



Project no. 033572

CASPAR

Cultural, **A**rtistic and **S**cientific knowledge for **P**reservation, **A**ccess and **R**etrieval

Instrument: Information Society Technologies

Thematic Priority: 2.5.10 Access to and preservation of cultural and scientific resources

CASPAR SYSTEM DEVELOPMENT – OVERALL MASTER PLAN



Document identifier:	CASPAR-D1302-TN-0101 -1_1
Submission Date:	30-11-2007
Due Date:	20-11-2007
Work package:	1300
Partners:	CASPAR Architecture Team
WP Lead Partner:	ENG
Document status	FINAL

Delivery Type

Report





Author(s) **CASPAR Architecture Team:**
 STFC - Registry
 FORTH - Knowledge Manager
 ENG - Preservation Orchestration Manager
 STFC - RepInfo Toolbox
 IBM - Preservation Data Store
 MW - Data Access Manager and Security
 MW - Digital Rights Manager
 CNR - Finding Aids
 ACS - Virtualisation
 STFC - Packaging
 UU - Authentication Management

Approval David Giaretta

Summary CASPAR Implementation Plan – Addendum to the D1301 “Overall Component Architecture and Component Model”

Keyword List CASPAR Key Components – CASPAR Implementation Plan

Availability **PUBLIC**

Document Status Sheet

Issue	Date	Comment	Author
0_1	13 November 2007	Initial draft bringing together contributions	Luigi Briguglio, Pasquale Andriani
0_1_3	14 November 2007	Revision based on feedback from FORTH, CNR, MW and UU	Luigi Briguglio
0_1_4	14 November	Revision based on feedback from IBM, CNR, FORTH, STFC (Packaging), ACS, UU	Pasquale Andriani
0_2	16 November 2007	Revision and refinement of all contributions. Release to CASPAR Architecture Team for peer-review	Luigi Briguglio
0_3	19 November 2007	Revision with CASPAR Architecture Team feedback	CASPAR Architecture Team
0_4	20 November 2007	Overall revision and final release to DG	Luigi Briguglio
1_0	20 November 2007	Final version	David Giaretta
1_1	30 Nov 2007	Rename “System Development – overall master plan” following suggestion from PO.	David Giaretta





Project information

Project acronym:	CASPAR
Project full title:	Cultural, Artistic and Scientific knowledge for Preservation, Access and Retrieval
Proposal/Contract no.:	IST-2006-033572

Project Officer: Carlos Oliveira

Address:	<p>INFSO-E3 Information Society and Media Directorate General Content - Learning and Cultural Heritage</p> <p>Postal mail: Bâtiment Jean Monnet (EUFO 1167) Rue Alcide De Gasperi / L-2920 Luxembourg</p> <p>Office address: EUROFORUM Building - EUFO 1167 10, rue Robert Stumper / L-2557 Gasperich / Luxembourg</p>
Phone:	+352 4301 33052
Fax:	+352 4301 33190
Mobile:	
E-mail:	Carlos.Oliveira@ec.europa.eu

Project Co-ordinator: David Giaretta

Address:	<p>STFC (formerly CCLRC), Rutherford Appleton Laboratory</p> <p>Chilton, Didcot, Oxon OX11 0QX, UK</p>
Phone:	+44 1235 446235
Fax:	+44 1235 446362
Mobile:	+44 (0) 7770326304
E-mail:	d.i.giaretta@rl.ac.uk





CONTENT

1	INTRODUCTION	5
1.1	HOW TO READ THIS DOCUMENT	5
1.2	GLOSSARY.....	6
2	OVERALL MASTER PLAN.....	7
3	IMPLEMENTATION PLAN - PHASE 1.....	9
3.1	REGISTRY.....	9
3.2	KNOWLEDGE MANAGER.....	10
3.3	PRESERVATION ORCHESTRATION MANAGER.....	11
3.4	REPINFO TOOLBOX.....	11
3.4.1	<i>Virtualisation</i>	11
3.5	PRESERVATION DATA STORE	11
3.6	DATA ACCESS MANAGER AND SECURITY	14
3.7	DIGITAL RIGHTS MANAGER.....	15
3.8	FINDING AIDS.....	15
3.9	PACKAGING	16
3.10	AUTHENTICITY MANAGEMENT.....	16
4	IMPLEMENTATION PLAN - PHASE 2.....	18
4.1	REGISTRY.....	18
4.2	KNOWLEDGE MANAGER.....	18
4.3	PRESERVATION ORCHESTRATION MANAGER.....	18
4.4	REPINFO TOOLBOX.....	19
4.4.1	<i>Virtualisation</i>	19
4.5	PRESERVATION DATA STORE	19
4.6	DATA ACCESS MANAGER AND SECURITY	20
4.7	DIGITAL RIGHTS MANAGER.....	20
4.8	FINDING AIDS.....	21
4.9	PACKAGING	21
4.10	AUTHENTICITY MANAGEMENT.....	21
5	IMPLEMENTATION PLAN - PHASE 3.....	23
5.1	REGISTRY.....	23
5.2	KNOWLEDGE MANAGER.....	23
5.3	PRESERVATION ORCHESTRATION MANAGER.....	23
5.4	REPINFO TOOLBOX.....	24
5.4.1	<i>Virtualisation</i>	24
5.5	PRESERVATION DATA STORE	24
5.6	DATA ACCESS MANAGER AND SECURITY	24
5.7	DIGITAL RIGHTS MANAGER.....	25
5.8	FINDING AIDS.....	25
5.9	PACKAGING	26
5.10	AUTHENTICITY MANAGEMENT.....	26
6	IMPLEMENTATION PLAN - PHASE 4.....	27
6.1	REGISTRY.....	27
6.2	KNOWLEDGE MANAGER.....	27
6.3	PRESERVATION ORCHESTRATION MANAGER.....	27
6.4	REPINFO TOOLBOX.....	27
6.4.1	<i>Virtualisation</i>	28
6.5	PRESERVATION DATA STORE	28
6.6	DATA ACCESS MANAGER AND SECURITY	28
6.7	DIGITAL RIGHTS MANAGER.....	29
6.8	FINDING AIDS.....	29
6.9	PACKAGING	29
6.10	AUTHENTICITY MANAGEMENT.....	30
	GLOSSARY	31





1 INTRODUCTION

The objective of this document, as “addendum” to the **D1301**, is to present the **CASPAR** overall implementation plan. This plan follows the **CASPAR** Overall Component Architecture and Component Model [**D1301**], and it’s produced with additional input from the “Specification Refinement” of each **CASPAR** Key Component.

As shown in [**D1301**], the **CASPAR** team has adopted an iterative process for the development of the **CASPAR** software platform. In that deliverable, an overall plan for the development has been produced. That overall plan is shown in chapter 2.

This document provides the details (i.e. inputs, outcomes, objectives, test and conditions, dates) to show how we expect to reach and verify the success of each iteration.

It is expected that revisions of this document will be produced following each iteration, to add further clarification and details for the coming iterations. At any point the immediately following iteration will be quite clear but any following iterations will have less detail; this will allow for the fact that our understanding of the requirements will develop as we deliver each iteration.

It is worth remembering that the key components, for the most part, provide a domain independent infrastructure to support digital preservation. The individual tools, to which, for example, the RepInfo Toolkit provides access, have the most immediate link to specific domains, although even there a good number of these tools will each be usable across many domains. Development of the individual tools will be linked to a sub-set of scenarios related to those in [**D4101**].

In the final 6 months of the project there will be increasing integration of all components into the running of the testbeds, implementing the scenarios described in [**D4101**], demonstrating full system integration and providing evidence for the validation metrics described in the [**DoW**].

The deliverable is completed by:

- *References*, which provides a reading list used during the preparation of this deliverable.
- *Glossary* represents the important phase of fixing terminology encountered during this first phase of the **CASPAR** Project.

1.1 HOW TO READ THIS DOCUMENT

This document assumes that the reader is familiar with the Open Archival Information Systems (OAIS) Reference Model. It is also essential to have read the **CASPAR** Overall Component Architecture and Component Model [**D1301**], **CASPAR** Guidelines [**D1202**] and the **CASPAR** Conceptual Model [**D1201**]. Some familiarity with the **CASPAR** Description of Work [**DoW**] would be useful.





Applicable documents and reference documents

Applicable documents

- [A1] Description of Work, April 2006
- [A2] Risk Form

Reference documents

- [R1] CASPAR proposal, Sept 2005
- [DoW] CASPAR Description of Work
- [D1101] D1101 Review of the State of the Art
- [D1201] D2101 CASPAR Conceptual Model
- [D1202] D1202 CASPAR Guidelines
- [D1301] D1301 CASPAR Overall Component Architecture and Component Model
- [D4101] D4101 User Requirements and Scenario Specifications

1.2 GLOSSARY

[Ax]	Applicable Document
[Rx]	Reference Document
CASPAR	Cultural, Artistic and Scientific knowledge for Preservation, Access and Retrieval
DoW	Description of Work
EC	European Commission
EPM	Executive Project Management
IPC	IP Coordinator
IST	Information Society Technologies
PACP	Partner Administrative Contact Point
PO	Project Officer
PPR	Project Progress Report
PQE	Project Quality Engineer
PTCP	Partner Technical Contact Point
RegRep	Registry/Repository
R&D	Research and Development
SQE	Stream Quality Engineer
ST	Stream
TN	Technical Note
WP	Work Package
WPL	Work Package Leaders





2 OVERALL MASTER PLAN

The following table represents the overview of the **CASPAR** Implementation Plan, its 4 main phases and the involved **CASPAR** Key Components defined in [D1301]. The broad scheme is to develop individual prototypes of components, followed by adding increasing functionality, deployment and interoperation of these components.

Key Component	Phase			
	1	2	3	4
Registry	Initial implementation with RepInfoLabel and initial load of RepInfo. This is being done in association with the UK Digital Curation Centre (DCC)	Trial federation capabilities. Working with Gap Manager	Release 2 and interface to a number of other registry systems	Release 3 of registry with installation kit and interoperability with other components
Knowledge Manager	Basic SWKM Services rel. 0.1	Provide Intelligibility-aware service	Second version of GapManager plus integration with POM	KM Utilities to support knowledge evolution
Preservation Data Store	Implementation on Object Storage Device (OSD) and eXtensible Access Method (XAM) interface	Integration with Packaging and possibly other CASPAR components	Suggest translation mechanisms of existing simple file systems and simple existing archive data stores to PDS interface	Enable scalability and prototype policy interface for some of the features needed in preservation.
RepInfo (incl. Virtualisation) Tools	Low level RepInfo for Scientific data and initial Virtualisation for simple objects	Integration of tools to view RepInfo and RepInfo creation for a variety of digital objects	Tools for description of semantic information. Initial “expert system” for providing suggestions	Virtualisation tools and guide integrated. Deployment in a variety of environments.
DAMS	Simple user management and user authentication.	Extended user management, management and retrieval of policies.	Enforcement of policies and simple preservation of DAMS.	Management, authorization and preservation of property-based Authorized Communities.
DRM	Simple Copyright core ontology.	License definition, distribution and enforcement, simple licensing ontology.	Rights enforcement, simple preservation of DRM and simple legal framework ontology	Complete rights ontology and complete preservation of DRM.
Finding Aids	Complete specification refinement (architectural document). Design of a first Finding Manager that uses SWKM to store, and query DescInfo objects. Finding Manager mockup.	Implementation of FM based on SWKM. Integrate FM with DescInfos from test bed communities. Integrate FA with PDS and/or Packaging in order to store true DescInfo objects associated to true AIPs.	Develop Finding Registry component.	Intelligent Finding Aids using DAMS, DRM, and Authenticity.
Packaging	Develop simple packaging tool to implement XFDU packaging.	Integration with PDS using package Testbed information objects.	Implementation of Validation and Transformation of Information Packages.	Complete integration with other components, Preservation of PACK.
Orchestration	Prototype a simple notification and alert system	Integration with DAMS and installation of Classification schemes for registration of interests	Integration with KM	Full alert system release.
Authenticity	Models for authenticity protocols and refinement specification of Authenticity Management	Refinement and Customisation of Authenticity Steps	Validation and test of authenticity procedures in specific testbed(s).	Release of Authenticity tools





According to that overall implementation plan, the CASPAR Architecture Team has added further details and dates, allowing to provide the following implementation Gantt diagram.

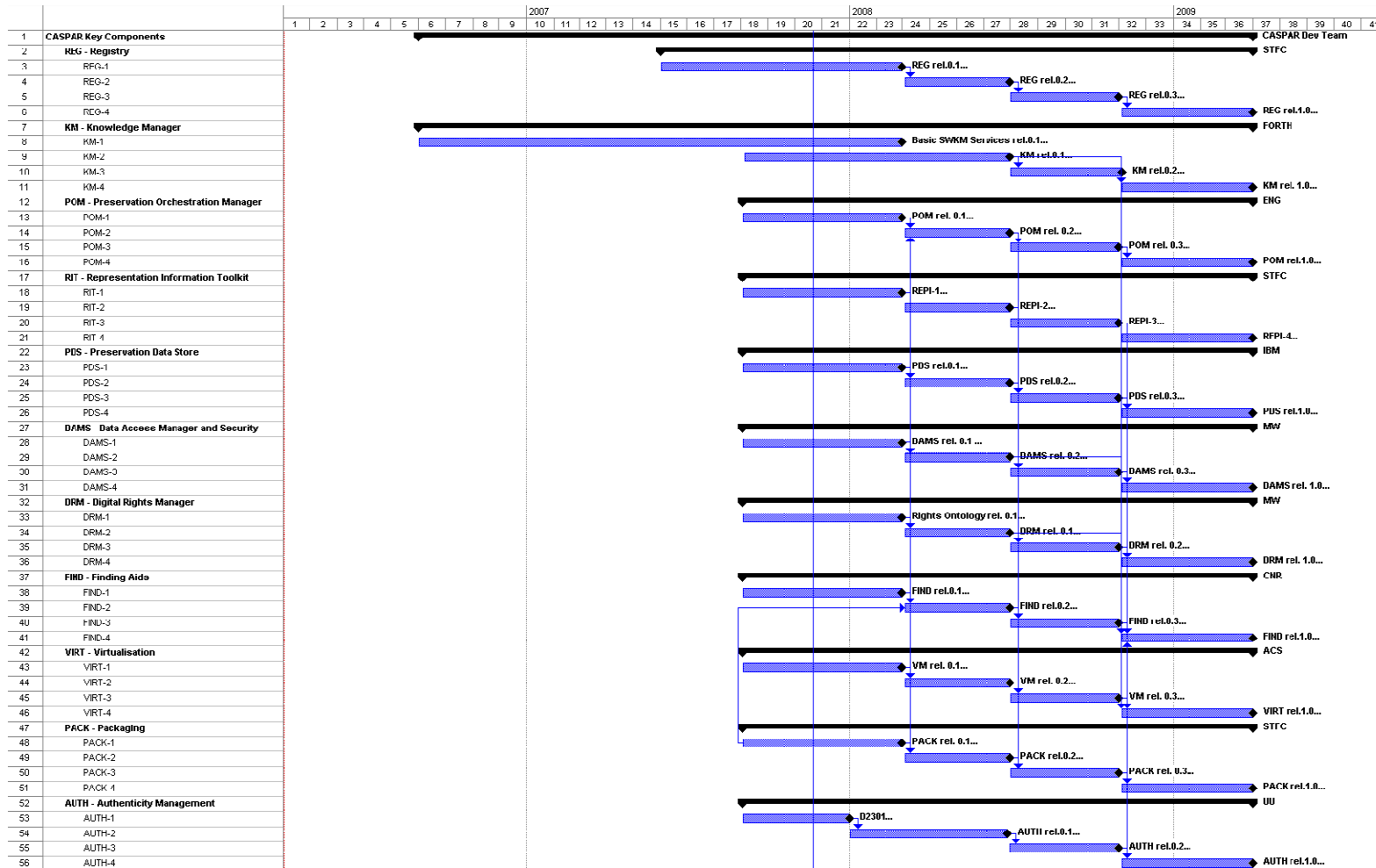


Figure 1 Schedule of iterations





3 IMPLEMENTATION PLAN - PHASE 1

3.1 REGISTRY

Input(s)	Existing REGISTRY Prototype based on ebXML, D1300 Component requirements		
Objective(s)	Align and develop REGISTRY prototype for CASPAR use		
Outcome(s)	<ul style="list-style-type: none"> ● CASPAR Registry 1.0 API ● Registry client API ● Refined registry label creation and editing tool using JGraph to make it more intuitive. ● Simple Registry maintenance tool. ● Registry simple Ingest tool 		
Description of Work	Take existing Registry Repository implementation and realign and re-factor for CASPAR		
Test and Condition	Test against other Registry Ingest and access components <ul style="list-style-type: none"> ● Retrieve RepInfo given a CPID ● Submit RepInfo to Registry ● Search for RepInfo – simple text search of name and/or description, and search by one or more Classification schemes ● Manage registry e.g. users, Classification schemes 		
Start Date	M18	End Date	M23

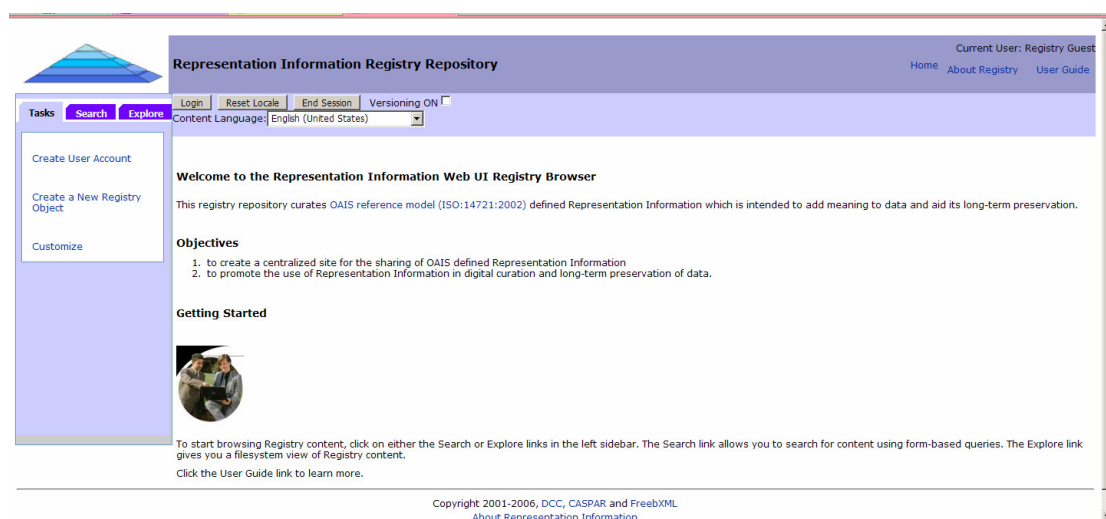


Figure 2 Initial tailoring of freebXML registry





Representation Information Registry Repository

Current User: Registry Guest

Home About Registry User Guide

Tasks Search Explore

Log In Reset Locale End Session Versioning ON

Content Language: English (United States)

Select Predefined Query: Basic Query

Basic query for selecting registry objects by name, description and classification

Object Type: RepresentationInformation

Name: [input field]

Description: %EAST%

Status: [dropdown]

Select Classification Node...

Search Clear Help

Use % as wildcard character

Registry Objects

Registry Objects Help

Apply Approve Deprecate Undeprecate Bookmark Relate Delete Deletion Option: Delete Object and Repository Item

Set or Change Status Select Status

Results 1 - 6 of 6

Bookmark And Relate Help

Pick	Details	Object Type	Name	Description	Version	Version Comment	Content Version	Content Version Comment
<input type="checkbox"/>	Details	DescriptionLanguageSpecification	EAST	The Data Description Language Enhanced Ada Subse T (EAST) Specification, CCSDS 644.0-B-2 Blue Book, November 2000.	1.1		1.1	
<input type="checkbox"/>	Details	DigitalFileType	Quake I save: ddm4 East side invertationa	Quake I save: ddm4 East side invertationa	1.1		1.1	
<input type="checkbox"/>	Details	FileDescription	DPS Ionosonde MMM EAST Description.	DPS Ionosonde MMM EAST Description.	1.1		1.1	
<input type="checkbox"/>	Details	FileDescription	DGS-256 Ionosonde MMM East Description.	DGS-256 Ionosonde MMM East Description.	1.1		1.1	
<input type="checkbox"/>	Details	FileDescription	DA_TC_3DP_PMOM_DESC.eas.txt	EAST description for Wind 3DP PMOM (Proton and helium on-board calculated MOMents) data files	1.1		1.1	
<input type="checkbox"/>	Details	FileDescription	BADC MST EAST DESCRIPTION	BADC MST NASA AMES CARTESIAN VERSION 2 EAST	1.1		1.1	

Figure 3 Search for RepInfo

3.2 KNOWLEDGE MANAGER

Input(s)	Analysis of the problem of managing Semantic Web metadata. Availability of ontologies for descriptive metadata (like CIDOC CRM).		
Objective(s)	Basic SWKM Services rel. 0.1		
Outcome(s)	<p>Basic SWKM Services rel. 0.1</p> <p>Options:</p> <ol style="list-style-type: none"> 1) Full installation 2) Installation of only the client. <p>The high level knowledge management services (e.g. RepInfoGap Manager) will be based on the these. In addition, the rest CASPAR components may exploit these basic services.</p>		
Test and Condition	<p>I1TC1 - Ability to build a SW repository offering persistence, validation and query and update services.</p> <p>I1TC2 - Ingestion test: one partner from the testbeds (e.g. IRCAM) uses these services in order to feed the repository with some data (e.g. with the CIDOC CRM ontology and descriptions of some indicative objects with respect to that ontology).</p> <p>I1TC3 - The component Access Manager uses these services to offer some content-based access services. E.g. provides some provenance queries assuming the CIDOC CRM ontology.</p> <p>Each of the above test scenarios will be checked by writing a number of JUnit tests. In this way it will be possible to run them periodically and automatically. This is important for ensuring the correctness of the code as the project proceeds and new functionality is added.</p> <p>More detailed test cases are described in a D2101B (http://wiki.casparpreserves.eu/bin/view/Main/D2101B)</p>		
Start Date	6	End Date	23





3.3 PRESERVATION ORCHESTRATION MANAGER

Input(s)	POM Specification: D1301, D1201, D1202, Development Guidelines; Framework Specification		
Objective(s)	Prototype a simple notification and alert system		
Outcome(s)	POM rel. 0.1 <ul style="list-style-type: none"> ● Simple notification and alert system ● Publish/Subscribe design pattern implementation 		
Test and Condition	<p>The following POM v.0.1 functionality are tested: registration of publishers and subscribers; simple demo scenario with CASPAR Key Component in the role of publisher and a human user in the role of subscriber (human user is still not a registered user managed by the DAMS). In this version, Topic and Expertise are assumed to be predefined.</p> <p>POM rel.0.1 is tested by using JUnit (unit test for each operation). Client Test Scenarios (i.e. client examples code) are defined in order to show conditions to test/use the POM functionalities.</p>		
Start Date	M18	End Date	M23

3.4 REPINFO TOOLBOX

Input(s)	Existing tools such as EAST toolset, DEBAT, DRB Selection of datasets from science testbeds Information on datasets to be described		
Objective(s)	Use of tools to describe existing real datasets		
Outcome(s)	Creation of RepInfo – Structure RepInfo and simple data dictionaries		
Test and Condition	<p>Use information about datasets to create RepInfo using one or more of the tools</p> <p>Test by confirming that the data descriptions allow us to extract correct data values, by comparing with native tools.</p>		
Start Date	M15	End Date	M23

3.4.1 Virtualisation

Input(s)	Virtualisation Specification: D1301, D1201, D1202		
Objective(s)	Develop the preliminary version of Virtualisation		
Outcome(s)	VM rel. 0.1 <ul style="list-style-type: none"> ○ Data Object description as Table and Image 		
Test and Condition	<p>Create Data Object description layered on top of data element descriptions, such as EAST, with additional description as table or image</p> <p>Test by confirming that a table can be imported into generic table application, and image into a generic image application.</p>		
Start Date	M18	End Date	M23

3.5 PRESERVATION DATA STORE

Input(s)	In the case of ingestAIP - an AIP and its packaging name		
----------	--	--	--





	In the case of accessAIP – the AIP ID to be retrieved		
Objective(s)	Implementation of AIP ingest and access on top of Object Storage Device (OSD) and eXtensible Access Method (XAM) interface		
Outcome(s)	<p>PDS rel. 0.1 which includes:</p> <p>In the case of ingestAIP – the AIP ID for the ingested AIP or null if the ingest process failed.</p> <p>In the case of accessAIP – the AIP for the given AIP ID or a notification that the given AIP ID was not found</p>		
Description of Work	<p>In the case of ingestAIP - the packaging component will call the ingestAIP service with an AIP and its packaging name. PDS will generate internal data structures for it, generate AIP ID, compute fixity, and add technical “ingest” provenance event. Then, the internal data structures are mapped to XAM XSets objects and from XSets to OSD objects. The OSD objects are stored in our Linux-based object storage.</p> <p>In the case of accessAIP - the packaging component will call the accessAIP service with an AIP ID. PDS will validate the AIP ID, validate fixity, and add technical “access” provenance event. Then, PDS will look up for the XAM XSets and OSD objects doe the given AIP and retrieve the data accordingly.</p> <p>See the pictorial description in figures 1, 2, 3 and 4.</p>		
Test and Condition	We’ll provide a based application to test the PDS first phase functionality. This test application is easily accessed (URL) and can run on any platform.		
Start Date	M18	End Date	M23

Figure 1 below depicts the web-based application that can be used to test the PDS ingest AIP functionality in phase 1.



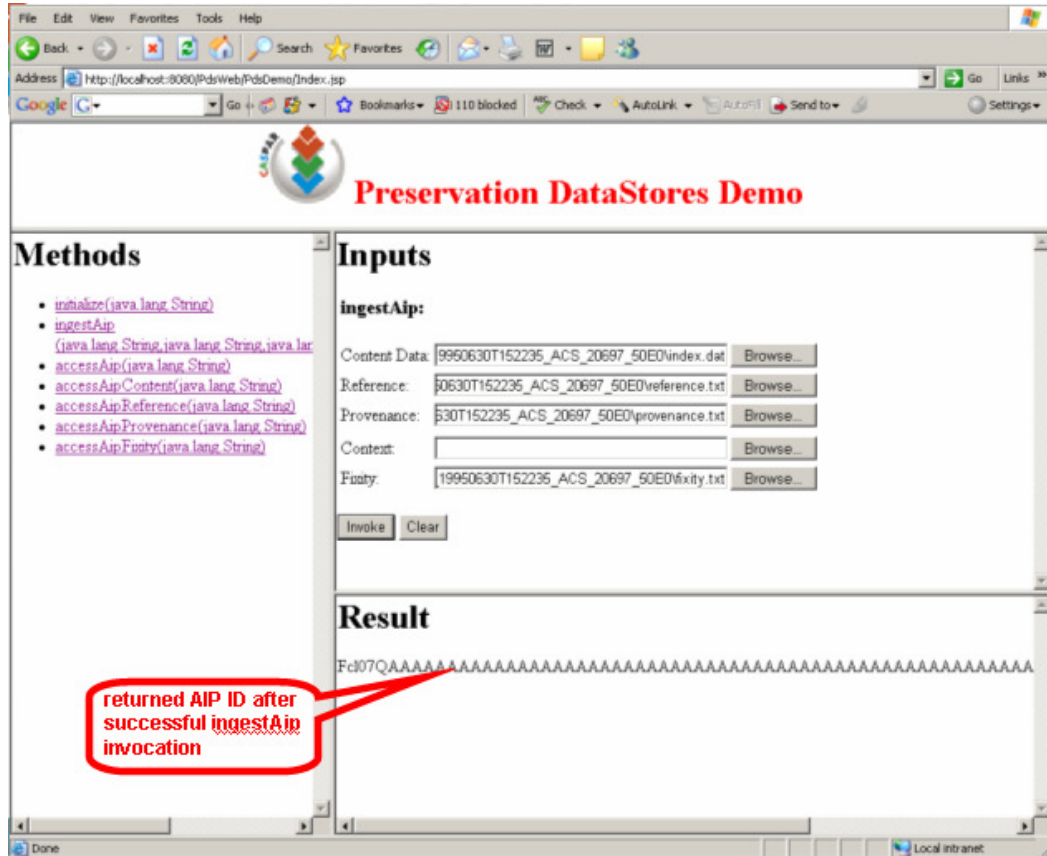


Figure 4: Ingest AIP Demonstration

Figure 2 below depicts the PDS functionality performed in a phase 1 Ingest AIP operation.

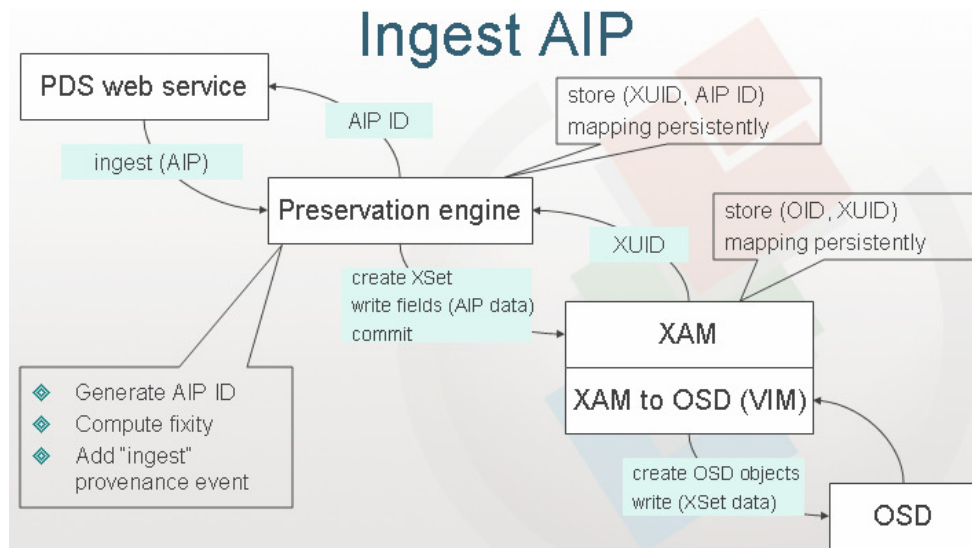


Figure 5: Ingest AIP Functionality

Figure 3 below depicts the web-based application that can be used to test the PDS access AIP functionality in phase 1.



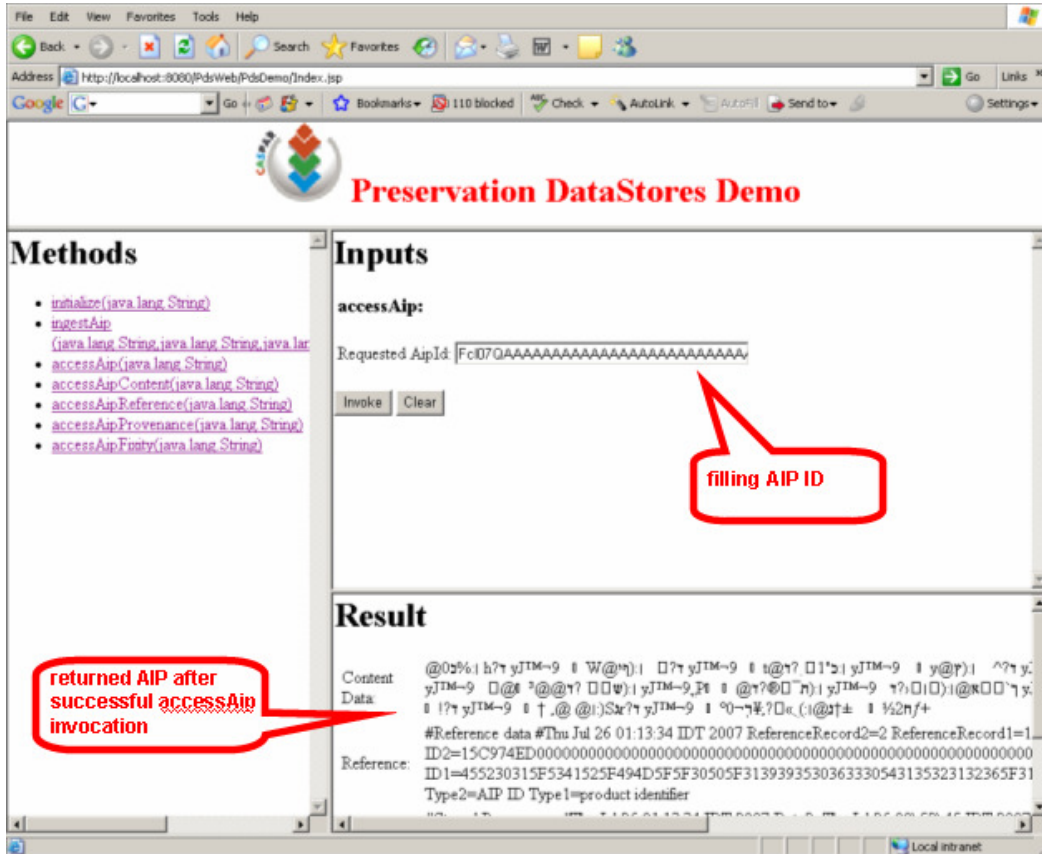


Figure 6: Access AIP Demonstration

Figure 4 below depicts the PDS functionality performed in a phase 1 Access AIP operation.

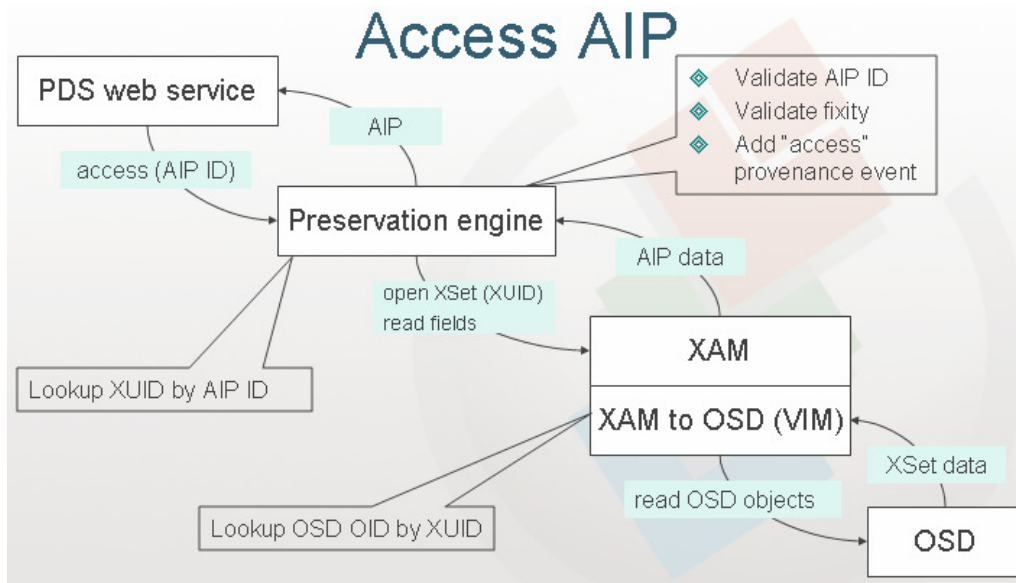


Figure 7: Access AIP Functionality

3.6 DATA ACCESS MANAGER AND SECURITY

Input(s)	DAMS Specification: D1301, D1201, D1202, Development Guidelines; Framework Specification
Objective(s)	Simple User Management and User Authentication





Outcome(s)	DAMS rel. 0.1 <ul style="list-style-type: none"> • Draft User Management (SimpleUser implementation, without AuthorizedCommunity) • User Authentication (UserPassword implementation) 		
Test and Condition	<ul style="list-style-type: none"> • User registers to the system • User retrieves the information of his profile • User updates his profile • User fails in authenticating with wrong credentials • User succeeds in authenticating with correct credentials <p>Either with real or simulated interaction with the DCProfileManager:</p> <ul style="list-style-type: none"> • User selects his DCProfile(s) • User updates his DCProfile(s) • Users of a certain DCProfile can be retrieved 		
Start Date	M18	End Date	M23

3.7 DIGITAL RIGHTS MANAGER

Input(s)	CIDOC-CRM, FRBR-OO, State of Art on copyright domain descriptions and applicable ontologies		
Objective(s)	Simple copyright ontology		
Outcome(s)	Rights Ontology rel. 0.1 <ul style="list-style-type: none"> ○ Draft Copyright core ontology 		
Test and Condition	<ul style="list-style-type: none"> ○ It is possible to describe the status of rights related to an intellectual work, in terms of instances of the Copyright core ontology concepts: i.e. the people that hold some kind of rights and which type of rights; to be evaluated by the testbeds ○ The concepts related to the intellectual property rights of a creative work must integrate and extend all the other concepts that contribute in documenting the work; this will be guaranteed if the rights ontology is consistent with the semantics of the adopted CIDOC-CRM and FRBR-OO 		
Start Date	M18	End Date	M23

3.8 FINDING AIDS

Input(s)	Report on the conceptual model of the CASPAR Finding Aids. FA specification based on SWKM: D1301, Development guidelines, Framework specification. SWKM documentation.		
Objective(s)	Complete specification refinement (architectural document). Design of a first Finding Manager that uses SWKM to store, and query DescInfo objects. Finding Manager mockup.		
Outcome(s)	FA mockup		
Description of Work	Interfaces will be fully defined and class diagram completed. A first Finding Manager mockup will be implemented based on defined design.		





Test and Condition	The following functionality will be tested: creation of DescInfo schema, storage of DescInfo, dummy AIPs discovery based on stored DescInfos.		
Start Date	M18	End Date	M23

3.9 PACKAGING

Input(s)	PACK Specification: D1301, D1201, D1202, Development Guidelines; Framework Specification Testbed data sets, Testbed Representation Information		
Objective(s)	Develop simple packaging tool to implement XFDU packaging		
Outcome(s)	<ul style="list-style-type: none"> • PACK rel. 0.1 • simple packaging toolkit system for package Construction • Packages constructed with Scientific testbed data and repInfo • Simple Integration with PDS 		
Test and Condition	<p>A) Purpose: Test the construction of simple scientific Testbed packages. Test creates packages which are complete, they must physically and logically store all the correct information objects and conform to the XFDU schema. Simple testbed packages will reference external information objects</p> <p>Method:</p> <ol style="list-style-type: none"> 1) Packages will be visually inspected by eye 2) Packages Manifests will be validated against the XFDU schema 3) Information Object checksums will be compared to those featuring in the XFDU manifest <p>B) Purpose: Test simple integration with PDS component. AIP package sent to PDS for ingest, an AIP id will be returned on successful ingest.</p> <p>Method:</p> <ol style="list-style-type: none"> 1) Packaging Component will call the PDS Components ingestAIP operation as a service, simple packages will be ingested into the PDS system 2) Packaging Component will expect returned AIP id on successful ingest 		
Start Date	M18	End Date	M23

3.10 AUTHENTICITY MANAGEMENT

Input(s)	AUTH Specification: D1301, D1201, D1202, Development Guidelines; Framework Specification		
Objective(s)	Provide specifications for tools that ensure integrity and authenticity verification of contents and contextual information along the preservation process, based on refinement specification of Authenticity Management.		
Outcome(s)	Authenticity Protocols Models and Specifications <ul style="list-style-type: none"> • Complete Specification D2301 		





Test and Condition	Release of D2301		
Start Date	M18	End Date	M23





4 IMPLEMENTATION PLAN - PHASE 2

4.1 REGISTRY

Input(s)	CASPAR Registry 1.0 PRONOM registry Gap Manager prototype		
Objective(s)	Federation with other registries Enhanced tools for RepInfoLabel creation and RepInfo submission to Registry		
Outcome(s)	CASPAR Registry 1.1		
Test and Condition	Create RepInfoLabel with CPID to more than one registry. Submit search to one Registry and obtain CPID to results in another. Provide dependencies to Gap Manager		
Start Date	M24	End Date	M27

4.2 KNOWLEDGE MANAGER

Input(s)	SWKM rel.0.1, plus the results of the first year of the project regarding intelligibility-aware services (as described in several research papers)		
Objective(s)	Provide Intelligibility-aware service		
Outcome(s)	KM rel.0.1 <ul style="list-style-type: none"> Revised version of the Basic SWKM Services First version of the Gap Manager RepInfo Gap Manager is instrumental for providing intelligibility-aware services		
Test and Condition	I2aTC1 - Ability to support some forms of knowledge evolution (recall the example of Yugoslavia). I2bTC2 - Ability to implement the examples described in this deliverable regarding intelligibility gaps. Definition of modules, profiles, dependencies, etc. Demonstrating examples that users with different profiles get different responses. More detailed test cases are described in a D2101B (http://wiki.casparpreserves.eu/bin/view/Main/D2101B) Import dependencies from Registry		
Start Date	M24	End Date	M27

4.3 PRESERVATION ORCHESTRATION MANAGER

Input(s)	POM rel. 0.1, DAMS rel. 0.1		
Objective(s)	Integration with DAMS and installation of Classification schemes for registration of expertises		





Outcome(s)	POM rel. 0.2		
	<ul style="list-style-type: none"> Integration of registered user for the subscription of expertises 		
Test and Condition	<p>The following POM v.0.2 functionality are tested: registered users (human users managed by the DASM) may have the roles of publisher/subscriber. The registration sequence (explained in the registration sequence diagram) is tested (not including the usage of KM).</p> <p>POM rel.0.2 is tested by using JUnit (unit test for each operation). Client Test Scenarios (i.e. client examples code) are defined in order to show conditions to test/use the POM functionalities.</p>		
Start Date	M24	End Date	M27

4.4 REPINFO TOOLBOX

Input(s)	RepInfo tools		
	Example datasets, including ones from Cultural Heritage testbeds		
Objective(s)	Initial RepInfo toolbox with dropdown menus providing access to set of individual tools, and advice on which may be effective		
Outcome(s)	GUI which provides access to a set of tools, creating RepInfo for selected datasets		
Test and Condition	<p>Create RepInfo using GUI.</p> <p>Confirm that the RepInfo allows the display and processing of the dataset</p>		
Start Date	M24	End Date	27

4.4.1 Virtualisation

Input(s)	VM rel. 0.1, RepInfoToolbox rel. 0.1		
	Selected datasets including a variety of images and tables		
Objective(s)	Integration with RepInfoToolbox and developing of advice on possible virtualisation techniques		
Outcome(s)	<p>VM rel. 0.2</p> <ul style="list-style-type: none"> Preliminary integration with RepInfo Toolbox (user interface definition) Preliminary advisory mechanism 		
Test and Condition	<p>Create virtualisation description of dataset from RepInfo Toolbox.</p> <p>Test by confirmation that advice can help produce data virtualisation description which allows selected datasets to be produced and examined in generic application, including semantic relationships.</p>		
Start Date	M24	End Date	M27

4.5 PRESERVATION DATA STORE

Input(s)	PDS rel. 0.1, PACK rel 0.1		
Objective(s)	Integration with Packaging and possibly other CASPAR components		





Outcome(s)	PDS rel. 0.2 in which PACK is calling PDS to ingest and access AIPs		
Test and Condition	Ability to ingest and access scientific test bed data by the PACK component		
Start Date	M24	End Date	M27

4.6 DATA ACCESS MANAGER AND SECURITY

Input(s)	DAMS rel. 0.1, KM rel 0.1		
Objective(s)	Extend User Management and Policy Management and Retrieval		
Outcome(s)	DAMS rel. 0.2 <ul style="list-style-type: none"> • Management of UserAuthorizedCommunity • Definition and Management of Access Control policies (realizing RBAC) • Retrieval of Access Control policies 		
Test and Condition	<ul style="list-style-type: none"> • A newly registered user is by default assigned to the ‘Registered User’ Authorized Community • Policy 1: The authorization manager creates a policy that grants the permission to manage Authorized Communities to the users that belong to the predefined Authorized Community ‘User manager’ • Policy 2: An Owner of an Information Object creates a policy that grants the permission to manage License Offers associated to his information object to the users that belong to the predefined Authorized Community ‘Owner’ of that object • The DRM component retrieves all policies that regulate access to its own services 		
Start Date	M24	End Date	M27

4.7 DIGITAL RIGHTS MANAGER

Input(s)	DRM Specification: D1301, D1201, D1202, Development Guidelines; Framework Specification		
Objective(s)	Definition of license offers, instantiation and verification of license instances Simple licensing ontology		
Outcome(s)	DRM rel. 0.1 <ul style="list-style-type: none"> • Rights Definition • Rights Distribution • Rights Verification Rights Ontology rel. 0.2 <ul style="list-style-type: none"> • Draft licensing concepts ontology 		
Test and Condition	DRM rel. 0.1 <ul style="list-style-type: none"> • User creates a License Offer for a certain information object • A License Instance is generated for a given consumer • A check is made to verify if the consumer has a valid License Instance for the requested object Rights Ontology rel. 0.2 (licensing concepts)		





	<ul style="list-style-type: none"> The above license offers and license instances, expressed in one of the chosen REL formalisms, can be expressed also in terms of the ontology 		
Start Date	M24	End Date	M27

4.8 FINDING AIDS

Input(s)	FA mockup, Basic SWKM Services rel.0.1. PDS ver.0.1, and Packaging ver .0.1		
Objective(s)	<p>Implementation of FM based on SWKM.</p> <p>Integrate FM with DescInfos from test bed communities.</p> <p>Integrate FA with PDS and/or Packaging in order to store true DescInfo objects associated to true AIPs.</p>		
Outcome(s)	FA ver.0.1		
Test and Condition	<p>Instantiate a FA with all information about AIPs available for test. When new AIPs are stored FA interface will be used to store corresponding DescInfo objects, and/or an ad-hoc procedure will be used to update FA repository with already stored AIPs.</p>		
Start Date	M24	End Date	M27

4.9 PACKAGING

Input(s)	PACK rel. 0.1, PDS rel 0.1		
Objective(s)	Integration with PDS using package Testbed information objects		
Outcome(s)	<ul style="list-style-type: none"> PACK rel. 0.2 Demonstration of Packaging tools using Testbed data Implementation of additional Packaging Interface operation Enhance Integration with PDS 		
Test and Condition	<p>Purpose: Test the ingest of more complex Testbed packages, packages will reference internal and external information objects. Where internal means physically stored within the package and external means referenced by URI both internal and externally referenced information objects are aggregates of the same Information Package.</p> <p>Method:</p> <ol style="list-style-type: none"> 1)Packaging Component will call the PDS Components ingestAIP operation as a service, simple packages will be ingested into the PDS system 2)Packaging Component will expect returned AIP id on successful ingest <p>Purpose: Test the unpackaging of packages. Correct and complete packages will be unpackaged into their individual information objects of which they were composed.</p> <p>Method: Packaging will provide test script to demonstrate the unpackage XFDU packages</p>		
Start Date	M24	End Date	M27

4.10 AUTHENTICITY MANAGEMENT





Input(s)	AUTH Specification: D2301, CASPAR Key Components Specs and APIs, Testbed Steps Models		
Objective(s)	Refinement and Customisation of Authenticity Steps		
Outcome(s)	AUTH rel.0.1 <ul style="list-style-type: none"> • Refinement of Authenticity Steps. Identification of requirements on other CASPAR Key Components. • Analysis and Customisation of Authenticity Steps for specific testbed(s), based on feedback 		
Test and Condition	Testbed Steps models define tests to evaluate and validate authenticity steps		
Start Date	M24	End Date	M27





5 IMPLEMENTATION PLAN - PHASE 3

5.1 REGISTRY

Input(s)	CASPAR Registry 1.1 RepInfo Gap Manager		
Objective(s)	Further refinement of Registry API – more User Tools and Registry management tools. Add interfaces to other Registry implementations. Enhanced interoperation with Gap Manager		
Outcome(s)	CASPAR Registry 1.2		
Test and Condition	Demonstrate Registry tools, including revised RepInfoLabel tool with built-in RepInfo examination and verification. Demonstrate that other registries can be accessed via standard API. Demonstrate exports to and updates of dependencies with Gap Manager		
Start Date	M28	End Date	M31

5.2 KNOWLEDGE MANAGER

Input(s)	KM rel. 0.1, SWKM, POM rel.0.2 Registry information		
Objective(s)	Integration with Preservation Orchestration Manager rel.0.2		
Outcome(s)	KM rel. 0.2 Second version of Gap Manager		
Test and Condition	I3TC1 - The registry feeds the KM Repository (with modules and dependencies). I3TC2 - The component Preservation Orchestration Manager uses the RepInfoGapManager services.		
Start Date	28	End Date	31

5.3 PRESERVATION ORCHESTRATION MANAGER

Input(s)	POM rel. 0.2, KM rel. 0.2		
Objective(s)	Integration with Knowledge Manager rel. 0.2		
Outcome(s)	POM rel. 0.3 <ul style="list-style-type: none"> • Use of simple Knowledge Management techniques for routing alerts • Integration of dependencies for the notification 		
Test and Condition	The following POM v.0.3 functionality are tested: the complete registration sequence. The dependencies from KM are used to send alerts of “warning of implied changes”. Topic and Expertise are managed as different concepts, and Expertise is used to alert “change-dependencies”. POM rel.0.3 is tested by using JUnit (unit test for each operation). Client Test Scenarios (i.e. client examples code) are defined in order to show		





	conditions to test/use the POM functionalities.		
Start Date	M28	End Date	M31

5.4 REPINFO TOOLBOX

Input(s)	Test datasets from science, cultural heritage and performing arts, together with text descriptions. Add tools to create RepInfo such as emulation		
Objective(s)	Show that user can be guided through the process of selecting and using appropriate tools for a wider variety of digital objects, including software.		
Outcome(s)	Appropriate RepInfo created for the test digital objects		
Test and Condition	Confirm that the RepInfo that has been created is correct and useful – note that some RepInfo will support automatic processing and others may not e.g. documentation.		
Start Date	M28	End Date	M31

5.4.1 Virtualisation

Input(s)	VM rel. 0.2, KM rel 0.1, RepInfoToolbox rel. 0.2		
Objective(s)	Integration with Knowledge Manager		
Outcome(s)	VM rel. 0.3 1. Use of KnowledgeManager to suggest possible Semantic/Ontology information 2. Integration refinement with the RepInfoToolbox (User interface refinement)		
Test and Condition	Confirm that the virtualisation RepInfo created following the suggestions of the tool is correct.		
Start Date	M28	End Date	M31

5.5 PRESERVATION DATA STORE

Input(s)	PDS rel. 0.2, some existing archive or file system		
Objective(s)	Suggest translation mechanisms of existing simple file systems and simple existing archive data stores to PDS interface.		
Outcome(s)	PDS rel. 0.3		
Test and Condition	Show how PDS can work with existing archive or file system		
Start Date	M28	End Date	M31

5.6 DATA ACCESS MANAGER AND SECURITY

Input(s)	DAMS rel. 0.2, POM rel 0.2		
Objective(s)	Enforcement of policies and simple preservation of DAMS		
Outcome(s)	DAMS rel. 0.3 <ul style="list-style-type: none"> • Enforcement of Access Control policies • Draft preservation of DAMS (import/export of users and policies) 		





Test and Condition	<ul style="list-style-type: none"> • Enforcement of policy 1: <ul style="list-style-type: none"> ○ A ‘registered user’ fails to create a new Authorized Community ○ A ‘user manager’ succeeds now in creating a new Authorized Community • Enforcement of policy 2: <ul style="list-style-type: none"> ○ A ‘registered user’ fails to create a new License Offer ○ The owner of the information object succeeds in creating a new License Offer • Users are exported into a (still to define) format, and then imported • Policies are exported into a (still to define) format, and then imported 		
Start Date	M28	End Date	M31

5.7 DIGITAL RIGHTS MANAGER

Input(s)	DRM rel. 0.1, rights ontology rel 0.2		
Objective(s)	Rights enforcement and simple preservation of DRM Simple legal framework ontology		
Outcome(s)	DRM rel. 0.2 <ul style="list-style-type: none"> • Rights Enforcing (Watermarking implementation) • Draft preservation of DRM (import/export of licenses) Rights Ontology rel. 0.3 <ul style="list-style-type: none"> • Draft legal framework ontology 		
Test and Condition	DRM rel. 0.2 <ul style="list-style-type: none"> • At dissemination time, a watermark is inserted into an object (the content file) carrying information about the license instance • License offers and instances can be exported/imported into/from equivalent expressions in terms of the rights ontology Rights Ontology rel. 0.3 (legal framework concepts) <ul style="list-style-type: none"> • It is possible to describe the relationships between intellectual property rights concepts and elements of the legal framework, in terms of the ontology 		
Start Date	M28	End Date	M31

5.8 FINDING AIDS

Input(s)	FA ver.0.1
Objective(s)	Develop Finding Registry component based on UDDI or DILIGENT.
Outcome(s)	FA ver.0.2
Test and Condition	A set of Finding Manager will be registered into an appropriate number of Finding Registries. Synchronization between them will be tested as well as query and browse available FMs into a CASPAR installation.





Start Date	M27	End Date	M31
------------	-----	----------	-----

5.9 PACKAGING

Input(s)	PACK rel. 0.2, PDS rel 0.2		
Objective(s)	<ul style="list-style-type: none"> • PACK rel. 0.3 • Implementation of additional Packaging Interface operation • Integration with FindingAid • Simple support for intelligibility aware packages, DIP constructed according to the selected DC profile 		
Outcome(s)	Implementation of Validation and Transformation of Information Packages		
Test and Condition	<p>Purpose: Test integration with Finding Aid component, provide Finding Aid with descriptive information for an AIP.</p> <p>Method: Packaging component will call Finding Aid component to set the descriptive information for the given AIP</p>		
Start Date	M28	End Date	M31

5.10 AUTHENTICITY MANAGEMENT

Input(s)	AUTH rel.0.1, Testbed Steps Models		
Objective(s)	Validation and test of authenticity procedures in specific testbed(s).		
Outcome(s)	AUTH rel.0.2 <ul style="list-style-type: none"> • Validation and test of authenticity procedures in specific testbed(s). 		
Test and Condition	Mapping Authenticity steps with a process language for specific testbeds		
Start Date	M28	End Date	M31





6 IMPLEMENTATION PLAN - PHASE 4

6.1 REGISTRY

Input(s)	CASPAR Registry 1.2		
Objective(s)	Deployment of Registries in multiple configurations, showing interoperability, and working with non-CASPAR registries		
Outcome(s)	CASPAR Registry 2.0		
Test and Condition	Use RepInfoLabel tool to <ul style="list-style-type: none"> • Create RepInfoLabels using a variety of Registries • Submit RepInfo to several different registries 		
Start Date	M32	End Date	M36

6.2 KNOWLEDGE MANAGER

Input(s)	KM rel. 0.2, SWKM		
Objective(s)	Support of knowledge evolution is important and it should be demonstrated.		
Outcome(s)	KM rel.1.0 <ul style="list-style-type: none"> • Diff over Knowledge Bases 		
Test and Condition	I4TC1 - It takes as input two versions of one ontology (say V1 and V2 of CIDOC CRM). The tool should be able to identify the changes and derive a set of change operations which could be applied on a knowledge repository on CIDOC CRM V1 in order to reach CIDOC CRM V2.		
Start Date	M32	End Date	M36

6.3 PRESERVATION ORCHESTRATION MANAGER

Input(s)	POM rel. 0.3, DAMS rel. 0.3, KM rel. 0.3		
Objective(s)	POM rel. 1.0		
Outcome(s)	<ul style="list-style-type: none"> • Full alert system release (integration of advanced functionality such as: import/export, persistent messages management and message preservability). 		
Test and Condition	Import/Export functionality are tested for Topics and Expertises. Solutions from the state of the art are evaluated in order to integrate advanced functionality. POM rel.1.0 is tested by using JUnit (unit test for each operation). Client Test Scenarios (i.e. client examples code) are defined in order to show conditions to test/use the POM functionalities.		
Start Date	M32	End Date	M36

6.4 REPINFO TOOLBOX

Input(s)	Advice from Orchestration manager about RepInfo needed. Variety of digital objects including data and software
----------	---





Objective(s)	Show that appropriate RepInfo can be created as required		
Outcome(s)	RepInfo can be created, new versions created and RepInfo network extended, as required		
Test and Condition	Provide a variety of advice from Orchestration Manager. Test by confirming that appropriate RepInfo can be created (where appropriate) and response provided to Orchestration Manager.		
Start Date	32	End Date	36

6.4.1 Virtualisation

Input(s)	VM rel. 0.3, KM rel. 0.3, RepInfoToolbox rel. 0.3		
Objective(s)	VM rel. 1.0		
Outcome(s)	<ul style="list-style-type: none"> Release of the object virtualisation component. 		
Test and Condition	<p>The Virtualisation assistant and associated tools can be deployed and integrated with RepInfo Toolbox</p> <ul style="list-style-type: none"> The VA is correctly operating in the environment; The RepInfoToolBox correctly calls the VA services; The VA (through its [VA]KM) correctly calls the caspar-KM keyComponent. 		
Start Date	M32	End Date	M36

6.5 PRESERVATION DATA STORE

Input(s)	PDS rel. 0.3		
Objective(s)	Enable scalability and prototype policy interface for some of the features needed in preservation.		
Outcome(s)	PDS rel. 0.4		
Test and Condition	<p>Show scalability and ingest/access of multiple objects.</p> <p>Show policy interface prototype.</p>		
Start Date	M32	End Date	M36

6.6 DATA ACCESS MANAGER AND SECURITY

Input(s)	DAMS rel. 0.3, POM rel. 0.3		
Objective(s)	Management and Authorization of <i>external groups</i> (PropertyAuthorizedCommunities), possibly involving human intervention		
Outcome(s)	<p>DAMS rel. 1.0</p> <ul style="list-style-type: none"> Management of PropertyAuthorizedCommunity (i.e. ExternalGroup) Complete Management of policies (with ExternalGroup) Advanced Authorization Management (with ExternalGroups and human intervention in the authorization process) Complete preservation of DAMS (including AuthorizedCommunity definition) 		





Test and Condition	<p>The owner of an object wants to restrict the access to it:</p> <ul style="list-style-type: none"> • he creates a PropertyAuthorizedCommunity (e.g. his childs) • he creates a policy that grants the permission to update License Offers (of his owned information object) to the users that belong to that PropertyAuthorizedCommunity • a ‘registered user’ fails in updating the license offer • a child of the owner succeeds in updating the license offer <p>Preservation of DAMS</p> <ul style="list-style-type: none"> • after the reception of an alert by the POM, the DAMS preservation manager (human) updates the definition of an Authorized Community and informs the issuers of the implicated policy rules (see Fig. 15) 		
Start Date	M32	End Date	M36

6.7 DIGITAL RIGHTS MANAGER

Input(s)	DRM rel. 0.2, Rights Ontology rel. 0.3, Knowledge Manager rel. 0.3, POM rel. 0.3		
Objective(s)	Complete preservation of DRM Complete rights ontology		
Outcome(s)	<p>DRM rel. 1.0</p> <ul style="list-style-type: none"> • Complete preservation of DRM (including Legal Framework management) <p>Rights Ontology rel. 1.0</p> <ul style="list-style-type: none"> • Refined overall rights ontology 		
Test and Condition	<p>DRM rel. 1.0</p> <ul style="list-style-type: none"> • After the reception of an alert by the POM, the DRM preservation manager (human) updates the Legal Framework ontology, and if necessary, he updates the affected license offers (see Fig. 14) <p>Rights Ontology rel. 1.0</p> <ul style="list-style-type: none"> • Addressed open issues identified in the validation of the rights ontology by the testbeds 		
Start Date	M32	End Date	M36

6.8 FINDING AIDS

Input(s)	FA ver.0.2 and DAMS ver.0.2 / DRM ver.0.1 / Authenticity ver.0.2		
Objective(s)	Integrate Finding Aids with DAMS, DRM, and Authenticity		
Outcome(s)	FA ver.1.0		
Test and Condition	<p>Access to information will be determined according to the rest of the CASPAR system.</p> <p>This way just users with appropriate roles will get access to Finding Aids functionality, e.g. look for AIP ID from querying DescInfo stored objects, store / update DescInfo objects, discovery of Finding components.</p>		
Start Date	M32	End Date	M36

6.9 PACKAGING





Input(s)	PACK rel. 0.3, PDS rel 0.3 FindingAid rel. 0.3		
Objective(s)	Complete integration with other components, Preservation of PACK		
Outcome(s)	<ul style="list-style-type: none"> ● Information Package Comparison & versioning ● Final integration of package Manager with all components Preservation of PACK		
Test and Condition	Confirm that appropriate packages can be created, submitted to storage and subsequently retrieved and unpacked.		
Start Date	M32	End Date	M36

6.10 AUTHENTICITY MANAGEMENT

Input(s)	AUTH rel.0.2, Testbed Steps Models		
Objective(s)	Release of Authenticity tools		
Outcome(s)	AUTH rel.1.0 <ul style="list-style-type: none"> ● Verification of overall consistency and completeness 		
Test and Condition	Further mapping and validation of authenticity steps with a process language on further specific testbeds		
Start Date	M32	End Date	M36





GLOSSARY

This glossary is largely taken from the OAIS Reference Model.

A

Access Aid	A software program or document that allow Consumers to locate, analyze, and order Archival Information Packages of interest.
Access Collection	A collection of AIPs that is defined by a Collection Description but for which there is no Packaging Information for the collection in Archival Storage.
Access Method	A method for retrieving an Archival Information Package based on its name or identifier, which is available to authorized users.
Access Software	A type of software that presents part of or all of the information content of an Information Object in forms understandable to humans or systems.
Access	The CASPAR entity that contains the services and functions which make the archival information holdings and related services visible to Consumers.
Ad hoc Order	A request that is generated by a Consumer for information CASPAR has indicated is currently available.
Administration	The CASPAR entity that contains the services and functions needed to control the operation of the other CASPAR functional entities on a day-to-day basis.
Archival Information Collection (AIC)	An Archival Information Package whose Content Information is an aggregation of other Archival Information Packages.
Archival Information Package (AIP)	An Information Package, consisting of the Content Information and the associated Preservation Description Information (PDI), which is preserved within CASPAR.
Archival Information Unit (AIU)	An Archival Information Package whose Content Information is not further broken down into other Content Information components, each of which has its own complete Preservation Description Information. It can be viewed as an "atomic" AIP. An example of an AIU would be a table of numbers representing temperatures in a certain region with all the associated documentation describing how and where the temperatures were measured, what instruments were used to make the measurements, who made the measurements, why they were made, what processing has been performed on the measurements and who has had custody of these measurements since they were first created, how the measurements relate to other information, how the measurements can be uniquely referenced by others, etc.
Archival Storage	The CASPAR entity that contains the services and functions used for the storage and retrieval of Archival Information Packages.
Archive	An organization that intends to preserve information for access and use by a Designated Community.
Associated Description	The information describing the content of an Information Package from the point of view of a particular Access Aid.

C

Client	An application which exchanges information with another application (see also Consumer).
--------	--





Collection Description	A type of Package Description that is specialized to provide information about an Archival Information Collection for use by Access Aids.
Common Services	The supporting services such as inter-process communication, name services, temporary storage allocation, exception handling, security, and directory services necessary to support CASPAR.
Consumer	The role played by those persons, or client systems, who interact with CASPAR services to find preserved information of interest and to access that information in detail. This can include other CASPARs, as well as internal CASPAR persons or systems.
Content Data Object	The Data Object, that together with associated Representation Information, is the original target of preservation.
Content Information	The set of information that is the original target of preservation. It is an Information Object comprised of its Content Data Object and its Representation Information. An example of Content Information could be a single table of numbers representing, and understandable as, temperatures, but excluding the documentation that would explain its history and origin, how it relates to other observations, etc.
Context Information	The information that documents the relationships of the Content Information to its environment. This includes why the Content Information was created and how it relates to other Content Information objects.
Context	The information that documents the relationships of the Content Information to its environment. This includes why the Content Information was created and how it relates to other Content Information objects.

D

Data Dictionary	A formal repository of terms used to describe data.
Data Dissemination Session	A delivered set of media or a single telecommunications session that provides data to a Consumer. The Data Dissemination Session format/contents is based on a data model negotiated between CASPAR and the Consumer in the Request Agreement. This data model identifies the logical constructs used by CASPAR and how they are represented on each media delivery or in the telecommunication session.
Data Management Data	The data created and stored in Data Management persistent storage that refer to operation of an archive. Some examples of this data are accounting data for Consumer billing and authorization, policy data, Event Based Order (subscription) data for repeating requests, preservation process history data, and statistical data for generating reports to archive management.
Data Management	The CASPAR entity that contains the services and functions for populating, maintaining, and accessing a wide variety of information. Some examples of this information are catalogs and inventories on what may be retrieved from Archival Storage, processing algorithms that may be run on retrieved data, Consumer access statistics, Consumer billing, Event Based Orders, security controls, and CASPAR schedules, policies, and procedures.
Data Object	Either a Physical Object or a Digital Object.
Data Submission Session	A delivered set of media or a single telecommunication session that provides Data to CASPAR. The Data Submission Session





format/contents is based on a data model negotiated between CASPAR and the Producer in the Submission Agreement. This data model identifies the logical constructs used by the Producer and how they are represented on each media delivery or in the telecommunication session.

Data	A reinterpretable representation of information in a formalized manner suitable for communication, interpretation, or processing. Examples of data include a sequence of bits, a table of numbers, the characters on a page, the recording of sounds made by a person speaking, or a moon rock specimen.
Derived AIP	An AIP generated by extracting or aggregating information from one or more source AIPs.
Descriptive Information	The set of information, consisting primarily of Package Descriptions, which is provided to Data Management to support the finding, ordering, and retrieving of CASPAR information holdings by Consumers.
Designated Community	An identified group of potential Consumers who should be able to understand a particular set of information. The Designated Community may be composed of multiple user communities.
Digital Migration	The transfer of digital information, while intending to preserve it, within CASPAR. It is distinguished from transfers in general by three attributes: (a) a focus on the preservation of the full information content, (b) a perspective that the new archival implementation of the information is a replacement for the old, (c) an understanding that full control and responsibility over all aspects of the transfer resides with CASPAR.
Digital Object	An object composed of a set of bit sequences.
Dissemination Information Package (DIP)	The Information Package, derived from one or more AIPs, received by the Consumer in response to a request to CASPAR.

E

Edition	An attribute of an AIP whose information content has been increased or improved from a source AIP and is therefore a candidate to replace the source AIP.
Event Based Order	A request that is generated by a Consumer for information that is to be delivered periodically on the basis of some event or events.

F

Federated Archives	A group of archives that has agreed to provide access to their holdings via one or more common finding aids.
Finding Aid	A type of Access Aid that allows a user to search for and identify Archival Information Packages of interest.
Fixity Information	The information which documents the authentication mechanisms and provides authentication keys to ensure that the Content Information object has not been altered in an undocumented manner. An example is a Cyclical Redundancy Check (CRC) code for a file.

G

Global Community	An extended Consumer community, in the context of Federated Archives, that accesses the holdings of several archives via one or more common Finding Aids.
------------------	---

I





Independently Understandable	A characteristic of information that has sufficient documentation to allow the information to be understood and used by the Designated Community without having to resort to special resources not widely available, including named individuals.
Information Object	A Data Object together with its Representation Information.
Information Package	The Content Information and associated Preservation Description Information which is needed to aid in the preservation of the Content Information. The Information Package has associated Packaging Information used to delimit and identify the Content Information and Preservation Description Information.
Information	Any type of knowledge that can be exchanged. In an exchange, it is represented by data. An example is a string of bits (the data) accompanied by a description of how to interpret a string of bits as numbers representing temperature observations measured in degrees Celsius (the representation information).
Ingest	The CASPAR entity that contains the services and functions that accept Submission Information Packages from Producers, prepares Archival Information Packages for storage, and ensures that Archival Information Packages and their supporting Descriptive Information become established within CASPAR.
K	
Knowledge Base	A set of information, incorporated by a person or system, that allows that person or system to understand received information.
L	
Local Community	The original Designated Community, in the context of Federated Archives, served by an archive.
Long Term Preservation	The act of maintaining information, in a correct and Independently Understandable form, over the Long Term.
Long Term	A period of time long enough for there to be concern about the impacts of changing technologies, including support for new media and data formats, and of a changing user community, on the information being held in a repository. This period extends into the indefinite future.
M	
Management	The role played by those who set overall CASPAR policy as one component in a broader policy domain.
Member Description	An Associated Description that describes a member of a collection.
Metadata	Data about other data.
Migration	Transfer of data from one medium to another, or of software from one hardware platform to another, or of servers from one operating system to another.
O	
Open Archival Information System (OAIS)	An archive, consisting of an organization of people and systems, that has accepted the responsibility to preserve information and make it available for a Designated Community.
Order Agreement	An agreement between the archive and the Consumer in which the





physical details of the delivery, such as media type and format of Data, are specified.

Ordering Aid	An application that assists the Consumer in discovering the cost of, and in ordering, AIPs of interest.
Overview Description	A specialization of the Collection Description that describes the collection as a whole.

P

Package Description	The information intended for use by Access Aids.
Packaging Information	The information that is used to bind and identify the components of an Information Package. For example, it may be the ISO 9660 volume and directory information used on a CD-ROM to provide the content of several files containing Content Information and Preservation Description Information.
Physical Object	An object (such as a moon rock, bio-specimen, microscope slide) with physically observable properties that represent information that is considered suitable for being adequately documented for preservation, distribution, and independent usage.
Preservation (format)	File prepared for long-term preservation
Preservation Description Information (PDI)	The information which is necessary for adequate preservation of the Content Information and which can be categorized as Provenance, Reference, Fixity, and Context information.
Preservation/technology watch	Constant review of developments that affect the long-term maintenance and functioning of a digital archive
(Preservation) Processing	Preparation of files for preservation and dissemination
Producer	The role played by those persons, or client systems, who provide the information to be preserved. This can include other CASPARs or internal CASPAR persons or systems.
Producer/Depositor/Author	Individuals, organisations or systems that transfer information for long-term preservation
Provenance Information	The information that documents the history of the Content Information. This information tells the origin or source of the Content Information, any changes that may have taken place since it was originated, and who has had custody of it since it was originated. Examples of Provenance Information are the principal investigator who recorded the data, and the information concerning its storage, handling, and migration.

R

Reception	Process under which an archive is transferred for long-term preservation
Records	Information identified upon its creation as having value as evidence in 'recording' an activity or decision and so requiring a 'lifespan' for its management either through to destruction or appraisal and retention as part of an archive
Reference Information	The information that identifies, and if necessary describes, one or more mechanisms used to provide assigned identifiers for the Content Information. It also provides identifiers that allow outside systems to





	refer, unambiguously, to a particular Content Information. An example of Reference Information is an ISBN.
Reference Model	A framework for understanding significant relationships among the entities of some environment, and for the development of consistent standards or specifications supporting that environment. A reference model is based on a small number of unifying concepts and may be used as a basis for education and explaining standards to a non-specialist.
Refreshment	A Digital Migration where the effect is to replace a media instance with a copy that is sufficiently exact that all Archival Storage hardware and software continues to run as before.
Refreshment	Systematic copying of files between used and unused media. A Digital Migration where the effect is to replace a media instance with a copy that is sufficiently exact that all Archival Storage hardware and software continues to run as before.
Repackaging	A Digital Migration in which there is an alteration in the Packaging Information of the AIP.
Replication	A Digital Migration where there is no change to the Packaging Information, the Content Information, and the PDI. The bits used to represent these Information Objects are preserved in the transfer to the same or new media instance.
Repository	Site at which digital records are stored: can be a data warehouse for records or a digital archive for archives
Representation Information	The information that maps a Data Object into more meaningful concepts. An example is the ASCII definition that describes how a sequence of bits (i.e., a Data Object) is mapped into a symbol.
Representation Network	The set of Representation Information that fully describes the meaning of a Data Object. Representation Information in digital forms needs additional Representation Information so its digital forms can be understood over the Long Term.
Representation Rendering Software	A type of software that displays Representation Information of an Information Object in forms understandable to humans.
Result Set	The set of descriptive records for those AIPs in CASPAR which match the criteria stated in a Consumer query, or other results from a search on Data Management.
Retrieval Aid	An application that allows authorized users to retrieve the Content Information and PDI described by the Package Description.
Retrieval	Recovery of files for use
Risk analysis	Assessment of effects of adverse events on long-term preservation
S	
Search Session	A session initiated by the Consumer with the archive during which the Consumer will use the archive Finding Aids to identify and investigate potential holdings of interest.
Secure (transfer/environment)	Protected from interference by non-authorised parties
Structure Information	The information that imparts meaning about how other information is organized. For example, it maps bit streams to common computer types





	such as characters, numbers, and pixels and aggregations of those types such as character strings and arrays.
Submission Agreement	The agreement reached between CASPAR and the Producer that specifies a data model for the Data Submission Session. This data model identifies format/contents and the logical constructs used by the Producer and how they are represented on each media delivery or in a telecommunication session.
Submission Information Package (SIP)	An Information Package that is delivered by the Producer to CASPAR for use in the construction of one or more AIPs.
T	
Transfer	Movement of files between organisations
Transformation	A Digital Migration in which there is an alteration to the Content Information or PDI of an Archival Information Package. For example, changing ASCII codes to UNICODE in a text document being preserved is a Transformation.
U	
Unit Description	A type of Package Description that is specialized to provide information about an Archival Information Unit for use by Access Aids.
V	
Version	An attribute of an AIP whose information content has undergone a transformation on a source AIP and is a candidate to replace the source AIP.

