



## Report on the Planets Functional Model

---

**Sub Project:** PP

**Work Package:**7

**Document Type:** Deliverable

Deliverable ID or referral: PP/7-D3 +D4

Planned:

Delivered: 8-5-2009

---

### Document owner

Person	Institution
Barbara Sierman	KB-NL

### Document approval

Person	Institution	Role	Approval date
Hans Hofman	NANETH	PP Sub project leader	25-06-2009
Manfred Thaller	UzK	Reviewer	04-06-2009

### Version information

Person	Institution	Version	Description	Date
Barbara Sierman and Paul Wheatley	KB-NL/BL	1.0	First version for review	10-5-2009
Barbara Sierman and Paul Wheatley	KB-NL/BL	1.1	Changes as a result of discussion with reviewers	4-12-2009

## Contents

1	Introduction .....	3
2	The Planets Functional Model .....	4
2.1	Planets Functional View (version 2) .....	4
2.2	The Planets Sub-Functions .....	6
2.3	Repository and external functional entities .....	9
2.4	Overview of Changes to the Planets Functional Model .....	10
3	Scenarios.....	12
3.1	Introduction and aims .....	12
3.2	Preservation Watch Scenario: Monitoring and Risk Analysis .....	13
3.3	Preservation Planning Scenario: Planning and Creation of Preservation Plan .....	16
3.4	Preservation Action Scenario: Migration.....	18
4	Conclusions .....	20
5	Further Work.....	22

## **1 Introduction**

This document is an updated version of the Report on Comparison of Planets with OAIS<sup>1</sup>, originally published as PP7/D1<sup>2</sup> in 2007.

In the Planets project an array of technologies, services, tools and reports have been developed to fulfil the key functions of Preservation Planning. This in turn has led to a greater understanding of the processes that these tools have been developed to execute. This learning process is expected to continue as the Project moves into its final year and applies Planets results in the preservation of partner institutions' content.

The Preservation Planning module in OAIS is the least well tested (via practical implementation) of the main functional OAIS components. Progress as part of the Planets Project has provided an opportunity to compare and contrast Project developments with the thoughts laid out in the OAIS model some ten years ago. The aim of this comparison is to identify key omissions in both the Planets work and in OAIS, to provide some validation where there are matching concepts, and to identify opportunities to enhance the communication of what are often complex preservation concepts to end users.

In the first iteration of this document a Planets Functional Model was developed. This was performed by the members of the PP7 work package with a limited consultation of the Planets Sub projects where developments were at an early stage. Some two years later however, the outline of the Planets results has become clearer, the work packages have delivered several iterations of key components and it became time to check the validity of the Planets Functional Model with the other members of the Planets project. The results of these discussions have been incorporated in a new version of the model (chapter 2).

To support the discussion with our Planets colleagues, three Scenarios describing typical preservation processes were captured. These will be explained later in this report (chapter 3).

The validity of the Conclusions in PP7/D1 were considered and revised (chapter 4).

Note that it was originally planned to produce a separate deliverable (PP/7-D3) describing just the Scenarios, but it was felt beneficial to combine that detail with this report.

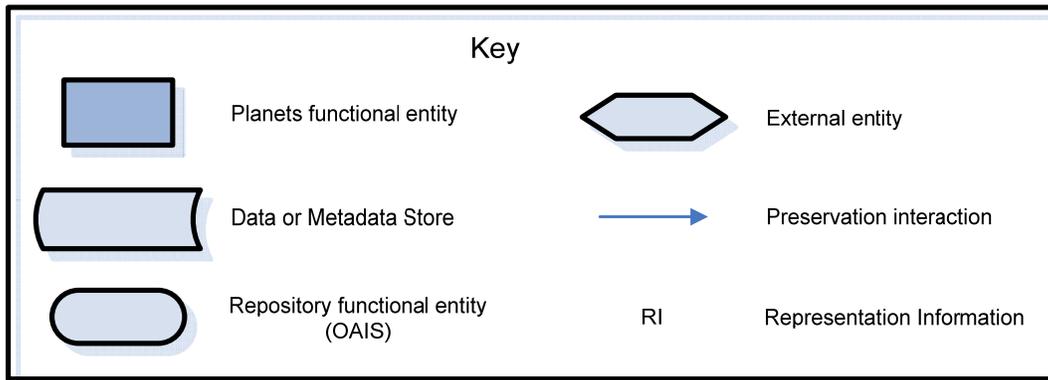
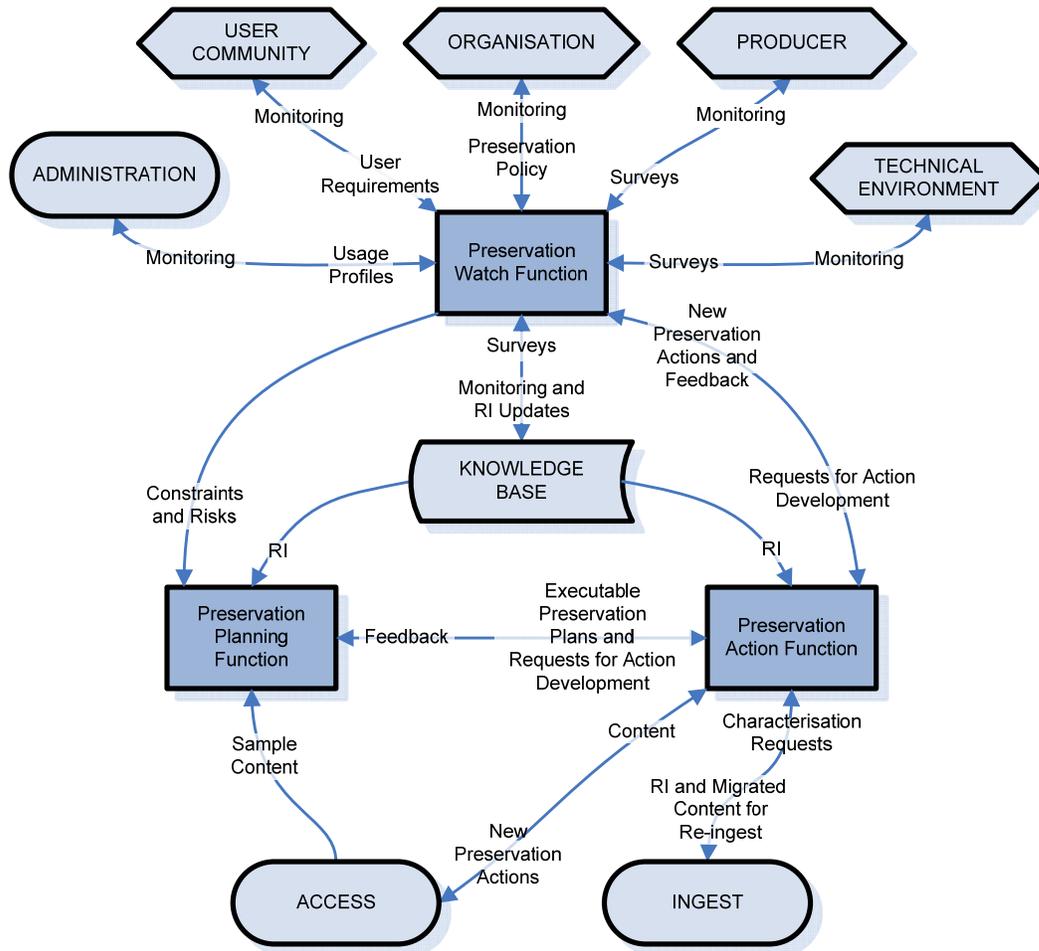
---

<sup>1</sup> Note that a revised draft of OAIS (Pink book) was given a public release at the time of finalising this report. Unfortunately it was not possible to take this revision into account, but this will be addressed in future deliverables.

<sup>2</sup> [http://www.planets-project.eu/docs/reports/Planets\\_PP7-D1\\_ReportOnComparisonOfPlanetsWithOais.pdf](http://www.planets-project.eu/docs/reports/Planets_PP7-D1_ReportOnComparisonOfPlanetsWithOais.pdf)

## 2 The Planets Functional Model

### 2.1 Planets Functional View (version 2)



## 2.2 The 3 Planets Functions

The model identifies 3 key preservation functions: Preservation Watch, Preservation Planning and Preservation Action. All functions are described briefly here, and further detail is provided in section 3.

Preservation Watch monitors a variety of internal and external entities, including the content preserved in the repository itself (via ADMINISTRATION). Where potential changes in the entities are identified (e.g. a new tool is available, a platform is no longer supported, or a new use case becomes popular), the resulting preservation risk is assessed. Critical or imminent risks are passed to Preservation Planning for further analysis and action. Information gathered from the entities is also used to provide Representation Information Updates to the KNOWLEDGE BASE<sup>3</sup>. For example, the addition of a Knowledge Base entry describing a new tool that renders a particular file format.

Preservation Planning assesses Constraints and Risks received from Preservation Watch. It evaluates available preservation options (informed by Representation Information from the KNOWLEDGE BASE) and then trials them on Sample Content. It assesses the results of those trials and identifies the most appropriate options. Plans for implementing the selected preservation options are then created and passed to Preservation Action for implementation.

Preservation Action performs actions on Content to ensure its continued accessibility and sends appropriate feedback to Preservation Watch. Where Preservation Planning requires that a migration be performed, a Preservation Plan is passed to Preservation Action, describing the required process. The Content is passed from Access. As part of the preservation plan a Characterization of the Content will be performed, both before and after the migration. The appropriate tools and services to perform the migration action are identified from the KNOWLEDGE BASE and will be deployed in the appropriate environment (see section 2.3.3 Deploy Action). The preservation action will then be executed and evaluated.

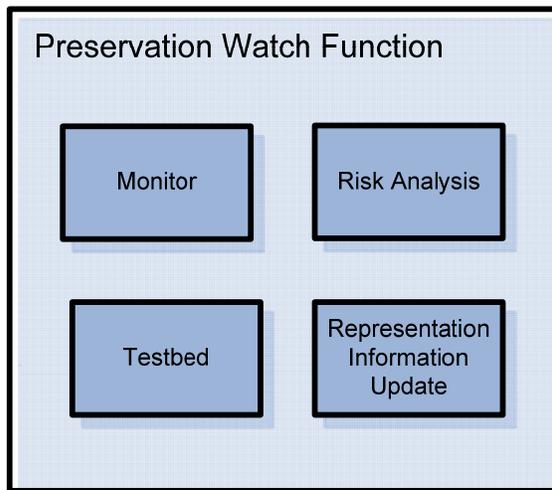
---

<sup>3</sup> The KNOWLEDGE BASE is an umbrella term, representing various repositories for preservation metadata. This will include Representation Information about formats, tools and environments; Testbed experiment results data; and risk register data (see for a more detailed description 2.4)

## 2.3 The Planets Sub-Functions

This section of the document provides descriptions of the Planets Sub-Functions.

### 2.3.1 Breakdown of the Preservation Watch Function



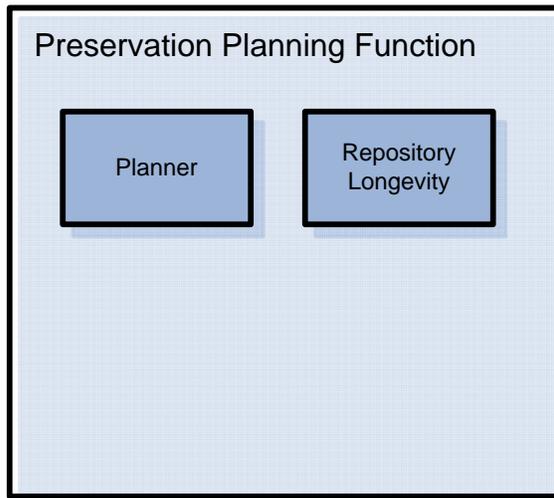
Monitor provides the role of collating preservation related information from a variety of internal and external entities. Monitoring will be scheduled on a regular basis but may also be triggered by alerts from Technology Watch services or other sources.

Risk Analysis provides an assessment of this information, relaying critical risks to Preservation Planning. For example a particular tool becomes obsolete, leaving content in the repository without a rendering mechanism. A further Preservation Planning exercise would then be triggered, to consider remaining alternatives.

Representation Information Update provides updates to Representation Information in the KNOWLEDGE BASE, typically enriching information describing file formats, tools or environments. For example, adding a reference to a new PDF rendering tool to the KNOWLEDGE BASE entry describing the PDF format.

The Testbed is a controlled environment that facilitates the execution of experiments that assess the capabilities of preservation tools and services.

### 2.3.2 Breakdown of the Preservation Planning Function

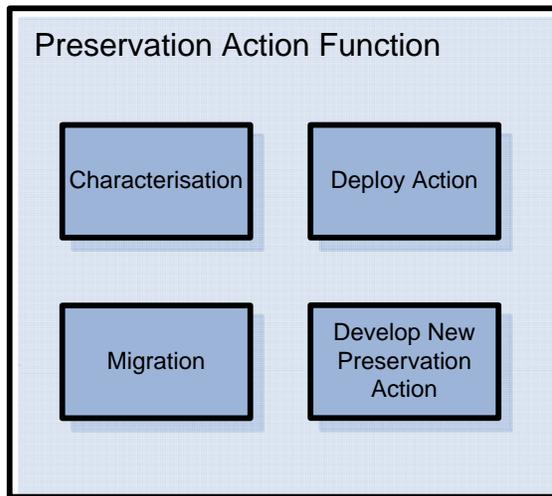


The Planner is the decision analysis function within Preservation Planning which analyses and selects appropriate preservation solutions.

The Testbed results are stored in the KNOWLEDGE BASE and these will inform Preservation Planning activities.

Repository Longevity provides a focus on review and necessary action to ensure longevity of the Open Archival Information System itself. This includes a focus on areas such as packaging designs, repository technologies and preservation of the KNOWLEDGE BASE and other databases.

### 2.3.3 Breakdown of the Preservation Action Function



A Preservation Action can take several forms and may even combine several stages or techniques. Migration transforms digital objects from format to format before they are re-ingested to the repository or served on demand to the user. Emulation mimics the original environment in which the digital object was created at the point of access for the user. The techniques can also be combined to extend the life of obsolete migration tools.

Characterisation identifies particular properties within content and extracts their values. These value-property pairs (called characteristics) can then be utilised to perform particular preservation processes. Characteristics can be interpreted in order to construct a Content Profile of preservation relevant characteristics of content (such as file format) which will in turn support activities such as: the monitoring of content (via Preservation Watch); the selection of content in a repository on which to perform Preservation Actions; and the selection of compatible preservation tools (eg. that perform rendering, migration). Characterisation can also be performed to support the validation of a Preservation action, for example, by extracting characteristics before and after a Migration and then comparing them for evidence of loss.

Deploy Action provides the ability to deploy specific preservation tools and services within the Open Archival Information System environment, as well as performing preparatory activities necessary to preservation action at the point of execution (for example building an emulator image).

Develop New Preservation Action focuses on developing new Preservation Action tools or services.

## 2.4 Repository and external functional entities

The text above references a number of repository functions. These are described below:

Functional entities	Description
INGEST	The function which manages ingest of digital objects into the repository (see OAIS).
KNOWLEDGE BASE	<p>The KNOWLEDGE BASE is used here as an umbrella term for a repository of a variety of key information which will inform preservation processes conducted within the OAIS repository. It provides the information required to understand and make use of a particular digital object, including information about how to render, interpret or re-use the digital object (which broadly equates to the OAIS described concept of a Representation Network).</p> <p>It may also contain a variety of other preservation information including risk registers for the organisation and the results of experiments or other evaluative activities on digital objects or tools,</p>
ACCESS	Provides access to digital objects from the archival store (see OAIS).
ORGANISATION	The ORGANISATION that is the custodian of the digital objects being preserved. This organisation might also act on behalf of other content owners as a service provider.
USER COMMUNITY	The community or communities of external actors that are expected to be users of the digital objects.
TECHNICAL ENVIRONMENT	An abstract entity representing the technical environment that the ORGANISATION, the PRODUCERS and the USER COMMUNITY operate within. Key elements are computing platforms, application software and file formats.
PRODUCER	The Producers of the digital objects being preserved. This might represent the actual creator and/or the actor supplying the digital objects to the repository.

ADMINISTRATION	The services and functions needed to manage the operation of the other functional entities on a day-to-day basis.
----------------	---

## 2.5 Overview of Changes to the Planets Functional Model

The PP7 Work package group had extensive discussions with key Planets Sub-Projects. This informed the production of the revised Functional Model, described above. This section of the document provides an overview of the main changes.

- Characterisation was merged with the Preservation Action Function.
  - o Based on the discussions and the development of the scenarios, it became clear that characterisation was better placed as an integral part of the Preservation Action function.
  - o As a consequence INGEST will consult the Preservation Action Function for Characterisation requests (and not via Characterisation), so a new interaction was added here.
  - o INGEST will receive both metadata and migrated content for re-ingest from the Preservation Action Function.
- Representation Information Network is changed into KNOWLEDGE BASE and placed more centrally in the diagram. The KNOWLEDGE BASE includes the Representation Information Network as originally intended, but will also encompass registries of other preservation related metadata (for example a risk registry or a set of experiment results).
  - o Preservation warnings from RIN changed to Surveys for consistency with other similar interactions.
  - o Removed Representation Information updates from Preservation Planning to the KNOWLEDGE BASE.
  - o An interaction was added from KNOWLEDGE BASE to Preservation Action as Requests for Action Development. This previously came directly from Preservation Planning.
- The interaction from Preservation Watch to Preservation Planning was changed from Planning Requirements and Alerts into Constraints and Risks.
- Access: Digital Objects was changed to the more specific term: Sample Content.

## Planets

### PP7/D3-4 Report on the Planets Functional Model

---

- The linkage from Preservation Action to ACCESS was changed from New Access Modules to New Preservation Actions, which is less specific and more suitable for an implementation neutral model.
- Preservation Planning will send Preservation Plans to Preservation Action. In the new diagram the request for development of new preservation actions will come from Preservation Planning, and also from Preservation Watch. After the Preservation Action Function realized this new preservation action, feedback is sent to the Preservation Watch Function, where the information is sent to the Knowledge Base.
- The layout of the Model was revised to improve readability

### **3 Scenarios**

#### **3.1 Introduction and aims**

Three Scenarios describing key preservation processes were produced with the aim of communicating the role of preservation functions and facilitating discussion across the Planets Project.

PP/7 has attempted to draw out key high level characteristics which have been developed with the benefit of the hands on work performed by Planets, while at the same time, not allowing low level detail to cloud the critical functions and relationships present.

The Scenarios offered a means of communicating the work of PP/7 and in particular a means of validating the Planets Functional Model and ensuring its relevance and accuracy to work across the Planets Project.

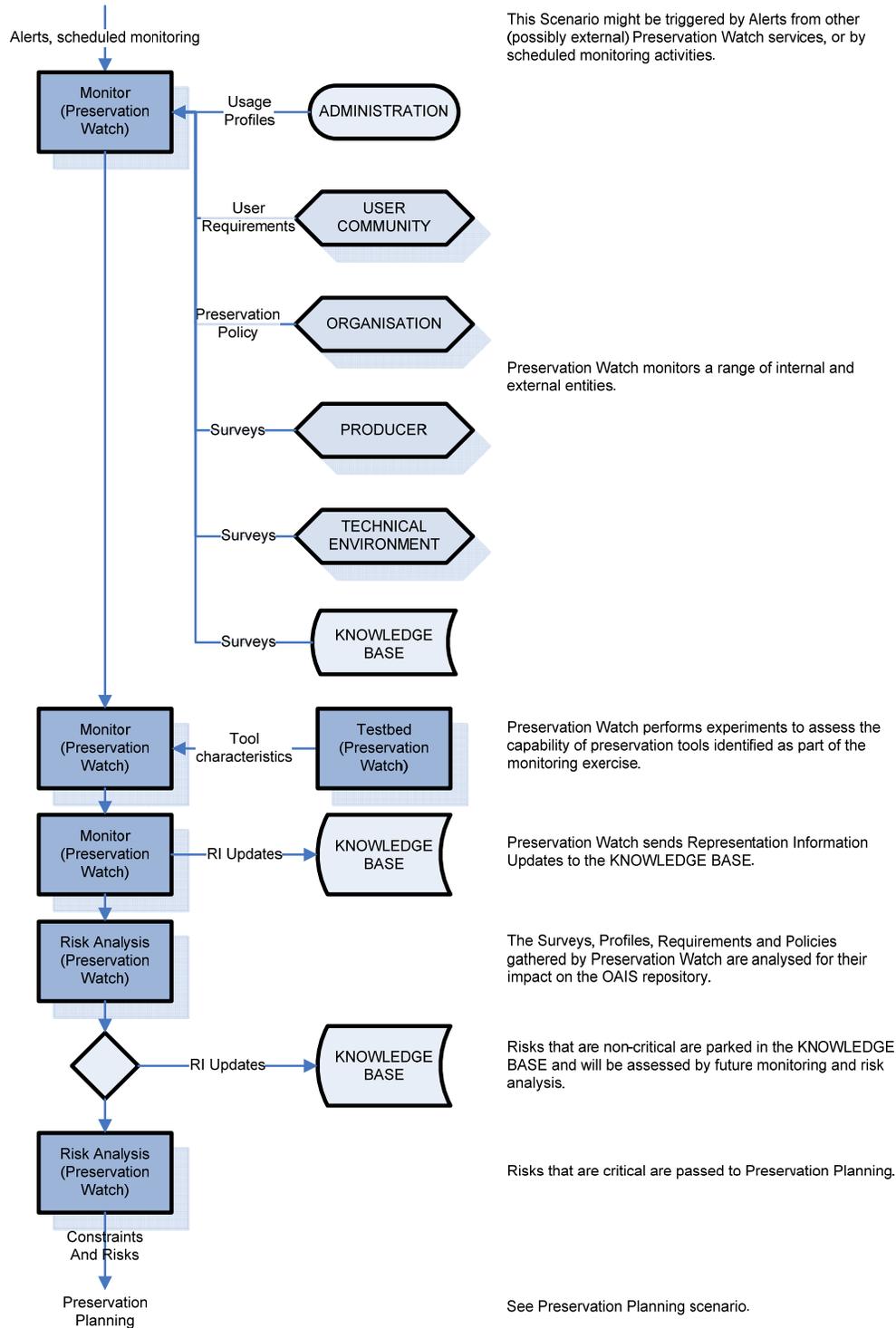
The scenarios were discussed with:

PP group, PC, PA representatives and project management on January 20<sup>th</sup> 2009 in The Hague

TB, IF and PP4 representatives on March 16<sup>th</sup> 2009 in Vienna

These meetings, particularly those made as part of the Preservation Planning Sub Project meeting, were felt to be particularly productive and the Scenarios providing useful talking points for the meeting participants. They also enabled the critical boundaries and linkages between existing work to be considered in a new light.

### 3.2 Preservation Watch Scenario: Monitoring and Risk Analysis



The Preservation Watch Scenario represents an ongoing monitoring process. Based on certain rules, translated into a pre-arranged schedule, monitoring activities are conducted. Monitoring might also be triggered by alerts from an external service, such as a File Format Obsolescence Service.

Monitor will survey various internal and external entities and will gather a variety of information which will inform preservation activities. This might for example include a change to the TECHNICAL ENVIRONMENT, such as a new preservation tool becoming available, or an existing tool or platform reaching the end of its life and no longer being supported by the author or producer. Gathered surveys may trigger further investigation into a particular preservation issue or perhaps a particular file format or preservation tool. The Testbed function is able to conduct this research and experimentation, returning results to Monitor.

Information gathered as part of Surveys or subsequent experiments may be suitable for updating the KNOWLEDGE BASE, and this is conducted by the Representation Information Update function. Such updates might include more detail on a particular file format, the addition of a new rendering tool for a particular file format, or a change in the Designated Community which requires an update to the information describing it.

The Risk Analysis function analyses information from gathered Surveys and assesses it for impact on the OAIS repository. Critical risks are then passed on to Preservation Planning for further assessment and preparation of a plan for the treatment of at risk content. Non-critical risks are parked in the KNOWLEDGE BASE.

### **3.2.1 Notes, Issues and Additional Information**

The monitored entities include:

- ADMINISTRATION

Information on collection and usage profiles

- Organization

Changes in its own organization, like a change in a collection profile, might lead to different preservation policies. Apart from its own organization, the Preservation Watch function might monitor an external organization in case parts of the repository are outsourced, or the repository is a service provider for other repositories.

- KNOWLEDGE BASE

Preservation Watch Function monitors periodically the information in different registries regarding the file formats and for example significant properties stored in the digital archive and decides whether there is a risk. The KNOWLEDGE BASE might contain information on earlier identified risks, information on experiments with tools and services of other organisations (Testbed) etc. The KNOWLEDGE BASE could be internal and / or external.

- USER COMMUNITY

Preservation Watch Function monitors the USER COMMUNITY, potentially accessing the digital objects and identifies possible risks, like for example a change in the expected DIP format

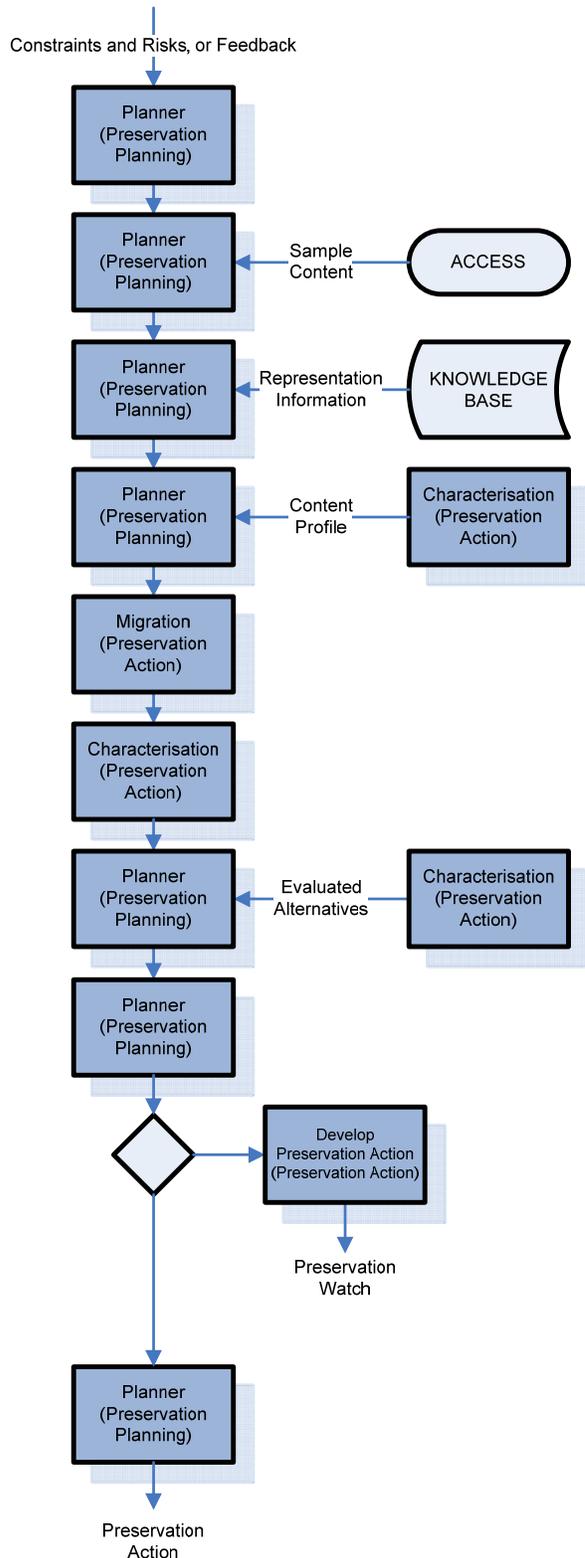
- PRODUCER

Preservation Watch Function monitors the PRODUCER, these are all organizations depositing material into the digital archive. Changes in this area might offer a risk or at least require adaptation

- TECHNICAL ENVIRONMENT

Preservation Watch Function monitors the TECHNICAL ENVIRONMENT in general: developments in the technical area, that might lead to new file formats, new digital object formats, new publishing methods.

### 3.3 Preservation Planning Scenario: Planning and Creation of Preservation Plan



This scenario might be triggered by Constraints and Risks from Preservation Watch or by Feedback from Preservation Action resulting from an action that did not meet quality levels identified in a previously created Preservation Plan.

Planning Requirements are defined.

Sample Content is selected and provided by ACCESS.

Appropriate alternatives (preservation tools or services) are selected from available tools and services described in the KNOWLEDGE BASE.

The Sample Content is Characterised and Content Profiles are returned.

The Experiments are executed on the Sample Content.

The actioned Sample Content is Characterised. Content Profiles are returned.

The Content Profiles are compared and assessed by Characterisation.

The Evaluated Alternatives are assessed by the Planner against the Planning Requirements.

If none of the evaluated alternatives meet the requirements a request is made to Preservation Action to develop a new Preservation Action.

When complete, a revised Preservation Planning process will begin.

The Executable Preservation Plan is constructed and passed to Preservation Action.

See Preservation Action scenario(s)

Preservation Planning receives a notification from Preservation Watch, describing critical risks and constraints for further assessment. Planning Requirements are then defined and the background and context to the planning exercise is captured. Appropriate Sample Content are selected and obtained from ACCESS. Preservation alternatives are selected, utilizing information from the KNOWLEDGE BASE. These alternatives will be explored in experiments conducted on the Sample Content. The Sample Content is characterised to produce a Content Profile and then the experiments are executed. Content produced as output from the experiments is re-characterised and compared with the original Content Profiles. These results are then assessed by the Planner against the original requirements for the planning exercise. If the results satisfy the requirements, an Executable Preservation Plan is created and passed to Preservation Action for execution. If the requirements are not satisfied, further assessment may be required and a request to Preservation Action to develop a new action may be necessary.

Preservation Planning may also be triggered by Feedback from Preservation Action resulting from an action that did not meet quality levels identified in a previously created Preservation Plan.

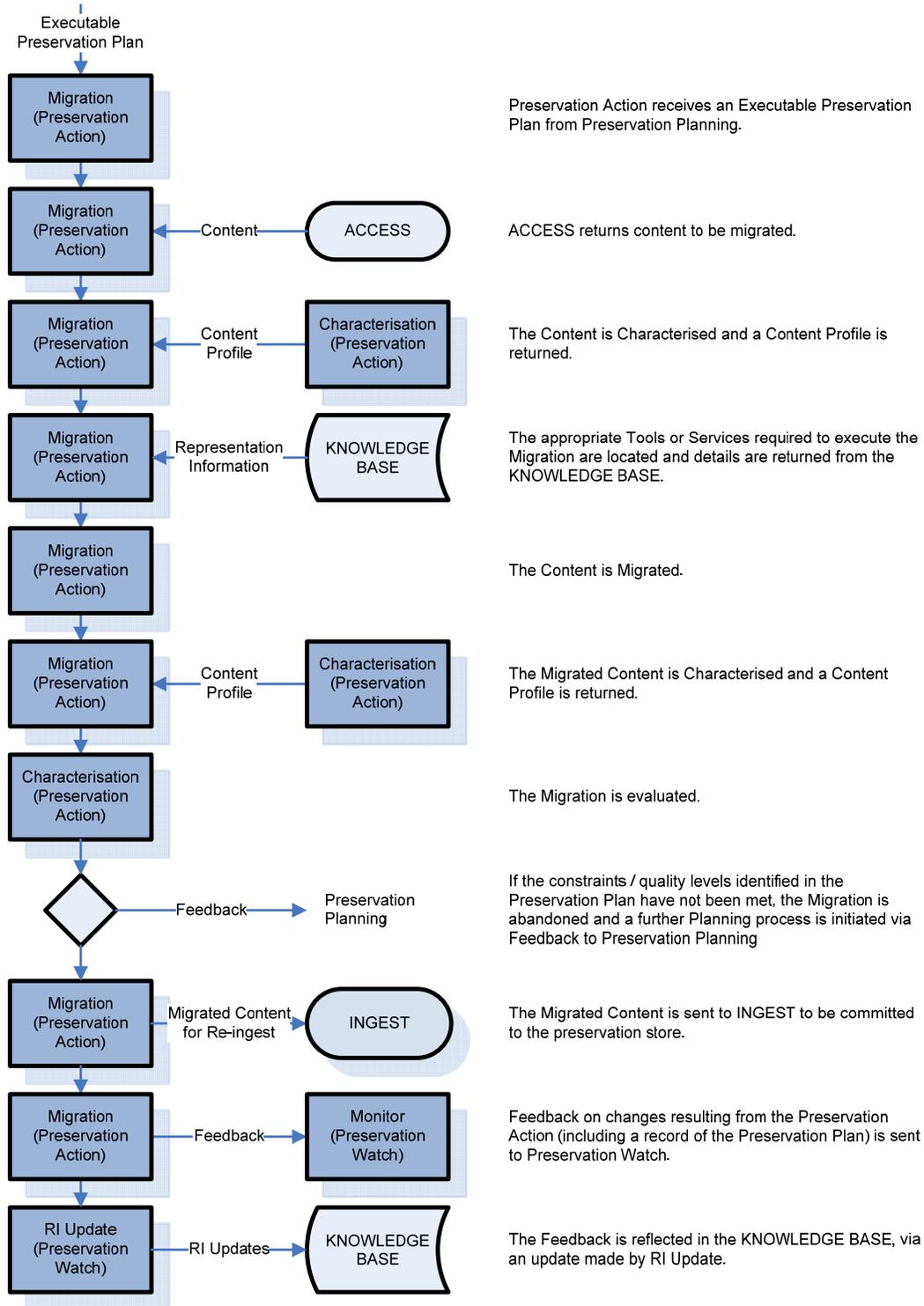
### **3.3.1 Notes, Issues and Additional Information**

The Constraints and Risks passed from Preservation Watch to Preservation Planning might include a wide range of information that will guide the planning process. Details including related preservation policies, an indication of the affected objects (this will help to create a representative subset of sample content) and additional constraints such as restrictions on the costs of the planned Preservation Action may be present.

The planning process shown in this scenario is likely to involve a number of cycles as detail is laid out and earlier stages need to be revised. This is not shown in the diagram to ensure clarity of communication of the core process.

In future it might be expected that organisations will be supported in creating preservation plans guided by existing templates in which common requirements and processes are made easily repeatable, perhaps with minor modifications. This will be one of the goals of Plato, the Preservation Planning tool of Planets.

### 3.4 Preservation Action Scenario: Migration



Preservation Action receives an Executable Preservation Plan from the Preservation Planning Function. In this case a Migration is planned. The Preservation Action Function obtains the set of objects identified in the Plan from ACCESS. The Content is then characterised and a Content Profile is generated. The Tools or Services required to execute the Migration are located and are deployed ready for use. The Content is then migrated and a further Content Profile is generated. This is compared and evaluated against acceptable quality levels. If these have not been met a further Preservation Planning process is required. Otherwise, the migrated Content is passed to INGEST in order to be re-ingested into the preservation store. Feedback (including the Executable Preservation Plan) is sent to Preservation Watch and is updated in the KNOWLEDGE BASE.

#### **3.4.1 Notes, Issues and Additional Information**

An issue raised during the discussions was the question of whether this scenario was applicable to more complicated Preservation Actions, for example the migration of complex objects that contain multiple file formats.

## **4 Conclusions**

The conclusions described in chapter 6 of PP/7 D1 are re-iterated here with some minor changes, which are noted.

1. Although the OAIS model is a widely adopted ISO standard, it is clear that there are a number of issues with the Preservation Planning concepts and their articulation. These issues are highlighted when a comparison is made with work based on more recent digital preservation thinking, such as Planets. OAIS lacks explicit definitions of key terms (for example: Prototyping) and does not include sufficient description or explanation of the Preservation Planning functions.

*Still valid – minor changes to text.*

2. While the main focus on preservation is found in chapter 4.1.1.6, key preservation concepts are discussed in other chapters. At the time OAIS was authored, these concepts were based predominantly on theory and consequently were not clearly defined. Further research and practical experience have progressed preservation practice since this time. It should now be possible to draw together what were theoretical concepts into clearly articulated functions and processes supported by practical experience. Ideally this should also extend to the requirements for OAIS compliance.

*Still valid – minor changes to text.*

3. Key preservation processes within OAIS currently lie outside of the scope of the Preservation Planning functional entity, for example Migration (in ADMINISTRATION). Further consideration needs to be given to this arrangement and the possibility of moving these to the Preservation Planning function.

*Still valid – no changes.*

4. The interactions between preservation functions and mechanisms for managing preservation metadata (Representation Information Networks) could usefully be described in OAIS, as these appear to be quite integral to the key preservation functions.

*Still valid – minor changes to text.*

5. Important preservation functions appear to have been omitted from OAIS. In some cases this is because implicit notions have not been made explicit. For example Characterisation can be seen as a result of the OAIS rule “Obtain sufficient control of the information provided to the level needed to ensure Long term Preservation”.

*Still valid – no changes.*

6. OAIS places too much emphasis on migration, and not enough on other preservation strategies. Subsequent phases of PP7 may be able to explore how other preservation strategies such as migration on request or emulation can be adequately represented within the OAIS model. Modelling

preservation actions that occur at the point of access requires further thought, and careful coordination with the Preservation Action sub project.

*Still valid – no changes (note that this is expected to be addressed in the OAIS Pink Book revision). Development of a preservation on access scenario is in progress and will be included in the next iteration of this work.*

7. The use of the term “standards” within section 4.1.1.6 of OAIS requires clarification.  
*Still valid.*
8. The Preservation Watch Function brings together several monitoring functions and in this way will be able to combine this information into an appropriate set of requirements for Preservation Planning purposes. In the OAIS model such a coordinating function seems not to be present. Further investigation of this possibility, informed by the practical experiences of the PP work-packages, is required.  
*The discussions with the Planets colleagues from different institutions highlighted that Preservation Watch is seen as a valuable, coordinating function for digital preservation and is a vital function in the model.*
9. The Planets scope is tightly focused on supporting the preservation of digital objects, while the OAIS model also pays attention to the preservation of the repository itself.  
*This function lies outside the scope of the Planets Project, but is now present in the model as Repository Longevity.*
10. The OAIS model refers to (re-)defining packaging designs as a result of Preservation Planning actions. This is an area which has not been addressed in Planets.  
*This topic will be raised at Planets program management for future development.*
11. The Planets Functional View may place too much emphasis on Characterisation as a key preservation function. Further thought needs to be given to incorporating Characterisation functions elsewhere. Characterisation of newly ingested objects could be moved to an expanded Generate Descriptive Info function within INGEST.  
*Characterisation is now part of Preservation Action.*
12. Extraction of significant properties metadata to facilitate evaluation of Preservation Actions could be subsumed into the Preservation Action function (preliminary thinking from PA4 work-package matches this conclusion).  
*These concerns were shared by our colleagues and the model was updated to incorporate Characterization into Preservation Action.*
13. Planets should consider identifying a wider prototyping capability, in line with an element of the OAIS Monitor Technology function. This might draw together elements of Preservation Watch, Preservation Planning and the Planets Testbed facility.  
*This is has now been incorporated as the Testbed sub function.*

## **5 Further Work**

A further revision of the Planets Functional Model and comparison with OAI is scheduled for the end of the Planets Project. This will be informed by further consultation across the Project. We expect that this timing will allow valuable practical experiences from applying Planets tools and services to partners' content to be taken into account (for example via input of initiatives of the IPAL(Implementing Planets into Archives and Libraries) group in Planets and similar initiatives for other environments. It is expected that an additional part of this final PP/7 report will work backwards from the Planets Functional Model and show where Planets tools and services (such as the Planets Core Registry, the Plato tool, the Planets Testbed and the Characterization Tools) might support the identified preservation functions in the Planets Functional Model.