



Project no.507618

DELOS

A Network of Excellence on Digital Libraries

Instrument: Network of Excellence

Thematic Priority: IST-2002-2.3.1.12

Technology-enhanced Learning and Access to Cultural Heritage

DELIVERABLE 5.0.1

“Final status report for Work Package 5”

JPA2 Period January 2005 – February 2006

Due date of deliverable: 31st January 2005

Actual submission date: 10th March 2006

Start Date of Project: 01 January 2004

Duration: 48 Months

UKOLN, University of Bath

Final

Project co-funded by the European Commission within the Sixth Framework Programme
(2002-2006)

Dissemination Level: [PU (Public)]

Final Status Report for Work Package 5

Authors: Liz Lyon, Manjula Patel, UKOLN, University of Bath, Stavros Christodoulakis, Polyxeni Arapi, Nektarios Moumoutzis, TUC, Sarantos Kapidakis, IU, Martin Doerr, FORTH, Dolores Iorizzo, Imperial College, Doug Tudhope, University of Glamorgan.

Document status

Date	Version	Modifications
03-Mar-2006	Draft	Circulated to Cluster partners.
09-Mar-2006	Draft	Minor revisions (Lyon).
09-Mar-2006	Draft	Revisions (TUC)
10-Mar-2006	Draft	Minor revisions (Lyon)
10-Mar-2006	Final	Approved by reviewers

TABLE OF CONTENTS

Preliminary Status Report for Work Package 5	ii
1 Executive Summary	1
2 Introduction and Scope	1
3 Information Repositories and Open Archives	2
3.1 <i>Main achievements</i>	2
4 e-Learning Applications and Digital Libraries	4
4.1 <i>Main achievements</i>	4
4.1.1 Interoperability framework and architecture	6
4.1.2 The instructional ontology	12
4.1.3 The Learner Information Model (LIM)	15
5 Ontology-driven interoperability	17
5.1 <i>Main achievements</i>	18
6 Management, Dissemination and Outreach	22
6.1 <i>Publications and Presentations</i>	23
7 JPA3 plans	28
7.1 <i>Interoperability of eLearning Applications with Digital Libraries proposal</i> ...	28
7.1.1 Development of an Earth Sciences' digital library according to the T5.4 eLearning environment for the provision of eLearning in Geography	28
7.1.2 Semantic description and interoperability of 3D objects and 3D scenes with eLearning and other digital library applications	29
7.2 <i>Ontology-driven interoperability proposal</i>	30
8 Integration: intra-cluster and inter-cluster	31
9 Conclusions	31
10 Partners	32
10.1 <i>Athens University of Economics and Business</i>	33
10.2 <i>ETH, Swiss Federal Institute of Technology, Zurich (Switzerland)</i>	33
10.3 <i>FORTH, Crete (Greece)</i>	33
10.4 <i>Imperial College, London</i>	33
10.5 <i>Ionian University</i>	33
10.9 <i>MTA Sztaki DSD</i>	34

10.6	<i>Netlab Knowledge Technologies Group, Lund University (Sweden)</i>	34
10.7	<i>NTNU</i>	34
10.8	<i>School of Electronics and Computer Science, University of Southampton (UK)</i>	34
10.9	<i>School of Informatics, University of Edinburgh (UK)</i>	34
10.10	<i>Technical University of Crete (Greece)</i>	34
10.11	<i>UKOLN, University of Bath (UK)</i>	35
10.12	<i>University of Athens</i>	35
10.13	<i>University of Glamorgan</i>	36
10.14	<i>UNIMI, University of Milan (Italy)</i>	36
11	References	36

1 Executive Summary

The work of the cluster has progressed during the reporting period with the following outcomes:

- Three new partners have joined the cluster: the University of Glamorgan, UK, Imperial College, London and the Athens University of Economics and Business.
- In Task 5.1, the 9th Delos Thematic workshop on *Digital Repositories: Interoperability and Common Services* was held in Heraklion, Crete on 11-13 May 2005.
- The Digital Repositories Report (D5.1.1) has been delayed due to staff ill health and conflicting core funder (JISC) commitments at [UKOLN](#). It is now rescheduled and due June 2006.
- Six internal deliverable reports (on *Interoperability framework and architecture*, *Semantic mapping between SCORM and MPEG7 concepts*, *Investigation of content packaging options*, *Personalization*, *Learning Styles*, and *Instructional ontology*) have been produced relating to aspects of interoperability between Digital Libraries and e-Learning applications.
- The *GraphOnto* tool has been identified as a key integration element for the KESI cluster and for the wider task activities of the Delos network.
- In Task 5.5, the CIDOC-CRM / FRBR Harmonisation work has produced a core ontology and draft detailed model.
- Work has begun on the mappings and preparatory modelling for the cultural heritage demonstrator.
- A total of 30 published papers and invited (high-profile) international conference presentations have been produced. The cluster work has been promoted in the US.
- The strategy for JPA3 activity has been largely to build on existing work.

2 Introduction and Scope

This report aims to summarise and synthesise the progress during the second Joint Programme of Activities (JPA2) to date (January 2005-february 2006) of the Knowledge Extraction and Semantic Interoperability (KESI) cluster.

Currently, the KESI cluster has the following goals for JPA2:

- (i) To complete foundation work carried out during the first 12 months, describing the state-of-the art in the three task areas of information repositories, knowledge extraction and semantic interoperability
- (ii) To develop prototypes which demonstrate the application of data models, metadata standards, metadata schema profiles and schema mappings for enhanced semantic interoperability

- (iii) To work in the domains of e-Learning and cultural heritage to inform the user requirements, design, testing and evaluation of the prototype systems.

Whilst providing current factual descriptions of the activity underway in each of the three Task areas in JPA2 (Information Repositories and Open Archives, e-Learning and Digital Libraries, Ontology-driven Interoperability), the report also aims to be reflective and critical in approach. Within the scope of the KESI cluster, it attempts to address some of the relevant emergent issues both in the European context and in the wider Digital Library (DL) environment but also within the Delos Network itself. These include dissemination and practical take-up of KESI work, promoting links between Delos activity and other global DL initiatives and facilitating intra- and inter-cluster integration. An indication is also given of preliminary plans for work during JPA3.

The Report is structured with a section on each of the three Task areas followed by sections on Management, Dissemination and Outreach, Intra- and Inter-cluster integration, JPA3 plans and finally, some Conclusions are presented. There is also an Executive Summary.

3 Information Repositories and Open Archives

This task T5.1 has been carried forward from JPA1. It has the objective of delivering a “Digital Repositories Forum” which will support discussion of particular topics which will contribute to an evaluative study on the development and implementation of community repositories to support research (institutional, national, e-prints, subject/disciplinary, species, scientific and other e-data) and learning & teaching (institutional, national learning objects and other materials). The study will also be informed by a Thematic Workshop. The Task Leader is UKOLN with partners including Southampton, Imperial, [FORTH](#) and [TUC](#) (See Partners section 10)..

3.1 Main achievements

Through the KESI cluster partner University of Southampton, Delos sponsored the Berlin 3 Open Access conference 28th February- 1st March 2005 : *Progress in Implementing the Berlin Declaration on Open Access to Knowledge in the Sciences and Humanities*. More details of the programme, international speakers and participants can be found at <http://www.eprints.org/events/berlin3/index.html>.

The highlight in 2005 from this Task was the 9th Delos Thematic workshop on *Digital Repositories: Interoperability and Common Services* held in Heraklion, Crete on 11-13 May 2005. This Workshop was planned and organised by UKOLN jointly with the Preservation cluster with a joint Programme Committee. The keynote was given by Sandy Payette, Co-Director Fedora Project, Cornell University, USA “*Rethinking the role of repositories in scholarly communication*” with submitted research papers and breakout groups contributing to Forum discussion. The Proceedings have been published and are available on the cluster Web site <http://delos-wp5.ukoln.ac.uk/> which is linked to the Delos site. Details of individual papers are given in Section 6.1. The Workshop was attended by 33 delegates from a wide range of institutions (including from two national libraries), different backgrounds (including from the UK National eScience Centre and

Poznan SuperComputing Centre) and nations, including two US representatives from the California Digital Library.

Work has also begun on the Report (D5.1.1) associated with this Task but it has been delayed for reasons stated in the Periodic Management Report provided to the Commission¹. One of the reasons for the delay has been the launch of the new £4M Digital Repository programme (see http://www.jisc.ac.uk/index.cfm?name=programme_digital_repositories) funded by the UK Joint Information Systems Committee (JISC), who are one of the core funders of UKOLN. In this role, UKOLN took responsibility for hosting and providing programme support for the projects and this has had a significant adverse impact on effort available to complete this Report. However this has been addressed, and the current position is that a Revised Table of Contents has been agreed (see below), existing contributions will be updated and the report is now due in June (Month 30).

Table 1. D5.1.1 Revised Table of Contents (February 2006)

Section	Title	Contributors	Status
1	Introduction (typology, characteristics, publishing context)	UKOLN	
2	Infrastructure & architectures (SOA context and DR applications)	UKOLN, Southampton, Imperial	
3	Relation to scholarly knowledge cycle/research cycle	UKOLN, FORTH	
4	e-Learning repositories	TUC	
5	Multimedia repositories	Imperial, Southampton	Needs update
6	Access, Use and Semantic Interoperability	FORTH, UKOLN,	Needs update
7	Summary and Recommendations	UKOLN	

¹ WP5 – Note 1

Deliverable D5.1.1 was originally planned to be included in JPA2. When we were asked by the Commission to reduce the number of JPA2 deliverables, this one was selected for cancellation, mainly due to a serious illness of Rachel Heery during 2004 and 2005 (off work for long periods). However, given the relevance of the subject, it was decided to complete the study in any case, even if it was not formally part of JPA2. Recently, UKOLN has taken on responsibility for the support of the JISC Digital Repository Programme and this activity (clearly beneficial to the DELOS Digital Repository work) is suggesting to further delay the production of the Report to month 30, in order to include some results of the cooperation with JISC. This report is a key piece of work, very topical and likely to be influential in the DL community and it should as complete as possible.

This major new JISC initiative is highly significant in many ways: it will provide and indeed is already providing, a wealth of information and experience which will inform the content of the Task Report, it signals the degree of perceived importance attached to the development of digital repositories within the education sector in the UK and it has influenced the development of a new European DL proposal *Digital Repository Infrastructure Vision for European Research (DRIVER)*, which aims to develop and deliver an infrastructure of networked repositories in Europe, and includes some KESI cluster partners.

There is much related research and development underway at UKOLN and Southampton (see Publications and Presentations Section 6.1), and the final report has the potential to be influential both in Europe and beyond, given the high levels of interest and investment in this field.

4 e-Learning Applications and Digital Libraries

Task 5.4 is exploring the interoperability of e-Learning applications and Digital Libraries looking particularly at data models, standards and workflow. The aim is to study the major standards for e-Learning (e.g. SCORM) and audio-visual (A/V) content description (e.g. MPEG7 etc.), produce mappings and following this analysis, develop an integration framework or architecture and build a demonstrator based on this architecture. The demonstrator will be tested on different classes of users: Digital Library users and distance learners. The Task leader is TUC with partners Ionian University and UKOLN.

4.1 Main achievements

The Task has sought to answer the following questions:

- What are the major architectural requirements and workflows for effectively supporting eLearning applications running over digital libraries?
- What are the major interoperability requirements for DL and eLearning standards?
- What are the management requirements and tools for audiovisual material and 3D object representations, which form the basis for many collections of learning resources?

The task is focussing on the design and implementation of appropriate tools which can be deployed across the wider DL practitioner community. Initial work has addressed developing models for an architectural framework and for workflow, producing mappings and transformations between relevant metadata standards, implementing the *GraphOnto* tool, implementing aspects of the architecture and documenting the issues in a series of reports.

A number of *Internal Deliverables* have been written covering the objectives of the task, achieving better collaboration between the partners, better organization and flexibility. These Reports are available in the private area of the cluster Web site at <http://delos-wp5.ukoln.ac.uk/> (username: private, password: D310s-WP5)

- *Interoperability framework and architecture*: In this deliverable from TUC, an interoperability framework and a corresponding architecture is presented for the integration of eLearning applications on top of digital libraries. See Section 4.1.1.
- *Semantic mapping between SCORM and MPEG7 concepts*: the major eLearning and audiovisual standards are studied and a mapping between them is presented by IU and TUC. To illustrate the equivalence between the two models, an ontology is being developed and coded in OWL.
- *Investigation of content packaging options – Final model for the storage of audiovisual objects along with educational information Content packaging*: In this deliverable from UKOLN and TUC, the major content packaging schemes are studied and a framework for the storage of audiovisual objects along with educational information in the digital library part and also for their delivery to the eLearning applications according to the architecture is developed. See Section 4.1.1.
- *Personalization*: A study of existing models for the representation of Learner Profiles (e.g. IEEE LIP, PAPI and ELENA Project) has been completed by UKOLN and a Learner Information Model (LIM) is proposed. See Section 4.1.3.
- *Learning Styles*: In this deliverable produced by IU and TUC, several existing approaches of categorization of learning styles are studied.
- *Instructional ontology*: In this deliverable a workflow model for the construction of abstract training scenarios (learning designs) represented in an instructional ontology (OWL) is presented by TUC, according to the interoperability framework proposed. This allows reusability of the learning designs and provision of real personalized learning experiences: learning objects are bound to the learning experience on run-time, according to the needs of the learner expressed in the learner profile. See Section 4.1.2

This Task is also working to extend *GraphOnto*, which is an interactive ontology editor and ontology mappings tool for OWL. GraphOnto which has been developed by TUC, is used in several DELOS tasks (T3.6, T3.9, T3.10, T3.11, T5.4 and T5.5) and is offered through the DELOS demonstrators and testbeds Web site: <http://astral.ced.tuc.gr/delos/> (Universities outside DELOS have also obtained it). The tool is being extended to provide full ontology mappings and query mappings. Wider use of the tool will clearly be advantageous to integration efforts across the DELOS network (see Section 8 Integration).

In associated work, models for supporting semantic 3D information to be used in a variety of eScience applications have been derived and functionality requirements investigated. Two ontologies for 3D scenes, based on formal and *de facto* standards have been developed. They are available on DELOS WP3 testbeds and demonstrators site for downloading: <http://astral.ced.tuc.gr/delos/>. This work is described in a paper from TUC that has been accepted in the IEEE Virtual Reality international conference (see Publications Section 6.1). This work is complementary to the work in the Task 3.8 “Description, matching and retrieval by content of 3D objects”, which does not consider semantic descriptions.

Additional successfully-funded research proposals have developed from the Task activities. These include:

- An IU, TUC proposal to the Greek National Educational Radio-Television (ERT) has been approved for using selected material for the provision of eLearning experiences to schools. A rich catalogue with available audio and/or audiovisual material has been organized and an investigation of material for e-learning suitable to the age of the pupils in an audiovisual form has been done. A part of this material will be used in our demonstration architecture.
- An IST EU STREP Project (LOGOS) has been approved, starting in February 2006, based on the architectural concepts developed in this task.

Work is continuing on the demonstrator Deliverable D5.4.2 which is due month 30: details of Task dissemination are presented in Section 6.

In the next sections, some of the main achievements of T5.4 are presented. Detailed information can be found in the internal deliverables in the private area of the cluster Web site at <http://delos-wp5.ukoln.ac.uk/> (username: private, password: D3l0s-WP5).

4.1.1 Interoperability framework and architecture

Digital Libraries are an important source for the provision of eLearning resources (McLean, 2004). However, digital library metadata standards and eLearning metadata standards have been developing independently, presenting interoperability issues between digital libraries and eLearning applications. This is a complex and multi-level problem, which can be seen as a stack of conceptual layers where each one is built on top of the previous one (left part of Figure 1): There are different data representations, objects, concepts, domains, contexts and metacontexts in the layer stack that should be efficiently managed in a standard way. Metadata models are languages that are used to represent the knowledge in a particular application area. Each metadata model is shown as a vertical bar on this stack to cover a specific region that represents the parts that the model tries to capture and describe in a standard way. If one places different metadata models besides this stack, he may identify gaps and intersection regions so that being apparent where the interoperability problems among these models occur.

The right part of Figure 1 shows such a picture in the case of MPEG7 (MPEG7, 2001, 2003) and SCORM (SCORM, 2004), the major metadata standards in the audiovisual and eLearning domains respectively. It is apparent from this graphical presentation that MPEG7 and SCORM are not completely overlapping meaning that we need additional models to provide interoperability mechanisms between them. Of course, interoperability problems exist also in the overlapping areas. But in these areas solving the problem of interoperability is easier and can be solved with standard methods (e.g. by means of mappings). The major problems arise in the areas with no overlaps between the two metadata standards.

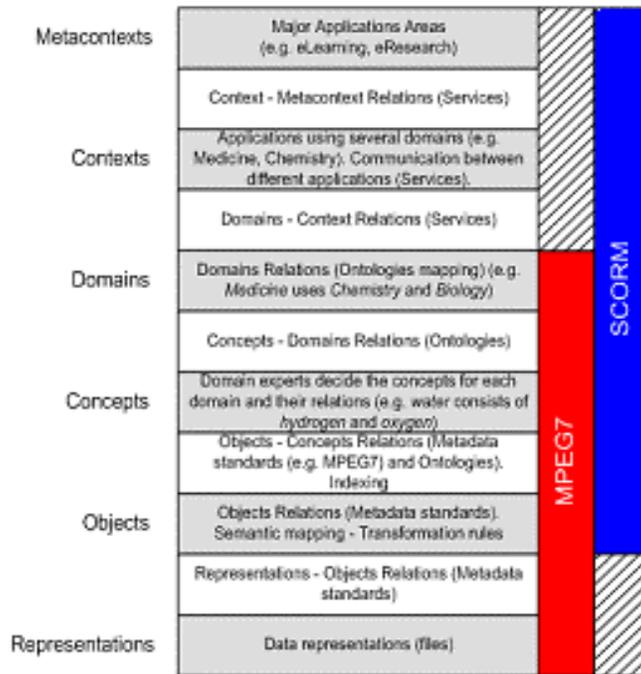


Figure 1. The multilevel problem of interoperability

A digital object can be described in many ways and delivered to many applications as illustrated in the upper part of Figure 2. However, performing just a transformation between the source and target metadata schemes is not always a panacea. As shown in Figure 1, standards do not always completely overlap. In the non-overlapping areas the interoperability problem cannot be simply solved using mappings.

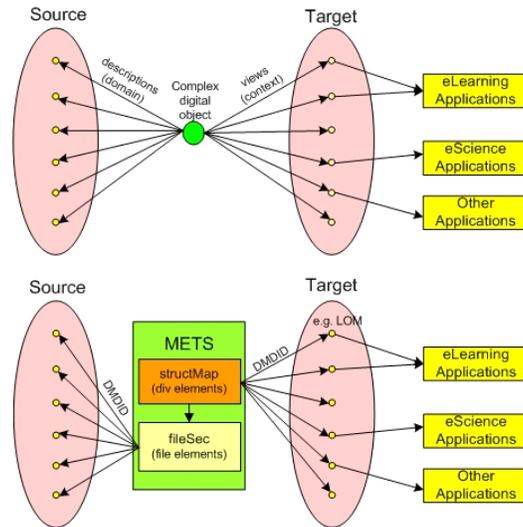


Figure 2. Using METS to support multiple-contexts views of digital objects

For example, SCORM contains an educational part that cannot be mapped to MPEG7 elements. Very often we want A/V digital objects that reside in a digital library and are described with MPEG7 to be used in eLearning applications. However, the MPEG7 descriptions do not say anything about the educational use (e.g. learning objectives) of

the digital objects. On the other hand, MPEG7 offers a comprehensive set of audiovisual Description Tools, which can not be represented in SCORM. In order to overcome these shortcomings we have to use a higher level metadata model that is able to encapsulate both SCORM and MPEG7 metadata descriptions in the context of a digital library. This model should be essentially a wrapper that will allow for the use of MPEG7 metadata of existing A/V objects and parts of them together with the necessary LOM (IEEE LOM, 2002) metadata to specify the educational characteristics of these objects and their parts.

In general, we need a way to have multiple descriptions (source metadata (domain), target metadata (context) pairs) for a digital object showing possible views of the object. Context and domain information should reside in different levels, where context information is described using domain information.

A flexible model that satisfies the above needs is the Metadata Encoding and Transmission Standard (METS) (METS, 2005). METS is a widely-accepted standard designed specifically for digital library metadata. METS is a flexible, but tightly structured, container for all metadata necessary to describe, navigate and maintain a digital object: Descriptive, Administrative and Structural metadata. Each type of metadata is described in a separate section, which is linked to its counterparts by internal identifiers. These metadata (any preferred scheme) may be held physically within the METS file, or in external files and referenced from within the METS document.

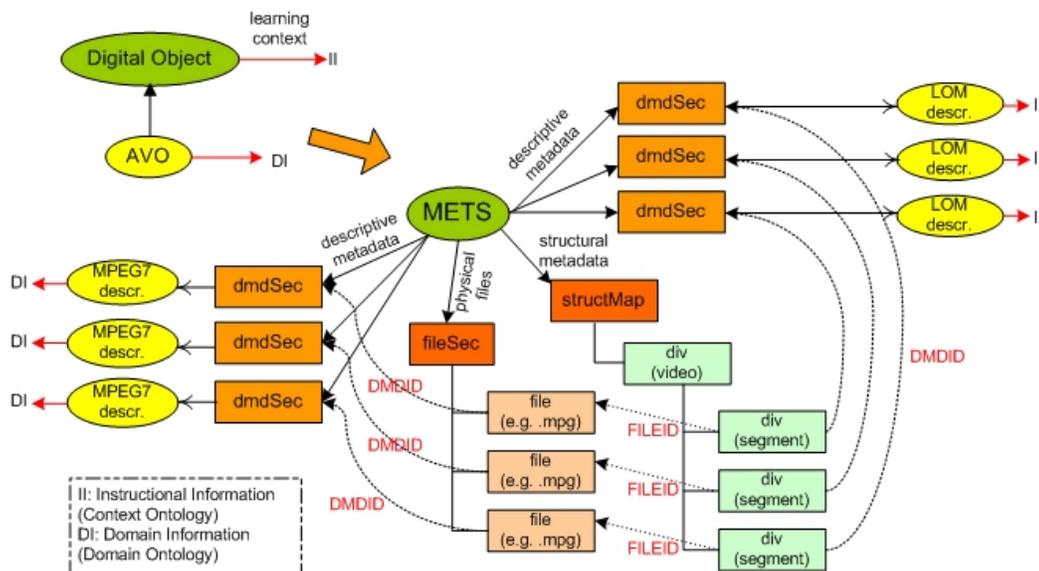


Figure 3. Combining METS, LOM and MPEG7 to build audiovisual learning objects

Using METS we can create different views of a digital object pointing to both source metadata description and target metadata description (context) in different levels. The methodology is illustrated in the lower part of Figure 2. Using the DMDID attribute of the <div> elements (structMap section describing the structure of the digital object) we can point to an appropriate metadata scheme creating a context (view) of this object and its parts. E.g., we can use LOM to describe the educational characteristics of each object and its parts, so that being able for them to be searched and retrieved by eLearning applications (learning context) (Figure 3). In parallel, using the DMDID attribute of the <file> elements (fileSec section, where all files comprising this digital object are listed,

we can point to a source metadata scheme that describes the low features or the semantics of this object (e.g. using MPEG7).

The ASIDE architecture: An Architecture for Supporting Interoperability between Digital Libraries and ELearning Applications

The architecture presented here addresses the identified interoperability problems in a layered architecture where eLearning (and other) applications are built on top of digital libraries and utilize their content. The ASIDE architecture offers a generic framework for the automatic creation of personalized learning experiences using reusable A/V learning objects. It is service-oriented and conforms to the IMS Digital Repositories Interoperability (IMS DRI) Specification (IMS DRI, 2003). The IMS DRI specification provides recommendations for the interoperation of the most common repository functions enabling diverse components to communicate with one another: search/expose, submit/store, gather/expose and request/deliver. These functions should be implementable across services to enable them to present a common interface. Figure 4 illustrates the architecture components, which are the following:

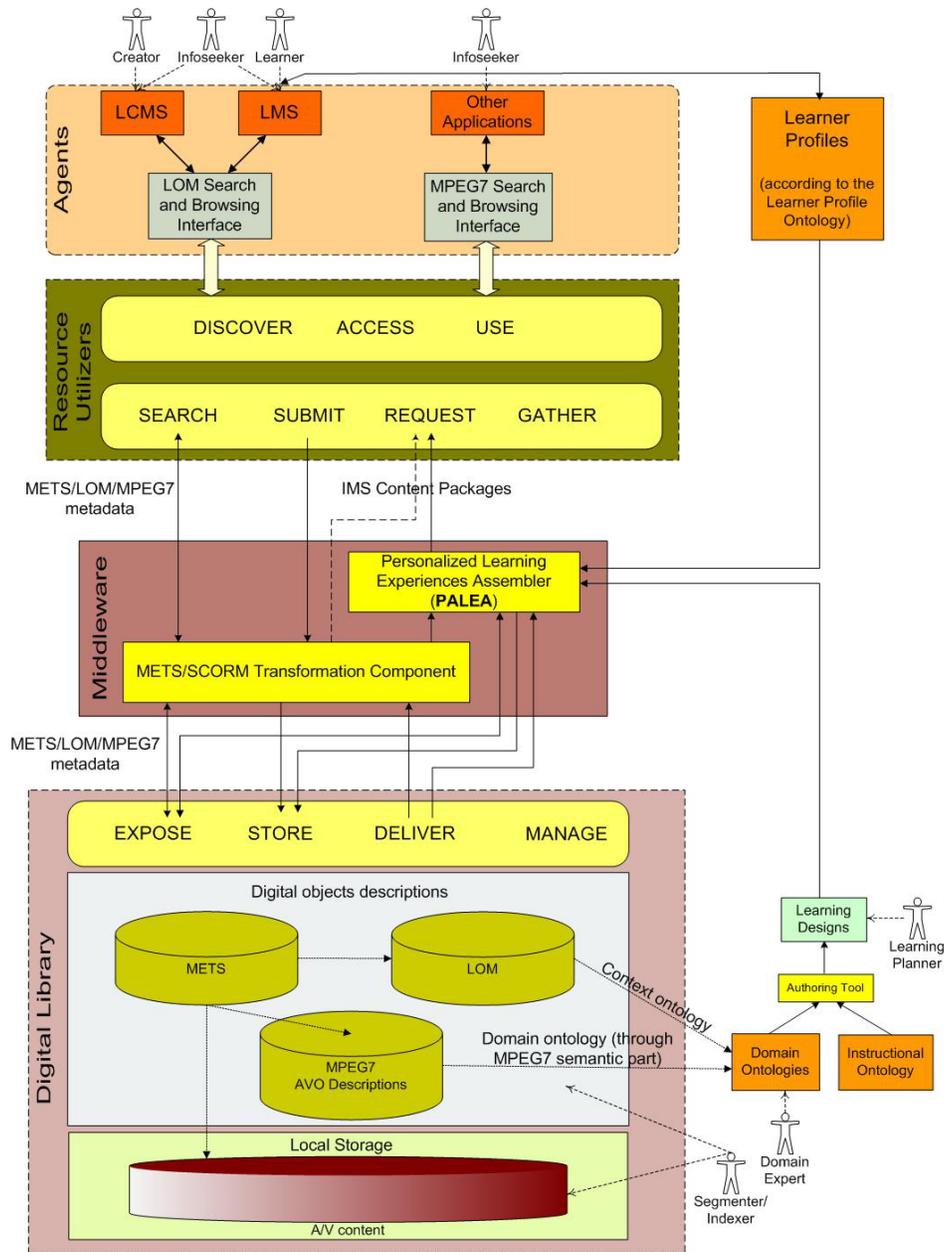


Figure 4. The interoperability architecture

- The **Digital Library**, where digital objects are described using METS+LOM (eLearning context), and MPEG7 (A/V descriptions) building this way interoperable A/V learning objects, which can be transformed to SCORM and delivered to eLearning applications (METS/SCORM transformation component). Some important elements used in the LOM descriptions are: educational objectives expressed as {verb (Bloom's Taxonomy (Bloom & Krathwohl, 1965))+subject (term from a Domain Ontology)} using the *classification* part of LOM, *context*, *typicalAgeRange* and *difficulty*. Regarding the MPEG7 descriptions, the methodology described in (Tsinaraki, Polydoros & Christodoulakis, 2004) is used for extending MPEG7 with domain-specific knowledge descriptions expressed in OWL (OWL, 2004)(domain ontologies).

- **Applications** (Software Agents in terms of IMS DRI, like Learning Content Management Systems, Learning Management Systems etc.) that discover, access and use the content of the A/V content of the digital library through appropriate services (resource utilizers). The generated personalized A/V learning experiences are delivered to the applications in the form of SCORM packages. Any SCORM-compliant system can recognize and “play” these packages.
- The **Middleware**, which is responsible for the assembly of personalized learning experiences. The middleware consists of the following parts:
 - The **METS/SCORM transformation component**, which is responsible for the transformation of the METS descriptions pointing to LOM and MPEG7 descriptions to SCORM Content Packages (SCORM, 2004). This includes not only simple transformation from METS XML file to SCORM manifest file, but also the construction of the whole SCORM package (PIF). Moreover, the mime-type of the files is taken into account and, if needed, intermediate html pages are constructed with links to these files (e.g. in case of video files).
 - The **Personalized Learning Experiences Assembler (PALEA)**, which, taking into account the knowledge provided by the Learning Designs (abstract training scenarios) and the Learner Profiles described later, constructs the personalized learning experiences and delivers them in the form of IMS Content Packages. Before transforming the resulted learning experience to a SCORM package, it is stored as METS+LOM+MPEG7 description in the digital library according to the interoperability framework, being ready and available in an interoperable way for later requests. The dashed arrow in the left side of PALEA indicates that using this component is optional, and that digital library services can be directly accessed (e.g. a teacher wants to find appropriate learning objects to construct manually a learning experience).
- **Ontologies** providing knowledge to the PALEA for the automatic construction of personalized learning experiences:
 - **Domain Ontologies** that provide vocabularies about concepts within a domain and their relationships.
 - **Instructional Ontology** (see section 4.1.2) that provides a model for the construction of abstract training scenarios. These are pedagogical approaches (instructional strategies/didactical templates), which can be applied to the construction of learning experiences.
- **Learning Designs** are abstract training scenarios in a certain domain built according to the model given in the instructional ontology.
- The **Learner Profiles** constructed using the vocabulary given in the **Learner Profile Ontology**, which represents a learner model for the creation of learner profiles. Elements from IEEE PAPI (IEEE PAPI, 2002) and IMS LIP (IMS LIP, 2005) specifications have been also used in this model. Some important elements

of this model are: learner goals, competencies, previous knowledge, educational level and learning style.

The interoperability architecture is being implemented using the following technologies: Web services, Java™ 2 Platform, Standard Edition, v1.5, Berkeley DB XML, Jena API, SPARQL RDF Query Language (Prud'hommeaux & Seaborne, 2005) and XQuery for querying the XML-based metadata descriptions of the digital objects stored in the digital library.

4.1.2 The instructional ontology

Nowadays, the need for e-learning systems supporting a rich set of pedagogical requirements has been identified as an important issue in the field of distance learning (Capuano et al., 2005). Several initiatives take place in order to meet this need. The most important of these initiatives seems to be IMS Learning Design (IMS LD, 2003) that provides a framework to depict pedagogies.

IMS Learning Design specification (IMS LD, 2003) is a development of the Educational Modelling Language (Hummel, Manderveld, Tattersall & Koper, 2004) (designed by the Open University of the Netherlands (OUNL) to enable flexible representation of the elements within online courses; not just the materials but also the order in which activities take place, the roles that people undertake, key criteria for progression, and the services needed for presentation to learners. The learning design specification does not detail how the course material itself is represented but rather how to package up the overall information into a structure that is modelled on a play, with acts, roles (actors) and resources.

The IMS Learning Design specification supports the use of a wide range of pedagogies in online learning. Rather than attempting to capture the specifics of many pedagogies, it does this by providing a generic and flexible language. This language is designed to enable many different pedagogies to be expressed. It allows different pedagogical approaches to be integrated into a single 'learning design' where different approaches may be appropriate for different types of learners. The approach has the advantage over alternatives in that only one set of learning design and runtime tools then need to be implemented in order to support the desired wide range of pedagogies.

The IMS Learning Design specifications (Figure 5) are structured in three levels. Level A includes activities, roles and environments. Activities (learning activities or support activities) can be grouped into activities structures and executed into specific environments. An environment is formed by learning objects and services provided to users during activity execution. Users are classified into roles (learners, teachers, tutors, etc.). Nowadays, learning objects are educational contents by which learners acquire knowledge and services are functionalities invoked during learning process in order to communicate with tutors or other learners. Level B adds properties (storing information about a single person or a group) and conditions (setting constraints upon the flow of activities) to the first level. Level C adds notifications (mechanism to handle messages passing between users) to the framework.

themselves, as IMS Learning Design (IMS LD, 2003) imposes. This ontology borrows some elements and ideas from the IMS Learning Design Specification and LOM and its purpose is to overcome the limitations that current eLearning standards and specifications impose. In the ASIDE architecture the final decided learning experiences along with the appropriate learning resources are finally packaged to a PIF file (SCORM package) and delivered to the eLearning applications.

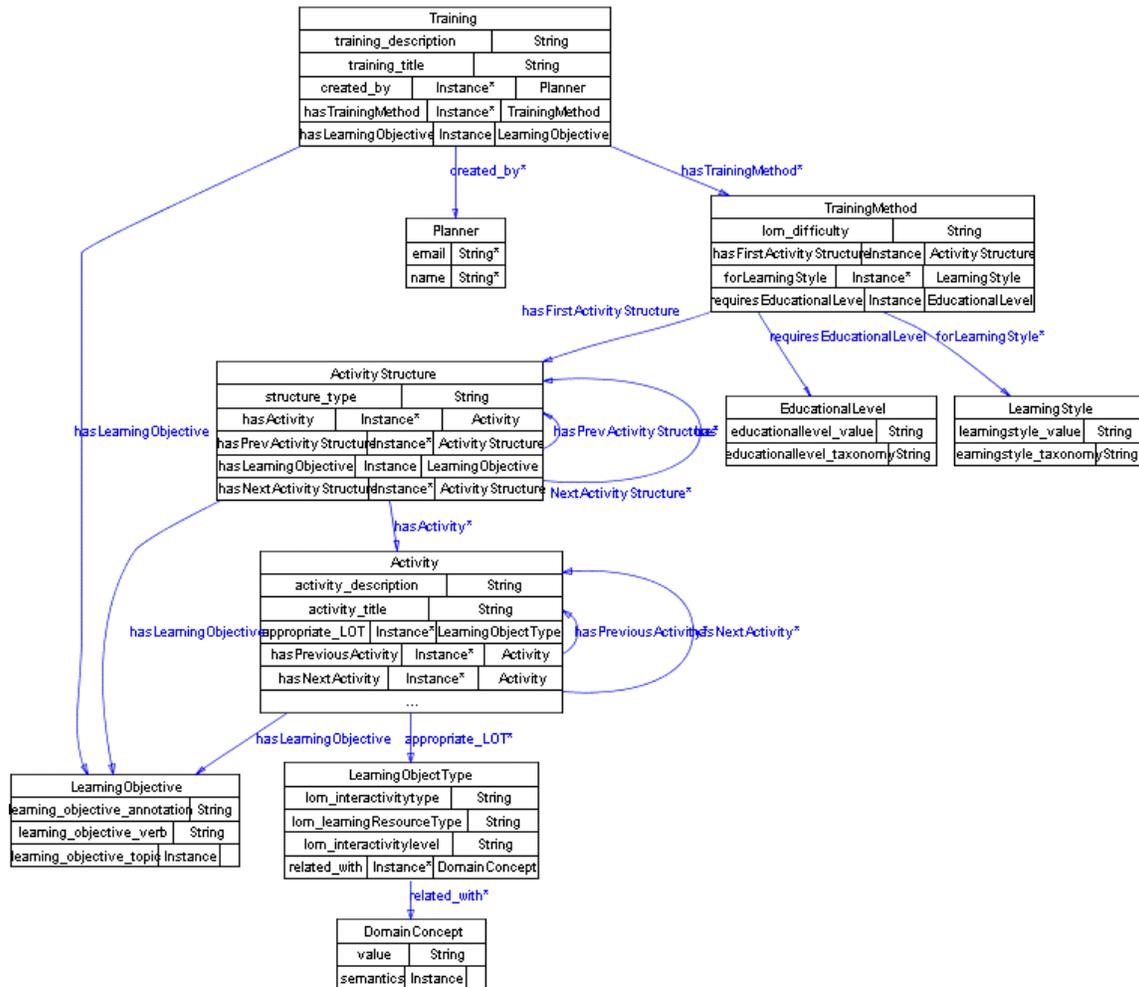


Figure 6. The instructional ontology

A *Training* is a collection of abstract training scenarios regarding one domain. The same subject can be taught in several ways (*TrainingMethods*) depending on the *LearningStyle* and the *EducationalLevel* of the Learner. There are several categorizations of Learning Styles and Educational Levels, thus these elements are flexible so that being able to point to values of different taxonomies. A *TrainingMethod* consists of a hierarchy of *ActivityStructures* built from *Activities* (elements taken from IMS LD) forming an arbitrarily complex structure of activities (sequence or selection). Since this model is RDF-based (OWL), existing *ActivitiesStructures* or paths of *ActivitiesStructures* can be reused in many Learning Designs. Each *Training*, *ActivityStructure* and *Activity* has a *LearningObjective*. Learning Objectives are treated here in a more formal way (as in SeLeNe project (Keenoy, Levene & Peterson, 2004)), than pure text descriptions. Thus,

each *LearningObjective* has: (a) a *learning_objective_verb*, taken from a subset of the outcome-illustrating verbs which characterise each type of learning objectives specified by a committee of college and university examiners in 1956 (known as "Bloom's Taxonomy (Bloom & Krathwohl, 1965)). This subset has been selected for the description of Learning Objectives by the SeLeNe project. (b) a *learning_objective_topic* that indicates the topic that the learning objective is about, referenced as an entry in the RDF binding of a subject taxonomy or ontology (context ontology e.g. ACM Computing Taxonomy (ACM, 1998)), and (c) *learning_objective_annotation* that indicates additional textual description of the learning objective; for example, to specify areas within the topic at a greater level of detail than is catered for by the subject taxonomy (or ontology). The *LearningObjectType* class is used to describe the desired Learning Object characteristics (requirements) without binding specific objects with *Activities* on design time. If more than one entries are used per *Activity*, the interpretation is "OR". Via the *related_with* property we can further restrict the preferred learning objects connecting them with *DomainConcepts* or individuals from a domain ontology.

4.1.3 The Learner Information Model (LIM)

With the huge range of digital educational resources becoming widely available and the opportunities afforded by the Internet and the Web, there is currently a trend towards the constructivist approach to eLearning (Burr, 1995). This approach puts the student at the centre of the learning process with the objective of tailoring educational resources to the individual learner's requirements and goals.

In addition, there is widespread recognition that learning is a life-long process and that learning technology systems need to track, manage and exchange information about their students. Learner information comes from three broad sources: personal information (address, telephone number etc.); preferences (operating system, network connection, desktop configuration, learning style etc.); and academic information (courses completed, grades etc.).

However, in order to match and tailor resources to an individual's requirements, it is also necessary to describe resources in an appropriate way, using for example the IEEE Learning Object Metadata (LOM) Standard (IEEE LOM, 2002). In this review, we will concentrate on learner information whilst recognising that there is a complementary issue with regard to the description of educational resources.

Much of the work on personalisation of eLearning has been based on metadata about a user or learner. The techniques fall into two broad categories:

- **Adaptive hypermedia:** mostly related to adapting user interfaces, navigation, content selection and presentation according to the user's performance in a particular domain.
- **Filtering and recommendation:** based on interests, preferences, likes, dislikes and goals a user has. Resources are recommended according to features extracted from a description of the resource or according to its ratings by users with a similar profile.

Given the open and heterogeneous nature of the Web, there has been much emphasis on interoperability and reuse of educational resources. To this end, there have been two major attempts to standardize a learner profile, IEEE Personal and Private Information (PAPI) (IEEE PAPI, 2002) and the IMS Learner Information Package (LIP) (IMS LIP, 2005).

According to Dolog & Nejd (Dolog & Nejd, 2003) these standards have been developed from different points of view. The PAPI standard reflects ideas from intelligent tutoring systems where the performance information is considered as the most important information about a learner. PAPI also stresses the importance of inter-personal relationships. On the other hand the LIP standard is based on the classical notion of a CV and inter-personal relationships are not considered.

One way forward is therefore to draw on both of the standards as a means of developing an adequate learner model, as in the ELENA Project (ELENA Project).

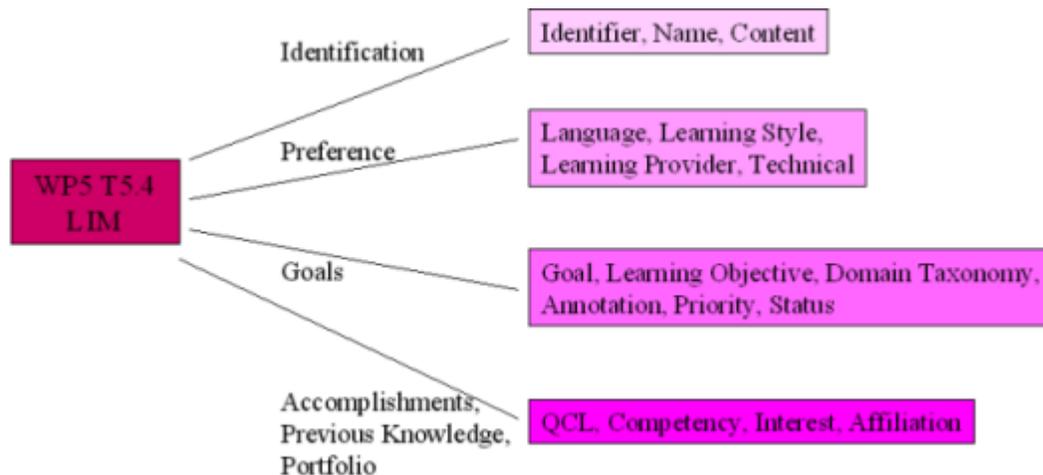


Figure 7. The Learner Information Model

Since the intention is to use SCORM in the Task 5.4 prototype framework, the User Model is based largely on IMS LIP, drawing on PAPI to fill in any gaps. SCORM uses IMS Content Packaging, which is also recommended by IMS LIP for the aggregation and packaging of learner information. Additional application specific elements are also required; these are drawn together using the concept of Application Profiles (Heery & Patel, 2004; Baker, Dekkers, Heery, Patel & Salokhe, 2001). In broad terms, application profiles are a type of metadata element set that draws on metadata terms from extant vocabularies and customises them for a local or specific application.

Figure 7 summarizes the T5.4 Learner Information Model, which is currently encoded in OWL.

5 Ontology-driven interoperability

Task 5.5 builds upon earlier Delos work and in particular upon the *Semantic Interoperability in Digital Library Systems* Report (D5.3.1) <http://delos-wp5.ukoln.ac.uk/project-outcomes/> produced during JPA1. It is concerned with the interoperability of data models and core ontologies which underpin the development of diverse digital library systems. The Task will investigate and develop methods for the integration of heterogeneous data types, models, upper level ontologies and domain specific knowledge organisation systems. The main objectives are:

- To achieve harmonisation of the FRBR-CIDOC Conceptual Reference Model (CRM) core ontologies.
- To produce mappings of the CRM core ontology to the metadata schema of TEI and Dublin Core.
- To develop and evaluate a demonstrator for a case study on the mapping of a Knowledge Organisation System (KOS) to a core ontology.

The lead partner is FORTH with Athens University of Economics & Business, Glamorgan, Imperial College, Ionian University, Lund, MTA Sztaki, NTNU, and TUC.

The potential impact of this Task both within Delos and across the wider community is considerable. There are three major aspects: metadata integration in key application areas, metadata integration in general, and core ontology harmonization.

The task is aiming to integrate the core ontology (ISO21127) describing the data structure semantics of museum / cultural heritage documentation with the one describing the future semantics of general library metadata (FRBR). The combined core ontology will allow for high precision schema integration of museums, libraries and archives (“MLA”) information. The widely accepted Dublin Core Metadata Initiative has limitations in its scope and level of granularity. The limits of applicability of the CRM for metadata integration in general have not yet been encountered (smaller extensions notwithstanding). It has the potential to become a core ontology for Digital Library metadata in general. This will be demonstrated by the CRM Core metadata elements, which are compatible with Dublin Core, but implement the key element of the CIDOC CRM and the ABC Model: explicit *event representation*.

The demonstrator will show the solution to a generic problem of metadata and schema integration in a particular case study using cultural heritage data. The more usual approaches to metadata interoperability, such as metadata registries and application profiles, work only with nearly homogeneous materials. The problem will become more acute when future digital libraries include all kinds of scientific data and multimedia material and the “simpler” approaches will reach their limits. Potential beneficiaries of this work include all library and digital library developers who want to integrate highly heterogeneous resources, such as the TEL, MICHAEL and BRICKS Projects.

The methodology of harmonizing core ontologies has been developed in the previous DELOS Project, with the harmonization of ABC and CRM. The problems and method

have already been published². Task 5.5 will carry out the task of harmonizing/merging the CIDOC CRM and FRBR through a series of meetings with invited experts and also students participating as part of their professional development.

5.1 Main achievements

From month 13 onwards, substantial progress has been achieved for Task 5.5.

Harmonization FRBR-CIDOC CRM:

Three working meetings led by FORTH have been held for the elaboration of a common ontology for FRBR and CIDOC CRM:

- a) Imperial College London, February 14-16 2005.
- b) FORTH, Heraklion, Crete, July 4-6 2005.
- c) Germanische Nationalmuseum, Nuremberg, Germany, November 16-18, 2005.

These meetings were attended by leading experts from IFLA and CIDOC:

Patrick LeBoeuf, Bibliotheque nationale de France (BNF), Service de normalisation documentaire (Standardization Department). Chair of the IFLA Working Group on the FRBR Review. (Funded by DELOS).

Maja Žumer, University of Ljubljana, Library Science and FRBR expert. (Funded by DELOS).

Christian-Emil Ore, University of Oslo, Chair of CIDOC.

Stephen Stead, Consultant, Vice Chair of CIDOC.

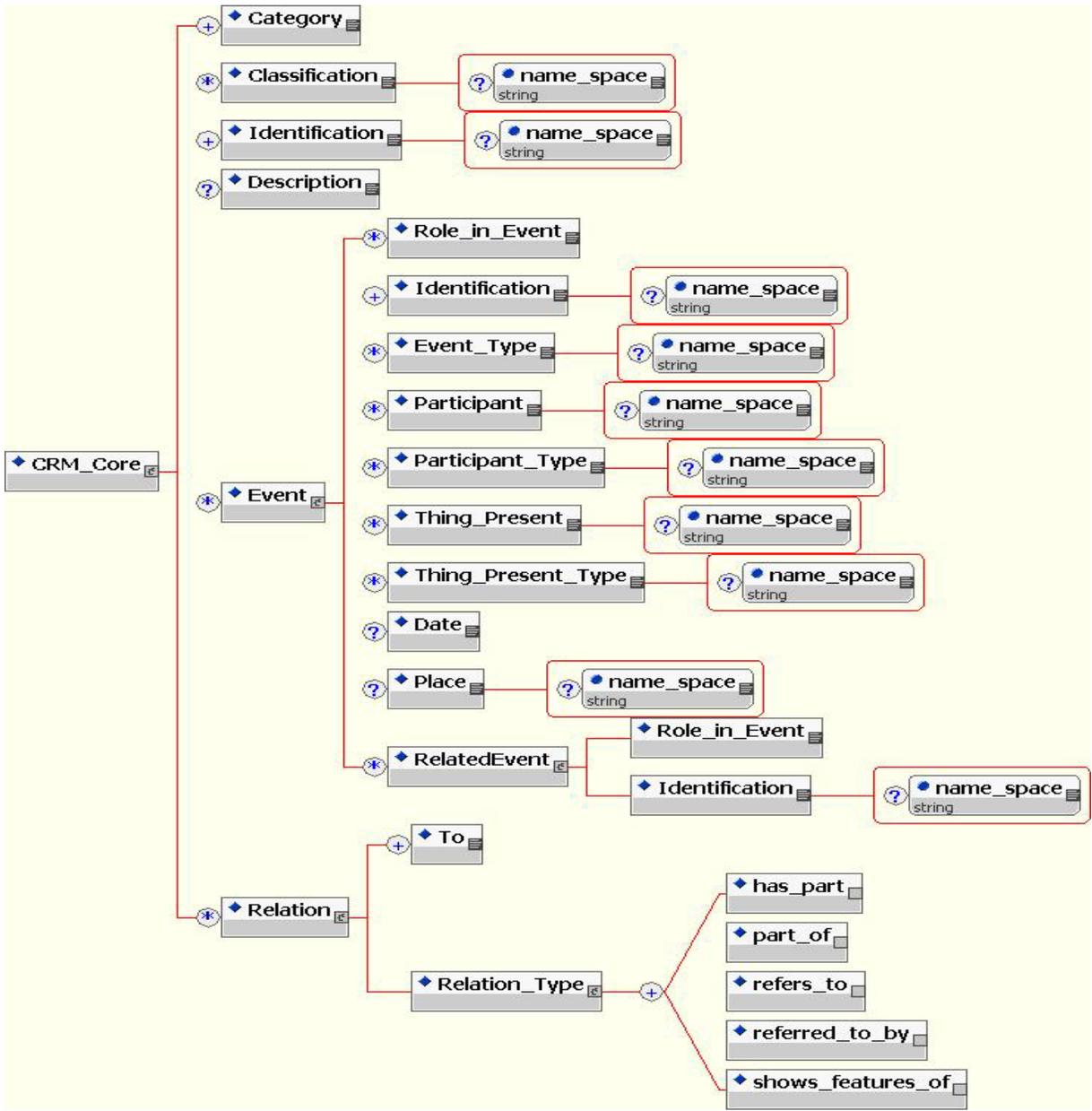
In the first two meetings, the conceptualization behind FRBR was analyzed as a graphical model extending the CIDOC CRM (ISO21127). It comprises the entities and relationships defined in FRBR, dealing with the notions of:

- a) Conception of a work.
- b) First realization of a work in a symbolic form.
- c) Reuse of work and continuation of work.
- d) The contribution and product of the publisher.
- e) The equivalence of electronic and material publishing.

The third meeting covered making textual definitions of classes and part of the relationships, as well as elaboration of the introduction. This harmonization process has resulted in a CRM Core ontology described in model form in Figure 5.

Figure 5. CRM Core ontology. (From Martin Doerr, FORTH).

² Martin Doerr, J. Hunter, Carl Lagoze, "Towards a Core Ontology for Information Integration", 2003, In *Journal of Digital information*, volume 4 issue 1, April 2003



Detailed reports of each meeting have been distributed to the partners of Task 5.5. The work represents a considerable intellectual achievement, since FRBR has not analyzed these processes to the level of detail necessary to produce a formal ontology, nor has there been any other work addressing these issues in a more formal way. In particular it has not modeled any process responsible for the existence of objects of bibliographic concern. The model now comprises 44 classes and 56 properties. It follows strictly the methodology used for ISO21127. It could be formulated without difficulties as a specialization of the CIDOC CRM classes. Considerable effort was needed to detect and resolve inconsistencies or fuzzy concepts in FRBR. See Figure 6. There have also been discussions to clarify concepts in the transition from the old to the new bibliographic practice which is underlying FRBR. Preliminary results have been presented at ECDL2005.

To be acceptable in the user communities, the model must be defined in a textual form with an introduction, description of its purpose and modeling methodology, and explicit scope notes and examples for each class and property. The class scope notes have been drafted and are expected to be finished and accepted by the working group by month 24. The work will continue in JPA3.

To produce mappings of the CRM core ontology:

The combined FRBR-CIDOC CRM model has been formally defined in OWL and TELOS. The GraphOnto tool from TUC has been functionally extended to be able to describe mappings between multiple ontologies. It is being evaluated by WP4 partners and based on this feedback, TUC will implement a second version of the enhanced GraphOnto with ontology mapping functionality that will be available by the month 30. TUC is also integrating a reasoning functionality to GraphOnto in order to support queries on top of ontologies. This functionality will be used to develop a retrieval demonstrator.

A mapping of Dublin Core to the CIDOC CRM has been defined and Ionian University (IU) has started working in two main areas: evaluation of the GraphOnto tool in cooperation with TUC to test the ontology uploading, creation and edit functions and experimentation with CRM. The research team has studied the CRM and now is developing a description of an excavation in Corfu, with the usage of the GraphOnto tool.

IU focuses on the mapping of DC-types to CRM: a first draft will be available by the month 24. See Figure 7. In addition, a mapping of a subset of EAD elements (version 2.0) to CRM will be tested. In particular IU plans to deal with the mapping of the <EADheader> as well some basic compound elements included in the <archdesc> element such as the <did>, <dsc>, <bioghist> and <scopecontent>.

SZTAKI is investigating new methods for collaborative support of the ontology mapping process. SZTAKI also experiments with query facilities in the scenario of multiple mapped ontologies. The CORES registry operated by SZTAKI contains more than 30 metadata schemas (element sets, application profiles, etc.) which can be browsed and searched in a connected way. This registry was connected with an OAI service provider with 500.000 records. The connection enables automatic adaptation of the service provider to schema changes. Furthermore, SZTAKI implemented a web-based search

interface which is capable to construct and execute queries combining the use of several metadata schemas.

Figure 6. CIDOC CRM-FRBR harmonization. (From Martin Doerr, FORTH).

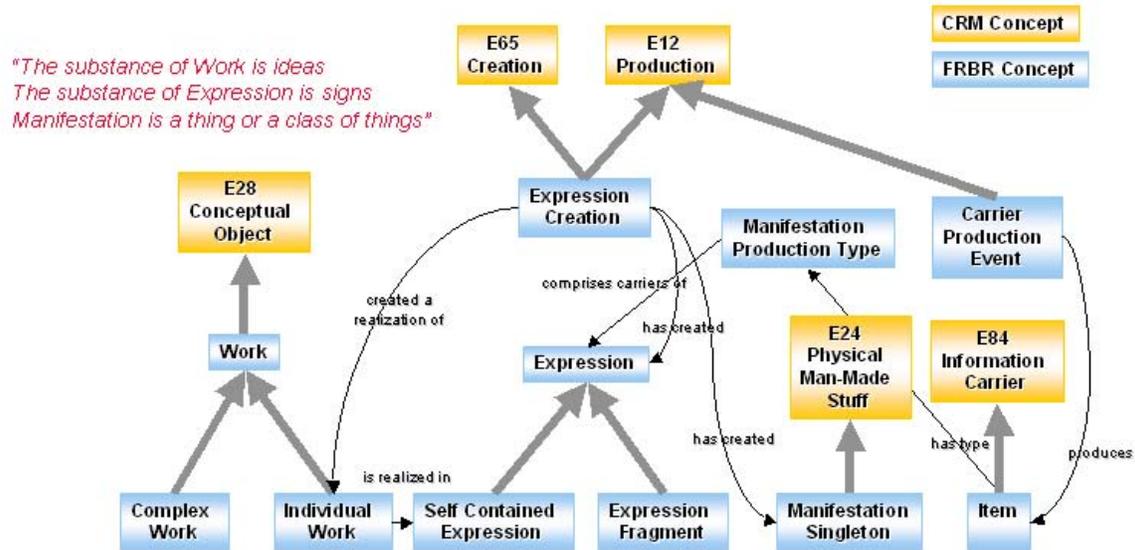
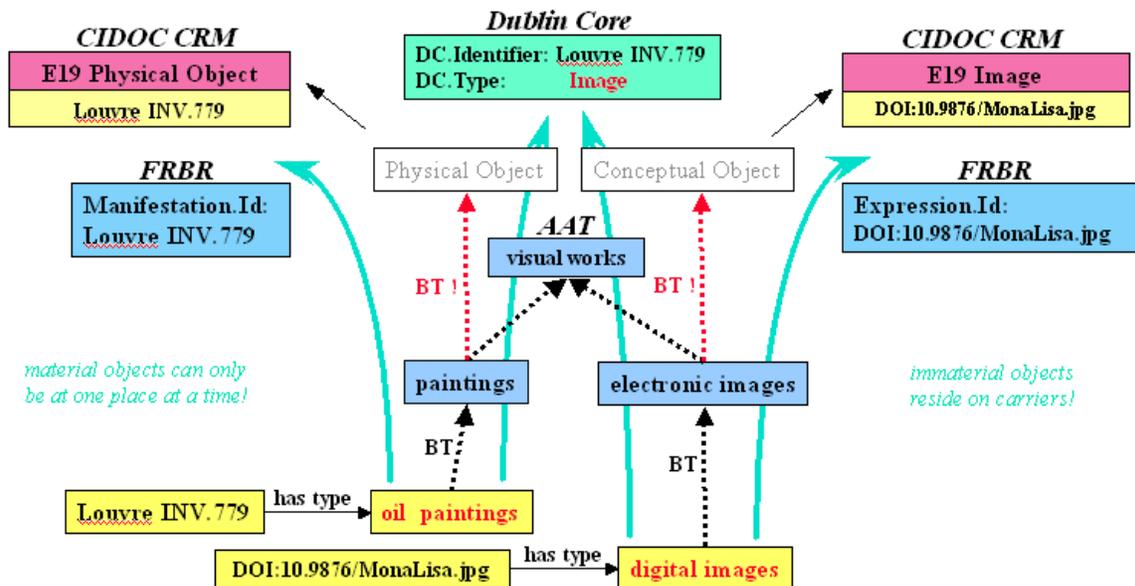


Figure 7. Resolving schema heterogeneity. (From Martin Doerr, FORTH).



Case study and demonstrator:

A case study for deriving an elaborate data structure and a core data structure for describing cultural objects from the CIDOC CRM core ontology will be available at month 24. These are the first steps to demonstrate in practical examples the utility of mapping to common core ontologies, and of merging core ontologies from related domains. The case study scoping work has been led by the University of Glamorgan through leverage of the ongoing English Heritage (EH) Revelation project and their application/extension of the CRM core schema to EH archaeology (CRM-EH). English Heritage have created a new Environmental Archaeology Thesaurus (EAT). While this is a new thesaurus, it closely matches index terms used in the EH Environmental Archaeological Bibliography (EAB), which will be the database for the case study. Lund have made preparations for the retrieval demonstrator including investigation of different technical alternatives.

All the above results can be demonstrated or provided in the form of documents. They will become part of the Deliverable D5.0.2 due in month 30.

Athens University of Economics & Business contribution has been delayed due to continuing processing of its new membership in DELOS.

Work continues and the Final status report for this Work Package due at month 30 will include the mappings and the demonstrator.

6 Management, Dissemination and Outreach

UKOLN is continuing to act as cluster co-ordinator. A number of new partners have been formally accepted by the Commission and welcomed into the cluster. They include the University of Glamorgan, Imperial College, London and the Athens University of Economics and Business. Each has brought additional relevant skills and expertise to the Network. Negotiations with another prospective partner DTSC, Australia (Dr Jane Hunter) have been delayed due to Hunter moving institutions to the University of Queensland during 2005.

The cluster has operated effectively and had three face-to-face meetings in 2005: in February, (Imperial College, London), May, (Crete) and September, (Vienna). The first cluster meeting of 2006 has been held on February 21st in London. In parallel work on the JPA2 tasks has been pursued through separate Task meetings and more details are given below.

Two Task 5.4 coordination meetings have taken place:

- Crete-Heraklion (on 13/5/2005) hosted by FORTH, and
- Vienna (on 23/9/2005) hosted by TUV.

However, several other informal meetings have taken place between the Greek partners (TUC and Ionian University).

A researcher exchange has been achieved: a MSc student of the Ionian University (Haroula Consolaki) has been moved to the Technical University of Crete for better collaboration between the two partners.

Two Task 5.5 coordination meetings have taken place:

- Crete-Heraklion (on 13/5/2005) hosted by FORTH, and
- Vienna (on 23/9/2005) hosted by TUV.

Three CRM-FRBR Harmonization Workshops have taken place as described in Section 5.1.

The Cluster Co-ordinator attended Scientific Board Meetings in Vienna and Nice and telcons, the Advisory Board meeting in Nice and the Project Review in Luxembourg in 2005. The Cluster Co-ordinator was unable to attend the All Tasks meeting held in Paris in January/February 2006 because the meeting dates were changed at relatively short notice and there were existing diary commitments.

The cluster Web site <http://delos-wp5.ukoln.ac.uk/>

has been maintained and updated with information relating to the new tasks. A private area has been set up for the sharing of draft documents, presentation slides and other work items. All items for dissemination have been linked to the cluster pages and deliverables made available.

6.1 Publications and Presentations

9th Thematic Workshop on Digital Repositories: Interoperability & Common Services, The Foundation for Research and Technology - Hellas (FORTH), Heraklion, Crete, 11-13 May, 2005. Proceedings published and on Cluster Web <http://delos-wp5.ukoln.ac.uk/>

Workshop Papers:

R. John Robertson, Jane Barton

"Optimising Metadata Workflows in a Distributed Environment" [[PDF](#)]

Chryssoula Bekiari, Panos Constantopoulos, Martin Doerr

"Information design for cultural documentation" [[PDF](#)]

José Borbinha, Jorge Machado

"DEPTAL: A Framework for Institutional Repositories" [[PDF](#)]

G. Knight, A. Wilson

"SHERPA DP: Establishing a persistent preservation environment for institutional repositories" [[PDF](#)]

Panos Constantopoulos, Martin Doerr, Meropi Petraki

"Reliability modelling for long-term digital preservation" [[PDF](#)]

Carl Rauch, Franz Pavuza, Stephan Strodl, Andreas Rauber

"Evaluating preservation strategies for audio and video files" [[PDF](#)]

Franz Pavuza

"Signal Ingest in Uncompromising Linear Video Archiving: Pitfalls, Loopholes and Solutions" [[PDF](#)]

Ed Crewe

"Extending the Open Journals OAI repository with RDF aggregation and querying (African Journals Online)" [[PDF](#)]

Zeno Tajoli (short paper)

"METALIS, an OAI Service Provider" [[PDF](#)]

Boon Low, John MacColl

"Searching Heterogeneous e-Learning Resources" [[PDF](#)]

Cezary Mazurek, Marcin Werla (short paper)

"Digital Object Lifecycle in dLibra Digital Library Framework" [[PDF](#)]

Other papers and presentations

Arahova, A., Kapidakis, S.: Empowering Our Libraries, Empowering Our Education System: Using the Research Results for Implementing Not the best, but the Most Effective Policy for School Libraries. IFLA (International Federation of Libraries Associations) Conference, August 14-18, 2005, Oslo, Norway.

Ardö, A. Focused crawling in the ALVIS semantic search engine [[PDF](#)]
[ESWC2005](#) - 2nd Annual European Semantic Web Conference in Heraklion, Crete May 29 to June 1, 2005

Binding C.

"XML Techniques for the Representation and Interchange of Thesaurus Data" [[HTML](#)]
D-Lib Magazine Volume 11 Number 1, January 2005.

Christodoulakis S, Polyxeni Arapi, Nektarios Moumoutzis, Manjula Patel, Sarantos Kapidakis, Antonia Arahova, Lina Boudouri
Interoperability of eLearning Applications with Audiovisual Digital Libraries [[PDF](#)]
the 9th European Conference on Research and Advanced Technology for Digital Libraries (ECDL 2005) September 18-23, 2005 Vienna, Austria

Simon Coles, Jeremy Frey, Michael Hursthouse, Mark Light, Leslie Carr, David DeRoure, Christopher Gutteridge, Hugo Mills, Ken Meacham, Mike Surrige, Liz Lyon, Rachel Heery, Monica Duke, and Michael Day.

The 'end to end' crystallographic experiment in an e-Science environment: From conception to publication.

In: [UK e-Science All Hands Meeting 2005](#), Nottingham, UK, 19-23 Sep 2005. Southampton, UK. (In Press)

Paper and Powerpoint presentation available from the University of Southampton e-Prints Service: [[entry in repository](#)]

Ding, H., Sølberg, I. Choosing Appropriate Peer-to-Peer Infrastructure for Your Digital Libraries [[HTML](#)]

ICADL 2005, LNCS 3815, pp.457-462.

- Doerr, M. Waking from a Dogmatic Slumber - A Different View on Knowledge Management for DLs
 DELOS NoE Brainstorming Meeting, Juan-les-Pins, France, December 5-6, 2005
 Presentation: [[Powerpoint](#)] [[HTML](#)]
- Monica Duke, Michael Day, Rachel Heery, Leslie A. Carr, Simon J. Coles
 Enhancing access to research data: the challenge of crystallography
[JCDL 2005](#) Digital Libraries: Cyberinfrastructure for Research and Education, Denver, Colorado, USA June 7-11, 2005.
 Paper (preprint): [[PDF](#) | [MS Word](#)] Abstract: [[HTML](#)] Abstract (and full text for subscribers) from: [[ACM Portal](#)]
 Presentation: [[Powerpoint](#)]
- Csaba Fülöp, Gergő Kiss, László Kovács, András Micsik: "Using a Metadata Schema Registry in the National Digital Data Archive of Hungary", the 9th European Conference on Research and Advanced Technology for Digital Libraries (ECDL 2005) September 18-23, 2005 Vienna, Austria.
- Golub, K. and A. Ardö Importance of HTML structural elements and metadata in automated subject classification [[PDF](#)]
 Proceedings of ECDL 2005 - the 9th European Conference on Research and Advanced Technology for Digital Libraries
 Vienna, Austria, 18-23 September 2005. P. 368-378
- Golub, K. Automated subject classification of textual Web pages, for browsing ([PDF](#))
 Thesis for the degree of Licentiate in Philosophy, Swedish intermediate degree between Master's and Doctoral degrees
 Lund: Department of Information Technology, Lund University. ISBN: 91-7167-034-3.
 [Formatted for 220 x 155 mm], 2005
- Golub, K. and Larsen, B.
 Different approaches to automated classification: Is there an exchange of ideas? [[PDF](#)]
 Proceedings of ISSI 2005 - the 10th International Conference of the International Society for Scientometrics and Informetrics
 Stockholm, Sweden, 24-28 July 2005, Vol. 1. Heery, R. and Anderson, S.. Digital repositories review. Report to accompany JISC Digital Repositories Programme call, February 2005. Available from:
 <http://www.jisc.ac.uk/index.cfm?name=programme_digital_repositories>
- Golub, K. Automated subject classification of textual Web documents
 Accepted for publication in Journal of Documentation. 2005.
- Hursthouse, M.B., S. J. Coles, J. G. Frey, L. Carr, C. Gutteridge, L. Lyon, R. Heery, M. Duke, M. Day
 ECRYSTALS(.CHEM.SOTON.AC.UK): Open Archive Publication of Crystal Structure Data [poster]
[XX Congress of the International Union of Crystallography](#), Florence, Italy, 23-31 August 2005

[Software demo: Grid Computing and Crystallography](#)

Poster presentation: [[Abstract](#)]

- Kalogerakis, V., Christodoulakis, S., Moumoutzis, N.: Coupling Ontologies with Graphics Content for Knowledge Driven Visualization. IEEE Virtual Reality International Conference, March 2006, Virginia, USA (to appear).
- Koch, T., Golub, K. and Ardö, A. Users browsing behaviour in a DDC-based Web service: A Log Analysis.
Accepted for publication in Cataloging & Classification Quarterly. 2005.
- Koch, T. The e-Science project eBank UK.
NetLab Society Seminar, Lund, Sweden, 1 September 2005
Presentation: [[Powerpoint](#)]
- Lyon, L. Digital Libraries and e-Research: new horizons, new challenges?
8th International Bielefeld Conference, Bielefeld, Germany, February 2006
Presentation: [[Powerpoint](#)] [[HTML](#)]
- Lyon, L. Adding value to open access research data: reflections on the process of data curation. Third European Conference on Research Infrastructures, Nottingham, UK, 6-7 December 2005.
Presentation: [[Powerpoint](#)] [[HTML](#)]
- Lyon, L. Adding value to open access research data: the eBank UK Project.
[CERN workshop on Innovations in Scholarly Communication](#) (OAI4), Geneva, Switzerland, 22 October 2005.
Presentation: [[Powerpoint](#)] [[HTML](#)]
- Lyon, L. Developing e-Infrastructure to support new research and learning paradigms.
[Building the Info Grid](#). Digital Library Technologies and Services - Trends and Perspectives. DEFF conference, Copenhagen, 26-27 September 2005
Presentation: [[Powerpoint](#)] [[HTML](#)]
- Lyon, L. Data, information and knowledge repositories: developing infrastructure to support the e-Research landscape.
euroCRIS Strategic Seminar, Brussels, September 2005
Presentation: [[Powerpoint](#)] [[HTML](#)]
- Lyon, L. From research data to new knowledge: a lifecycle approach.
Making the strategic case for institutional repositories. CNI-JISC-SURF Conference, Royal Netherlands Academy of Arts and Sciences, Amsterdam, 10-11 May 2005.
- Lund, H., Larsen, B., Voel Jensen, R.E., Ardö, A., Golub, K. and Ingwersen, P.
Capturing Contexts for Web Filtering in the Humanities [PDF]
Second Workshop on Information Retrieval in Context (IRiX), SIGIR 2005.
- Sølvberg, I., Larsen, R., and Mitra, S. Digital Libraries and e-Learning
Panel discussion InfoVision 2005, Bangalore, 15-16 September, 2005

Tentori, A. Concerted Information Management in a Virtual Library[PDF]
Master's Thesis, Department of Information Technology, Lund University, March 7, 2005.

Tsinaraki, C., Polydoros, P., Christodoulakis, S.: GraphOnto: A Component and a User Interface for the Definition and Use of Ontologies in Multimedia Information Systems. In Proc. of AvivDiLib 2005, pp. 99-102.

Tudhope D., Binding C.
Towards Terminology Services: experiences with a pilot web service thesaurus browser
Proceedings of the International Conference on Dublin Core and Metadata Applications, (DC 2005), 269-273.

Tudhope, D. The FACET project and implications for the Semantic Web
7th NKOS workshop on Next Generation Knowledge Organization Systems: Integration Challenges and Strategies, Denver, June, 2005

Tudhope, D. Applications of NKOS: some examples and questions
NKOS Special Session at Dublin Core Metadata Initiative Conference (DC 2005), Madrid, September, 2005.

KOS issue forthcoming in NRHM 2006 - The New Review of Hypermedia and Multimedia 2006 (1). Special Issue Call for Papers on: Knowledge Organization Systems and Services. Guest Editors: Douglas Tudhope, University of Glamorgan, UK (dstudhope@glam.ac.uk), Marianne Lykke Nielsen, Royal School of Library and Information Science, Denmark (mln@db.dk).

The Task 5.5 work was presented in the DELOS poster session held in conjunction with ECDL 2005.

NKOS Workshop led by Doug Tudhope has been accepted for ECDL 2006.

KESI cluster Ontology-driven Semantic Interoperability work was also promoted in the US in December 2005 by Dolores Iorizzo, Imperial College:

1) Cornell University: Presentation of results to Computer Science Group working in Digital Libraries and discussion held with Carl Lagoze to look at a way of integrating our work with the National Science Digital Library and FEDORA. The meeting was very successful and Carl and Dolores did a joint presentation about future work funded by NSF-DELOS in Washington DC.

2) OCLC: Presentation of results and CIDOC CRM-FRBR harmonisation to members of OCLC (including Lorcan Dempsey, Ed O'Neill and Stu Weibel). There was general agreement to do an analysis of the FRBR+CRM, and also take-up of CRM CORE as a possible enriched Dublin Core (i.e. testing of TEL, CRM CORE and DC to determine best results).

3) NSF: Presentation of results to NSF Directorates (hosted by Steve Griffin and Lee Zia) on WP5 and our vision for the future. Positive outcomes included an invitation to discuss

international cooperation between NSF-DELOS on global Cyberinfrastructure, and possible international conference event in summer 2006.

This type of dissemination is particularly valuable since it raises awareness of Delos in the wider community and facilitates broader take up of Delos standards activity.

7 JPA3 plans

The strategy for scoping JPA3 activity has been largely to build on existing work.

7.1 Interoperability of eLearning Applications with Digital Libraries proposal

The architectural aspects of integrating eLearning applications on top of digital libraries that have been identified will be dealt-with within JPA2 in order to produce a prototype by the end of JPA2. However, many more issues are expected to appear during the design and implementation that will need further investigations. In addition, two very important issues for the integration of applications on top of digital libraries have appeared and some preliminary investigation has been done. The first is the development of an Earth Sciences' digital library according to the T5.4 eLearning environment for the provision of eLearning in Geography, and the second is the support of interoperability and semantics for 3D objects. The latter is very important both in eLearning and eScience applications (chemistry, biology, archeology, etc.). Note that MPEG7 does not really cover the semantics of 3D objects where other standards are more pertinent (like X3D and Maya). An initial approach for both areas has been done already within JPA2 and reports are in preparation. The initial investigations showed a very reach set of research issues related to the task as well.

7.1.1 Development of an Earth Sciences' digital library according to the T5.4 eLearning environment for the provision of eLearning in Geography

According to the ITF report, the ITF identified 3 activities that are not direct evolution of the current JPA, but would be worth investigating due to their relevance in the future scenarios. From these, two are satisfied by this task:

1. audio DLs: this activity should investigate the management of digital libraries of audio objects, with respect to the classical problems: storing, accessing, indexing, retrieving, as well as to the definition of typical services related to this kind of medium, which has also a high potential commercial value.
2. e-science digital libraries: although DELOS addresses e-health, there is a wider range of science-related digital libraries, including, e.g. bio-informatics and earth-science, that are becoming increasingly important in the scientific and scholar fields.

IU and TUC have already succeeded in negotiations for using selected material from the Greek National Educational Radio-Television whose purpose is to be used at schools and a rich catalog with available audio and/or audiovisual material has been organized. An

investigation of material for e-learning suitable to the age of the pupils in an audiovisual form has been done.

During JPA3, the task will:

- Select an appropriate geography domain ontology (<http://reliant.tekknowledge.com/DAML/Geography.owl>) which is based on:
 1. CIA World Fact Book for geographic regions.
 2. WFB climate data interpreted using Koeppen Climate Classification system summary at <http://www.fao.org/WAICENT/FAOINFO/sustdev/EIdirect/climate/EIsp0002.htm>
 3. Sea Level definitions at <http://www.pol.ac.uk/psmsl/puscience/index.html> .

Population of the classes of the geography domain ontology, based on the courses to be developed.

- Create a digital library of geography learning material, based on the interoperability architecture developed in JPA2, by digitizing and segmenting selected material from the Greek National Educational Radio-Television archives. Indexing of the material will be made using the enriched geography domain ontology.
- According to the instructional ontology developed in JPA2 abstract training abstract training scenarios (learning designs) for geography teaching (based on secondary school curricula) will be developed. The scenario nodes (activities) will be connected with the enriched geography domain ontology.
- Creation of Model for Usage-Oriented Evaluation of eLearning applications with audiovisual digital libraries, that support the modular development of personalized learning experiences inside a virtual environment according to specific measurement strategy.

7.1.2 Semantic description and interoperability of 3D objects and 3D scenes with eLearning and other digital library applications

As the hardware and software capabilities in graphics are continuously improving the number of 3D models available in online repositories is growing dramatically. These objects may be associated with domain semantics and then used in elearning experiences in various fields such as science, culture, history etc. The semantic enrichment of scenes can play an extremely important role in enabling the viewers to query, understand and interact with the usually complex and incomprehensible visualized information, in simple, intuitive and user-friendly ways and allowing them to identify 3D objects or sets of them based on their graphical and semantic properties and relationships with other objects in the scene at a time. Interactive queries, such as “what is this object which I clicked upon with my mouse?”, “what is the functionality, the

behavior and the role of the component represented by this object?” or “show all the components of the same type or of the same hierarchical level of abstraction and hide all the others from the scene” should be answered with the appropriate visual and textual response. Such queries could allow users to explore the external and inner parts of the models and understand their behavioral and functional patterns, which usually carry mappings with processes and events of the real world. The mere visual display of complex and large amounts of information is not sufficient itself to answer such queries especially when viewers do not possess sophisticated knowledge of the domain related to the visualized information. Such functionality should be supported by 3D repository services in coherent manner.

Designers are also interested in generating, updating or deleting graphics content from the scenes based on the semantics of the visualized information. For example, user commands, like “apply a half-transparent material to all the objects representing the concept X (e.g. a window)” or “apply a transformation (translation, rotation, scaling etc.) to all the objects representing the concept Y according to its properties” or even “create a 3D object for every instance of concept Z accordingly”, could significantly facilitate or even automate the manipulation of the graphics content according to the specific domain knowledge that is coupled with the scenes. The above manipulation commands can also be personalized with the help of user profiles. A simple example of personalization of the scenes would be: “apply a specific background stored in the user’s profile to the scene” or “apply this material taken by the same user’s profile to every object representing the concept W”. User profiles combined with domain knowledge could automatically adapt the content of the scenes to designers’ preferences.

7.2 Ontology-driven interoperability proposal

This proposal continues WP5 Task 5.5 “ontology-driven interoperability” from JPA2. The proposal addresses the cluster’s key aim of achieving semantic interoperability at both data and metadata levels. Knowledge Organization Systems (KOS), such as classifications, gazetteers and thesauri provide a controlled vocabulary and model the underlying semantic structure of a domain for purposes of retrieval. Ontologies provide a higher level conceptualisation with more formal definition of roles and semantic relationships. The objective of this project is the investigation and development of methods for the integration of heterogeneous data types, models, upper level ontologies and domain specific KOS. This effort will be driven by a domain overarching core ontology starting from the CIDOC CRM (ISOCD21127).

The challenge for the next generation of information access systems is the ability to retrieve complementary objects and deep paths of relevant relationships that cross multiple document and resource boundaries. Complementary information can only be identified against an application and domain overarching core-ontology, that allows for relating, mediating or translating the elements of necessarily heterogeneous data and metadata schemata employed in multiple applications and domains.

The main focus of the work proposed for JPA3 is:

1. To provide a core ontology that covers the basic conceptualization of documentation in libraries, archives and collections of cultural and scientific

heritage. This will be based on completion of the FRBR-CIDOC CRM harmonization work.

2. To demonstrate, how complementary information documented in metadata structures of very different form and level of detail can be retrieved coherently under the core ontology, and how domain ontologies and indexing languages must fit under the core ontology in order to do so. This work will be based on further developing and extending the demonstrator to include the English Heritage archaeological 'grey literature' reports and the Archaeobotanical Computer database (ABCD). This will allow investigation of the core ontology to facilitate cross-domain linking and search from Environmental Archaeology to Economic Botany via biological taxonomies, and linking from datasets to literature.

8 Integration: intra-cluster and inter-cluster

Organisation and delivery of the Workshop in Task 5.1 has proceeded jointly with the Preservation Cluster and this has been very productive. There are other areas of synergy around characteristics of a trusted digital repository, curation and preservation requirements for repositories and audit and certification. There are also potential links to digital library architectures and reference models work (WP1) when considering the role of federated repositories as content providers in service oriented architectures and frameworks. These points will be fully investigated in the rescheduled D1.4.0.

Tasks 5.4 and 5.5 have much commonality since both are concerned primarily with metadata and semantic interoperability. The management of complex objects is a key issue for developers of digital libraries, e-learning materials and e-research data sources. In the Delos context, there are strong links between WP3 and WP5 through investigations of content packaging for multimedia objects and associated metadata standards such as METS, MPEG 21 DIDL etc.

Task 5.4 is also working to extend *GraphOnto*, the interactive ontology editor and ontology mappings tool for OWL. GraphOnto which has been developed by TUC, is used in several DELOS tasks (T3.6, T3.9, T3.10, T3.11, T5.4 and T5.5) and is offered through the DELOS Web site (Universities outside DELOS have also obtained it). The tool is being extended to provide full ontology mappings and query mappings. Wider use of the tool will clearly be advantageous to integration efforts across the Delos network.

Following on from the All Tasks meeting in Paris in January/February 2006, this idea has been explored further in a paper from TUC which highlights common components and Task inter-relationships focussing on semantic multimedia management. GraphOnto is identified as the common tool for describing all components and/or applications in the eight named tasks. It is also proposed that the future focus of the tasks should be towards "providing service-oriented interfaces to the outside world and compatibility with the OSIRIS platform".

9 Conclusions

In general, the activities of the cluster have progressed well. A number of new partners have been formally accepted by the Commission and welcomed into the cluster. They

include the University of Glamorgan, Imperial College, London and the Athens University of Economics and Business. Each has brought additional relevant skills and expertise to the Network. Negotiations with another prospective partner DTSC, Australia (Dr Jane Hunter) have been delayed due to Hunter moving institutions to the University of Queensland during 2005.

The partners are working effectively together and communication channels have been adequate with a mix of face-to-face meetings and email discussion. Some cluster partners have been more active than others but these organisations have been more deeply involved/lead other clusters and inevitably resources must be managed within the available funding constraints.

Dissemination activity has continued to grow and can be expected to increase as more research outputs and deliverables become available. A total of 30 published papers and invited (high-profile) international conference presentations have been produced. Awareness of ontology-driven interoperability work has been raised in the US

In Task 5.1 Information Repositories and Open Archives (from JPA1), the 9th Delos Thematic workshop on *Digital Repositories: Interoperability and Common Services* was held in Heraklion, Crete on 11-13 May 2005 with Sandy Payette, Co-Director Fedora Project, Cornell University, USA giving the Keynote speech. Digital Repositories are the focus of much investment and development activity around the globe and it is anticipated that the delayed Deliverable D5.1.1 (now due June 2006), will be received with much interest from the community. The delays have resulted from staff ill health and conflicting core funder (JISC) commitments at the partner site.

Task 5.4 which focuses on e-Learning Applications and Digital Libraries has produced six internal deliverable reports (on *Interoperability framework and architecture*, *Semantic mapping between SCORM and MPEG7 concepts*, *Investigation of content packaging options*, *Personalization*, *Learning Styles*, and *Instructional ontology*) relating to aspects of interoperability. These are informing the development of the demonstrator. It has also been identified as a key integration element for the KESI cluster and for the wider task activities of the Delos network.

Task 5.5 Ontology-driven Interoperability is building on the JPA1 Report *Semantic Interoperability in Digital Library Systems* Report (D5.3.1) through the CIDOC-CRM / FRBR Harmonisation work, which has produced a core ontology and draft detailed model; mappings from CRM to Dublin Core and some preparatory modelling for the cultural heritage demonstrator.

Finally, JPA3 work is continuing, building largely on existing activity.

10 Partners

This section gives some general information about KESI cluster partners. The cluster has also developed collaboration with British Telecom, UK through David Alsmeyer and with Dr Jane Hunter at the University of Queensland, Australia. It is hoped that the latter organization will become a formal partner during 2006.

10.1 Athens University of Economics and Business

10.2 ETH, Swiss Federal Institute of Technology, Zurich (Switzerland)

http://www.ethz.ch/index_EN

10.3 FORTH, Crete (Greece)

<http://www.forth.gr/>

FORTH is leading the CIDOC CRM SIG development team and has initialized together with BNF the dialogue with FRBR. It has led the Harmonization of CIDOC CRM and ABC Harmony under the previous DELOS Project. FORTH is contributing to WP5, Task 3 Semantic Interoperability and WP2, Task 2. FORTH develops RDF Technology. FORTH develops advanced cultural information systems and digital library system components.

10.4 Imperial College, London

Imperial College is leading the Newton Project. In the first five years of its existence, the Newton Project has produced the first ever comprehensive [catalogue](#) of Newton's non-'scientific' papers and has placed online nearly 50 per cent of the two and a half million words he devoted to the subject of [theology](#), as well as a large selection of his [personal](#) and [scientific](#) papers. At the beginning of 2004, a closely linked sister project in the US received funding to begin similar work on the [alchemical papers](#). Contact person is: Dolores Iorizzo worked for many years on the [Commentaria in Aristotelem Graeca](#) Project at King's College, London. She has held lectureships in Philosophy and Medical Ethics at King's College, London and Imperial College, London. She has also worked extensively as a Project Coordinator for global development economics projects. Her main area of research is in the transmission of Aristotelian and Stoic philosophical and medical theories into the Early Modern period. Now she is member of the Newton Project Staff.

10.5 Ionian University

IU participates in Cluster 2 dealing with semantic interoperability and information integration - the 3 faculty and 4 PhD students have recently written (with TUC) a state of the art review (activity 2, Task 2.2) on semantic interoperability. Participates in Cluster 7 (DL Evaluation) where it deals with interaction and DL services evaluation. IU has developed and maintains Cluster 7's website (<http://dlib.ionio.gr/wp7>) and discussion forum (<http://dlib.ionio.gr/delosforum>).

10.6 MTA Sztaki DSD

MTA SZTAKI DSD:

- Participation in all of previous DELOS projects
- Development of DL reference model (previous Delos NoE)
- Implementation of the MetaLibrary service for DL evaluation and testbed collection
- Participated in CORES EU project, operating public registry for metadata schemas
- Implementation of a national OAI-based harvester and service provider
- Participating in INFRAWEB project for implementing a framework for Semantic Web Services

10.7 Netlab Knowledge Technologies Group, Lund University (Sweden)

Knowledge Discovery and Digital Library Research Group (KnowLib), Lund University (Sweden)

<http://www.it.lth.se/knowlib/>

Research focuses on networked knowledge organization. KnowLib research group continues work in automated classification and focused Web-crawling (FP6 ALVIS, <http://combine.it.lth.se/>), develops intelligent DL components, eg DELOS harvesting/search, and News. Active in international NKOS efforts. Integration efforts in DELOS include active participation in Cluster 5, and developing components for the DELOS Web-server.

10.8 NTNU

NTNU has experience from FRBR related projects and has current activities on the implementation of a datamodel for FRBR and the application of FRBR in a large scale library catalogue. NTNU is WP7 participant and will coordinate a proposed metadata evaluation JPA with high relevance to this project proposals.

10.9 School of Electronics and Computer Science, University of Southampton (UK)

<http://www.ecs.soton.ac.uk/>

10.10 School of Informatics, University of Edinburgh (UK)

<http://www.inf.ed.ac.uk/>

10.11 Technical University of Crete (Greece)

<http://www.music.tuc.gr/Research/Projects.htm>

The activities of TUC include research, development, training and technology transfer in the area of multimedia information systems. The staff's research interests include Multimedia Information Systems, Very Large Data Bases, Multimedia Communication

Systems, Collaborative Environments, Information Retrieval, Human-Computer Interaction, Electronic Commerce, Tourism and Cultural Systems and Applications.

For this reason the laboratory maintains strong links with other universities, research institutes and high technology companies, all over the world, and actively participates (or has participated) in numerous EU research and development projects (IST, ESPRIT, ACTS, Leonardo Da Vinci, RACE, AIM, DELTA, LINGUA, INCO, STRIDE, SPA etc.).

A second activity is to train graduate and undergraduate students of the Technical University of Crete in advanced technology related to the area of Information Systems. Many members of TUC are also associated with the Technical University of Crete and university students have easy access to the advanced research facilities of TUC and to the experience of its personnel.

A third area of TUC activities consists of technology transfer and collaboration with leading Greek and European companies. TUC has already established strong links with the leading Greek forces in the area of communications and computer technology. These links are maintained through joint participation in EU and National (competitive) projects.

10.12 UKOLN, University of Bath (UK)

<http://www.ukoln.ac.uk/>

UKOLN is a centre of expertise in digital information management, providing advice and services to the library, information, education and cultural heritage communities by:

- Influencing policy and informing practice
- Promoting community-building and consensus-making by actively raising awareness
- Advancing knowledge through research and development
- Building innovative systems and services based on Web technologies
- Acting as an agent for knowledge transfer

Research activity includes work on digital curation, institutional repositories, terminologies, metadata, schemas and registries. Development activity focuses on service-oriented architectures, development of innovative software demonstrators, and advice on digital library standards and protocols. Guidance is also given in areas of policy relating to semantic interoperability, collection level descriptions, bibliographic management, accessibility and Web use.

10.13 University of Athens

The Information Systems and Data Bases Division of the Department of Informatics has broad and intense activity in most areas of information systems and data base systems involving 10 faculty members. Of particular relevance to DELOS is research in formal and methodological aspects of conceptual modeling, information integration and interoperability, faceted ontologies, term composition, model mapping, semantic

similarity, sense disambiguation, concept extraction, temporal aspects of document ranking, question answering systems and NLP techniques for IR, and semantically extended web services. A DL working group has been formed by 4 faculty members of the Division. Close collaboration with FORTH. Until the beginning of 2004 Constantopoulos has led the Information Systems Lab of FORTH and its participation in DELOS 1 and in preparing for DELOS 2.

10.14 University of Glamorgan

The Hypermedia Research Unit has a track record in hypermedia and knowledge organisation systems, projects including semantic hypermedia and spatial ontologies, with three PhDs awarded in this area. The EPSRC-funded FACET project investigated the automatic expansion of faceted search queries, using the semantic relationships inherent in a thesaurus. The project was in collaboration with the J. Paul Getty Trust, who provided the Art and Architecture Thesaurus (AAT) - the primary thesaurus used in the project, and the UK National Museum of Science and Industry (NMSI). An extract of the NMSI Collections Database acted as a test-bed for the project. The EPSRC's final assessment rated the project as 'Outstanding' in Communication of Research Outputs and Cost Effectiveness ('Tending to Outstanding' overall). A case study of the project was featured in a recent thematic issue of the EC project DigiCult, a technology watch for cultural and scientific heritage.

10.15 UNIMI, University of Milan (Italy)

<http://www.unimi.it/engl/>

<http://dakwe.dico.unimi.it/>

Our research activities cross the areas of Artificial Intelligence, Database Systems, and Mobile Computing. Reasoning techniques and well-founded logical approaches are applied to data and knowledge management. A theoretical line of research investigates time related aspects in data and knowledge management. A more applicative line of research investigates the application of knowledge-based techniques to different problems in mobile computing.

11 References

ACM. (1998) ACM Computing Classification System. Available at: <http://www.acm.org/class/1998/>

Arapi, P. , Moumoutzis, N. & Christodoulakis, S. (2005). DELOS WP5 -Task 5.4 : Interoperability of eLearning applications with AV digital libraries. Working Draft. 11th June 2005

Baker, T., Dekkers, M., Heery, R., Patel, M. & Salokhe, G. (2001). What terms does your metadata use? Application profiles as machine-understandable narratives. Journal of Digital Information, Vol.2, No.2. Available at: <http://jodi.ecs.soton.ac.uk/Articles/v02/i02/Baker/baker-final.pdf>

- Bloom, B.S. & Krathwohl, D.R. (1965). *Taxonomy of Educational Objectives: The Classification of Educational Goals: Handbook I, Cognitive Domain*. Longman, New York
- Capuano, N., Gaeta, M., Lannone, R. & Orciuoli, F. (2005). Learning Design and run-time resource binding in a distributed e-learning environment. Proceedings of the 1st International Kaleidoscope Learning Grid SIG Workshop on Distributed e-Learning Environments. Vico Equense (Naples), Italy, 14 March 2005
- Christodoulakis, S., Arapi, P., Moumoutzis, N., Patel, M., Kapidakis, S., Arahova, A., Bountouri, L. (2005). Interoperability of eLearning Applications with Audiovisual Digital Libraries. DELOS poster session in conjunction with ECDL 2005, September 2005, Vienna, Austria
- ELENA Project. Available at: <http://www.elena-project.org/en/index.asp?p=1-1>
- Heery, R. & Patel, M. (2004). Application profiles: mixing and matching metadata schemas. *Ariadne*, No. 25, September, last updated October 04, 2004 <http://www.ariadne.ac.uk/issue25/app-profiles/intro.html>
- Hummel, H. G. K., Manderveld, J. M., Tattersall, C. & Koper, E. J. R. (2004). Educational Modelling Language and Learning Design: new opportunities for instructional reusability and personalised learning. *Int. J. Learning Technology*, 1(1), pp. 111-126
- IEEE LOM. (2002) IEEE 1484.12.1-2002 Learning Object Metadata Standard. Available at: <http://ltsc.ieee.org/wg12/>
- IEEE PAPI. (2002). Personal and Private Information (PAPI), Draft 8 Specification. Available at: <http://edutool.com/papi>
- IMS DRI. (2003). IMS Digital Repositories specification V1.0. Available at: <http://www.imsglobal.org/digitalrepositories/>
- IMS LD. (2003). IMS Learning Design specification V1.0. Available at: <http://www.imsglobal.org/learningdesign/index.cfm>
- IMS LIP. (2005). IMS Learner Information Package Specification V1.0.1 Available at: <http://www.imsglobal.org/profiles/>
- Jena Semantic Web Framework, <http://jena.sourceforge.net/>
- Keenoy, K., Levene, M. & Peterson, D. (2004). SeLeNe –Self E-Learning Networks (IST-2001-39045); WP4 Deliverable 4.2, Personalisation and Trails in Self e-Learning Networks, Jan 2004
- McLean, N. (2004). The Ecology of Repository Services: A Cosmic View. Keynote Address. 8th European Digital Library Conference ECDL 2004. UK
- METS. (2005). Metadata Encoding and Transmission Standard (METS) Official Website. Available at: <http://www.loc.gov/standards/mets/>
- MPEG7. (2001). ISO/IEC JTC 1/SC 29/WG 11/N3966, Text of 15938-5 FCD Information Technology – Multimedia Content Description Interface – Part 5 Multimedia Description Schemes. MPEG Group, Singapore 2001

- MPEG7. (2003). ISO/IEC JTC1/SC29/WG11N5525 MPEG-7 Overview (version 9). International Organisation for Standardisation (ISO), Martines, J. (Ed.), Pattaya 2003. Available at: <http://www.chiariglione.org/mpeg/standards/mpeg-7/mpeg-7.htm>
- OWL. (2004) Web Ontology Language (OWL). Available at: <http://www.w3.org/TR/owl-ref/>
- Prud'hommeaux, E.& Seaborne, A. (2005). SPARQL Query Language for RDF, W3C Working Draft, 23 Nov. 2005
- SCORM. (2004). Sharable Content Object Reference Model (SCORM) 2004 2nd edition. Available at: <http://www.adlnet.org>
- Tsinaraki, C., Polydoros, P., Christodoulakis, S (2004). Integration of OWL ontologies in MPEG-7 and TVAnytime compliant Semantic Indexing, CAiSE 2004