Planning the Future with Planets

Preservation Planning Process An Overview

Presented by **Hans Hofman**
Nationaal Archief Netherlands
14 April 2008

Overview

- What is the problem? Defining the issues
- The scope and role of Preservation Planning
- The Planets approach
  - The organisational/ business context
  - Usage requirements and collection profiles
  - Essential characteristics of digital objects
  - Preservation plans and preservation actions
- The planning process

Why Preservation Planning?

- Who are the stakeholders?
- What is the issue? Why important?
- What are the objectives?

Stakeholders

- Memory institutions (‘content holders’)
- (Scientific) data centres
- Government organisations (record creators)
- Business companies (record creators, intellectual capital)
- Individuals

The issue / challenge

- The enormous and rapidly increasing amount of digital information
  - Fragile resources
- The rapid evolution in technology
- The risk of obsolescence and therefore corruption and/or loss of valuable information
- ***(Pro-)active*** and ongoing attention / maintenance required
- Potential solutions: still fragmented
  - infrastructure
  - not comprehensive

Objectives of Preservation Planning

- To identify and analyse the organisational context
  - including a risk assessment
  - define a framework for preservation / policy
- To 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
- To enable formulation, evaluation and execution of high-quality and cost-effective preservation plans that suit the organisational (e.g. repository) needs
- To develop a sustainable, well-documented and re-iterative process for preserving digital objects

What is Preservation Planning?

- Defining criteria for preservation of specific digital objects (in the collection) based on risk analysis in a defined (business) context
- Following a systematic and structured workflow/ procedure in making decisions about the best possible approaches given the organisational context and the content of the repository
- Defining one or more appropriate preservation plan(s)
  - depending on the types of objects and available strategies
- Execute those plans when needed, and
- Be able to assess the quality of the results
- Document all these steps in order to be accountable

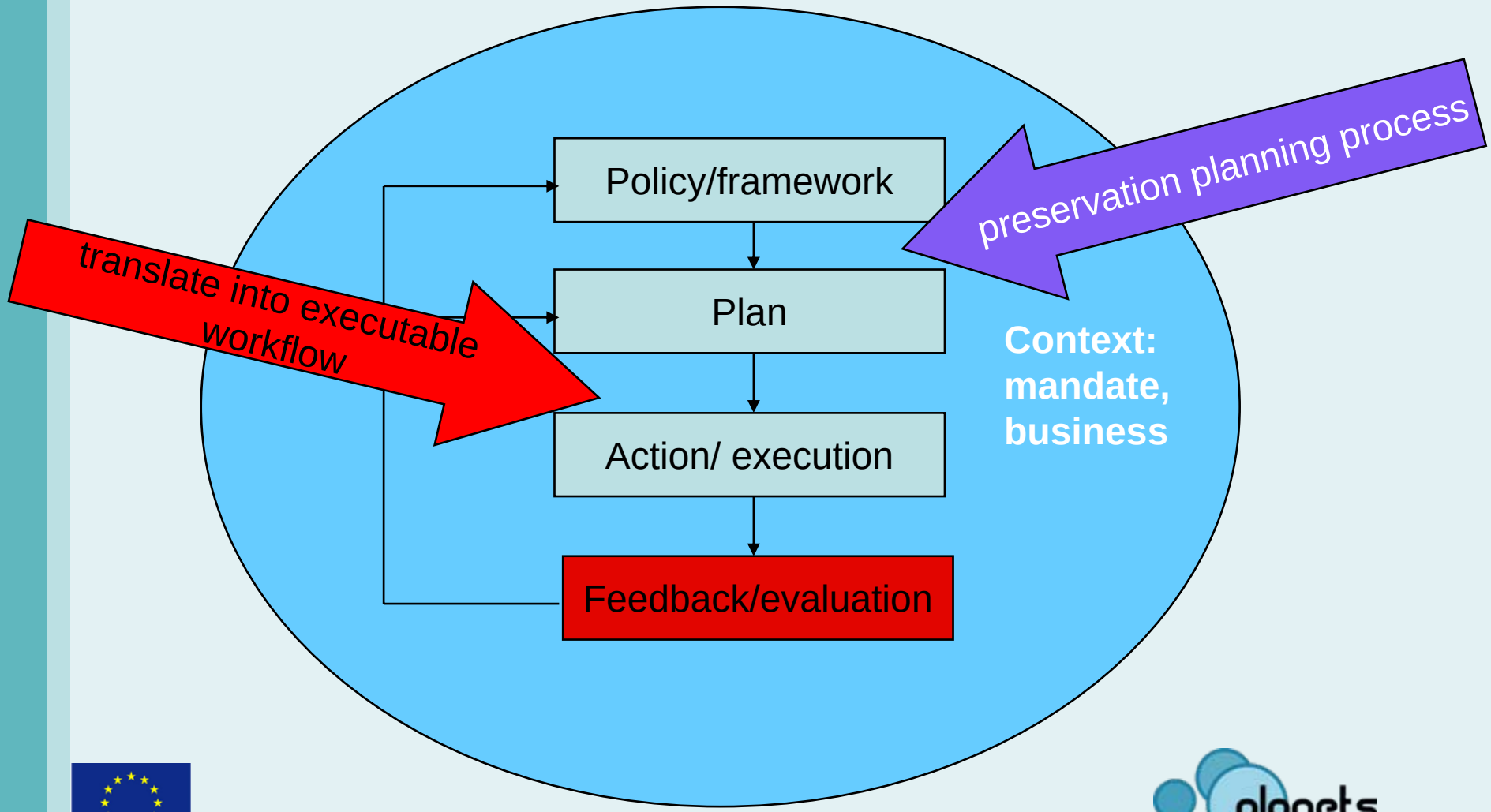
- Issue: scope includes preservation policy?

Preservation policy: what is it?

- Framework for maintaining digital objects over time
- Context (positioning the institution, mandate, legal and social context, user needs, …)
- Scope and objectives
- Principles
- Concepts
- Roles and responsibilities
- Strategies, rules, standards, and procedures
- ….

- Policy: a framework identifying the organisational setting and providing high-level guidance for preservation (planning and actions)
  - explicit ‘statement’

From preservation policy to action



Terminology

- Preservation vs. curation
  - recordkeeping, archiving
- Framework, policy, strategy
- Plan, action, method
- Object, deliverable unit, record, collection

The Planets approach: method

- Understanding the parameters of successful preservation strategy in different contexts, i.e archives, libraries and data centres, e-government, …
  - Understanding organisational policies
    - What is happening in practice?
    - What would a policy encompass?
    - Policies may exist, but are hardly implemented
  - Understanding usage of digital objects
    - Carrying out probes and interviews with selected researchers in different institutions
      - Collect information and experiences from researchers in several iterations
      - Let users talk as much as possible
    - Analysis and structuring of results (‘affinity analysis’)
      - Identification of areas of user interests
      - Interviews with futurologists

Planets approach (2)

- Collection profile
  - 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?

Planets Approach (3)

- Usage requirements for digital objects
  - Performance
  - Usability, presentation
  - Authenticity
  - Understandability

  - Rights
  - Costs

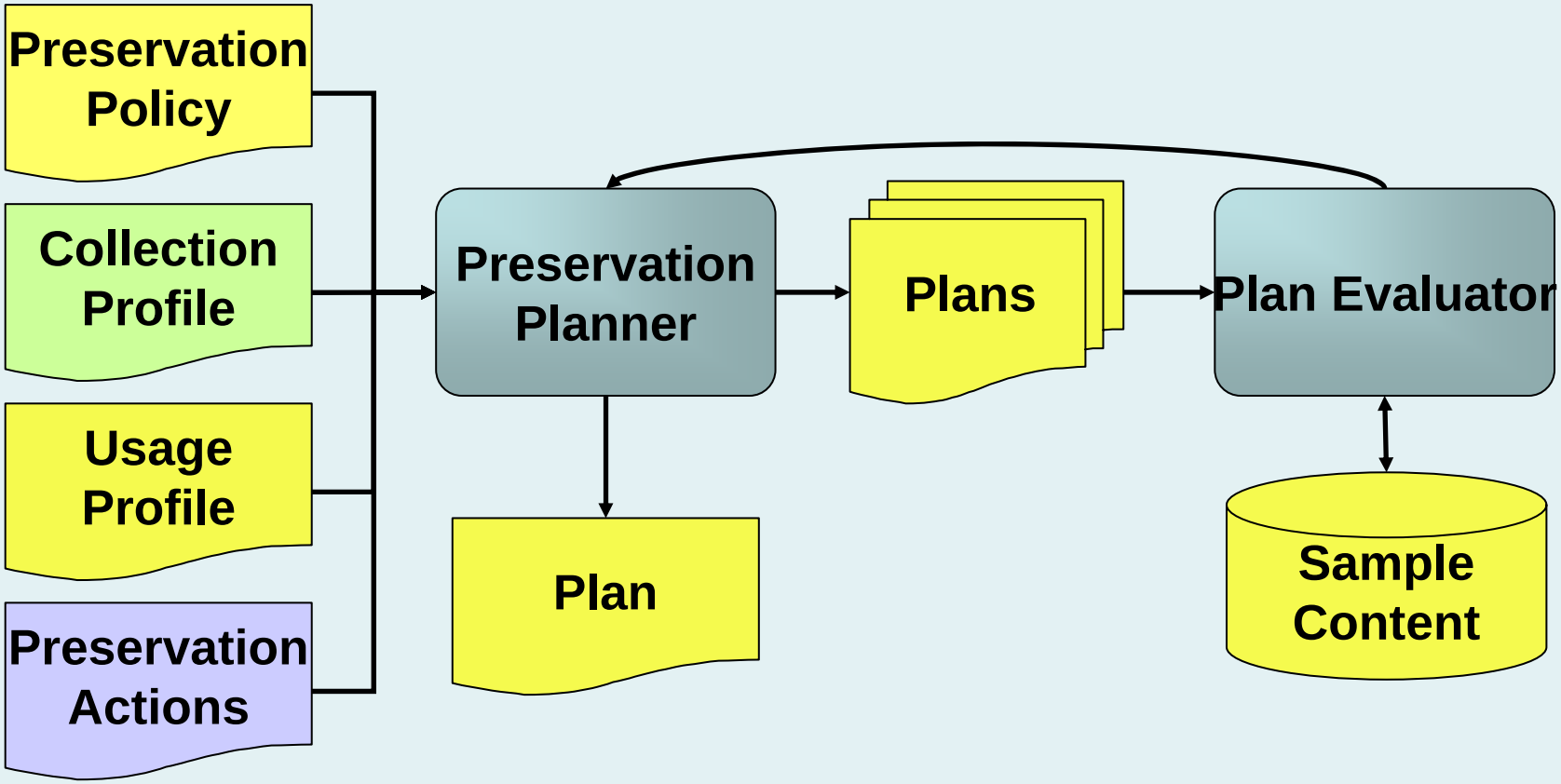
Planets approach (4)

- Providing a decision-support workflow and mechanism capable of translating those parameters into best available strategy
  - Conducting case studies to identify sets of essential characteristics for different digital objects (objective trees)
  - Developing a decision support tool (software): Plato
    - 1st version, end of November 2007 (not public)
    - 2d version publicly available, second half of 2008
  - Systematic procedure for evaluating preservation strategies
- Based on evaluation of strategies create a preservation plan

From policy to action: Preservation Plan as a result

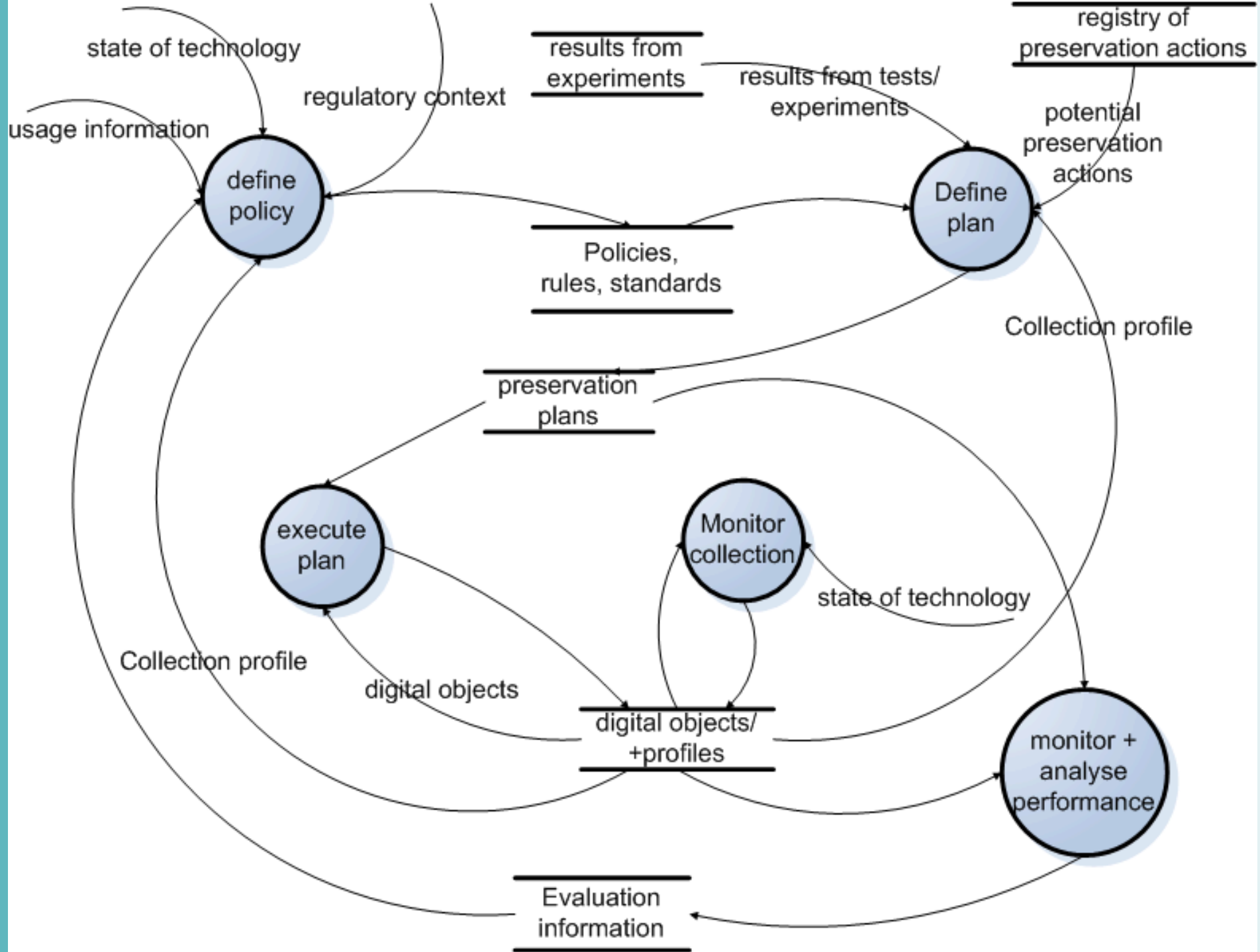
- Preservation context
  - legal and regulatory environment
  - mandate
- Roles and responsibilities
- Selected preservation strategy
  - based on evaluation
  - evaluation result (which alternatives have been compared, the outcome)
- Costs of applying the preservation plan
- Trigger(s) to initiate the execution of the plan
- more…?

Preservation planning



Planets functions

- Technology watch
  - Risk assessment service
  - Recommender services
  - Trigger for adapting preservation plans
- Comparison of (available) preservation strategies based on organisational policies, usage information and collection profile(s)
  - ‘Utility analysis’: identifying essential characteristics (tomorrow)
- Developing and updating preservation plans according to (new) monitoring information and the available strategies evaluation
  - The plans will trigger preservation actions (preferredly in an automated fashion)
  - Tools and services registry
- Validation framework (+ metrics) for evaluating the results of preservation actions
- Testbed
  - experimenting, documented and comparable results.
- Characterisation
  - File format registry


state of technology
regulatory context
usage information
define policy
results from experiments
registry of preservation actions
results from tests/ experiments
potential preservation actions
Define plan
Policies, rules, standards
Collection profile
preservation plans
execute plan
Monitor collection
state of technology
Collection profile
digital objects
digital objects/ +profiles
monitor + analyse performance
Evaluation information

Requirements for objects

- Objects:
  - documents, data sets, websites, databases, CAD-drawings, video, audio, …
- Authenticity
- Reliability
- Integrity
- Usability
- Accuracy
- …
- Tension between preservation and user requirements
  - not always compatible

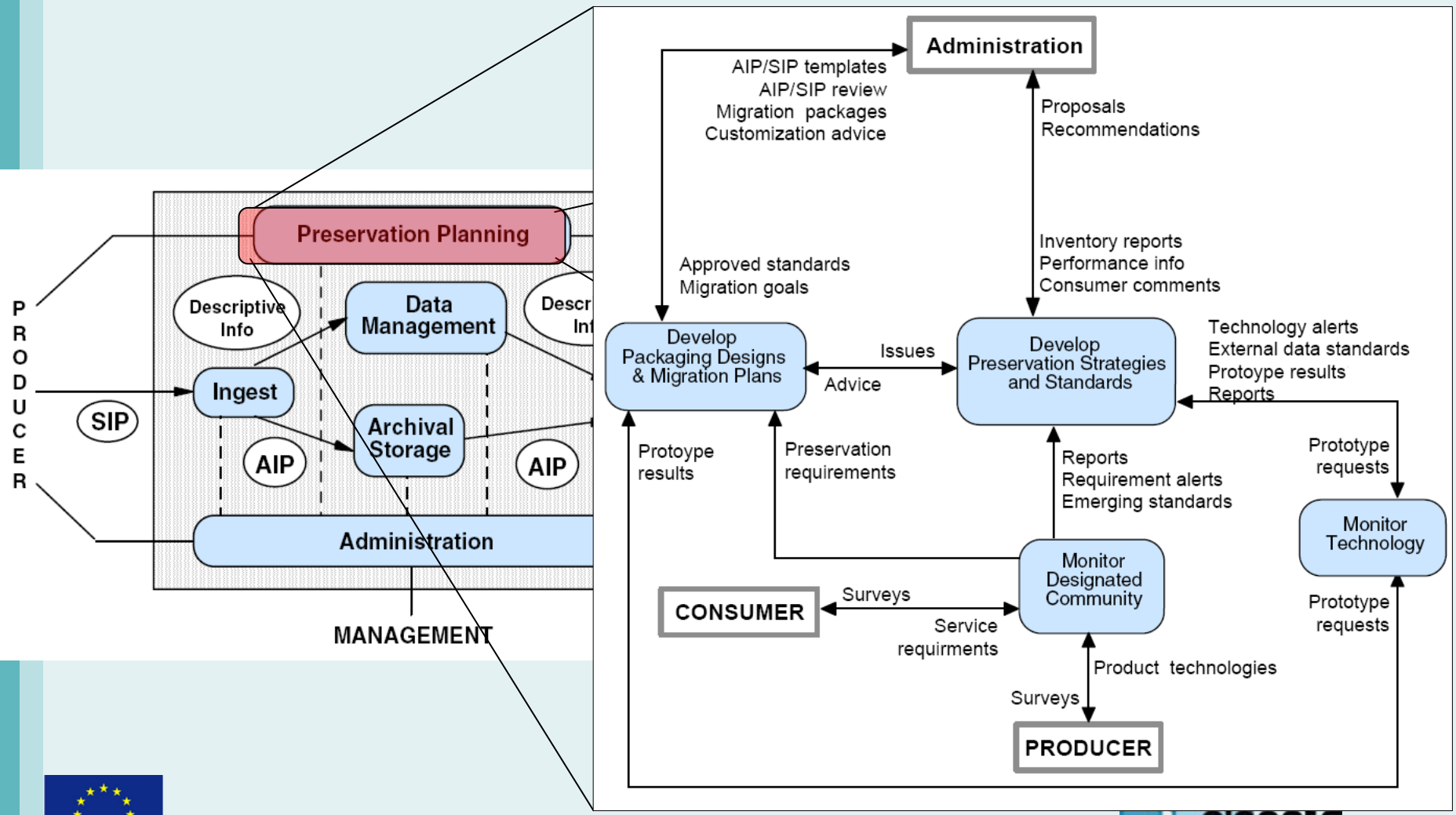
Requirements for objects (2)

- **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.

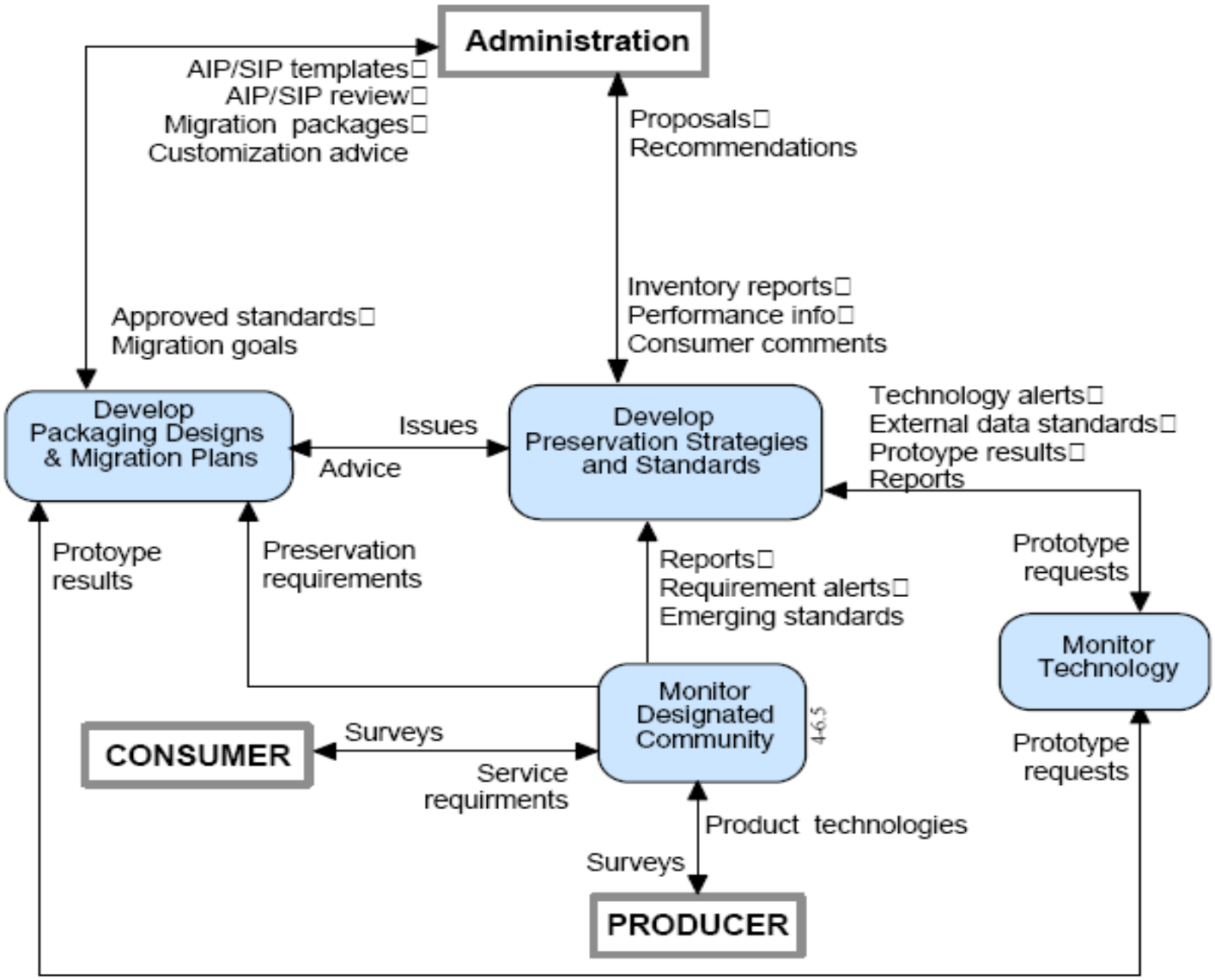
Relationship to OAIS model

+ Potential triggers to start preservation planning

OAIS Model


PRODUCER
SIP
Preservation Planning
Descriptive Info
Data Management
Ingest
Archival Storage
AIP
AIP
Administration
MANAGEMENT
Administration
AIP/SIP templates
AIP/SIP review
Migration packages
Customization advice
Proposals
Recommendations
Approved standards
Migration goals
Inventory reports
Performance info
Consumer comments
Develop Packaging Designs & Migration Plans
Issues
Advice
Develop Preservation Strategies and Standards
Technology alerts
External data standards
Protoype results
Reports
Protoype results
Preservation requirements
Reports
Requirement alerts
Emerging standards
Prototype requests
Monitor Technology
Monitor Designated Community
CONSUMER
Surveys
Service requirments
Prototype requests
Product technologies
Surveys
PRODUCER

The OAIS Preservation Planning Function



OAIS Functions

- Develop preservation strategies and standards
- Develop Packaging designs and migration plans
- Monitor designated community/-ies
- Monitor technology

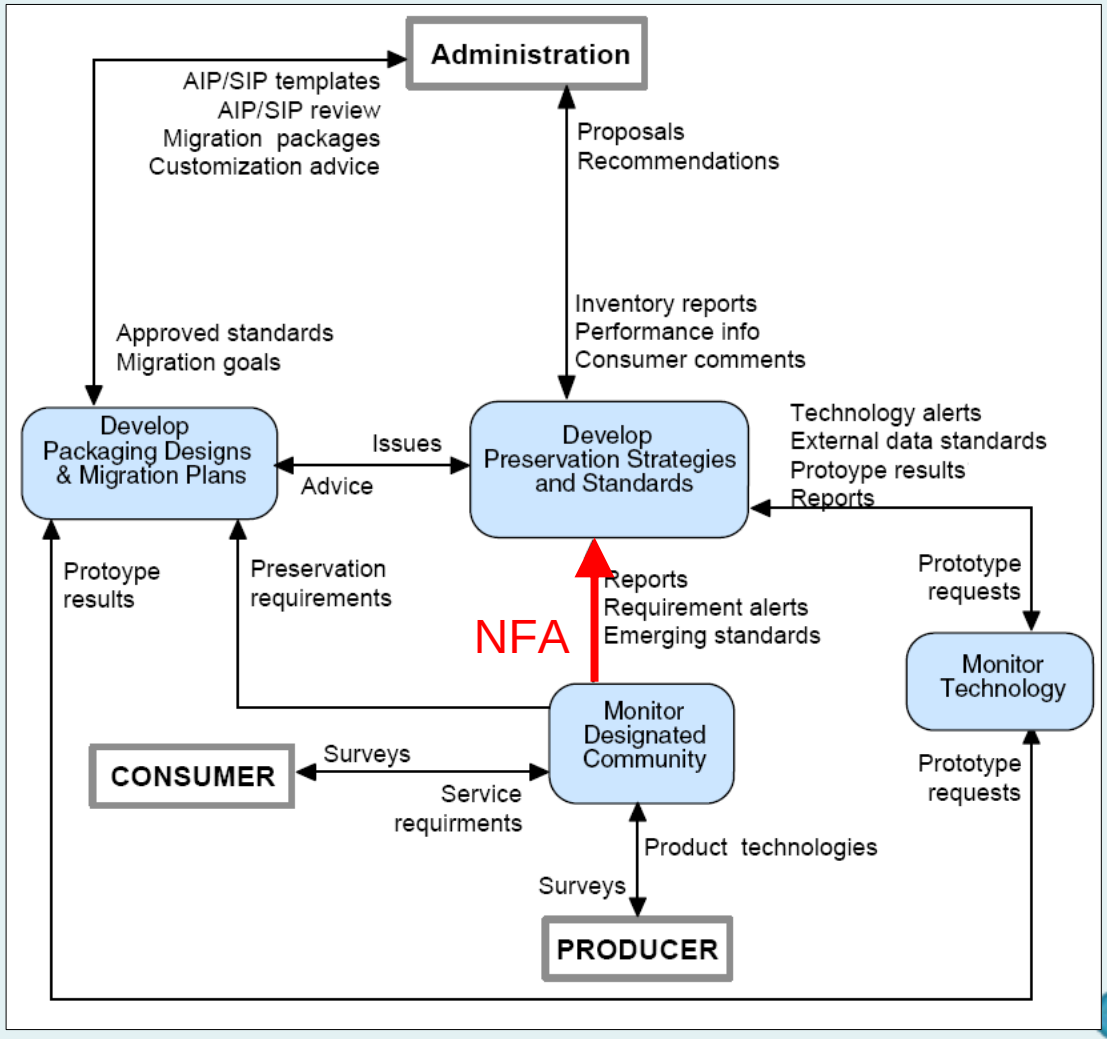
Alerts (1)

Alerts that can trigger a preservation planning activity

- New Format Alert (NFA)
  - New format is accepted in the repository
  - New preservation strategy has to be identified and evaluated
  - Raised by:
    - Monitor Designated Community
      - Changes in Designated Community
      - Frequent unanticipated submissions

Change in the Designated Community

For example use of a new file format



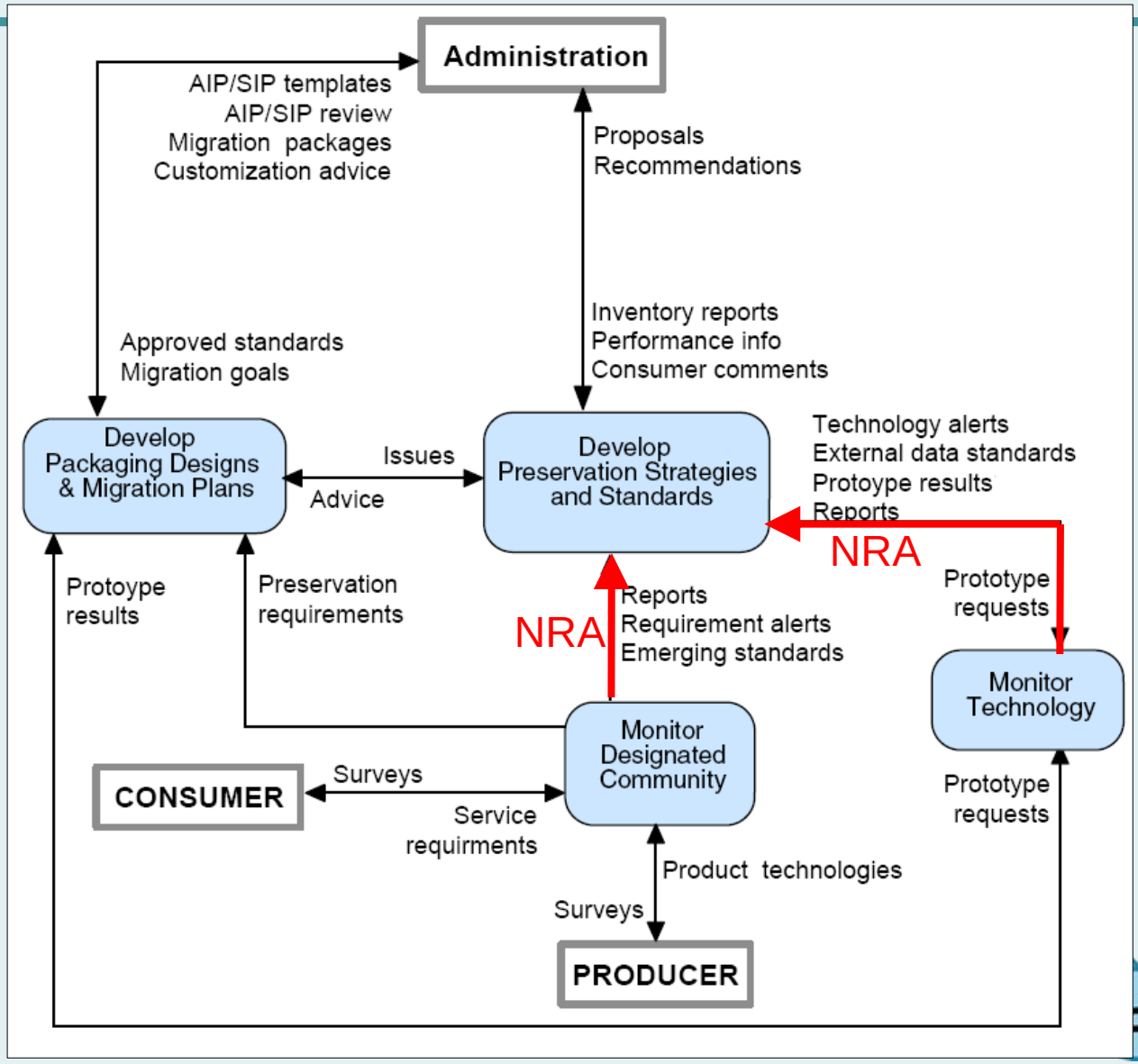
Alerts (2)

Alerts that can trigger a preservation planning activity

- New Requirement Alert (NRA)
  - Changes and development in technology
  - Change in preservation requirements
  - Re-evaluation of existing preservation strategies

  - Raised by
    - Monitor Designated Community
    - Monitor Technology

New Requirements Alert

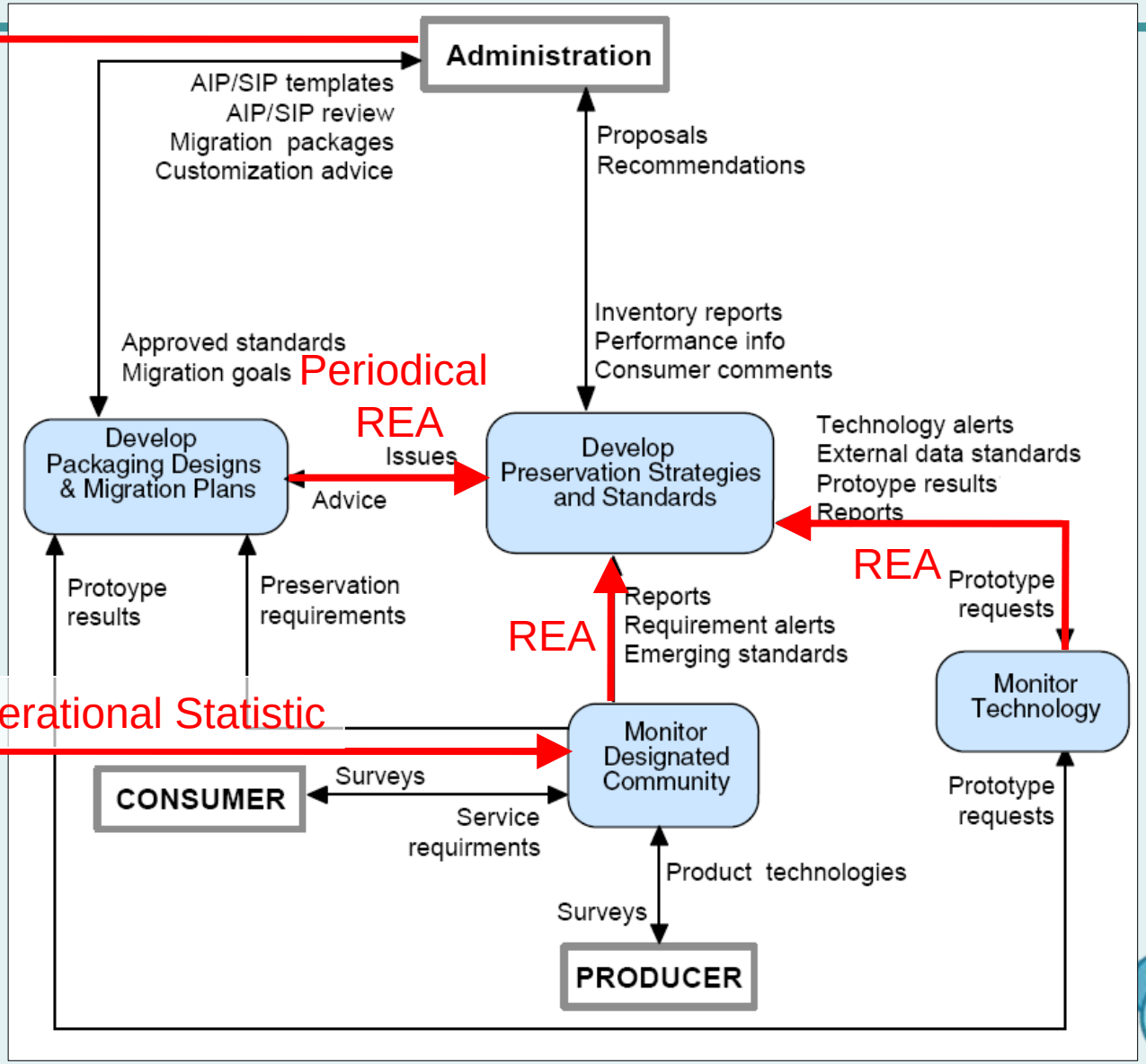


Alerts (3)

Alerts that can trigger a preservation planning activity

- Revision Alert (REA)
  - Periodical reviews help to improve and further develop existing preservation strategies
  - 3 kinds:
    - Periodical review
    - Review depending on operational statistics
    - Review for new developments

Revision Alert



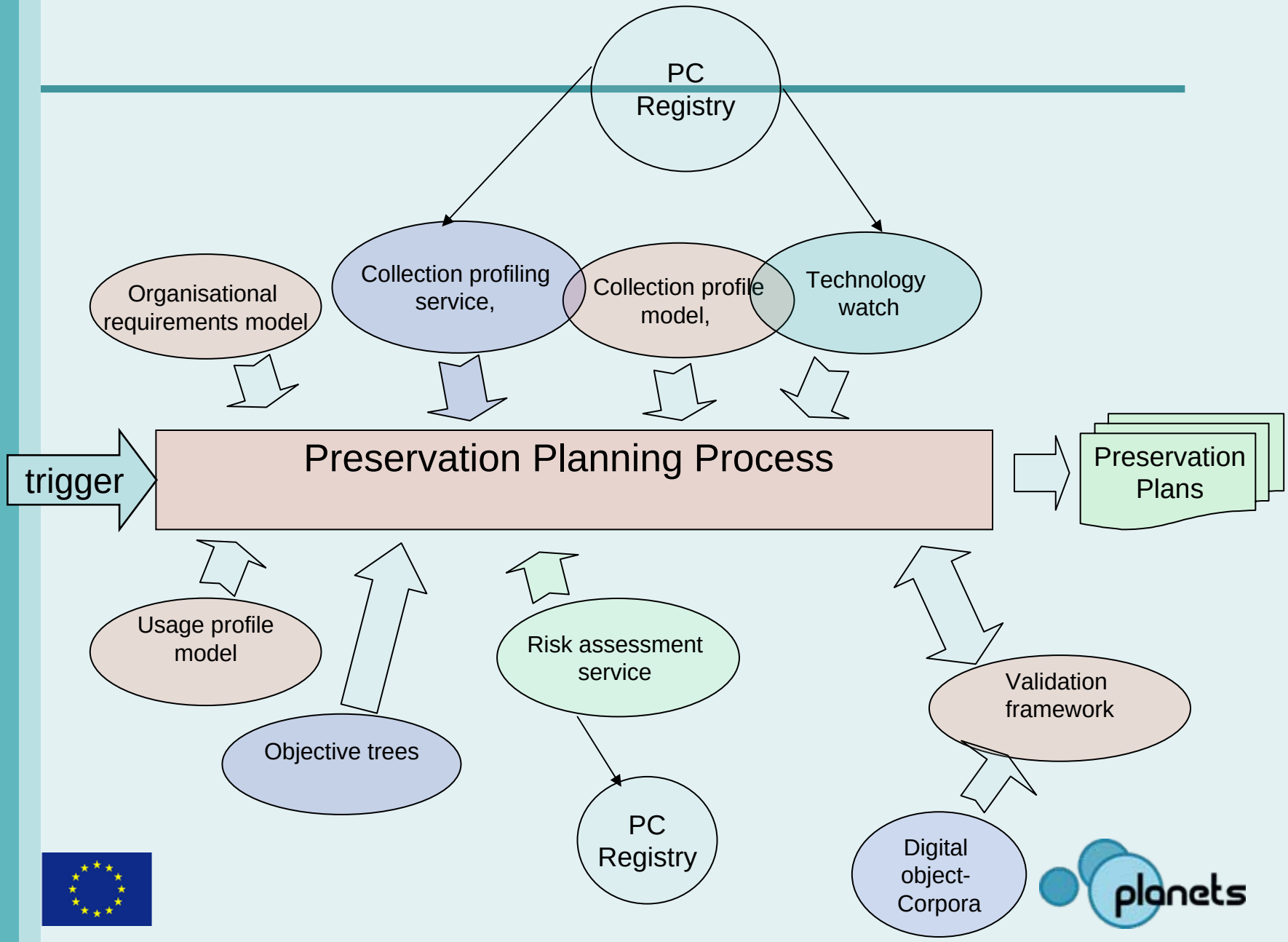
Preservation Planning: Summary

Steps

- Understanding of context
  - analysis of organisational needs, user needs, legal requirements
- Identify criteria for preservation
  - how long, restrictions of formats, standards, …
- Determine what to keep/maintain
  - essential characteristics (objective trees), characterisation of computer files
- 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/perfotrmance
- Re-iterate when technology changes or review when policy and/or collection and/or usage changes
- Automated process support (?)

Components of preservation planning

- Policy framework
- Collection profile:
  - profile model
  - profiling service
  - technology watch + risk assessment
- Usage requirements:
  - usage model
  - identification of requirements
- Essential characteristics
  - objectives trees, database of tree-modules to be collated in required tree
- Available tools
  - tools registry
  - quality of tools: testbed documentation
- Preservation actions
  - available services, services registry
- Validation framework
  - metrics


PC
Registry
Organisational requirements model
Collection profiling service,
Collection profile model,
Technology watch
trigger
Preservation Planning Process
Preservation Plans
Usage profile model
Objective trees
Risk assessment service
PC
Registry
Validation framework
Digital object-Corpora
planets