RELEVANT METADATA TO PRESERVE "ALIEN" AIP

Angela Di Iorio
Fondazione Rinascimento Digitale
Via Maurizio Bufalini Firenze

Maurizio Lunghi
Fondazione Rinascimento Digitale
Via Maurizio Bufalini Firenze

ABSTRACT

This article describes the development of Archives Ready To Archival Information Packages (AIP) Transmission a PREMIS Based Project (ARTAT). Following the project approach, the starting phase consisted of prototyping a layer conveying preservation metadata, which can be encoded from the existing archival systems, and exchanged with other repositories. This layer called Preservation Metadata Layer (PML) uses PREMIS semantics as the common language to overcome archival systems differences, and to transmit out of its original context, relevant preservation information about content objects comprising an AIP. Since a repository, following the OAIS reference model, usually provides resources with metadata container objects, the experiment performed an analysis on commonly used container formats, in order to enable the traceability of semantics from a local to extra-local level, and the technological understandability of alien AIPs. The analysis has allowed the definition of a PML data model, laying the production of prototypes. The adoption of common semantics, like PREMIS, supports the opportunity of preserving correctly alien AIPs, coming from different technological environments, and hopefully enables the overcoming of obstacles to the interoperability among diverse archival systems.

1. INTRODUCTION

This article describes the development of the project named Archives Ready To AIP Transmission a PREMIS Based Project (ARTAT) [3], that took place, from March to April 2010. The goal of ARTAT is to experiment with the adoption of a common preservation metadata standard as an interchange language in a network of cooperating organizations that need to exchange digital resources with the mutual objective of preserving them in the long term. The project in pursuing its initial objectives, has experimented with the definition of a Preservation Metadata Layer (PML) following the PREMIS standard Data Dictionary (DD) specifications [8] that will integrate repositories' preservation metadata. The exported repositories' AIPs [2] including a PML will be received by selected repositories and ingested into their archival systems. Hopefully, because of the common PREMIS knowledge base, the receiving repositories will be able to locate information objects and data objects contained in the AIPs transmitted by the originating repositories.

To date, the project consisted of testing the Preservation Metadata Layer prototype produced from representative samples, selected from the initial participant repositories.

The critical path analysis, which was conducted on the PML prototypes, will be traced in order to support the ultimate objective of transmitting resources destined for preservation in a repository other than the originating repository. It is assumed that the devised layer will be agnostic about the originating archival systems, as well as about the receiving archival systems. A successful transmission can be accomplished as long as both of the repositories in the transfer can manage XML conforming to the PREMIS framework.

The milestones that will be explained below aim to test the feasibility of inventing a layer which contains all relevant information for the receiving repository to offer long term preservation services in the foreseeable future.

More information about aim, objectives, tools and methodologies of the project are available on the Fondazione Rinascimento Digitale website which supported this project and the Italian PREMIS community (http://www.rinascimento-digitale.it/projects-artat.phtml).

2. ARTAT PROJECT OVERVIEW

The ARTAT [1] project started in March 2010 with interviews conducted with the first three participants:

- ICCU's MAGTECA [2] an institutional repository which collects resources from geographically dispersed Italian cultural heritage institutions
- Magazzini Digitali (MD) [3] a project undertaken by Fondazione Rinascimento Digitale and National Library of Florence to preserve Italian doctoral theses for the long term
- The digital repository of the Library & Archive of the British School at Rome [4]

The interviews followed the inquiry phase that were reported in the project workplan.

The project aims to provide existing digital repositories with a layer of preservation metadata that is exchangeable with other repositories. The focus is not on changing existing archival systems, but rather on creating the ideal conditions for exchanging resources, strengthening their own management with a view to long

1 http://www.rinascimento-digitale.it/projects-artat.phtml
2 www.internetculturale.it
3 http://www.rinascimento-digitale.it/magazzinidigitali.phtml
4 http://digitalcollections.bsrome.it/

term preservation, and enabling opportunities for offering preservation services to third parties.

The experiment's approach is to define and test a preservation metadata layer, encoded according to the PREMIS standard.

The export of a repository's AIPs with a PML provided should enable selected repositories to receive and ingest them into their own repository systems.

The building of PML will originate from the archival management system and, through a controlled data flow, will feed the PML exchanged with other receiving repositories.

The approach taken from the beginning of the project has two phases: the first inquiry phase, where participants are interviewed about their repositories' architectures and their management of preservation metadata, and the second PML production phase, which experiments with the translation of metadata contained in AIPs into the PML layer encoded in PREMIS semantics.

3. INQUIRY PHASE RESULTS

3.1. General consideration about initial application of the questionnaires

The initial inquiry phase was conducted in March 2010 with the initial participating repositories and concluded at the beginning of April 2010. The information was obtained through interviews generally guided by semistructured questionnaires. The initial questionnaires, which focused on archival systems and preservation metadata management, were dramatically reduced during the interviews, because it became clear that in spite of all technological differences, nearly all systems mainly contain metadata useful to the preservation but they do not manage it as preservation metadata. A side effect of the interviews was to make repositories' managers aware of the risks of the lack of management of preservation metadata.

The results of the questionnaire as well as the report of information gathered during the prototyping will be published on the ARTAT website for the preservation community.

A review of the questionnaire will be conducted and submitted to the future project's partners.

3.2. Repositories technologies overview

The information gathered from the inquiry phase regarding the metadata schemas managed is summarized in Table 1. This is the basis from which we have started to address the problem of differences in standards adoption, as well as to find a solution in overcoming the interoperability issues that in practice limit the understandability of AIPs, exchanged by repositories.

Knowledge about the metadata container standard adopted by the repositories is an important starting point of the experiment. Analyzing the application and composition of the containers used, and the comprehensiveness of information gathered inside, is important in order to structure correctly the PML description.

Institution/ Project	Metadata type	XML Schema name	Version
ICCU	Container	MAG	1.0-2.01
	Descriptive	DC simple	1.1
	Technical	MIX	0.1 draft
MD	Container	MPEG21-DIDL	-
	Descriptive	DC simple	1.1
	Technical	Jhove	1.5
	Technical	MIX	0.2
BSR	Container	METS	1.9
	Descriptive	MODS	3.3
	Descriptive	DC simple	1.1
	Technical	MIX	2.0

Table 1. Metadata schemas used by the interviewed repositories

The evidence of semantics adopted in metadata containers is useful to the likely exchange scenario, allowing the data conversion, from the repositories internal structure to PREMIS [5].

At the end of the repository's inquiry phase, repository managers were asked to submit a sample metadata object encoded in XML that is representative of their AIPs.

4. THE PRESERVATION METADATA LAYER (PML)

4.1. PML Target

The target of preservation is the information package defined as the AIP in the OAIS conceptual model. This package, in actual applications, consists of content and metadata. Without consideration for how the AIP is managed by the archival systems or whether the metadata encoded in XML is used to support one or more OAIS process (submission, archival, dissemination), the focus of PML is the XML metadata files. In particular, the PML target is all files that package different metadata categories together in a formally declared structure, and that usually are defined as metadata containers, like for example the METS files. More specifically "A *container* is the unit for aggregating the typed metadata sets, which are known as *packages"* [7]. In ARTAT approach these files will be considered as objects, conforming to the PREMIS data model specifications. As Metadata Container Object (MCO) is meant the container file object that can bind different types of metadata objects and content objects together, by means of the embedding or referencing mechanism.

The MCO samples coming from the participating archival systems will be submitted to the PML prototyping process. The outcomes will confirm the feasibility of translating the system internal AIP into an

"exchangeable AIP", which in ARTAT terms, means an AIP provided with a PML.

The PML is essentially a translation of the content and the relationships among the constituent objects (metadata and content) of an AIP.

4.1.1. Metadata Container Objects

Usually, metadata containers are used to package different types of metadata and can fulfill different OAIS functions. The interoperability difficulties that arise when containers are used in contexts outside of their original archival systems are well known. These difficulties are caused by differences in structural design and in different levels of granularity of metadata application.

An MCO can be used by repositories to support the various functions specified in the OAIS conceptual model [2]. An MCO is a composite that can contain a diverse set of structured information conforming with formally specified semantics. As such, it is a purpose-specific object type. Usually, metadata containers are intended to bundle various types of metadata that describe the resource from different points of view. For example, METS is a widely used XML container format that wraps metadata types in well-circumscribed sections. METS can contain information about objects, both content and metadata, which is embedded (mdWrap) or referenced (mdRef) in some way.

Finally, to support their implementation, container standards have bindings in XML schema which may organize information quite differently from its original structure.

The characteristics of MCOs in use can be a significant factor when an exchange involves different MCO standards. Consequently, exchange packages derived from local MCOs, need to be structured with a common and well defined set of information, overcoming the local coding practices and constraints [6].

4.2. PML Structure

In ARTAT, a Preservation Metadata Layer (PML) will be added to the AIP by the originating repository, which needs to translate its native information into PREMIS semantic units. PREMIS was not originally designed to be a transmission format, but in ARTAT it is used to exploit the knowledge base focused on preservation metadata and founded on a well known model.

The PML is composed of two parts: the PML core and the PML redundant part, which together describe technically and structurally the AIPs content.

The PML core is the part which essentially translates the container's relevant metadata into PREMIS semantic units. The translation consists of a mapping from the original administrative, technical, provenance, rights and structural information into the PREMIS framework.

The PML redundant part simply describes the content objects in PREMIS terms mapping information like *objectidentifier*, *compositionlevel*, *fixity*, *size*, *format*, *originalName*, and *storage* from the object's related metadata.

The PML consists of one or more PREMIS files connected by internal and external identifiers, and connected by reference, to the AIPs' metadata and content objects.

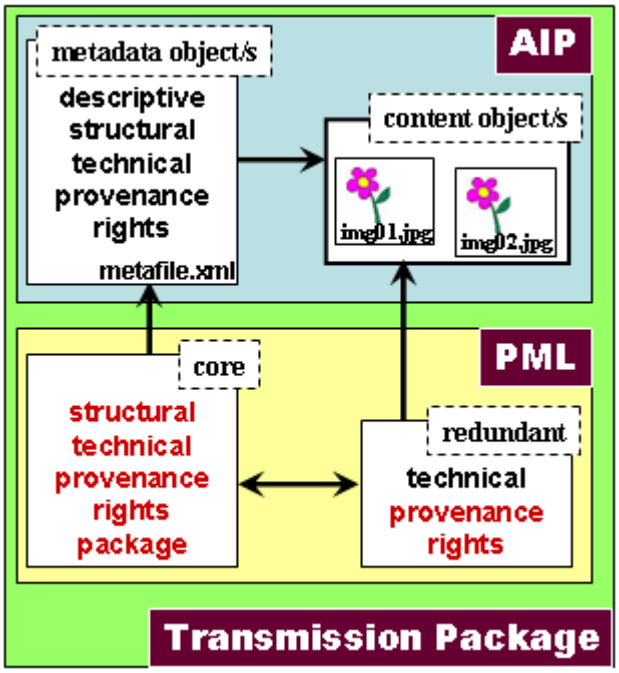

Figure 1. Transmission Package structure composed of Archival Information Package and Preservation Metadata Layer.

4.3. PML Coding and requirements

The PREMIS metadata standard was selected for the PML because it is strictly focused on preservation metadata and because it has been widely implemented in the international preservation community. The choice was made on the assumption that the standard is built on well-defined semantics and a well-known data model, so it ought to be conducive to interoperability at organizational and technological levels.

The ARTAT project defined three main requirements for the PML. The first requirement is PREMIS conformance, which requires: following the specifications of PREMIS Data Dictionary names and definitions for semantic units, adhering to Data Dictionary applicability guidelines, conforming to repeatability and obligation stipulations, and using mandatory semantic units as the minimum amount of metadata useful to preserve digital objects in the long-term. The second requirement is to provide PREMIS metadata as comprehensively as possible, in order to facilitate the receiving repository correctly understanding the PML, since the originating repository could have some missing or implicit preservation metadata. The third requirement is the independence of

the PML from the AIPs, making its reuse easier and its preservation feasible in different technological contexts.

4.4. PML application context

The cooperative context held by the agreement among different partners that manage diverse archival systems is the ideal application context where AIPs can be exchanged in order to share the preservation responsibility or also to provide or receive third party preservation services.

In this context the project predicted the transmission scenario (par.4.5) where AIPs are provided with PML by the originating repository which makes the "translation in" PREMIS code. The whole package, AIP and PML, is transmitted to the receiving repository system which acquires and "translates out" the PML and archives the objects as its own AIP.

The transmission package is the set of the XML formatted original AIP (content objects and metadata objects) and the preservation layer as PML (core and redundant) which is the translation part understandable by the different systems.

The cooperative context will be supported as much as possible by the adoption of common controlled vocabularies in order to translate the PML. The adoption of controlled vocabularies, as well as shareable nomenclature systems, for example the agent information, will facilitate the automatically encoding of the precompiled set of the PREMIS semantic units.

4.5. PML Transmission scenario

The AIP with metadata translated into a PML will constitute the transmission package. The transmission will happen in some formally established way, where the agreements' terms will be explored in further investigations.

The originating repository *A* which holds the AIPs performs the PML "translation in". The receiving *B* repository performs a PML "translation out", which consists of reading PML core metadata, detecting the MCO structure, the AIP's metadata, and the content objects and their relationships. Finally *B* pieces together the PML jigsaw, interpreting the original AIP and creating a new *B* MCO corresponding to *A*'s MCO in its own archival system, which will manage all of the original AIP objects plus the original *A* MCO. *B* MCO is connected to *A* MCO by means of the digital provenance information (events and agents) and objects' relationships.

This mechanism was devised to avoid the loss of information, which is natural when you make a mapping from one standard to another.

In the envisaged context, the receiving repository will act on the alien AIP just to complete the migration and to preserve its integrity and authenticity or to perform other predetermined preservation actions. In this scenario, it is not supposed to make any modification of original AIP or MCO but only integrate the AIP. All the events that have affected the AIP's objects will be recorded in some way as *B* MCO conforming to the PREMIS Data Dictionary specifications.

Two other possible transmission scenarios are:

- transmission back to the originating repository *A*: the transmission is performed in the same way, *B* makes a PML "translation in" of its AIP. The resulting PML should contain the same structure as the former transmission *A* original AIP and MCO, plus *B* MCO integrated with the events that occurred in the elapsed time. *A* makes a PML "translation out" of PREMIS information differences, which occurred in managing or updating actions;
- transmission forward to other receiving repositories *C*, *D*…: *B* MCO with relative AIP (objects plus original *A* MCO) is translated in PML; *C* translates out *B* MCO recording the *C* MCO digital provenance (from *B* MCO, from *A* MCO); *C* translates in its AIP, *D* translates out *C* MCO recording the *D* MCO digital provenance (from *C* MCO, from *B* MCO, from *A* MCO) and so on.

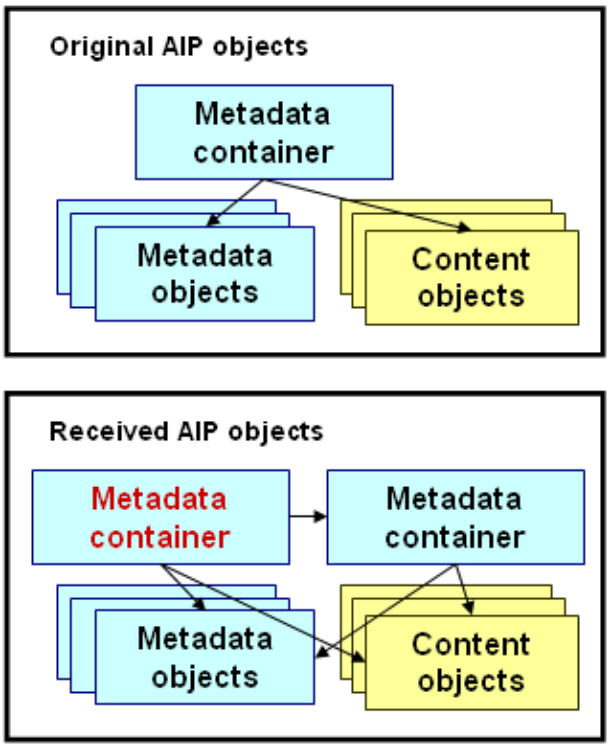

Figure 2. Differences in transmitting repositories of the Archival Information Packages.

5. METADATA CONTAINER OBJECTS ANALYSIS

The PML prototypes built from the sample files of metadata objects obtained from the participating repositories, were realized through the following milestones.

5.1. Samples' analysis process

An analysis of samples was performed in order to verify the existence of all necessary elements for building the PML encoded in PREMIS and to comply with requirements.

The sample files are encoded in three different metadata containers: the most common in the digital library community, METS[1]; the multimedia framework MPEG21-DIDL[2]; and the Italian application profile MAG[3].

Despite the containers' differences in the information framework architecture, the samples analyzed contain at least one descriptive section well circumscribed. The structural metadata are gathered in a formally defined section or in hierarchical elements, structurally added.

The administrative information usually is scattered in different sets that can be delimited in a fragmented way as technical, provenance, or rights. Despite the fragmentation, the presence of these sets of metadata should be considered obligatory in transmission contexts, even though the MCO XML schema doesn't require them as mandatory. Conforming to the obligation rules declared in the MCO schemas, a METS document can have only one structural section, MAG can have descriptive and only some of administrative metadata and in MPEG21-DIDL it is sufficient to declare only a didl:Item element. This is obviously not sufficient to describe a digital resource from a preservation point of view, but actually the repositories use containers in a sufficiently exhaustive way to describe their resources.

5.2. Samples' analysis results

Considering lessons learned in the transfer context of TIPR (par.6), and the necessary maintenance of metadata quality at a non-local level, the analysis has detected the existence of the mandatory PREMIS DD semantic units as well as the lack of or the inefficiency of information at a cooperative level. The following list is a draft of the information areas where ARTAT has to make metadata integration in order to cover cooperative needs:

- the object's identifier system, has to be refined and customized in order to identify unambiguously objects, agents, events in a nomenclature system recognizable by all ARTAT partners;
- the rights declared into three samples referred to access conditions for the resource as whole. The METS samples, the copyright information was replicated in both the descriptive section and in the METS rights section. The rights in the PML core will cover the rights and permissions about the transmission package since more detailed rights and permissions applied to the single objects will be replicated into the PML redundant. A shareable rights framework system has to be developed in order to supply the needs around third party preservation;
- events information is managed by archival systems but are not yet implemented in the MCO consequently events semantic units will be integrated at the first provision of the PML;
- the agents are not provided homogeneously but will be added automatically from the partners' nomenclature system.

6. LESSONS LEARNED FROM TIPR PROJECT

The goal of the Toward Interoperable Preservation Repositories (TIPR) [1] project is to experiment with the transfer of complex digital objects between dissimilar preservation repositories that need to be able to exchange copies of AIPs with each other. The ARTAT project has similar objectives but the application context is slightly different, because it cannot rely on the knowledge base of a single container format like METS. For this reason the PREMIS translation methodology has been adopted to overcome the interoperability issues due to the differences in the container adoption.

The development of ARTAT has taken into account the issues and the outcomes obtained by the transfer test of TIPR outlined in the referred article [1]. The TIPR requirements are: 1) based upon METS and PREMIS, 2) exchange package flexible, agnostic about the internal structure of AIPs, 3) exhaustive at package and representation level, 4) selected information must be understood by the receiving repository.

The TIPR approach is to define a common exchange package format, the Repository eXchange Package (RXP) where certain information critical to digital preservation must be, not only stored, but also understood based on the concept that a meaningful exchange can be achieved with semantic interoperability.

The information gaps that emerged from TIPR transfer tests results and ARTAT lessons learnt are:

- TIPR found information pertaining to the exchange package (history, description, and high level rights) must at this time be recorded at the intellectual entity level, because the highest level of object describable in PREMIS is a representation object. The PML core gathers events and rights at the exchange package level;
- both TIPR and ARTAT found problems with the unambiguous identification of entities;
- details about RXP composition by the source repository – relationships' information of PML core;
- how a packages will be transferred from source to target repository - devising partnership's agreement and transmission conditions applicable to the massive transmission of AIPs;
- actions to be performed - providing a common controlled vocabulary about actions that must be selected at PML production time and associated with agents;

[1] http://www.loc.gov/standards/mets/

[2] http://mpeg.chiariglione.org/standards/mpeg-21/mpeg-21.htm

[3] http://www.iccu.sbn.it/genera.jsp?id=267

- rights and permissions - rights framework system;
- archiving and preservation treatment - partnership's agreement level;
- financial and legal aspects of agreement - should be provided in ARTAT partnership agreement.

These lessons learned have affected the following PML data model.

7. PML DATA MODEL

The PML data modelling milestone consists of a selection of metadata elements from the PREMIS DD.

The data model in this context can be defined also as an obligation model, because it summarizes the mandatory elements necessary for AIP transmission.

Conforming to the PREMIS DD specifications, the mandatory semantic units pertaining to objects will be obligatorily used for the PML core and for every object's information, and replicated into the PML redundant part: *objectIdentifier*, *objectCategory*, *objectCharacteristics*, *storage*.

In the *objectCharacteristics* container, the optional semantic units *fixity* and *size* are considered mandatory for AIPs transmission, in the cooperative preservation context. These semantic units are considered useful because they allow the receiving repository to compare the original objects characteristics information to that processed by its own archiving system on the translated AIP.

Even though the actual prototyping did not use the digital signatures, this PREMIS metadata container might be considered mandatory for future transmission tests, to support the assessment process of the origin and the integrity of packages transmitted.

The semantic units pertaining to Agents are considered mandatory to identify the originating repository, as well as the receiving repository, in order to trace the chain of responsibility. All agents' semantic units will be supplied automatically, thanks to the ARTAT partners' nomenclature system.

The semantic units pertaining to Events *(eventidentifier, eventype, eventDateTime eventDetail, eventOutcomeInformation)* are all mandatory to describe the event history of the objects. The first version of the PML will include events' records will be produced, detailing this operation. Further events information should be provided if existing systems are integrated with events management functions.

Considering the transmission objective, the PML rights at PML core level will include the following semantic units: *rightbasis* (by default a "license" where all terms of the agreement are defined), *licenseInformation* which specifies metadata about license document and *rightsGranted* which specifies the actions that receiving repository can perform on AIPs.

The PML data model design and the anticipated transmission scenario, led the project to the early belief that significant properties and relationships are critical for conveying the structure of AIPs. The particular role played by these elements, will require more tasks focused on ascertaining the correct communication of the AIP's internal structure.

7.1. Significant properties of metadata container objects

Since the target of the PML core is the MCO, the actual literature about characterization of digital objects was consulted in order to identify the significant properties of the MCO. The latest outcomes from the INSPECT[1] project, which gathered and leveraged all the former projects on this topic like CEDARS[2], CAMILEON[3], DELOS[4], CASPAR[5], PLANETS[6] etc., were found to be extremely useful.

As defined by the INSPECT project significant properties are "The characteristics of digital objects that must be preserved over time in order to ensure the continued accessibility, usability, and meaning of the objects, and their capacity to be accepted as evidence of what they purport to record". It is is evident that MCOs are themselves digital objects that encompass all relevant information needed to make referred objects accessible, meaningful, authentic and reliable. In our context, where information has to be not only transmitted but also properly interpreted from other systems, it would be useful to subject the MCO to the INSPECT workflow analysis [4], in order to convey the significant properties to third parties.

The workflow consists of three sets of activities: Objects analysis, Stakeholder analysis, and Reformulation. MCO analysis and experiment will be detailed in the coming months, but a draft of the ongoing activities of this task is showed in Figure 3.

Some of the steps of objects analysis are listed here:

- *Identify the purpose of technical properties:* Considering the INSPECT categories, the *content* of MCO is XML text; the *context* is the environment, where the participants manage metadata and its exchange; the *rendering* is considered the recreation of an AIP in a recipient repository by means of a translated MCO, where metadata values and relationships among metadata objects and content objects are replicated in a new container; the *structure* is metadata which contains information about intra-relationships and inter-relationships; the *behaviour* is how the information object is connected to other metadata or content objects (i.e. the mdRef for external metadata files used in METS).

[1] http://www.significantproperties.org.uk/
[2] http://www.ukoln.ac.uk/metadata/cedars/papers/aiw02/
[3] http://www2.si.umich.edu//CAMILEON/
[4] http://www.delos.info/
[5] http://www.casparpreserves.eu/
[6] http://www.planets-project.eu

- *determine expected behaviours:* Limiting the analysis to the transmission context, where a source and a recipient have to exchange AIPs between their heterogeneous archival systems, the stakeholders involved in transmission of AIPs are repositories' systems that have to be able to make an interpretation of the alien AIPs and to ingest them as their own AIPs. This particular "user" with a well defined objective may wish to perform the following main activities: selecting information relevant to preservation, interpreting technically the selected information, and understanding the relational structure conveyed.

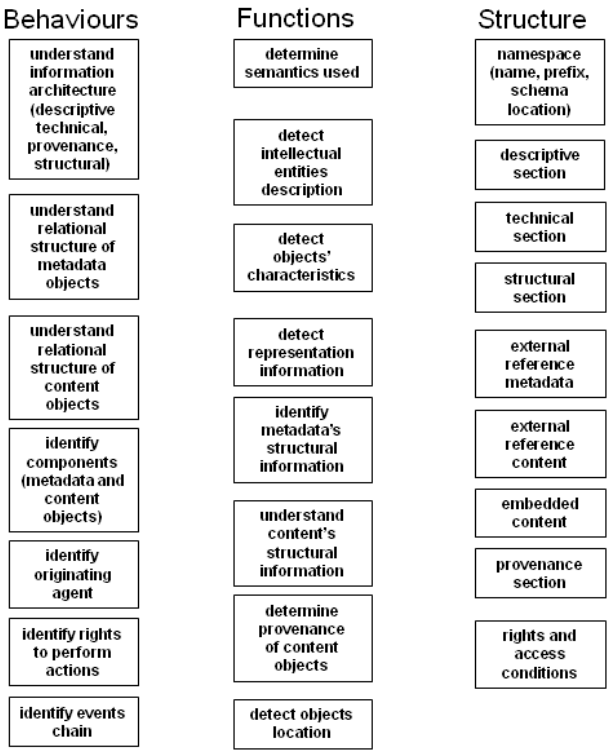

Figure 3. Draft of INSPECT workflow for Metadata Container Objects.

The premise underlaying the future experiment on the MCO are the following.

The hypothetical MCO should contain information about the schema used, validation outcome, authenticity, complex inter-relationships with other metadata container objects and intra-relationships with content objects and other metadata objects (i.e. technical metadata externally referred).

Furthermore, the need to determine two types of information has been recognized: 1) information created by the originating repository that is intended to transmit to the receiving repository; 2) information establishing the provenance of an AIP indicating its purpose and the processes through which it was created and transmitted.

The authenticity and integrity of the MCO has to be maintained, in order to demonstrate that the MCO exchanged is what it purports to be. Consequently the identification of the originating repository as well as the receiving repository/repositories are important information, because the MCO is used for a specific purpose. Also the digital provenance information guarantees the continued authenticity in the future.

The experiment has not yet been in practice performed, but will consist of the production of MCOs encoded in different metadata container standards. The MCOs will be submitted to the related participating repositories that manage the same container format, in order to test the feasibility of translation, and the exhaustiveness of significant properties as determined by the applied INSPECT framework analysis.

The outcomes of submission to repositories of the proposed MCO, resulting from the INSPECT analysis, will drive the revision of PML data model.

7.2. AIP's Relationships modelling

The MCO intra and inter-relationships with content and metadata objects will be described by means of a structured set of information. The *relationshipType* semantic unit has been defined for recording the conceptual connection among pieces of information: descriptive, structural, technical, provenance and rights.

In addition the value "referencing" was taken into account for outlining the simple reference to a content object or a metadata object.

At this time the following values have been defined for *relationshipSubType*:

relationSubType
external metadata/content
internal metadata/content
metadata wrapper

The Figure 4 shows graphically how the relationships [9] between metadata and content objects can be defined.

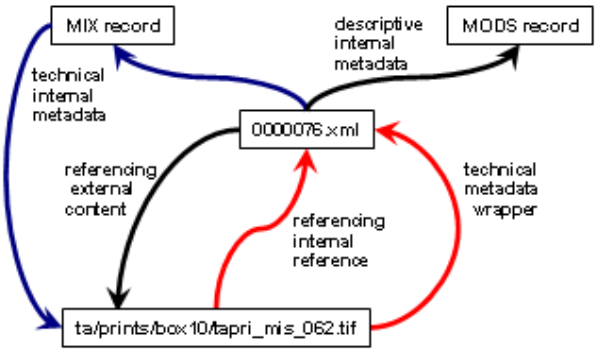

Figure 4. Relationships' prototyping.

To provide more information about the metadata schema used, the prefix and version has been tentatively added, but it is supposed that this information is related in some way to the significant properties of the MCO.

The figure below shows the corresponding simplified PREMIS code for the PML core and PML redundant.

PML core
objectIdentifier: 0000076.xml **relationshipType:** descriptive **relationshipSubType:**internal metadata:MODS 3.3 **relatedObjectIdentification:** MODS0000076 **relationshipType:** technical **relationshipSubType:** internal metadata:MIX 2.0 **relatedObjectIdentification:** tif-138 **relatedObjectIdentification:** ta/prints/box10/tapri_mis_062.tif **relationshipType:** referencing **relationSubType:** external content **relatedObjectIdentification:** ta/prints/box10/tapri_mis_062.tif

PML redundant
objectIdentifier: ta/prints/box10/tapri_mis_062.tif **relationshipType:** referencing **relationshipSubType:** internal reference **relatedObjectIdentification:** 0000076.xml **objectIdentifier:** ta/prints/box10/tapri_mis_062.tif **relationshipType:** technical **relationshipSubType:** metadata wrapper **relatedObjectIdentification:** tif-138

The prototypes and experiments will drive the refinement of significant properties and relationships data model.

8. FUTURE DEVELOPMENTS

Even though many information units still require more investigation and the PML data model is still far from being finalized, the PML prototypes implemented in XML PREMIS semantic units will be published on the ARTAT website in late July 2010. Depending on the availability of the repositories technologists, tests will be performed on the understandability of the prototypes transmitted into their systems. In autumn 2010, the project will publish results about all workflows tested on the first participants. Hopefully, in the next year the ARTAT framework will be ready to welcome new participants.

Other developments on controlled vocabularies, MCO significant properties and relationships modelling will be integrated during the developing activities, as well as the feasibility of adopting semantic web technologies which could empower the shared preservation metadata for project's partners' advantage.

9. AKNOWLEDGEMENTS

The authors would like to thank Laura Ciancio of ICCU and Luca Lelli MAGTECA technologist, Giovanni Bergamin of National Library of Florence and Raffaele Messuti of Magazzini Digitali and Alessandra Giovenco of BSR who collaborated enthusiastically during the inquiry phase.

10. REFERENCES

[1] Caplan, P., Kehoe, W., Pawletko, J. (2009). "Towards Interoperable Preservation Repositories (TIPR)", *Proceedings of the iPRES 2009: the Sixth International Conference on Preservation of Digital Objects*, San Francisco, USA, 2009. Retrieved from: http://escholarship.org/uc/item/5wf5g5kh

[2] CCSDS, January 2002. *Reference Model for an Open Archival Information System (OAIS). CCSDS 650.0-B-1, Blue Book* (the full ISO standard). http://public.ccsds.org/publications/archive/650x0b1.pdf

[3] Di Iorio, A. (2009). "A Translation Layer to Convey Preservation Metadata", *Proceedings of the iPRES 2009: the Sixth International Conference on Preservation of Digital Objects*, San Francisco, USA, 2009. Retrieved from: http://escholarship.org/uc/item/4219t4n1

[4] Grace, S. (2009). "Investigating the Significant Properties of Electronic Content over Time final report", King's Inspect Project, JISC, College London, The National Archives, http://www.significantproperties.org.uk/inspect-finalreport.pdf

[5] Guenther, R., Wolfe, R. (2009). "Integrating Metadata Standards to Support Long-Term Preservation of Digital Assets: Developing Best Practices for Expressing Preservation Metadata in a Container Format", *Proceedings of the iPRES 2009: the Sixth International Conference on Preservation of Digital Objects*, San Francisco, USA, 2009. Retrieved from: http://escholarship.org/uc/item/0s38n5w4

[6] Jackson, A. (2006), "Preliminary Recommendations for Shareable Metadata Best Practices White Paper", Digital Collections and Content Project Grainger Engineering Library, University of Illinois at Urbana-Champaign. http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.121.7095&rep=rep1&type=pdf

[7] Lagoze, C., Lynch, C., Daniel, R.Jr., (1996). "The Warwick Framework: A Container Architecture for Aggregating Sets of Metadata", Cornell Computer Science Technical Report TR96-1593. http://cs-tr.cs.cornell.edu:80/Dienst/UI/2.0/Describe/ncstrl.cornell%2fTR96-1593

[8] PREMIS Editorial Committee, 2008. *PREMIS Data Dictionary for Preservation Metadata version 2.0*. March 2008. http://www.loc.gov/standards/premis/v2/premis-2-0.pdf

[9] W3C Working Group Note (2006), "Defining N-ary Relations on the Semantic Web", www.w3.org/TR/swbp-n-aryRelations