

SIXTH FRAMEWORK PROGRAMME

PRIORITY IST-2002-2.3.1.12

Technology-enhanced Learning and Access to Cultural Heritage



Information Society
Technologies



Contract for:

NETWORK OF EXCELLENCE

Annex 1 - "Description of Work"

Detailed JPA for the period January 2006 – June 2007

Network acronym: **DELOS**

Network full title: DELOS: a Network of Excellence on Digital Libraries

Proposal/Contract no.: **G038-507618**

Related to other Contract no.: IST – 1999 - 12262

Date of preparation of Annex 1: **01 December 2003 (Final)**

Operative commencement date of contract: **01 January 2004**

DELOS NoE

Detailed Joint Programme of Activities – third period (JPA3, 1 January 2006 to 30 June 2007)

Table of Contents

UPDATED LIST OF PARTICIPANTS FOR JPA3	3
DETAILED JOINT PROGRAMME OF ACTIVITIES 3 (JPA3) – (MONTH 25 TO 42)	5
INTRODUCTION	5
<i>Continuity</i>	5
<i>Integration</i>	5
<i>Technology transfer</i>	7
<i>Co-operation with other projects</i>	7
WP1 - DIGITAL LIBRARY ARCHITECTURE	9
<i>Strategic Objective</i>	9
<i>How will JPA2 tasks continue in JPA3</i>	9
<i>New Task in JPA3</i>	9
WP2 - INFORMATION ACCESS AND PERSONALIZATION	11
<i>Strategic objectives</i>	11
<i>Workpackage JPA2 Activities</i>	11
<i>Plans for JPA3 Activities</i>	12
<i>Inter-Workpackage and Intra-Workpackage Integration</i>	13
WP3 - AUDIO/VISUAL AND NON-TRADITIONAL OBJECTS	13
<i>Strategic Objectives</i>	13
<i>Workpackage Joint Research Activities and Integration Activities in JPA3</i>	14
<i>Expected Results</i>	14
<i>Description of Tasks</i>	15
WP4 - USER INTERFACES AND VISUALIZATION	19
<i>Strategic Objectives</i>	19
<i>Workpackage Activities and Integration</i>	19
<i>Expected Results</i>	21
WP5 - KNOWLEDGE EXTRACTION & SEMANTIC INTEROPERABILITY	21
<i>Strategic Objectives</i>	21
<i>Framework of JPA3 Activities</i>	21
<i>Expected results</i>	24
<i>Inter-Cluster and Intra-Cluster Integration Commentary</i>	24
WP6 - PRESERVATION	25
<i>Strategic Objectives</i>	25
<i>Transition to JPA3</i>	25
<i>WorkPackage Activities and Expected Results</i>	25
<i>Inter-Cluster and Intra-Cluster Integration</i>	27
<i>Expected Results</i>	27
WP7 - EVALUATION	28
<i>Strategic Objectives</i>	28
<i>Workpackage Activities and Integration</i>	28
<i>Expected Results</i>	29
WP8 - DISSEMINATION AND TECHNOLOGY TRANSFER	30
<i>Strategic Objectives</i>	30
<i>Technology transfer to the library and the cultural communities</i>	32
WP9 - ASSESSMENT	35
WP10 - SCIENTIFIC COORDINATION	35
WP11 - ADMINISTRATIVE AND FINANCIAL MANAGEMENT	35

Contractor list**List of Participants**

CO = Coordinator; CR = Contractor

Role	N.	Participant legal name	Short name	Country	Enter	Exit
CO	1	GEIE-ERCIM	ERCIM	France	1	48
CO	2	Consiglio Nazionale delle Ricerche	CNR-ISTI	Italy	1	48
CR	3	Eidgenoessische Technische Hochschule Zurich	ETH Zurich	Switzerland	1	48
CR	4	University of Bath	UKOLN	UK	1	48
CR	5	National and Capodistrian University of Athens	UOA	Greece	1	48
CR	6	Technical University of Crete	TUC	Greece	1	48
CR	7	Università degli Studi di Firenze	UNIFI-MICC	Italy	1	48
CR	8	Fraunhofer-Gesellschaft zur Förderung der Angewandten Forschung	FHG/IPSI	Germany	1	48
CR	9	University of Glasgow	GU	UK	1	48
CR	10	Universität Duisburg-Essen	UNIDU	Germany	1	48
CR	11	Stichting Centrum voor Wiskunde en Informatica	CWI	Netherlands	1	48
CR	12	Risoe National Library	RISOE.DK	Danemark	1	12
CR	13	Foundation for Research and Technology – Hellas	FORTH	Greece	1	48
CR	14	Università degli Studi di Roma “La Sapienza”	ROMA1	Italy	1	48
CR	15	Brunel University	UBRUN	UK	1	12
CR	16	Institute of Communication and Computer Systems	ICCS	Greece	1	48
CR	17	Università degli Studi di Padova	UNIPD	Italy	1	48
CR	18	Università degli Studi di Milano	UNIMI	Italy	1	48
CR	19	Institut fuer Medizinische Informatik und Technik Tyrol	UMIT	Austria	1	27
CR	20	Max-Planck-Gesellschaft z.F.d.W. represented by: MPI fuer Informatik	MPII	Germany	1	48
CR	21	Kuratorium OFFIS E.V.	OFFIS	Germany	1	48
CR	22	Queen Mary & Westfield College, University of London	QMUL	UK	1	48
CR	23	University of Strathclyde	USG	UK	1	48
CR	24	Ionian University, Archive and Library Science Department	IU	Greece	1	48
CR	25	Université Paris-sud XI	UPSXI	France	1	48
CR	26	The University of Southampton	UOS	UK	1	48

Role	N.	Participant legal name	Short name	Country	Enter	Exit
CR	27	The University of Edinburgh	UEDIN	UK	1	48
CR	28	Institut fuer Informationsverarbeitung und Computergestuetze Neue Medien	IICM	Austria	1	48
CR	29	Technische Universität Wien	TUW	Austria	1	48
CR	30	Universita' degli Studi di Urbino Carlo Bo	UNIURB	Italy	1	48
CR	31	Norges Teknisk-Naturvitenskapelige Universitet NTNU	NTNU	Norway	1	48
CR	32	Lunds Universität	ULUND	Sweden	1	48
CR	33	Institut National de Recherche en Informatique et en Automatique	INRIA	France	1	48
CR	34	Aristotle University of Thessalonikis	AUTH	Greece	1	12
CR	35	Nationaal Archief	NANETH	Netherlands	1	48
CR	36	Univesität Bremen	TZI	Germany	1	48
CR	37	Österreichischen Akademie der Wissenschaften	OEAW	Austria	1	48
CR	38	University of Leeds	UNIVLEEDS	UK	0	0
CR	39	24Universität zu Köln	UCO	Germany	1	48
CR	40	Forma (Centro di) Formazione e Ricerche per Metodologie Applicate alla Conservazione, Gestione e Comunicazione di Beni Culturali	CF	Italy	1	48
CR	41	SICS, Swedish Institute of Computer Science AB	SICS	Sweden	1	48
CR	42	Universita' di Modena e Reggio Emilia	UNIMORE	Italy	1	48
CR	43	Masarykova Universita v Brne	MUNI	Czech Republic	1	48
CR	44	Universiteit van Amsterdam	UVA	Netherlands	1	48
CR	45	Universita' della Svizzera Italiana	UNISI	Switzerland	1	48
CR	46	Magyar Tudományos Akademia Szamitastechnikai es Automatizalasi Kutatointezet	MTA SZTAKI	Hungary	1	48
CR	47	Universita' degli Studi di Bari – Dipartimento di Informatica	UNIBA	Italy	1	48
CR	48	Health Information Technologies Tyrol	HITT	Austria	13	48
CR	49	Lancaster University	UoLanc	UK	13	48
CR	50	ATHENS University of Economics and Business	AUEB	Greece	13	48
CR	51	University of Glamorgan	UGLAM	UK	13	48
CR	52	University of Queensland	Uqueensland	Australia	0	0
CR	53	Austrian National Library	OENB	Austria	13	48
CR	54	Goettingen State and University Library	SUB	Germany	13	48
CR	55	Imperial College	Imperial	UK	13	48
CR	56	Institute of Knowledge Sharing	IKS	DK	13	48
CR	57	Center of Cognitive Systems Engineering	CSE	DK	13	48
CR	58	Centre Virtuel de Connaissance sur l'Europe	CVCE	Luxembourg	25	48
CR	59	Universite de Lille	USTL	France	25	48
CR	60	University of Konstanz	UKON	Germany	25	48
CR	61	University of Basel	UNIBAS	Switzerland	25	48
CR	62	University of Queensland	Uqueensland	Australia	25	48

For the new partners, see details in appendix 1

Detailed joint programme of activities 3 (JPA3) – (month 25 to 42)

Introduction

The DELOS Network of Excellence is carrying out a broad range of interrelated activities whose combined effect should be to contribute to the successful achievement of the objectives as defined in Section 2 of the Technical Annex.

During the first year, the DELOS activities (JPA1) focused on tightening up the Network by dedicating efforts to constructing a strong community sense among the Network members, and setting up the necessary infrastructure (Website, DELOS Digital Library, BSCW) to support collaboration and integration between the members.

The Joint Programme of Activities for the second period (JPA2) was defined as the result of a Call for Proposals among the DELOS members. At the end of an exhaustive selection process, the Scientific Board selected 25 of the 59 proposals received. JPA2 has increased the intra Task integration and internal cohesion of the Network as the Task activities are jointly carried out by an average of 4,2 research teams per Task. One of these Tasks has the objective of defining a reference model for digital library systems. This model should reflect the new vision for digital libraries defined by DELOS during the first year.

Both, the DELOS vision for digital libraries and the reference model constitute a reference framework within which the activities of JPA3 will be carried out in an integrated and coordinated manner. The key aspects of JPA3 will be: *continuity, integration, technology transfer and co-operation with other projects*. However, the first step towards the establishment of JPA3 has been an evaluation of the results of JPA2. Each Task Leader was asked to produce a self-assessment report regarding her/his Task. These reports were submitted to the "Integration Task Force" which made an assessment of the performance of each Task. In addition, the results of the activities carried out by the Tasks were presented to the Advisory Board, which formulated an overall judgment on the performance of the DELOS Tasks.

Continuity

JPA3 will represent an evolution of JPA2 based on the successful evaluation of JPA2 Tasks. All Tasks of JPA2 renewed, refocused and with expanded goals will continue under JPA3. In addition, five new Tasks will be started: Integrated Prototype of a Digital Library Management System (WP1 Task T1.8); Validation and Refinement of the Reference Model through Interaction with TEL (WP8 Task T8.2); Multi-lingual Information Access in TEL (WP8 Task T8.3); Personalization Capabilities in TEL (WP8 Task T8.4); User Interface Design for TEL, navigation and Visualization Services (WP8 Task T8.5).

Integration

A deeper exchange of working experiences and results between the Tasks will be achieved. In order to increase the level of integration between the Tasks in JPA3 an "Integration Task Force" has been established with the task of identifying areas of synergy between the Tasks, over and beyond the existing workpackage structure, and of producing recommendations for future integration activities. The Task force will remain in place also during JPA3, to facilitate and monitor the progress of the integration activities, periodically reporting the status of those activities to the Scientific Board. During JPA2, the Integration Task Force produced a report containing recommendations for future integration plans. More precisely, the ITF suggested the creation of "integration clusters", assessed on a topic basis, where each integration cluster would contain Tasks coming from the same or different Workpackages, based on the degree of relevance of each Task to the common topic. In particular, the

creation of the following clusters has been suggested. It has to be noted that some tasks appear in more than one integration cluster, as they can provide useful contributions to different topics.

Integration Cluster 1, the Reference Model. The reference model, i.e. a formal and conceptual framework describing the characteristics of a Digital Library Management Systems (DLMS), is going to be a fundamental piece of work, not only for its technical value, but also for its nature of being at the same time a technical tool, an integration tool, and a “policy tool”, providing a frame of reference to be used to aggregate the various efforts going within and outside of DELOS. Task T1.4 is therefore open to collaboration with external initiatives (such as the e-Framework and the DLF Service Framework), input from all the DELOS Workpackages and more specifically from all the DELOS Tasks dealing with the modeling of specific aspects of a digital library. They are: Task T4.7 (User Requirements-driven Support for a DL Design Framework), Task T4.10 (Design, Implementation and Evaluation of Multimedia Annotations for User Collaboration), Task T6.5 (Enabling the Integration of Digital Preservation Architectures) and Task T7.5 (A Digital Library Testbed Framework).

Integration Cluster 2, the Peer-to-Peer paradigm. Peer-to-peer (P2P) architectures easily allow for loosely coupled integration of different information providers, where data and documents cannot be integrated into a single source. This also includes the user clients so as to facilitate some collaborative data sharing among them (e.g. annotations, recommender systems). Several Tasks are addressing different aspects of peer-to-peer systems (both at the system level and at the application level) and a closer interaction between them will lead to mutual benefits. The tasks involved are: T1.6 (Management of and Access to Virtual Electronic Health Records), T1.7 (Integration of Data Stream Management into an eHealth Digital Library), T2.6 (Advanced Access Structures for Complex Similarity Measures), T2.7 (Application of the P2P Paradigm in Digital Libraries) and T2.8 (Personalized Query Routing in Peer-to-Peer Federations of Digital Libraries).

Integration Cluster 3, complex similarity measures. Complex similarity measures and indexing are becoming more and more relevant with the increase of non-alphanumeric content in digital libraries. Task T2.6 (Advanced Access Structures for Complex Similarity Measures) could obtain solid results in indexing and retrieval by similarity through interaction with Tasks T1.6 (Management of and Access to Virtual Electronic Health Records), T1.7 (Integration of Data Stream Management into an eHealth Digital Library), T3.6 (Video Annotation with Pictorially Enriched Ontologies), T3.8 (Description, Matching, and Retrieval By Content of 3D Objects), T4.5a (Visualization and relevance feedback).

Integration Cluster 4, personalized information access. Personalization is becoming increasingly relevant as the quantity of information available on line increases. Several Tasks are addressing tools and services that are at the base of personalization, such as content integration, use of context, personal ontology, smart browsing, relevance feedback, query refinement/enlargement. The tasks involved are: T2.8 (Personalized Query Routing in Peer-to-Peer Federations of Digital Libraries), T2.9 (Context-dependent Access to Digital Libraries), T2.10 (Modelling of User Preferences in Digital Libraries), T4.5a (Visualisation and relevance feedback), T4.7 (User Requirements-driven Support for a DL Design Framework), T4.8 (Task-centered Information Management).

Integration Cluster 5, ontology mapping, interoperability and metadata. The ability to map between core and domain ontologies is very important to achieve semantic interoperability both at the data and metadata level. Tasks T5.4 (Interoperability of eLearning Applications with Digital Libraries) and T5.5 (Ontology-driven Interoperability) are addressing exactly those topics and have direct connections with some of the work in WP3 through the common use of the *GraphOnto* tool, used in several WP3 and WP5 Tasks (T3.6, T3.9, T3.10, T3.11, T5.4 and T5.5), and which can be extended to provide full ontology mappings and query mappings. In addition to ontology mapping, the tasks participating in this cluster will also focus on metadata interoperability.

Finally, a global integration effort in JPA3 will be the implementation of a prototype of a digital library management system, which can show combined text and audio-visual search functionality, personalized browsing using new information visualization and relevance feedback tools, annotation and processing of retrieved information, integration and processing of sensor data stream, and finally, from a systems engineering point of view, allow simple configuration and adaptation while being reliable and scalable. The prototype will be built by integrating digital library functionality provided by

the DELOS partners into the OSIRIR/ISIS platform, i.e., a middleware developed by ETH. The platform provides generic (application independent) services that include the registration of services and processes, interface for application development, engine for decentralized execution of (search and maintenance) processes, service for load balancing and failure treatment. In order to meet this challenging objective a close collaboration between many partners is necessary, as for all building blocks that are candidates for inclusion in the prototype, a careful analysis of the systems architecture need to be performed to identify, from a technical perspective, if a building block can be integrated at all and whether the integration should be done in a loosely-coupled or in a tightly-coupled way. A more complete description of the effort is in WP1 Task T1.8.

Technology transfer

DELOS has started a series of actions specifically intended to transfer to some user communities the results achieved and the prototypes developed by the DELOS members. The focus in JPA3 is on the library and cultural heritage communities, through cooperation with the TEL Office (The European Library) and with MICHAEL (representing the community of the Ministries of Cultural Assets). It is also planned to continue the participation in joint activities and events with ELAG (European Library Automation Group) and to start participation in LIBER (Ligue de Bibliothèques Européennes de Recherche) and LIDA (Libraries in the Digital Age) events. Other opportunities for cooperation with museums and archives will be sought and considered during JPA3. Presently, a number of activities aiming at transferring digital library functionality developed by DELOS members to the TEL system are in plan and will be carried out jointly with the TEL Technical Working Group. The following four major areas of common interest have been identified.

The reference model. It will be possible to capitalize on the expertise of the TEL participants to bring into the reference model activities valuable input for its refinement; at the same time it is expected that the conceptual framework resulting from the reference model will provide valuable input to TEL for extensions and refinement of the architecture of The European Library system.

Multilingual capabilities. TEL users should be able to access and search the library in their own (or preferred) language, retrieve documents in other languages and have the results presented in an interpretable fashion (e.g. possibly with a summary of the contents in their chosen language). One of the first activities will be a feasibility study aimed at producing guidelines and strategy for the attainment of this long-term ambitious goal.

Personalization. A first step will be the development of personalization guidelines to identify those services more suitable for personalization, followed by a second step for integration and testing of existing prototype software and development of possible improvements over those services.

User interface and Visualization. The first step will be an evaluation of the existing TEL user interface, as well as the exploration of additional services, especially for supporting query formulation, collection navigation and results visualization. The activities in this area will focus on four topics: evaluation, support for query formulation, virtual collections and navigation, presentation/visualization of results.

For each area of interest a specific new Task has been defined in Workpackage 8, and the new Tasks are: T8.2 (Validation and refinement of the reference model through interaction with TEL), T8.3 (Multi-Lingual Information Access in TEL), T8.4 (Personalization capabilities in TEL), T8.5 (User interface design for TEL, navigation and visualization services).

Co-operation with other projects

In JPA3 DELOS will further develop the collaborative links with the group of projects and initiatives funded by the European Commission in the context of the 5th and 6th Framework Programmes (BRICKS, MINERVA, PRESTOSPACE, TEL, TEL-ME-MOR, MICHAEL, EPOCH, MUSCLE, DILIGENT, DRIVER, ENGAGE). As already started in JPA2, a number of joint events (working groups, technical workshops, panels, summer schools etc.) will be organized for the purpose of exchanging information, sharing working experiences, and explore possibilities of closer cooperation.

For continuity with the past, the activities of JPA3 will continue to be organized into topic-specific workpackages, where each workpackage is structured according to the Workpackage-Task hierarchy. As in JPA2, four workpackages cover aspects regarding DELOS management, assessment and dissemination. The Network is thus composed of the following Workpackages for the third 18-month period:

- WP1 Digital Library Architecture
- WP2 Information Access and Personalization
- WP3 Audio/Visual and Non-traditional Objects
- WP4 User Interfaces and Visualization
- WP5 Knowledge Extraction and Semantic Interoperability
- WP6 Preservation
- WP7 Evaluation
- WP8 Dissemination and Technology Transfer
- WP9 Assessment
- WP10 Administrative Management
- WP11 Scientific Management

WP1 - Digital Library Architecture

Strategic Objective

The overall goal of this workpackage is to analyze, develop, and integrate architectures and technology for digital libraries that enable the building of the next generation digital library management systems. While in the first phase we concentrated on establishing the basis via two workshops and two survey documents as expressed by the deliverables in JPA1 we have undertaken serious efforts in prototype development and integration of components in JPA2 and will intensify global integration in our plans for JPA3. The Architecture Work Package (WP1) by its very nature will take a lead in this integration effort. The ambitious objective is to have a demonstrator for future Digital Library Management Systems that not only shows new combined text and audio-visual search functionality and personalized browsing by new adaptable information visualization and relevance feedback tools at the interface but also proves that generic systems can be build that not only support finding of relevant information but also enable to annotate and process found information, to integrate sensor data stream processing, and – from a systems engineering point of view - allows simple configuration and adaptation while being reliable and scalable.

How will JPA2 tasks continue in JPA3

The WP1 tasks from JPA1 and JPA2 will continue but there will be a shift in focus from the stage of foundational research towards evaluation and towards transforming results into prototypes and building blocks that have the potential to be integrated into the global prototype system. Work on the DL reference model will continue in close relationship and mutual benefit with the global prototype development. An important new aspect will be the inclusion of representatives from the TEL project. Work on an annotation service (in combination with WP4) and its prototype system will be further developed and integrated into the global prototype. Upon recommendation of the first review two new tasks had been started on special purpose digital library architectures for e-Health, the distributed health record task and the data stream task. While we will continue the distributed health record task we will generalise the stream data task. The distributed health record task is important because of its rich nature of different kinds of data (multimedia, textual, structured, factual) and gives an excellent case study for the architectural and functional aspects of a future DL. Stream data originate from various sensors for context and situation awareness. Sensors for health monitoring fall into this category but are a very special case. Generally sensor data play an important role for personalization and adaptation of information personal needs. Also in the area of mobile information and interfaces, work on the “Active Paper” project will further be developed still as continuation of the JPA1 task but now with emphasis on its integration into the global prototype.

New Task in JPA3

Task 1.8 – Integrated prototype of a Digital Library Management System (DELOS DLMS).

A central task in JPA3 is the global prototype development. The objective is to build a joint prototype for the future Digital Library Management System that makes available results of many groups in DELOS. The prototype is based on the OSIRIS/ISIS middleware that has been started at ETH for ETHWorld, the virtual campus of ETH. The system was further developed and extended for data streams and for medical objects at UMIT and now will be further extended at the University of Basel. The system provides generic, i.e. application independent, services that include the registration of services and processes, interface for application development, engine for decentralized execution of (search and maintenance) processes, service for load balancing and failure treatment. As for Digital Library functions it has already rudimentary support for textual and audiovisual search facilities and combinations of it and it has primitive support for relevance feedback and visualization. Therefore it seems doable in a relatively small project to initiate collaboration with several partners of DELOS in order to replace existing components and in order to add missing functionality. The ambitious objective is to have a demonstrator for future Digital Library Management Systems that not only shows new

combined text and audio-visual search functionality and personalized browsing by new adaptable information visualization and relevance feedback tools at the interface but also proves that generic systems can be build that not only support finding of relevant information but also enable to annotate and process found information, to integrate sensor data stream processing, and – from a systems engineering point of view - allows simple configuration and adaptation while being reliable and scalable.

In order to meet this challenging objective close collaboration between many partners is necessary and will be realized by many short term exchanges of scientists. The coordination of this effort is in WP1 but effort from all WPs must be spent in order to be successful. In more detail and in order to give examples various building blocks from the following partners are already available and have a high potential to be integrated in a loose or even tight manner as described in the OSIRIS/ISIS document available from the WP1 website:

- From WP1, CNR and their OpenDlib prototype, we take sophisticated term extraction from text, text indexing and collection management. From Padua and partners in the task 1.5 annotation services are provided. From task 1.7 (Basel former UMIT), reliable sensor data management will be added to the existing prototype.
- From WP2 tasks, especially from the CNR MILOS system and from task 2.6 we can use various kinds of multimedia indexing and compare their performance when applied to different multimedia collections: Search process generation and personalization services are under development in WP2, UoA, and will be incorporated as a major improvement over the existing system
- Another major improvement will be brought in by tasks from WP3 for image feature extraction, for 3D shape recognition and for related special indexing techniques as well as for video retrieval. From Vienna audio feature extraction and audio retrieval will be incorporated.
- From WP4 and from partners in their tasks visualization services and visual relevance feedback will be incorporated as well as SOM visualization from Vienna. Active paper from ETH is another very useful services at the interface.
- From WP5 we expect services for various transformations between standards in order to increase interoperability and the use of ontologies. Specifically the Graphonto prototype system and natural language access from Crete are very promising services that improve the functionality
- From WP6 preservation services will be available.
- Multi-linguality will be provided by language translation services made available from WP7.
- Basel/Konstanz: will take care of the OSIRIS/ISIS infrastructure that must be extended and interoperability aspects as well as client services taken from the Daffodil prototype will be investigated jointly.

This list is not complete. It gives examples of building blocks. Other building blocks will be identified in the course of the project and after a “call for service integration”.

For all building blocks that are candidates for being included in the DELOS DLMS prototype, first a careful analysis of the systems architecture will be performed. The goal is to identify, from a technical perspective, if a building block can be integrated at all and whether this should be done in a loosely-coupled or in a tightly-coupled way. Also the efforts needed to prepare a building block for integration will be estimated. After this analysis, the actual integration efforts will take place in close collaboration between the organization responsible for a building block and the task coordinators.

WP2 - Information Access and Personalization

Strategic objectives

The main activities of WP2 are organized along three dimensions: access of information in a single provider; integration of information in multiple providers; personalization, i.e., customization of system behavior to the user.

Workpackage JPA2 Activities

During JPA2, the IAP cluster continued its cross-fertilization activities, their highlight being the co-organization (with the Architecture cluster) of the 8th DELOS Thematic Workshop on “Future Digital Library Systems – Architectures and Information Access”, in Schloss Dagstuhl, Germany, March 29-April 1, 2006. In addition to those, the IAP cluster started five new activities, which represent a logical continuation of the earlier work. These are briefly described below:

Task 2.6 - Advanced Access Structures for Complex Similarity Measures: The aim of this task is to broaden the filtering approach towards metric indexing in order to provide indexing proposals for metric spaces that (i) efficiently support nearest-neighbor, range, and ranking queries, (ii) operate on any kind of multimedia data, and (iii) are generic in the distance measure to be employed.

So far, an efficient indexing method has been developed, which adopts quantization-based approximation techniques, pivot-based clustering, and a novel principle of self-refinement to yield substantial cost reduction in similarity search atop metric distance measures. Both an indexing and a retrieval component for ranked queries have been implemented to support content-based image retrieval. Furthermore, DL-oriented similarity search techniques and scalable and distributed similarity search structures have been investigated.

Task 2.7 - Application of the P2P Paradigm in Digital Libraries: Application of the P2P Paradigm in Digital Libraries: The aim of this task is to support large-scale, decentralized sharing of data and services in a network of autonomous and heterogeneous DL nodes operating under the peer-to-peer (P2P) paradigm. Emphasis is on query reformulation and query roaming.

So far, the emphasis has been on view-based peer advertisements and queries expressed in tree (e.g., Xpath) and graph (e.g., RQL) query languages. In particular, several ways have been developed for organizing in a distributed way the knowledge about what DL nodes can actually contribute to a specific query/view. Efficient query routing algorithms has been devised by employing the intentional information of XML or RDFS views published by DL nodes.

Task 2.8 - Personalized Query Routing in Peer-to-Peer Federations of Digital Libraries: This task explores routing of various types of queries (SQL, XQuery, etc.) over a P2P network where, apart from DLs, user agents with powerful personalized tools may participate as peers as well. It aims at novel strategies for query routing that can effectively exploit profile information on the participating users.

So far, all relevant existing technologies have been brought together and aligned for the specific purpose of devising novel strategies for P2P query routing. These include preference models for personalized queries and query execution strategies, methods for similarity search in distributed systems, and strategies for information filtering in publish-subscribe systems. In addition, collaboration has started to investigate how to leverage existing prototype platforms (Minerva at MPII and Pepper at UNIDU) to integrate the two into a more powerful system as well as to integrate components developed by others.

Task 2.9 - Context-dependent Access to Digital Libraries: This task investigates how the notion of context can be incorporated into data stored in DLs. It includes laying the foundations for a context-aware DL system and defining a uniform mechanism for personalized/context-dependent access to DLs.

So far, the Context Relational model (CR model) has been defined, which extends the relational model by incorporating the notion of context. Several alternative techniques for storing context using

conventional tables have been investigated and how each one affects performance of context-aware operations. Implementation of a prototype that incorporates these techniques has already started.

Task 2.10 - Modeling of User Preferences in Digital Libraries: The aim of this task is to study the provision of services for the management of user preferences, focusing on query personalization (enhancing a query by incorporating into it user-specific preferences), user notification (notifying a user when an “event” of interest happens), and document customization (allowing the user to compose documents that will be “materialized” by the system).

So far, a formal framework for the definition of qualitative and quantitative preferences has been developed as a common basis for comparison between the two and as a foundation for their integration. Furthermore, several qualitative and quantitative models have been devised and/or expanded from earlier ones, capturing different aspects and forms of human preferences.

With respect to the three main directions of the cluster, one may place the above tasks as follows:

Information Access	Task 2.6
Information Integration	Task 2.7, Task 2.8
Personalization	Task 2.8, Task 2.9, Task 2.10

From this and by comparing the description of the tasks, one may additionally observe the following unifying themes for sets of tasks: emphasis on the peer-to-peer (P2P) architecture; profile-based customization of multiple behaviors (query reformulation, query routing, query processing, etc.), using profiles of the particular user or the particular context. In some sense, the IAP cluster activities are centered around 3 Ps:

P-roximity	of objects as a similarity measure
P-eer to Peer	as the distributed architecture
P-rofiles	of users and contexts as a customization approach

Plans for JPA3 Activities

All five tasks have made progress towards solving the hard problems within their purview. At the same time, they all have several issues that they have not been able to address or solve optimally. Hence, the plan for JPA3 is for the same tasks to continue, with renewed, refocused, and expanded goals. The particulars for each task are shown below:

Task 2.6 - Advanced Access Structures for Complex Similarity Measures: The plan is to refine and extend access structures for enhanced distributed search and increased search effectiveness. In particular, new directions will be explored to build access methods for image similarity search, such as access methods for region-based image retrieval.. Image segmentation, clustering, and feature weighting will be exploited to index images in a compact way. Furthermore, new structures will be developed to achieve not only scalable but also tunable performance of similarity search. All these will be complemented by joint experimentation among partners.

Task 2.7 - Application of the P2P Paradigm in Digital Libraries: The premise is that the full query processing and execution capabilities of all DL-node views will be returned by the query routing service. Mapping rules established between the individual peer schemas will be used to route queries in the P2P DL network. The use of such rules will deal with schema heterogeneity among DL nodes. Furthermore, the current prototype system of P2P DLs using Structured Overlay Networks will be extended to use mapping rules and so address query reformulation issues.

Task 2.8 - Personalized Query Routing in Peer-to-Peer Federations of Digital Libraries: A priority will be to integrate different profiling and routing strategies into one or both of the existing prototype systems Minerva (at MPII) and Pepper (at UNIDU), and coupling the two prototypes into a

single P2P federation. This will permit to extensive distributed experimentation and collection of realistic user-behavior information. Furthermore, the P2P paradigm for publish-subscribe applications will be explored and specific application scenarios from the e-health sector will be studied.

Task 2.9 - Context-dependent Access to Digital Libraries: A new direction that will be explored is the definition of dynamically definable/verifiable context, i.e., a model in which a context navigation tree will act as a guide to the appropriate query associated with the user's context. The user's query will be a possibly incomplete navigation tree which will assist the user in finding the information he/she is looking for. An appropriate information model based on context-aware navigation trees will be devised and techniques for context-dependent preferences, preferences as context, and other relationships between the two will be studied as well.

Task 2.10 - Modelling of User Preferences in Digital Libraries: The study of preference modelling will continue, exploring new forms of preferences, unifying the foundations of them, and devising techniques for the computation of personalized, preference-influenced queries. Of particular importance will be the study of the interplay between qualitative and quantitative preferences, focusing on identifying commonalities and differences, devising user models combining both types, and developing algorithms for personalizing queries using the combined profiles that are incrementally updateable. Finally, the use of profiles in both query-initiated (pull) and system-initiated (push) computational models will be studied as well.

Inter-Workpackage and Intra-Workpackage Integration

Inter-Workpackage integration is evident among most of the tasks, in terms of their themes, activities, and participating partners. There is interaction between Tasks 2.7 and 2.8 (P2P theme), Tasks 2.8 and Task 2.10 (preference modeling and use), and Tasks 2.9 and 2.10 (interplay of preference and context themes, common meetings).

Several of the Workpackage activities are also in collaboration or in communication with tasks in other Workpackages. Indicative of such interactions are Task 2.6 with WP3, Tasks 2.7 and 2.8 with WP1, and Task 2.10 with Tasks 4.5a and 4.8 of WP4.

WP3 - Audio/Visual and Non-traditional Objects

Strategic Objectives

The overall long term research objective for WP3 is to formalize and develop a *Semantic Multimedia Management Framework (SMMF)* supporting the Audio-Visual aspects of the overall DELOS vision for Digital Libraries as well as the global demonstrator for future Digital Libraries described in WP1. Core research targets for this framework include: Metadata Capturing for Audio-Visual Content, Universal Efficient Access and Interactions with Audio-Visual Libraries, and effective Management of the Audio-visual Content.

The tasks of WP3 are directed towards the above three targets. In particular the six tasks described below, along with those WP5 tasks (focusing on semantic interoperability) in which WP3 members participate, cover the three main objectives of WP3 as follows:

Metadata Capturing for Audio-Visual Content:

- T3.6: Video Annotation with Pictorially Enriched Ontologies
- T3.9: Automatic, Context-of-Capture Based Categorization, Structure Detection and Segmentation of News Telecasts
- T3.7: Multimedia Interfaces for Mobile Applications (in cooperation with WP4)

Universal Access and Interactions with Audio-Visual Libraries:

- T3.10: Content and Context-Aware Multimedia Content Retrieval, Delivery, and Presentation
- T3.8: Description, Matching and Retrieval by Content of 3D Objects
- T3.11: Natural Language and Speech Interfaces to Knowledge Repositories (in close cooperation with WP4, where part of this task is being brought forward as Task 4.9)

Management of the Audio-Visual Content:

- T2.6: Advanced Access Structures for Similarity Measures (WP2 task, in cooperation with WP3)
- T5.4: Interoperability of e-Learning Applications with Digital Libraries (WP5 task, in cooperation with WP2 and WP3)
- T5.5: Ontology Driven Interoperability (WP5 task, in cooperation with WP2 and WP3)

Workpackage Joint Research Activities and Integration Activities in JPA3

JPA1 and JPA2: During the first 24 months (JPA1 and 2) the work in WP3 aimed to develop a common understanding and a foundation of the work that had to be done in DELOS II in terms of identification of partners' expertise, preparation of state of the art reports, installation of communication channels, establishment of complementary research directions (with respect to DELOS II target objectives), and gradual identification of research inter-dependencies and integration points among the various tasks. Important progress of each one of these tasks (see below) separately has been made. Significant outcome of the cluster is the elaboration of the basis of a consensual Semantic Multimedia Management Framework (SMMF) as a result of both the individual task progress and several intra and inter-cluster integration activities in the scope of the strategic objectives of WP3. The SMMF will evolve in JPA3 also accommodating further task results. Preliminary results can be seen in a Working Document: "Delos Integration Proposal on Semantic Multimedia Management", available at the following web site: http://astral.ced.tuc.gr/delos/cls_resource_description.jsp?id=10420.

The SSMF identifies the following functionalities that are important for any future Audio Visual Library infrastructure:

1. *Audiovisual Content Metadata Management:* Functionality referring to the capturing of metadata for audiovisual content (semantic, structural, media, etc.). The metadata will be produced both manually and automatically. This will be achieved through the utilization of constructs provided by well accepted standards (MPEG 7/21) as well as domain knowledge captured in (OWL/RDF) domain ontologies.
2. *Audiovisual Content Access and Personalization:* Functionality referring to advanced techniques for selection, delivery and personalization of the multimedia content, based on well-structured user preferences. In this context, transcoding and transmoding process must be taken into account to cope with user requirements and resource availability. Another important objective here is the provision of multimodal query interfaces, search engines for efficient content retrieval.
3. *Openness & interoperability support for domain-specific applications:* Specification and formalization of the proper interface that will allow exploitation of audiovisual libraries by domain specific applications (e.g. e-Learning, cultural heritage, etc.)

A common requirement for the accomplishment of all the above functionalities is the provision of semantic multimedia content descriptions based on standards (MPEG-7/21) and domain ontologies (OWL/RDF). Such functionality is already provided into a large extent by the GraphOnto toolkit developed in JPA2 (tasks 5.4, 3.6, 3.10 and 5.5). Thus, this component becomes a central point of integration in the envisioned Semantic Multimedia Management Framework since it provides various functionalities that support several other parts of the Framework.

Expected Results

We expect that during the JPA3 the finalization of the cluster tasks and their integration activities both at design and technical point of view will result in the development of the SMMF for Future Digital Libraries. The framework is considered to be an integral part of future DL infrastructures as these are described in the Delos II vision for Digital Libraries.

A number of prototypes (as described in the respective tasks) are being developed. For the realization of the SMMF intra and inter-cluster integration of these prototypes will be achieved based on service-oriented principles. That is, when needed, components will provide their functionality as a set of well defined services that can be exploited (through loosely integration) by other components of the SMMF

that need it. This way, beyond the manageable deployment of individual components, more complex functions (processes) can be dynamically configured and executed on the selected platform to serve specific user requirements. Moreover, smooth integration of the entire framework into the global Delos Demonstrator described into WP1 (task 1.8) can be easily achieved.

Description of Tasks

Task 3.6 - Video Annotation with Pictorially Enriched Ontologies. The ultimate goal of this task (VAPEON) is to automatically extract high-level knowledge from video data, enabling the automatic annotation of videos. The task aims at analyzing methodological aspects, mostly related to the definition of an ontology of linguistic terms augmented with “visual concepts”, resulting in a pictorially enriched ontology. Visual concepts will be defined by means of global distinguishing features, as well as meaningful spatial and temporal segments. Implementation aspects will also be addressed, covering the design of automatic annotation and summarization engines, integration into standards, and the development of a prototype system for sport digital libraries. The main results of the activities carried out during JPA2 are:

- PE Ontologies for soccer and F1 videos have been defined in order to extend with visual prototype the linguistic concepts describing soccer highlights. A methodology for the creation of the PE Ontology has been defined and tested. A methodology for using the PE Ontology to perform automatic annotation of selected clips has been developed.
- The extensions of the GraphOnto component so as to support graphically the linking of pictorial information to linguistic ontologies have been designed and developed. The creation and population of an MPEG-7 repository is undergoing. Access to the repository will be provided through a Java retrieval API which is also under development.
- An effective shot detection algorithm and a software architecture for parallel and distributed video feature extraction have been also developed.

In JPA3, VAPEON will continue, completing and extending the work carried out during JPA2. In particular, the work will be finalized to the development of an integrated prototype system combining together the work developed by partners. In addition, a main objective of VAPEON in JPA3 is to integrate the prototype that will be developed at the task level into the DELOS global prototype system foreseen in JPA3 by the new WP1 task 1.8 (DelosDLMS). In this way, the demonstrator will be made available to the future Digital Library Management Systems.

Task 3.7 - Multimedia Interfaces for Mobile Applications. This task (MIMA) will investigate several strictly interrelated sub-problems in the field of multimedia access for video presentation on mobile devices. In particular, the main subjects of investigation will be: (i) automatic video extraction of meaningful objects and events according to user interests; (ii) user profiling and design of flexible small screen device interfaces; (iii) performance measures and quantitative/qualitative indexes of user experience and satisfaction. Based on the obtained results, the task will develop a prototype system (for sports and news videos) composed of three subsystems: Off-line video annotation, On-demand video summarization, and User interface. The main results of the activities carried out during JPA2 are:

- Automatic video annotation systems and visual features extraction have been developed. The goal is to detect highlights and extract objects to perform both annotation and drive video adaptation based on the events and objects that may be of interest for the user. Development of systems that extract visual features from the compressed MPEG stream have been carried on, in order to reduce the computational requirements of the annotation systems.
- Two possible applications of televiewing from mobile devices, dealing with both stored and online data, have been analyzed. In particular, the access to a large Video Digital Library of Formula 1 races and the remote transmission of distributed video surveillance stations have been considered. For the second problem, the extraction and tracking of the moving objects in the scene has been studied, in order to distinguish between two different important areas in the video: the background scene and the interesting actors. To this purpose annotated information

- regarding every object have been provided, thus allowing specific interest level tied to the user/client capabilities, preferences and to the particular application.
- Design of an abstract (UML based) model able to characterize the user interaction. The model allows for capturing the information that is exchanged between the user and the system and is based on a set of atomic interaction units (AIUs).

In JPA3, MIMA will continue, complete and extend the work carried out during JPA2. In particular, the work will be finalized to the development of an integrated prototype system combining together the work developed by partners. In addition, a main objective of MIMA in JPA3 is to integrate the prototype that will be developed at the task level into the DELOS global prototype system foreseen in JPA3 by the new WP1 task 1.8 (DelosDLMS). In this way, the demonstrator will be made available to the future Digital Library Management Systems.

Task 3.8 - Description, Matching, and Retrieval By Content of 3D Objects. The goal of this task (RERE3D) is to develop a system to support structural as well as view-based retrieval by content of 3D objects. It aims at the investigation of models for extraction of view-based and structure-based descriptors of 3D objects, models for indexing and similarity matching of structural and view-based descriptors, models and metaphors for querying archives of 3D objects. A theoretical study of these models will be followed by the design and development of a prototype system supporting structural and view-based retrieval of 3D objects. The main issues to be addressed are: (i) Content description (both view-based and structure-based); (ii) Indexing and similarity matching (by means of distance metrics); (iii) Querying and presentation (based both on an object image and on object parts). The main results of the activities carried out during JPA2 are:

- Several 3D retrieval engines based on state of the art approaches have been developed and integrated within a Web accessible interface which has been designed and implemented.
- A new solution which relies on Spin Image signatures and clustering to achieve an effective, yet efficient representation of 3D object content has been developed. Performance of retrieval based on Spin Image signatures has been compared to performance of state of the art prototype retrieval engines. Comparison is still in progress.
- Development of a 3D descriptor based on 2D depth buffer images and integration into the IKONA interface. These descriptors have been evaluated on the Princeton Shape Benchmark Database.
- Development of a fully compliant MPEG-7 description of 3D models, by means of the Shape Spectrum. Comparison between the Shape Spectrum descriptors is obtained by using the Earth Movers Distance, since histograms with different number of bins need to be compared.
- An extensive literature study on 3D reconstruction methods has been performed which clearly identified the possibilities and shortcomings of the state-of-the-art. Software tools for each of the steps in 3D structure reconstruction namely feature detection, matching, and building point cloud models have been developed. Methods for rectification and texture mapping are currently being developed.

In JPA3, RERE3D will continue, complete and extend the work carried out during JPA2. In particular, the work will be finalized to the development of an integrated prototype system combining together the work developed by partners. In addition, a main objective of RERE3D in JPA3 is to integrate the prototype that will be developed at the task level into the DELOS global prototype system foreseen in JPA3 by the new WP1 task 1.8 (DelosDLMS). In this way, the demonstrator will be made available to the future Digital Library Management Systems.

Task 3.9 – Automatic, Context-of-capture-based Categorization, Structure Detection, and Segmentation of News Telecasts. In Task 3.9, the major research directions for JPA3 will be the following, based on a continuation of the main interests of the task:

- Integration of browsing and access capabilities in digital libraries of segmented and categorized news stories by modeling user profiles and contexts including mobility support.

- Detailed modeling of the structure and content of special parts of news telecasts referring to interviews, debates and other parts apart from reports, commercials and presentations that have been modeled in JPA2.
- Using alternative classification hierarchies to categorize news stories apart from the standard hierarchies in order to support interoperability between different classification schemes.

For the recognizer modules, we also identify the following possible major improvements:

- Automatic structure and temporal layout detection in news broadcasts:
- In JPA 2 we have shown that it is possible to automatically detect certain structuring elements, namely frequently occurring anchor shots, with very little a-priori knowledge. Other structuring elements, e.g., the main title, do not occur frequently enough in one video to be detected automatically in the same fashion. Nevertheless, they do occur very frequently if the whole series is taken into consideration. So, by examining similar sequences in different videos of one series, it should be possible to automatically detect these structuring elements.
- Enhanced speech recognition tool:
- In JPA3, the keyword spotting algorithm will be improved, exploring more general approaches to the task of extracting high-level semantic features from the spoken audio track of the analyzed media. The keyword spotter approach followed in JPA2 shall be expanded in JPA 3, such that the keyword list can be adapted on the fly. To cope with unknown words, even with a very large vocabulary, the use of sub-word units (syllables, morphemes) as the base units in the speech recognizers will be explored. Such systems have been shown in past work to offer high promise of yielding useful semantic features.
- The expected results for this task in JPA3 are the following:
- An integrated browsing and access model that takes into account current standardization efforts (e.g. NewsML, MPEG7 and MPEG21) to represent user interests for news content, context, device capabilities and news consumption preferences and can support ubiquitous scenarios.
- Integrated model for the structure and categorization of news telecast content including the additional categories of news content of interviews, debates in extensible manner that could accommodate new kinds of informational broadcast videos such as talk shows.
- An interoperability model to support alternative news classification hierarchies and the mappings between them exploiting semantic web technologies.
- A demonstrator that will implement the above models.
- Considerable reduction of manual efforts for the creation and maintenance of the news broadcasts grammar and models by using the inherent visual similarity between different videos of one news broadcast series.
- Valuable insights into the specifications with which speech recognizers must be designed in order to be usefully integrated into a larger media analysis framework,

Task 3.10 - Content and Context Aware Multimedia Content Retrieval, Delivery and Presentation. CoCoMA (Content and Context Aware Multimedia Content Retrieval, Delivery and Presentation) focuses on the study of an integrated multimedia content delivery environment which integrates semantic content and context-based multimedia retrieval from digital libraries with the personalized delivery and consumption of the retrieved multimedia data. CoCoMA focuses on the use of the dominant multimedia content description standards MPEG-7/21 and their extension with semantic description capabilities. CoCoMA will build on the existing in DELOS II sports testbeds.

During JPA2, CoCoMA focused on the specification of an architecture and the implementation of a demonstrator for multimedia content retrieval and distribution, which takes into account the content semantics. The demonstrator includes advanced components developed by the partners for selection, delivery and personalization of the multimedia content, and includes an innovative semantic multimedia content description model and a semantic multimedia user preference model. The task proceeds very well and is on schedule: an overall architecture has been defined, based on existing advanced components by the partners, which are being extended and integrated, and the integration effort is in place. Good collaboration of partners, several visits and exchanges.

JPA3 will exploit better the strengths of the architecture and it will focus on the evolution and integration of the semantic multimedia content description model and the semantic multimedia user preference model with a semantic context model and the exploitation of those models in all the aspects of the multimedia content retrieval, distribution and personalization in multimedia content networks.

The CoCoMA task is significant and addresses the vision of Delos II, as it focuses in an important area (content delivery networks) and the approach taken is knowledge based. CoCoMA has been designed and is being implemented so that (most of) the functionalities provided will be easily integrated in a service-oriented infrastructure. In addition, CoCoMA will contribute in the development of the SMMF:

- The GraphOnto component, which is a central point of integration in the WP3 Semantic Multimedia Management Framework, is already used in CoCoMA for the creation, storage, and management of the semantic metadata for audiovisual content. This allows semantic-based authoring, through the integration, in the UNIMI authoring tool (MPAT), of the ontology-based object annotation capabilities provided by GraphOnto.
- Advanced retrieval support will be provided through the integration of the (low level feature-based) Content-Based Retrieval (CBR) functionality provided by the VizIR framework with the Semantic-Based Retrieval (SBR) capabilities provided by the MPEG-7/21 Repository underlying the GraphOnto component. Cooperation with T3.6 (VAPEON) regarding this issue is foreseen.
- Support for semantic-based retrieval and filtering through the development of a connector that allows the MM4U framework to utilize the querying capabilities of the MPEG-7/21 Repository and the user preference descriptions (standard MPEG-7/21 or semantic-based) stored in the repository.
- Ontology-based semantic adaptation support, which is under development in CoCoMA and will be exploited in conjunction with Task 3.7.

Task 3.11 - Natural Language and Speech Interfaces to Knowledge Repositories (in cooperation with WP4). This task supports the objective of multimodal interaction with Audiovisual Libraries by focusing on principles, methodologies and software for the automatic construction of natural language and speech interfaces to knowledge repositories. These interfaces include capabilities for the declaration and manipulation of new knowledge, as well as querying, filtering and ontology driven interaction formulation. The entire functionality realizes the multimodal interaction objective of the Semantic Multimedia Management Framework developed in WP3.

During JPA2 this task developed an architectural framework for addressing uniformly a range of problems in sentence analysis each of which traditionally had required a separate computational mechanism. A prototype system that is using MPEG-7 semantics and ontologies to provide disambiguated natural language queries that refer to repositories with MPEG-7 information and which is based on the above framework is under way. Progress has been achieved in the speech recognition aspects and the cooperation of the speech recognizer with the natural language processing component.

In JPA3 formal methods for using natural language as knowledge representation will be investigated in order to bridge the gap between the two communities. Reasoning systems using statistical and symbolic reasoning, logical models, Bayesian networks, fuzzy logic etc. that benefit from properties of NL to reason with NL will be investigated, designed and implemented. Extended models for using user preference models for the syntactic and semantic disambiguation of user expressions will be developed. The generality and the performance of the approach will be investigated in different type of applications (knowledge domains). In the speech recognition component a more in depth investigation of algorithms and alternatives will be performed and in the HCI subtask a detailed evaluation of the new prototypes and application implementations will be performed.

The OntoNL as a part of the SMMF can interact with other components based on the service oriented paradigm. That is, not only OntoNL will exploit services provided by other components (MPEG Repository, GraphOnto ontology manipulation API, etc.), but also part of its own functionality (POS tagging, grammatical relations annotation, sentence analysis, domain disambiguation and query enhancement using personalized information (user profiles) and OWL ontologies, NLI generation for

question-answering systems, search engine functionalities using natural language or keywords) will be provided as a set of well defined services that can be exploited in building further application that need NL-based interaction with Digital Libraries.

WP4 - User Interfaces and Visualization

Strategic Objectives

The ultimate goal of the User-Interface and Visualization cluster is to develop methodologies, techniques and tools to establish a theoretically motivated and empirically supported frame of reference for designers and researchers in the field of user interfaces and visualization techniques for digital libraries, so to enable future DL designers and developers to meet not only the technological, but also the user-oriented requirements in a balanced way.

Specific objectives of the WP are:

- To elaborate a common understanding of the role and scope of user interface research in the digital library area.
- To develop a theoretical framework for digital library user interface design.
- To develop user-centered methodologies, techniques and tools to be exploited by DL designers and developers.

During JPA1 the research concentrated mainly on the understanding of the role and scope of DL user interface research and the investigation of the requirements for a DL interface design that provides support to the user throughout the entire DL lifecycle. Based on this foundation, in JPA2 the cluster started to address the development of the above-mentioned theoretical framework, methodologies and tools. In particular, Task 4.7 is producing a design framework and interaction model, based on the revised lifecycle model and equipped with concrete examples of suitable interaction paradigms, both visual and non-visual, e.g., natural-language based. Such interaction paradigms are outcomes of other tasks, namely 4.5 and 4.9. The design framework will also consider more advanced task-oriented interaction mechanisms and semantic views of the available information as produced in task 4.8. Actual implementation and evaluation with real users of specific tools representing instantiations of some of the advanced functionalities covered in the framework, such as collaboration support, automatic linking, recommendation, visual exploration and enrichment are provided by tasks 4.10 and 4.5. All JPA2 tasks are expected to continue in JPA3 along the just mentioned research lines. Moreover, they are expected to produce component prototypes to be integrated in the DELOS DLMS, so also offering an example of innovative DL interface developed on the basis of the Delos frame of reference.

Workpackage Activities and Integration

During JPA1 many collaborations among members of the Work Package have been initiated, including PhD students' exchanges, giving rise to a strong collaboration and integration among members both in the production of the deliverable and in the organization of the Delos workshop on Audio-Visual Content and Information Visualization in Digital Libraries (AVIVDiLib'05), held in Cortona on May 4-6, 2005, jointly with WP3. During JPA2 several workpackage meetings, as well as task meetings, were held and further PhD students' exchanges are presently going on. At least the same level of collaboration and exchanges is foreseen for JPA3. Moreover, the intercluster collaboration is quite strong, since all but one of the tasks are conducted in cooperation with other clusters, as specified in the following.

Task 4.5 – Visualization in DL systems – aims at providing tools and techniques exploiting visual clues in making sense out of information, focusing on two specific application scenarios: content-based multimedia retrieval and small screen interfaces. More specifically, the task is divided in two subtasks:
4.5a: Relevance feedback – this subtask investigates novel methods of combining relevance feedback mechanisms with advanced visualizations of query results in multimedia retrieval. Main research directions pursued in this task are mainly in proposing new interactive visualizations and exploring

mathematical foundations for advanced clustering algorithms and feature weighting. Other research directions are in studying how to apply personalization techniques and in extending the proposed approach to distributed environments. Some results have already been achieved in the various areas but they need to be further refined and integrated into an overall architecture, as it is planned for JPA3. The collaboration among partners has been established through both meetings and people exchange and it is planned to increment it during JPA3. This subtask is in cooperation with WP1 and WP2.

4.5b: Multimedia interfaces for mobile applications – this subtask deals with multimedia access for video presentation on mobile devices. In particular, it investigates automatic video extraction, segmentation, summarization and annotation taking into account user preferences and needs through an adaptive user interface. The cooperating partners have complementary competencies and expertise, ranging from video compression to usability, and some of them have already a tradition of working together. Then, a strong cooperation among partners exists and has been producing first joint results that will eventually converge in a fully integrated prototype during JPA3. In the meantime, interesting individual outcomes have been achieved by the various partners in their specific areas of interest. This subtask is in cooperation with WP3.

Task 4.7 - User Requirements-driven Support for a DL Design Framework - builds over the results of JPA1 in identifying functional and non-functional requirements of DLs and aims at extending such results towards the systematic investigation of non-conventional interaction paradigms, and the correlation of such paradigms with different usage phases of digital libraries. During the last months the research focused on the development of the DL usage lifecycle, on accessibility issues and on two specific interaction paradigms: scatter-browsing and catalogue browsing. Further activities planned for JPA3 aim at extending the previous studies seeing traditional DLs as components of larger software environments, so moving to DUKEs. The collaboration among partners has been established through several meetings.

Task 4.8 - Task-centered Information Management – aims at proposing the task-oriented modeling of user interaction in digital libraries, with the final goal of realizing a system able to learn the way in which users carry out typical tasks and then almost automatically reproduce them. The task concentrated on personal DLs, and already got the planned results on several topics, including modeling and management of personal ontologies, user profiling, first draft of the task language. The design of a first system prototype has been also carried on. Planned activities for JPA3 will develop further the ontology and personalization related aspects but mainly concentrate on the task modeling and intent inferencing. During JPA2 a strong collaboration among the partners was achieved, also resulting in a first joint paper. The overall goal of this task will be accomplished by also integrating contributions from WP2.

Task 4.9 - Natural Language and Speech Interfaces to Knowledge Repositories - this task is in strict cooperation with WP3. Its main objective is to provide principles, methodologies and software for the automation of the construction of natural language and speech interfaces to digital libraries. Results achieved during JPA2 concern the theoretical foundations of the approach, in particular the sentence disambiguation, and the overall system architecture. The usability evaluation of the interface is going to start soon. Further extensions coupling the theoretical approach with more heuristic techniques are planned for JPA3. Initial collaborations among partners have been established and are going to be incremented during the forthcoming period.

Task 4.10 - Design, Implementation and Evaluation of Multimedia Annotations for Users' Collaboration - this task continues the already very successful collaboration among the involved DELOS partners on a comprehensive annotation model and web annotation system which will serve as design guidelines for the specification and implementation of an annotation service compliant with contemporary interface standards. The overall goal will be accomplished by integrating also contributions from WP1 and WP7. During JPA2 the architecture of the system has been defined and some components already implemented and integrated with both Bricks and Daffodil. Further extension of the system, also in the direction of becoming a basic service that could be integrated in

any DL, are planned for JPA3. Finally, the high cooperation among partners already produced joint publications.

Expected Results

More specifically, the major expected results for JPA3 are:

- Release of a complete design framework based on the revised lifecycle model and equipped with concrete examples of suitable interaction paradigms, both visual and non-visual, e.g., natural-language based. Such a design framework will also consider more advanced task-oriented interaction mechanisms and semantic views of the available information.
- Actual implementation and evaluation with real users of specific tools representing valuable instantiations of some of the advanced functionalities covered in the theoretical framework such as collaboration support, automatic linking, recommendation, visual exploration and enrichment.
- Integration of the above tools in the Delos reference architecture.
- Making DL's accessible to disabled users. Current research on DL's has shown, so far, little concern for accessibility. The main goal is to build new design techniques for accessibility, going beyond simple technicalities (e.g. addressed by current standards), going beyond web interfaces and specifically tailored to DL's; to build a set of tools (plug-ins), that (in a manner analogous to what screen-readers do for the Web) can be combined with existing SW and interfaces in order to enhance accessibility of DL's.
- Guidelines definition for User Interface Logic implementation. By using the results expected from the evaluation activities we are performing in the next months, we plan to study specific guidelines and heuristics that could be defined for user interface implementation on generic DL's.

WP5 - Knowledge Extraction & Semantic Interoperability

Strategic Objectives

The cluster has continued to build on the earlier foundation work and during JPA2 has made significant progress in the two task areas of *Interoperability of eLearning applications with Digital Libraries* and *Ontology-driven Interoperability*. Work on the third task area (*Digital Repositories*) has been progressing more slowly and will be completed during the next six months.

Over the next 18 months we will pursue the following strategic goals:

- (i) To build links between the digital repository work and the digital library reference model developments.
- (ii) To work more closely with e-Learning and cultural heritage user communities in the design, testing, evaluation and enhancement of the prototype systems and to facilitate their integration with the Delos Digital Library.
- (iii) To extend the functionality of the *GraphOnto* tool and promote its application and use in other Delos research tasks as an exemplar of integration within the Network.
- (iv) To investigate and develop methods and a demonstrator for the integration of heterogeneous data types, models, upper level ontologies and domain specific KOS .

Framework of JPA3 Activities

Task 5.1 Information Repositories & Open Archives. In JPA2 the Task delivered 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.

In JPA3, this Task activity will produce a report which will build on a significant base of existing work on Digital Repositories. It will draw on the workshop outcomes, on the experience of the major Digital Repositories Programme launched by the JISC in the UK in 2005 and on other key initiatives including FEDORA, DSpace and ePrints.org, and the DARENET work. It will cover both technical infrastructure, scholarly communications, semantic descriptions and socio-economic aspects and will highlight areas for future European research and development.

Task 5.4 - Interoperability of eLearning Applications with 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 digital libraries (e.g. METS, Dublin Core, etc.), eLearning (e.g. SCORM) and audio-visual content description (e.g. MPEG7), produce mappings and following this analysis, develop an integration framework and a service-oriented 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.

During JPA2 the task has sought to answer the following questions: (a) What are the major architectural requirements and workflows for effectively supporting eLearning applications running over digital libraries? (b) What are the major interoperability requirements for DL and eLearning standards? (c) 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 focused 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. The architectural components of this task follow the service-oriented approach and can be easily integrated in the SMMF (Semantic Multimedia Management Framework) envisioned in WP3.

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.

In addition, 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 complementary to the work in the Task 3.8 “Description, matching and retrieval by content of 3D objects”, which does not consider semantic descriptions.

During JPA3, task 5.4 plans to work on three major objectives. The first objective is to implement further extensions to the GraphOnto tool, which are needed for metadata management on top of integrated repositories of various kinds (audiovisual, 3D graphics etc.). This objective directly contributes to the development of the SMMF envisioned in WP3 by providing an enhanced core component for the management of ontologies for audiovisual and non-traditional objects.

The second objective is the development of an Earth Sciences’ digital library according to the T5.4 eLearning environment for the provision of eLearning in Geography. During JPA3, an appropriate geography domain ontology will be created and populated. Moreover, a digital library of geography learning material, will be created based on the interoperability architecture developed in JPA2, by digitizing and segmenting selected material from the Greek National Educational Radio-Television archives (approved cooperation). 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. Finally, a 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 will be developed and deployed.

The third objective is the support of interoperability and semantics for 3D objects. This 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). 3D 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 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. Manipulation commands can also be personalized with the help of user profiles. User profiles combined with domain knowledge could automatically adapt the content of the scenes to designers' preferences. This objective is complementary to Delos Task 3.8 and is related with the envisioned SMMF of WP3 by studying the semantic aspects of 3D objects and their integration in multimedia digital libraries.

An initial approach for the areas addressed by objectives 2 and 3 has been done already within JPA2 and reports are in preparation. The initial investigations showed a very rich set of research issues related to the task.

Task 5.5 Ontology-driven Interoperability. Task 5.5 “ontology-driven interoperability” is continued from JPA2. The work 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 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.
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. Extend this to cross-domain retrieval.

The Task will continue the collaboration with CIDOC CRM-SIG and IFLA-FRBR Review Group on the Harmonization or merging of FRBR and CIDOC CRM. IFLA has published another document, which complements FRBR: FRAR, “Functional Requirements for Authority Records”. CIDOC CRM and FRBR/FRAR are the only formulations of high-level conceptualizations of the cultural heritage and libraries communities which are officially supported by their international top organizations (ICOM and IFLA). Since the proposed work is carried out directly with their representatives, the impact is by far higher than that of usual research work. The work will cover the following steps:

- Finalizing the complete draft definition of the FRBR-CRM model
- Review by the communities of the FRBR-CRM model and revision. In parallel trying to establish a working group for CIDOC CRM-TEI harmonization.
- Harmonization of FRAR with the FRBR-CRM model.

In addition, the growing core ontology will be employed in order to show:

- Systematic ways to produce from a core ontology the necessary diversity of ergonomically optimized data structures without losing the meaning with respect to a common ontology.
- Mapping of existing data structures, core ontologies, domain ontologies and real data encoded in these structures and indexing languages.
- Collaborative editing of mappings and implementation of data examples on a suitable knowledge base or database in order to demonstrate the effect on retrieval.
- Study version control for collaborative editing of ontologies and ontology mappings.

The mapping work will be extended to include new datasets and also literature. It is anticipated that this will include an FRBR related dataset to show CRM-FRBR ontology potential. Discussions with English Heritage (EH) have identified EH archaeological ‘grey literature’ reports and the Archaeobotanical Computer database (ABCD) as very promising. ABCD is linked to (JPA2’s) EAB via plant names and we propose to investigate potential of the core ontology to facilitate cross-domain linking and search from Environmental Archaeology to Economic Botany via biological taxonomies. This will also include linking from datasets to literature.

Expected results

The major expected results for JPA3 are:

- An evaluative study on the development and implementation of community repositories to support research and learning & teaching (month 30).
- Status report on progress with eLearning and DLs, and Ontology-driven Interoperability Tasks (month 36).
- Final status report and demonstrators for eLearning and DLs and Ontology-driven Interoperability Tasks (month 42).

Inter-Cluster and Intra-Cluster Integration Commentary

Work on Task 5.1 has proceeded jointly with the Preservation Cluster and this has been very productive. There are also links to service-oriented architectures and reference models work (WP1) when considering the role of repositories as content providers in service oriented architectures and frameworks. These links will be pursued during JPA3.

Both Tasks 5.4 and 5.5 are exploring ontologies and the ability to map between core and domain ontologies. There are direct connections with some of the work in WP3 through the common use of the *GraphOnto* tool. *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. In addition, the

semantic 3D information 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.

WP6 - Preservation

Strategic Objectives

During the JPA3 (the 18 months running from January 2006 to June 2007) the Preservation Cluster will pursue the following goals, which will contribute to the overall strategic objectives of the Preservation Cluster:

- To integrate the preservation concepts that were developed by the cluster in JPA1 and JPA2 with the digital library reference model that is being constructed as part of WP1. (This particularly the case of deliverables related to *JPA2WP6Task5*).
- To continue to make progress on the semi-automation of the processes of ingesting material into preservation environments so as to improve construction of digital libraries;
- Complete the development of tools to support the application of utility analysis to selection of preservation approaches;
- Examine processes of integrating the preservation tools brought together in JPA1 and JPA2 to support automatic re-appraisal of holdings;
- Promote the adoption of preservation technologies in digital library development designs;
- Raise the profile of digital preservation issues within the Digital Library Community; and,
- Increase our collaboration with other international researchers conducting research within the digital libraries and preservation communities.

Transition to JPA3

Progress in Preservation in digital libraries depends very heavily on modelling (*JPA2WP6Task4* and *JPA2WP6Task6*), testbeds (*JPA2WP6Task1*, *JPA2WP6Task2*, *JPA2WP6Task6*), automation of processes from metadata extraction to selection of preservation processes (*JPA2WP6Task6*, *JPA2WP6Task7*) and workflow. JPA3 will build on continuity of JPA2 and through the integration of the JPA1 and JPA2 results with the new work planned in JPA3 we will be able to bring preservation capabilities more centrally into the framework of digital library systems. It thus contributes to the DELOS grand vision of a generic digital library system architecture. Work that has been done so far needs to be integrated, the processes of re-assessing (appraising) digital materials already ingested into digital libraries needs to be studied if the value of DLs within in the archival community is to be secured, and we need to put in place a participatory document corpus that can be used by researchers to test tools for extracting metadata (e.g. technical or semantic metadata). WP6 will continue the research at Cluster level in four ways all of which will lead to an integration of current tasks and the next stages of development:

- (a) continue work on the current development and validation of the preservation approach selection tools (Task 6.6);
- (b) continue work in the area of the (semi-)automated extraction of metadata and the evaluation of the results for their capability to improve the preservation process (Task 6.7);
- (c) commence work to investigate the possibilities of automating re-appraisal, building on our experience with Tasks 6.6 and 6.7.
- (d) commence work to define a framework for a community managed and developed corpus to support preservation experimentation with a substantial and well-defined corpus of digital objects (e.g. documents, data sets).

WorkPackage Activities and Expected Results

In the following we describe the activities to be carried out by the WP in the next 18 months and the expected achievement.

Task 6.6 - Digital Preservation Testbed and Evaluation Framework. In JPA3 we will continue work begun under JPA2 on developing and evaluating our digital preservation testbed framework. It implements the testbed and evaluation metric design work performed in the past period, to allow automating major parts of the process of deciding about and evaluating preservation solutions is to further integrate, automate, and evaluate a framework for digital entity preservation by integrating and combining the testbed framework and the evaluation metrics developed during the first JPA of DELOS, with a specific focus on evaluation of the resulting framework in a set of real-world case studies at preservation institutions. Software tool support is being developed to assist in the processing of the Utility Analysis to inform the selection of preservation methodologies. This tool is currently being expanded with the additional steps stemming from the integration of the Dutch testbed procedures (DELOS Digital Preservation Testbed). Following the conceptual integration of the Dutch Testbed with the Austrian Utility Analysis Tool (a software tool to support the process of the Utility Analysis by automating sensitivity analysis) the cluster is conducting case studies performed in the areas of audio materials of the Austrian Phonogram Archive, video records of the Austrian Phonogram Archive, documentary records of the Dutch National Archive. Further case studies, specifically for database preservation (in cooperation with CNR), collections of thesis documents (in cooperation with the Austrian National Library) and a special collection with the State and University library of Göttingen are underway. Activities that are being continued in JPA3 include integration of additional steps resulting from the integrated DELOS digital preservation testbed and the conducting of case studies to validate the work (migration of a database, thesis publication, and special collections). This work combined with that on preservation modelling (JPA2WP6Task6.5) will be fed directly into new JPA3 work on appraisal modelling.

Task 6.7 - Digital Preservation Automated Ingest and Appraisal Metadata. JPA3 will see a continuation of this task, which has been examining the viability of automating preservation processes as a mechanism for improving these processes and demonstrating the viability of automation more generally in the area of digital preservation and curation. To aid the process of ingest, selection and appraisal, for the preservation of digital material, the goal of JPA2 task 6.7 is to look at ways of automating the semantic metadata extraction process and create a prototype tool, integrate this tool with other metadata extraction tools and ingest processes for automatic population of document repositories. We are using linguistic and layout analysis techniques to automate this process of metadata extraction. The research within this task can be divided into six domains: (a) selecting metadata to be extracted and that can be extracted, (b) integrating previous and current related research, (c) designing a prototype metadata extraction tool, (d) implementing a prototype metadata extraction tool, (e) establishing a well-designed corpus of documents to validate the effectiveness of the prototype, and (f) testing and refining the prototype. Progress has been made with (a) to (c) above and progress on (d) will be continued in JPA3. The creation of a framework for a corpus (e) will be investigated JPA3WP6Task9. The work of this new task will enable us to refine and test tools created under JPA2WP6Task7 in a robust way as access to a substantial corpus of documents that we can use to test this tool and to compare it to the results achieved by others working in this area is lacking. In the task we considered the issues surrounding the development of a substantial corpus and have concluded that we should begin preliminary work on this in JPA3 (see below, JPA3WP6Task6.9).

Task 6.8: Investigation of the Automation of re-Appraisal. The usefulness of digital libraries to the archival community will be influenced by our ability to automate the process of re-Appraisal of holdings within a digital library at different points during the lifecycle of a digital object. Building on work completed in JPA1 Task 6.4, JPA2 Tasks 6.5, 6.6, and 6.7 this new task will investigate how the process of automatic re-appraisal (resulting in either the disposal or retention) of digital holdings might be effectively handled in the context of digital libraries. Organisations repeatedly take decisions on what information (or digital objects) should be preserved and for how long; they retake these decisions repeatedly about the same materials as they come up for re-appraisal. Criteria governing what to keep and what to discard are usually based upon such factors as organisational needs/objectives, juridical requirements, and information value that are relevant to the business context of the organization (whether a library, a public sector institution or a commercial company). This is happening in government organizations, business companies and memory organisations including digital libraries. The main reason for this is that preserving too much digital material makes it difficult

to manage it cost effectively. In the paper world this appraisal and selection process is conducted manually and requires an enormous amount of effort. In a digital environment new approaches are possible, one of them could be automating parts of the actual appraisal and selection process. Appraisal points arise at the time of each preservation action, such as ingest, migration and probably access. Based on these stages it will be possible to identify where automation may best support the process. Our modelling work will build on the efforts to extract information about the digital objects (either about content or context or both as addressed in Tasks 6.6 and 6.7). The new core work will involve analysing and modelling appraisal criteria, rules and processes, and experimenting with their automation. Crucial in this process will be a better understanding of the role of preservation metadata in the appraisal, selection and disposal of digital objects. As has been noted by the PREMIS working group (2004) our understanding of the useful metadata for digital preservation is limited and no studies have effectively demonstrated its value.

Task 6.9 – Development of a Open Testbed Document Corpus. The quantities of digital materials to be ingested in digital libraries make it impossible for each of them to be described manually. While for many user communities commonly available search techniques are sufficient to mine identify documents for some knowledge about such characteristics of a document as context, subject, technical characteristics, and genre are critical to ensuring that we can compare different metadata extraction approaches. In work on Task 6.7 (begun in JPA2 and completed in JPA3) we recognised that to do any meaningful work we needed a substantial, consistently designed and documented corpus. We looked at how this might be generated and concluded that the design and implementation modelling of such a corpus needed serious attention. Such factors as scale, detail of labelling, distribution of file types and technical characteristics need to be analysed in more detail to ensure that the corpus has maximum advantage to the preservation research community for tool testing, tool training, and comparative evaluation and benchmarking of methods and tools. The test bed will provide us with mechanism to experiment with any outputs from Task6.8 that the project agree should be taken forward beyond JPA3.

Inter-Cluster and Intra-Cluster Integration

The Cluster is working with other Clusters and especially cluster WP1 to ensure that our work on preservation design methodologies will have on the design of the digital library reference model (Task 1.4.1). DLs need to be preservation aware and the work that has been completed by the cluster in this area is ready now for integration with the work going on in other clusters—indeed we are now ready to integrate in a way that our cluster was not at an earlier stage. The Cluster has developed links with key research teams working in similar areas who are not part of DELOS such as National Archives in the Netherlands, Digital Curation Centre in the UK, InterPARES, National Library of New Zealand and it will continue to develop new links during JPA3. Third, the Cluster has grown in size in JPA2 and has perfected its mechanism for collaborative working. In JPA3 the Cluster intends to bring together all the strands of its work.

The Cluster will run its Summer School (originally planned for Cortona) from 4-10 June 2005 on Digital Preservation in the Context of Digital Libraries in San Miniato (Italy). Last year it attracted an exceptional audience from the DL community and we anticipate the same this year.

Expected Results

More specifically, the major expected results for JPA3 are:

- Continue to exploit the modeling methods and techniques that we developed in JPA1 and JPA2 and put to work most effectively in Task 6.5.
- Integration of the results of the work of the Preservation Cluster into the DELOS Reference Architecture.
- A Report on the results of our investigations of the semantic metadata extraction in the context of digital preservation.
- A framework for a community-driven corpus of digital materials to be used in digital preservation research.
- A model of the processes of re-appraisal.

- Continued development of awareness of the processes.

WP7 - Evaluation

Strategic Objectives

During the next 18 months (January 2006 – June 2007), the Evaluation Workpackage will pursue the following strategic goals:

- community creation for DL evaluation,
- developing new evaluation models, methods and testbeds,
- creating new evaluation toolkits for implementing these methods,
- continuing existing evaluation initiatives, with a special focus on those aspects which are relevant to the digital library application area. .

To achieve these goals, the following tasks are planned:

- T7.3: INEX: An Evaluation Campaign for XML Information Access in Digital Libraries
- T7.4: CLEF: An Evaluation Framework for Multilingual Information Access in Digital Libraries
- T7.5: Digital Library Testbed Framework.

These three tasks are continuations from the first JPA. Task 7.1 “Evaluation Forum” from JPA1 and JPA2 will be integrated into task 7.5, in order to become more focused.

Workpackage Activities and Integration

The main integration activities of this WP will involve research and application communities external to DELOS. Both the CLEF and INEX evaluation activities are involving a large number of research groups not only from Europe but also from the America, Asia and Australia. In this way, they are contributing to spreading DELOS knowhow and best practice in experimental evaluation to the wider interested communities. Task 7.5 will primarily aim at integrating evaluation activities within DELOS, but also try to attract researchers outside DELOS by providing specific services. The three tasks are described in more detail in the following.

Task 7.3: INEX: An Evaluation Campaign for XML Information Access in Digital Libraries. The aim is to provide the infrastructure and a framework to investigate the performance of systems and services that aim at providing effective access to XML-based DLs. To achieve this objective, two sets of complementary activities are planned: INEX campaigns on one side and research into appropriate evaluation methodologies and test-suites for both current and untested aspects of XML information access on the other. In the 2006 campaign of INEX, the following tracks are planned:

1. Relevance feedback,
2. Natural language queries,
3. Heterogeneous Collections,
4. Interactive retrieval,
5. Multimedia retrieval,
6. Document mining,
7. User-case studies,
8. XML Entity Ranking.

The first six tracks were already part of INEX in 2005. Like last year, the document mining track will be carried out in cooperation with the PASCAL NoE. The user case studies track will carry out a rather general case study on user requirements for XML retrieval. The XML Entity Ranking Track will regard variations of question answering, in order to retrieve instances of specific semantic types, or identify relationships between sets of instances.

Task 7.4: CLEF: An Evaluation Framework for Multilingual Information Access in Digital Libraries. The provision of multilingual information access functionality is a key issue for Digital

Libraries (DLs). Two sets of complementary activities are planned: CLEF campaigns on one side, and research into new evaluation methodologies for untested aspects of multilingual information retrieval systems on the other hand. The 2006 campaign of CLEF will consist of the following tracks:

1. Mono-, Bi- and Multilingual Document Retrieval on News Collections (Ad-Hoc)
2. Mono- and Cross-Language Information Retrieval on Structured Scientific Data (Domain-Specific)
3. Interactive Cross-Language Information Retrieval (iCLEF)
4. Multiple Language Question Answering
5. Cross-Language Retrieval in Image Collections (ImageCLEF)
6. Cross-Language Speech Retrieval (CL-SR)
7. Multilingual Web Track (WebCLEF)
8. Cross-Language Geographical Retrieval (GeoCLEF)

All eight tracks are continuations from 2005; GeoCLEF was run as a pilot study in 2006, but will be a regular track in 2006. Furthermore, ImageCLEF was redesigned in order to consider aspects of various specific image retrieval tasks.

Task 7.5: A Digital Library Testbed Framework. This task aims at providing a standard testbed framework for comparative evaluation of digital library systems. Currently, this task is working on the definition of an integrated digital library evaluation model and the specification of a corresponding logging standard. Based on this theoretical framework, the task is developing a testbed framework for digital library evaluation, by extending the Daffodil system (Daffodil is a front-end system consisting of a rich collection of services for accessing federations of digital libraries) by appropriate services and a logging function following the proposed standard. This testbed framework will be used for DL evaluation in several ongoing DELOS activities:

- In T1.5/4.10 (Design, Implementation and Evaluation of Multimedia Annotations for User Collaboration), the annotation service will be integrated into Daffodil and then be evaluated.
- In T4.7 (User Requirements-driven Support for a DL Design Framework), several new information access functions will become part of Daffodil, followed by an evaluation of the extended system.
- As part of the cooperation between DELOS and TEL, a comparative evaluation between the TEL user interface and a Daffodil-based system will be performed.

In addition, T7.5 will aim at developing evaluation methods for two new areas:

- Collaboration: Based on a study on the needs for collaboration, the Daffodil system will be extended by appropriate services, and then an evaluation of collaboration in DLs will be carried out.
- Context-aware proactive and interactive help systems in DLs: In addition to the pro-active functions already available, the Daffodil system will be extended by multimedia tutorials and contextual help systems, and then these features will be evaluated.

Expected Results

For the three tasks of WP7, the following results are expected during JPA3:

- The INEX 2006 testbed, resulting from the INEX 2006 evaluation campaign. The testbed will contain the collections, topics and assessments from the different INEX tracks, along with the results of the
- The CLEF 2006 testbed, resulting from the CLEF 2006 evaluation campaign. The testbed will contain the collections, topics and assessments from the different CLEF tracks, along with the results of the participating groups.
- A digital library testbed framework, comprising a large variety of services. Evaluation results for various DELOS activities will be produced based on this framework. In

addition, a scientific primary data repository for DL interaction logs will be set up, and appropriate tools for statistical analysis will be provided.

WP8 - Dissemination and Technology Transfer

Strategic Objectives

Also in JPA3 the dissemination activities will fall into different categories, depending on the main objective of a specific activity, its contents and the expected recipients. It has to be noted the central role that the new web site will have in terms of visibility (use of the DELOS Digital Library), in terms of dissemination (increased use of RSS feeds) and in terms of focal point for cooperation and integration. In addition to the actions described below, three major topics will be addressed specifically in JPA3.

The first action is the definition and the actuation of a "Marketing Plan" with a set of actions that DELOS will perform to increase its visibility and its outreach in communities not necessarily directly involved in digital libraries. The steps started in JPA2 to reach out to the library communities will be continued, and extended to include other communities. A preliminary plan will be in place by September 2006.

The second action will be the consolidation of the DELOS community, mainly through the DELOS web site, by increasing the relevance of the DELOS Digital Library, by increasing the content and the utilization of the W3D data base (Who, What and Where in DELOS, i.e. a data base of DELOS skills), and finally by actively promoting increased cooperation and exchange of ideas through the use of state of the art tools supporting blogs, discussion lists, annotation, etc.

The third action will be an increase in Technology Transfer activities through close cooperation with the library and cultural communities. Given the renewed interest of the European Commission for the cultural sector, including the launching of a European Digital Library, it is essential for DELOS to participate in these efforts by providing not only scientific contributions, but providing also tools and prototypes to be used in real-world applications. DELOS has started a series of actions specifically intended to transfer to some user communities the results achieved and the prototypes developed by the DELOS members. The focus in JPA3 is on the library and cultural heritage communities, through cooperation with the TEL Office (The European Library), and with MICHAEL (representing the community of the Ministries of Cultural Assets). It is also planned to continue the participation in joint activities and events with ELAG (European Library Automation Group) and to start participation in LIBER (Ligue de Bibliothèques Européennes de Recherche) and LIDA (Libraries in the Digital Age) events. Other opportunities for cooperation with museums and archives will be sought and considered during JPA3. Presently, a number of activities aiming at transferring digital library functionality developed by DELOS members to the TEL system are in plan and will be carried out jointly with the TEL Technical Working Group.

The activities described below under the headings Scientific Dissemination, Education and Training, and Visibility are all part of Task 8.1. With respect to the cooperation between DELOS and TEL, four new Tasks have been defined; another Task might be added during JPA3 related to the cooperation with MICHAEL. These activities are described in more detail in the next section.

Scientific Dissemination

Thematic Workshops. This series of workshops will continue, providing the opportunity to the DELOS community researchers to present results of on-going research activities and to exchange opinions and experiences in an informal and friendly environment.

Brainstorming Workshops. This series of workshops, which brings together top researchers from all parts of the world, will focus on specific technology issues (e.g. search engines) to produce reports on "future research direction" in the field of Digital Libraries that can provide input for the definition of future research programmes both to the EC and to national research funding agencies.

European Conference on Digital Libraries (ECDL). This conference has become the major European event on the subject of Digital Libraries. DELOS is coordinating the conference Steering Committee and is committed to maintaining its high scientific level, through the involvement of the most active research teams in Europe and in the rest of the world.

Education and Training

Summer Schools. This series of one-week high-level courses will continue, providing education in the domain of Digital Libraries and its underlying technologies. The schools are addressed to members of the research community (in the wide sense): primarily graduate students, but also young researchers and professionals involved in R&D in Digital Library related areas.

Awareness Events. This series of events (courses, tutorials, panels, demonstrations, etc) is organized in collaboration with application and industrial communities. These events will present Digital Library technologies and successful case studies in a customized manner to user communities (Electronic publishing, Libraries, Archives, Museums, Broadcasting, etc.) and to interested potential industry partners.

National/Regional Technology Transfer Workshops. This series of events is organized by Network members in cooperation with national application communities and industry, addressing themes of mutual interest. Particular attention will be given to the involvement of SMEs in each country.

Research Exchange Program. The exchange of researchers between organizations working on joint projects is one of the most effective ways to achieve integration of the working teams and exchange of skills and results. For "exchange of researcher" is meant the visit of a researcher from one DELOS member to another, for a period ranging from a few weeks to a few months. In JPA3 DELOS has increased the funding available to cover the additional expenses involved in an exchange, and a substantial increase in the number of exchanges is expected.

Visibility

Web site. The DELOS web site has been completely redesigned, to bring at the forefront of the site the DELOS Digital Library and to introduce the notion of DELOS Community. The home page provides direct access to a search function, which can search simultaneously within the DELOS web site, the DELOS digital library, and the set of the web sites of the participating members of the DELOS community. In JPA3 additional resources will be dedicated to its further development and maintenance in order to: (i) continue to receive presentations, papers, reports and documents produced by the community members in order to be inserted in the digital library; (ii) promote the inclusion of active community members in the skill database, called W3D (Who, What and Where in DELOS), which contains research interests and availability to participate in DELOS activities; (iii) promote the automatic exchange of news and events through the use of RSS feeds, increasing the number of sites harvesting the DELOS site.

Press Releases. In JPA3 it is planned to increase the number of press releases, possibly jointly with TEL or MICHAEL, to capitalize on the interest and visibility presently given to subjects related to the building of a European Digital Library.

Poster sessions. In JPA3 it is planned to select a few relevant events where the DELOS activities and the joint activities with TEL and MICHAEL can be presented through poster shows.

Electronic Newsletters. A new shorter format of the newsletter has been defined, and it is planned to increase its frequency, providing timely information and links to DELOS events and to related activities of importance to the digital library community.

DELOS Awards. DELOS has established an award in the amount of 4 KEUR, named "DELOS Research Exchange Award", to be given each year to the best paper authored by a young researcher at the ECDL Conference. The award will recognize the achievement of the young researcher by offering the author a chance to spend a period of time (at least one month) at a DELOS European partner research organisation.

Technology transfer to the library and the cultural communities

During JPA2 DELOS has concentrated its outreach efforts in the library and cultural sectors, establishing good links with TEL (The European Library) and with ELAG (European Library Automation Group), and strengthening its links with MINERVA and also MICHAEL. During JPA3 it is expected that the cooperation with ELAG, given that ELAG is a “group of professionals”, will continue primarily through the organization of joint events and participation in common initiatives. In addition, DELOS will establish a closer cooperation with TEL and MICHAEL that will also include the integration of DELOS-provided software and prototypes in their systems.

In particular, the cooperation with TEL has been structured into four “technology transfer” Tasks, described below, and will focus primarily on the integration of DELOS-provided software and prototypes into the existing European Library system, which presently provides access to the collections of 15 National Libraries in Europe, with 8 more libraries expected to be accessible by the end of 2006. Detailed plans for the cooperation with MICHAEL will be defined in the second quarter of 2006.

Task 8.2 – Validation and refinement of the reference model through interaction with TEL. An important task in the process of defining the DELOS digital library reference model is its validation against the requirements of concrete digital libraries and existing digital library management systems. During two joint workshops the main goals of this task have been defined: (i) the expertise of the TEL participants (especially members of the TEL Technical Working Group TWG) will bring into the reference model activities valuable input for its refinement; (ii) similarly, the conceptual framework resulting from reference model will provide valuable input to TEL for the extension/refinement of the architecture of The European Library.

TEL has a business process-oriented view on requirements for digital libraries. The DELOS view is bottom-up, trying to identify basic digital library services. A major outcome of the validation of the reference model will be an attempt to map the business process-oriented view to the reference model to basic digital library services.

In the short term, the following steps are planned.

- DELOS provides to the European Library Office (ELO) a draft version of the reference model document in the March/April time frame.
- The ELO will circulate the draft within the TEL community and will collect feedback from different groups, e.g. business level, technical level, service-oriented level.
- The TEL-TWG will attempt to describe the existing TEL system in terms of the reference model (including the identification of gaps, differences etc.).
- The results of these activities will be discussed at the forthcoming meeting of the TEL-TWG, where the input of TEL to the reference model will be consolidated and integrated.
- Finally, appropriate TEL representatives will participate in the DELOS Digital Library Reference Model workshop, scheduled to be held in late May 2006, which will bring together experts in the digital library area, with the aim of producing a revised version of the reference model.

The mid-term cooperation will capitalize on the results of an on-going task of DELOS (in the Architecture cluster) aimed at integrating different digital library services (provided by the DELOS partners) into the existing base of the OSIRIS platform (the DELOS Digital Library Management System). The DELOS DLMS should provide a concrete example of an implementation of the reference architecture, and be a test-bed for digital library functionality. Currently, TEL runs a metadata repository over the different national libraries giving links to documents in the partner organizations. The services and the infrastructure provided by the DELOS DLMS might be used by TEL for implementing value-added services of The European Library.

In order to obtain results based on real-world data, DELOS will harvest the available TEL metadata and make use of them in the DELOS DLMS, possibly also adding links to the TEL metadata in the services implemented within the integrated prototype. On the basis of those results, TEL will be able to

evaluate the possibility of testing, integrating or adding to the TEL architecture some of the services developed within the DELOS DLMS.

Task 8.3 – Multi-Lingual Information Access in TEL. The ultimate goal of this task is to provide the capability to users of TEL to access and search the library in their own (or preferred) language, retrieve documents in other languages and have the results presented in an interpretable fashion (e.g. possibly with a summary of the contents in their chosen language). The problem is complex and many factors are involved. These include: the number of languages involved, the current heterogeneous setup of TEL, the lexical tools and resources needed.

- *Number of Languages.* The number of different languages represented in TEL constitutes a major hurdle for Multi-Lingual Information Access (MLIA), as ideally it should be possible to launch a query in any one of the national languages of the TEL collections and retrieve relevant material in any one of the collections. Possible approaches to the problem might be the use of multilingual ontologies, metadata and subject authority data, similar document search.
- *Heterogeneous set-up.* A major problem is represented by the heterogeneous set up of TEL, as it is not clear whether the existing infrastructure is able to accept a cross-language query result.
- *Resources Needed.* Any cross-language strategy implies the acquisition and development of appropriate lexical tools and linguistic resources such as stemmers, morphologies, bilingual dictionaries, etc.

The implementation of MLIA in TEL is an ambitious task that can be considered a medium/long-term goal, to be achieved through a series of intermediate steps. These should be defined and an implementation plan stipulated in the short-term. The first step is therefore the establishment of a joint DELOS/TEL working group to perform a feasibility study aimed at producing the following output:

- Guidelines as to how the TEL infrastructure should be adapted to be ready for the requirements of multilingual access and output, given the current state of play to be able to access the data at the national libraries. Recommendations as to how the libraries should be delivering their data may form part of these guidelines.
- Guidelines for the preparation of multilingual textual resources, which will be included into TEL (either digitized from existing text or borne digital)
- Definition of strategies that should be adopted by TEL with respect to enabling TEL users to search in their preferred language and retrieve documents in other languages
- Identification of the most promising possible implementation directions: linking metadata, similar document search, etc.

It is planned to hold a meeting at the TEL Offices with the TEL Technical Developers and Architects to gain a better understanding of the technology and underlying problems. Based on the results of this meeting, a feasibility study will be undertaken by DELOS and, a detailed plan for the medium/long term activities will be defined by the end of the third quarter of 2006.

Task 8.4 – Personalization capabilities in TEL. The final objective of this task is to produce guidelines and prototype software for new added-value services of interest to the final users, initially selecting those services that present a lower-risk of failure when personalized. A first short-term objective (spanning approximately one year) is addressing the development of personalization guidelines, and a second medium-term objective (spanning approximately two years) will address integration and testing of existing prototype software and development of possible improvements.

The initial personalization topics to explore are listed below. The exploration will start with a study of the existing access logs from TEL, in order to perform an analysis and categorization of context, to derive specifications for new types of logged data and suggestions on the design of innovative personalized services:

- query expansion (e.g. given a query related to “stars”, distinguish between a hobby astronomer versus a cosmology researcher)
- profile building
- notification about new material based on profiles

- recommendations based on profile similarity
- annotation sharing based on profiles
- provision of added-value links and/or service (e.g. OpenURL, etc.), based on preferences and rights of user or organization.

For the short-term objective the expected results are a preliminary study of TEL logs in the second quarter of 2006 (the study is in any case a valuable input to DELOS, since it is based on real-world data), and a more comprehensive report before the year's end, containing suggestions for added-value personalized services based on their potential and user-perceived relevance. The first analysis will produce some recommendations for the TEL Office to implement over what they should be logging to obtain the most useful information about site use and to inform the analysis for personalisation possibilities on the site. It is hoped to implement any recommendations in the July v1.3 release of The European Library. This will then give better log files for analysis purposes. For the medium-term objective it is expected to have by mid-2007 an initial prototype of a toolkit for log analysis and personalization services, in order to perform jointly an evaluation of the prototype and produce by November 2007 a refined toolkit and a final report containing agreed guidelines for personalized services.

Task 8.5 – User interface design for TEL, navigation and visualization services. This task addresses the overall design of the TEL user interface, as well as the exploration of additional services, especially for supporting query formulation, collection navigation and results visualization. The activities in the short-medium term will focus on four topics:

Evaluation. As part of an on-going task in DELOS (in the Evaluation cluster), we are preparing a comparative evaluation between the current TEL interface and an appropriate variant of Daffodil (www.daffodil.de). This evaluation will follow both an analytical and an empirical approach. The goal of the analytical evaluation is to assess the functional similarities and differences between the two systems. The goal of the empirical evaluation is to evaluate how well each tool supports the users needs.

The analytical evaluation will consider the usability, functions for search, browsing and result display, and the feedback/help functions of both. For this purpose, the methodology and the questionnaires developed by DELOS WP4 will be used. The empirical evaluation will be based on a user-centered and qualitative approach. Its focus is on the users experience with the tools, considering user characteristics, preferences and strategies, the types of activities/tasks users perform, and the environment in which the search tool is used. The evaluation must take into consideration the current necessary portal nature of the site.

Support for query formulation. The Daffodil interface already provides functions that help the user in formulating better queries. Most basic, a built-in spell checker will flag search terms not contained in the dictionary, and will propose correct variants. For advanced query formulations, a syntax checker will point out syntactically incorrect formulations. Finally, there is a 'related term' tool that proposes (statistically) similar terms for any of the query terms entered. The comparative evaluation will show to what extent these tools are useful for the TEL users, and then possible integration into the TEL system will be discussed.

Virtual collections and navigation. It is planned to provide and test an add-on service for building virtual digital collections starting from a set of real ones. For the definition of the virtual collections, the service should distinguish between expert users (like technicians and librarians) and end users. The service relies on automatic batch techniques of indexing, clustering, and classification of existing collections, allowing a visual navigation of their content, for an easier definition of the virtual collections. The benefits expected are that the user can be presented with cross collection views, can deal with smaller set of more relevant data and therefore queries can be processed in a faster way.

Presentation/visualization of results. The Daffodil system already provides functions for relevance ranking or quick filtering of results, as well as extracting attributes like author names or frequent terms from the result set. In addition, it is planned to provide and test an add-on service that allows end users to interact with the query result in a more effective way. Several techniques can be used in this context:

real time indexing, cluster-gather algorithms, smart use of relevance factor, information visualization techniques. One of the objectives of this activity is to understand which of those techniques are more useful in the TEL environment, in addition to motivate the end users to explore more large datasets and have “more fun” while exploring TEL. Also expert users like librarians could use this service to better visualize (and hence understand) the results produced by the navigation service described above, in order to define more easily customized views for end users.

WP9 - Assessment

This Workpackage has two specific objectives:

- to review and assess the overall activity of the Network and to provide recommendations
- to monitor and facilitate the integration process of the Network

The first objective will continue to be achieved by the Advisory Board composed of external experts, which is responsible for a periodic assessment of the activities of the Network, as described in detail in Section 8 of the Technical Annex.

The second objective will be achieved by guaranteeing the successful conduction of the third Joint Programme of Activities. As already said, the third JPA will focus on the integration of demonstrators, tools and prototypes into an integrated Digital Library Management System (DELOS DLMS) showing a large collection of digital library services and being the first example of the DELOS reference architecture. Therefore, the successful conduction of this ambitious goal will mean the satisfactory integration of the research activities of a number of members, belonging to different Tasks and Workpackages. It will be the responsibility of the Integration Task Force to oversee the work progress and to guarantee the successful implementation of the integration activities aiming at setting up a prototype of the “reference architecture”. In addition, the Integration Task Force will facilitate and monitor the exchange of information and experience that will happen through the “Integration Clusters”, and will report periodically to the Scientific Board on the status of the integration activities.

WP10 - Scientific Coordination

As in the second period, during the third period the scientific coordination workpackage will direct and supervise the scientific work of the Network. In particular, it will (i) organize the Network as a whole, (ii) supervise the scientific progress of the Network, (iii) ensure that all deliverables are available on time, (iv) create and maintain the conditions necessary for successful collaboration, (v) represent the Network in concertation with other scientific events.

The coordination will rely on the following management structures:

- Advisor Board (AB)
- Scientific Board (SB)
- Workpackage Steering Committees (WPSCs)
- Members General Assembly (MGA)
- Scientific Coordinator
- Administrative Coordinator
- Network Representative Towards the Outside World
- Integration Task Force

WP11 - Administrative and Financial Management

As in the second period, during the third period the administrative and financial management workpackage will ensure a strong and coherent administrative and financial management of the Network. The administrative and financial coordinator will also handle the reimbursement of the

expenses incurred by the (invited) participants to events and meetings organized by WP8, WP9 and WP10 (e.g. workshops, national events, Advisory Board and Scientific Board meetings, etc.). For this reason, a large part of the money budgeted for those activities has been allocated to WP11 rather than to the WPs responsible for the events. It has to be noted that in JPA3 also the budget allocated to the implementation of the DELOS DLMS and to the technology transfer tasks will be allocated to WP11, in order to be administered by ERCIM. That will allow a greater flexibility in the allocation of funds to specific tasks (i.e. partner organizations) depending on the actual effort needed for the integration of a particular component or the transfer of a particular tool, as it is difficult to make an a-priori evaluation of that effort.

This Workpackage will, in particular, (i) handle all the administrative tasks connected with the NoE's activities, (ii) handle all the financial tasks connected with the NoE's activities, (iii) provide a global Intellectual Property Rights (IPR) frame for the whole participants, (iv) ensure institutional exchanges with the EC representatives, and (vi) promote the gender equality within the Network.