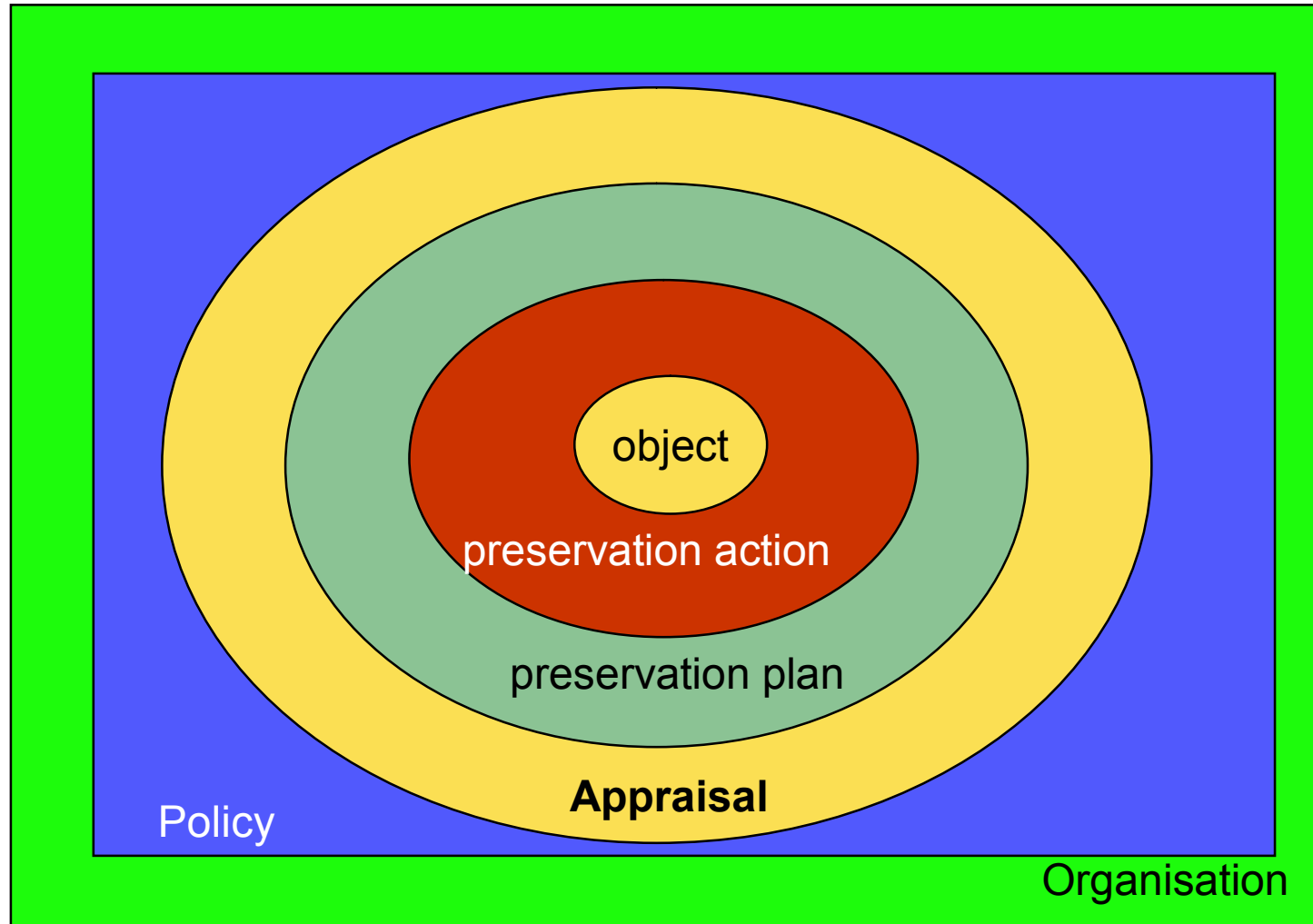
- appraising what to ingest
- selecting what to keep and why?
  - essential characteristics
  - during the whole existence of digital objects
- appraising metadata
- .....

- Stage 1: Understanding the organisational context
  - mandate/ legislation
  - user community
  - the organisational policy
  - method: *risk analysis*, for instance DRAMBORA interactive tool
- Stage 2: Understanding the business context
  - what information needs to be created and kept and for how long?
- Understanding the objects
  - (collection of) digital objects: characteristics
- Understanding the infrastructure
  - technology (past, present, future), infrastructure
  - people, knowledge, skills
- Available options
  - potential methods/ strategies
- Decision making process: *preservation planning*

# Objectives of Preservation Planning

- Identify and analyse the organisational context
  - including a risk assessment
  - define a framework for preservation / policy
- Support decision-making about digital preservation including
  - Identifying criteria for preservation within that context
  - Defining workflow for evaluating/ defining preservation plans
  - Developing methodologies for assessing the risks of applying different preservation strategies for different types of digital objects
- Enable formulation, evaluation and execution of high-quality and cost-effective preservation plans that suit the organisational (e.g. repository) needs
- Support the on-going evaluation of the results of executing preservation plans and provide a feedback mechanism
- Document the planning process carefully



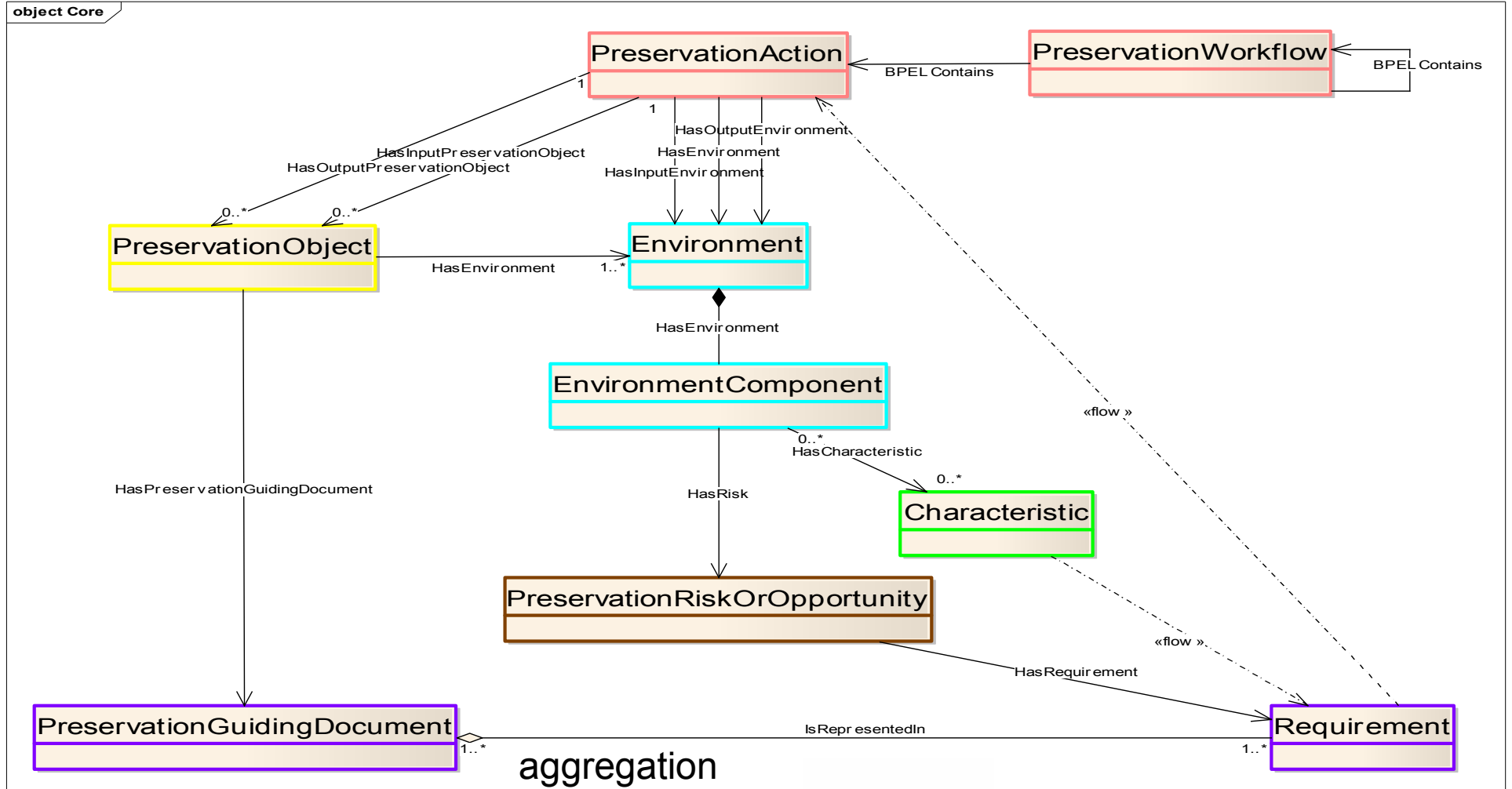


Ownership  
Awareness  
Responsibility

**TOOL: Table – Digital Preservation Policy: AREAS OF COVERAGE**

AREAS	Authority and responsibility
	Conversion and reformatting
	Appraisal, selection and acquisition
	Storage and maintenance
	Access and dissemination
	Implementation
	Standards
	Procedures
	Quality control, auditing and benchmarking
	Cooperation
	Technical infrastructure

# A generic model

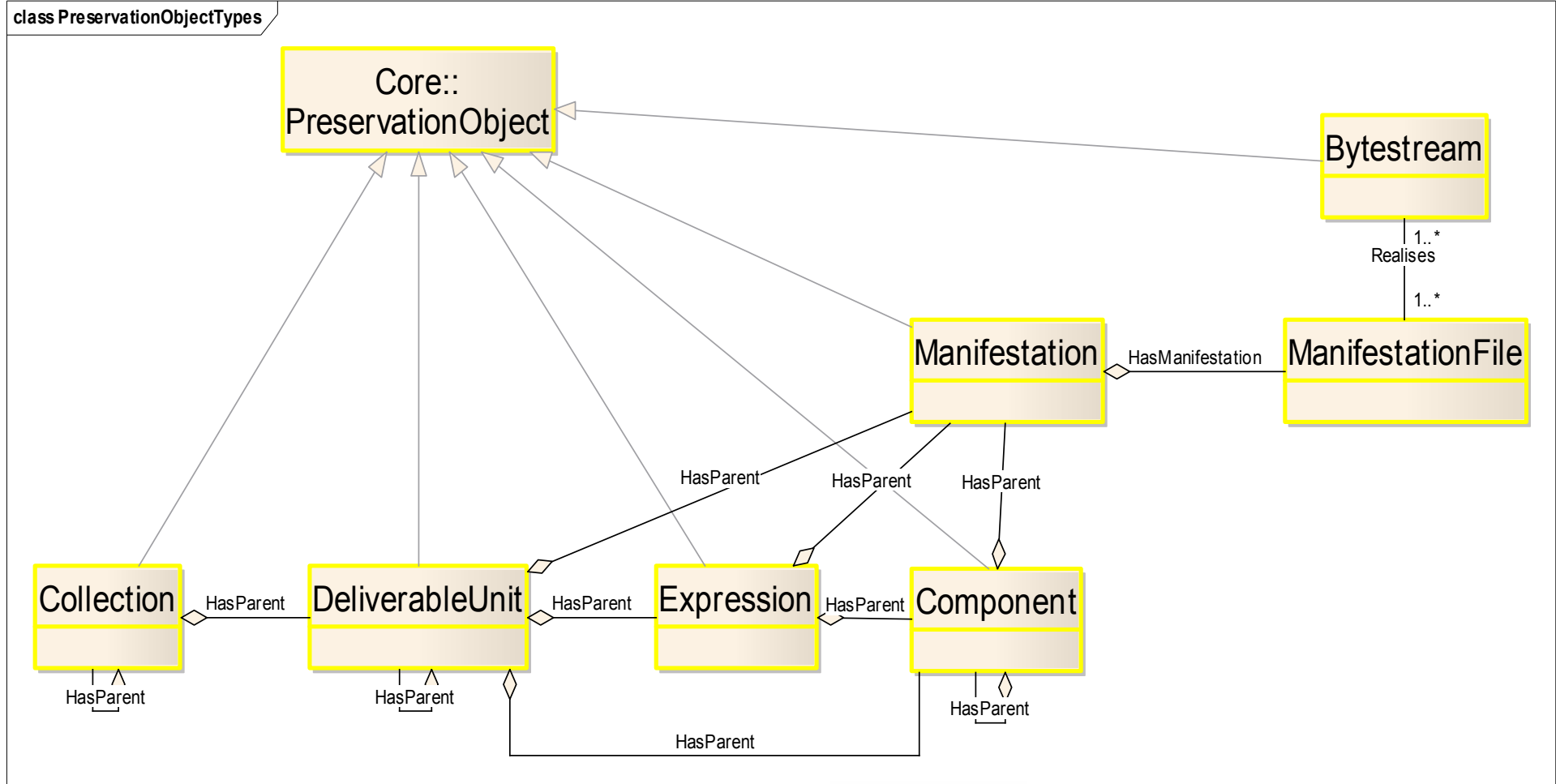


- What types of objects (both technical and intellectual aspects)?
- Technical: file formats
  - registries (e.g. PRONOM, ...)
- Intellectual: for instance documentary form, structure, look and feel, ‘behaviour’
  - objective tree ‘templates’
  - an (intellectual) object may consist of different computer files
    - what strategy then?

# Essential characteristics of 'digital objects'

- What needs to be preserved?
  - content
  - context
  - structure
  - form / appearance
  - (sometimes) behaviour
- What criteria for determining these essential characteristics?

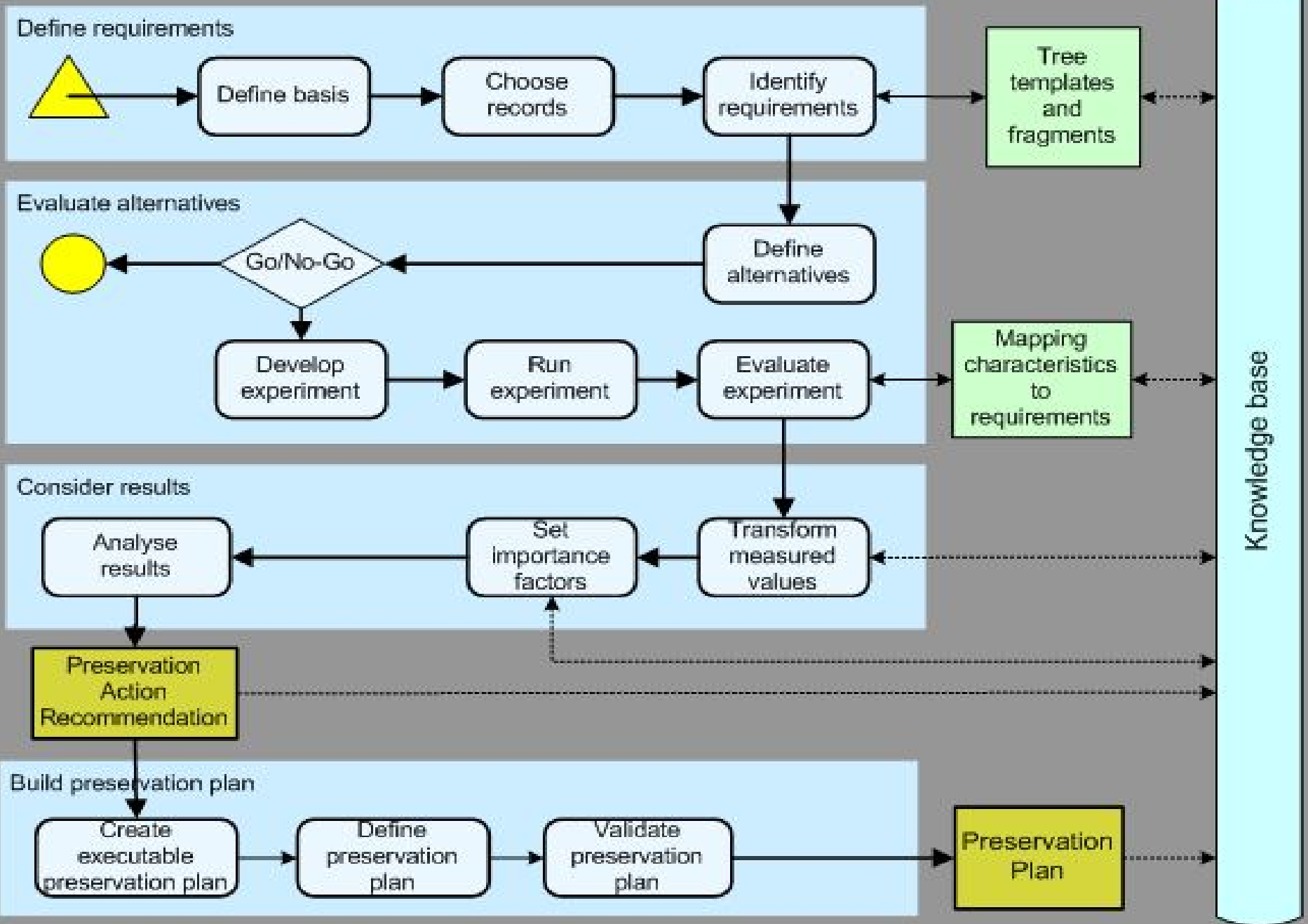
# The digital object



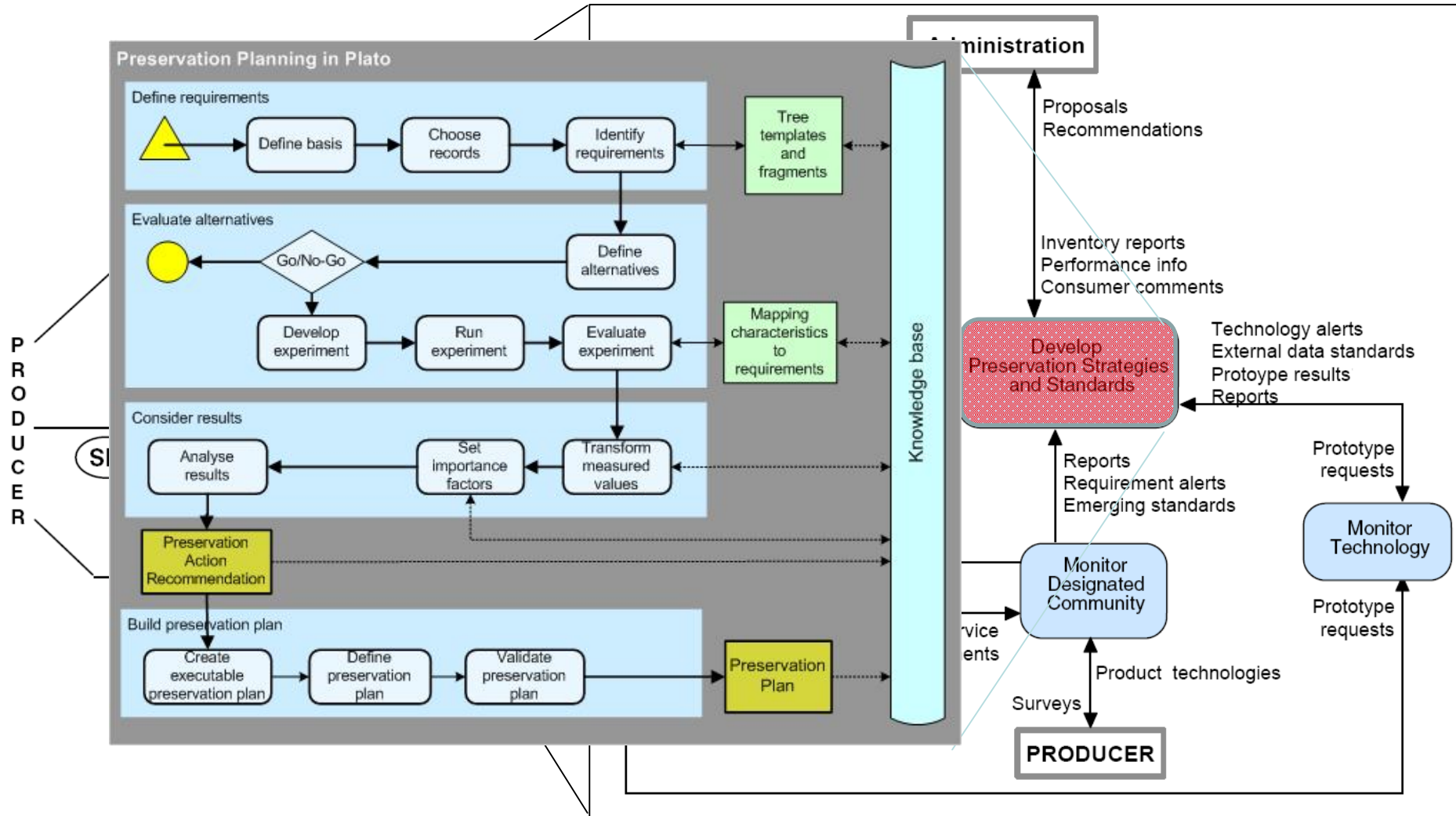
# Requirements for objects

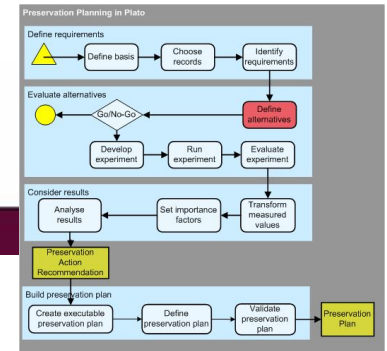
- **Authenticity**
  - to be what it purports to be,
  - to have been created or sent by the person purported to have created or sent it, and
  - to have been created or sent at the time purported
- **Reliability**
  - contents can be trusted as a full and accurate representation of the transactions, activities or facts to which they attest and can be depended upon in the course of subsequent transactions or activities
- **Integrity**
  - being complete and unaltered
- **Usability**
  - can be located, retrieved, presented and interpreted, so retrievable, readable, interpretable
- **Accuracy**
  - the degree to which data, information, documents or records are precise, correct, truthful, free of error or distortion or pertinent to the matter.

# Preservation Planning in Plato









# Service discovery and invocation

Create alternatives from applicable services

Sample record #1 has format JPEG File Interchange Format, 1.01.

You can look up services that are able to handle this object type in the following registries:

**Planets Preservation Action Tool registry**



Show Preservation Services

	Preservation Action	Target Format	Info
<input type="checkbox"/>	JPG > BMP	Windows Bitmap, version 3.0	JPG>BMP
<input checked="" type="checkbox"/>	JPG > TIF	Tagged Image File Format, version 3	JPG>BMP>TIF
<input type="checkbox"/>	JPG > TIF #2	Tagged Image File Format, version 3	JPG>TIF
<input checked="" type="checkbox"/>	JPG > TIF_2	Tagged Image File Format, version 3	JPG>TIF_2
<input type="checkbox"/>	JPG > PNG	Portable Network Graphics, version 1.0	JPG>PNG
<input type="checkbox"/>	JPG > JP2	JPEG 2000	JPG>JP2

Create alternatives for selected services

**Planets Service Registry**



Show Preservation Services

**CRiB Service Registry**



Show Preservation Services



# Summary

- Understanding of context
  - analysis of organisational needs, user needs, legal requirements
- Identify criteria for what to keep and thus preservation
  - how long, restrictions of formats, standards, ...
  - *Risk analysis !*
- Determine **what to keep/maintain**
  - high level: what kind of information resources, what do we want to document for next generations
  - essential characteristics (objective trees), characterisation of computer files
- Next step: how to keep it:
  - what kind of infrastructure and what strategies?
- Evaluate available strategies (actions) against criteria
  - identify best strategy
  - well-founded and documented decision
  - create/finalise preservation plan
- Execute plan when needed
- Evaluate what happened/performance
- Re-iterate when technology changes or review when policy and/or collection and/or usage changes
- Automated decision-making support (?)

Thank you for your attention!

**Questions?**

[www.planets-project.eu](http://www.planets-project.eu)

[hans.hofman@nationaalarchief.nl](mailto:hans.hofman@nationaalarchief.nl)

This work is licenced under the Creative

Commons Attribution-Non-Commercial-No Derivative Works 3.0 Netherlands License.

To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/3.0/nl/>

or send a letter to Creative Commons, 171 Second Street, Suite 300, San Francisco, California 94105, USA.