DELOS
Network of Excellence on Digital Libraries

Newsletter

Issue 2 : October 2004



# Contents of Issue 2

[PDF version of this issue]

## Feature Article



---

## Cluster Reports

All issues of the Newsletter will carry a report from the clusters operating within the DELOS Network of Excellence. These reports will seek to keep you informed of the developments being made by groups within the cluster and keep you up to date with current interests and the direction in which research and implementation work is proceeding.

The clusters reporting are as follows:



---

## Dissemination Event



---

## The latest news from round DELOS

Each issue of the Newsletter will carry the most current news items from the DELOS website. The full listing will grow over time.



Return to top

Publication Date: October 2004
File last modified: Wednesday, 27-Oct-2004 18:47:12 CEST

The Delos Newsletter is published by the DELOS Network of Excellence
and is edited by Richard Waller of UKOLN, University of Bath, UK.

DELOS Network of Excellence on Digital Libraries

Newsletter
Issue 2 : October 2004



# The Work and Vision of Work Package 1: Digital Library Architecture

**Hans Schek** and **Can Türker** together with representatives of the groups introduce the work, research and vision of the constituent groups within the DELOS Work Package 1: Digital Library Architecture (ARCH).

## Introduction

The feature article for this issue of the DELOS Newsletter is devoted to the first DELOS work package (WP1). Our goal is to introduce briefly the groups that are active in this work package. Each group has arranged its information into the following areas:

- General information about the group and its overall research direction
- Brief sketch of the group's vision of a future digital library architecture
- Previous research that is closely related to WP1 objectives
- Principal planned contributions to WP1

To give an overview, the groups and their particular interests are as follows:

- ETH Zurich, Switzerland (WP1 Leader):
  Hyperdatabase Technology: The Basis of Future Digital Library Infrastructure
- CNR-ISTI Pisa, Italy:
  Meeting the Demands of Virtual Organizations
- FhG/IPSI Darmstadt, Germany:
  Grid-based Infrastructures the Basis of Future Virtual Digital Libraries
- MPII Saarbrücken, Germany:
  Achieving Service-quality Collaborative Searching in a Peer-to-Peer Environment
- Masarykova Universita v Brne, Czech Republic:
  Working Towards a Uniform Computational Framework
- OFFIS Oldenburg, Germany:
  Super Peer Networks: Providing Scalability and Autonomy
- Technical University of Crete, Greece:
  Focusing on Service-oriented Architectures
- UKOLN, University of Bath, UK:
  Open Standards Key to the Creation of Long-Term Viable Applications and Resources
- UMIT Innsbruck, Austria:
  Designing an Infrastructure for Highly Networked Information in a Pervasive Computing Environment
- University of Athens, Greece:
  Information Access Methods: Respecting the Autonomy of Differing Digital Libraries
- Università degli Studi di Milano, Italy:
  Providing Ubiquitous Access through Personalized and Content-aware Presentation
- Università degli Studi di Padua, Italy:
  Flexibility of Implementation to Reflect Differing Architectural Paradigms

However, before we move to the individual groups, let us very briefly recall the objectives of this work package: they are to evaluate both conceptually and experimentally the impact of recent computing technologies on digital library architectures. These new directions can be summarized as:

1. Web services and service-oriented architectures
2. Grid middleware
3. Peer-to-peer data management

A thorough evaluation of existing approaches will reveal the advantages and disadvantages of these strategies.

# ETH Zurich: Hyperdatabase Technology: The Basis of Future Digital Library Infrastructure

## Introduction to the Group

The ETH database research group is headed by Prof. Hans-Jörg Schek. The group currently consists of two senior researchers and six doctorate students. Its research activities aim to realize the vision of a hyperdatabase as the key infrastructure for developing and managing future information systems. At its interface, a hyperdatabase supports component and service definition, specification of transactional processes encompassing multiple service invocations, service publication and subscription. A hyperdatabase performs metadata management, component and service discovery and tracking, scheduling, routing, and optimization of service requests, monitoring, flexible failure treatment, availability and scalability. A hyperdatabase particularly provides an effective and efficient infrastructure to manage and retrieve documents over large multimedia repositories. Due to the high cost of capturing the content of multimedia documents, the infrastructure is able to make use of a large number of machines running various types of components to store and analyze documents, to extract features from them, cluster them, and to maintain indexes over those features. For more information about the group and its research, see http://www.dbs.ethz.ch.

## Vision of a Future DLA

In our vision, digital library users will be able to gain access to a myriad of forms of knowledge from anywhere and at any time and in an efficient and user-friendly fashion. To realize this vision, a highly scalable, customizable and adaptive infrastructure is needed. For future digital libraries we visualize an infrastructure that is based on hyperdatabase technology. Such an infrastructure combines techniques from peer-to-peer data management, grid computing middleware, and service-oriented architectures.

Peer-to-peer networks allow for loosely coupled integration of digital library services and the sharing of information such as recommendations and annotations. Grid computing middleware supports the dynamic allocation and deployment of complex and computationally intensive digital library services such as the extraction of features from multimedia documents to support content-based similarity search. A service-oriented architecture provides common mechanisms to describe the semantics and usage of digital library services. Furthermore, it supports mechanisms to combine services into workflow processes for sophisticated search and maintenance of dependencies. As depicted in Figure 1, the digital library architecture envisaged consists of a grid of peers which provide various kinds of digital library services such as storage, extraction or retrieval services. These digital library services can be combined with processes. High scalability is achieved by executing the processes in a completely distributed, peer-to-peer fashion. For that, metadata about processes, services, and load of the peers is distributed and replicated over the grid. This is performed by a small hyperdatabase (HDB) layer atop each peer. This layer also takes care of peer-to-peer navigation and execution of processes. Figure 1 depicts the execution of the process "Insert Image".

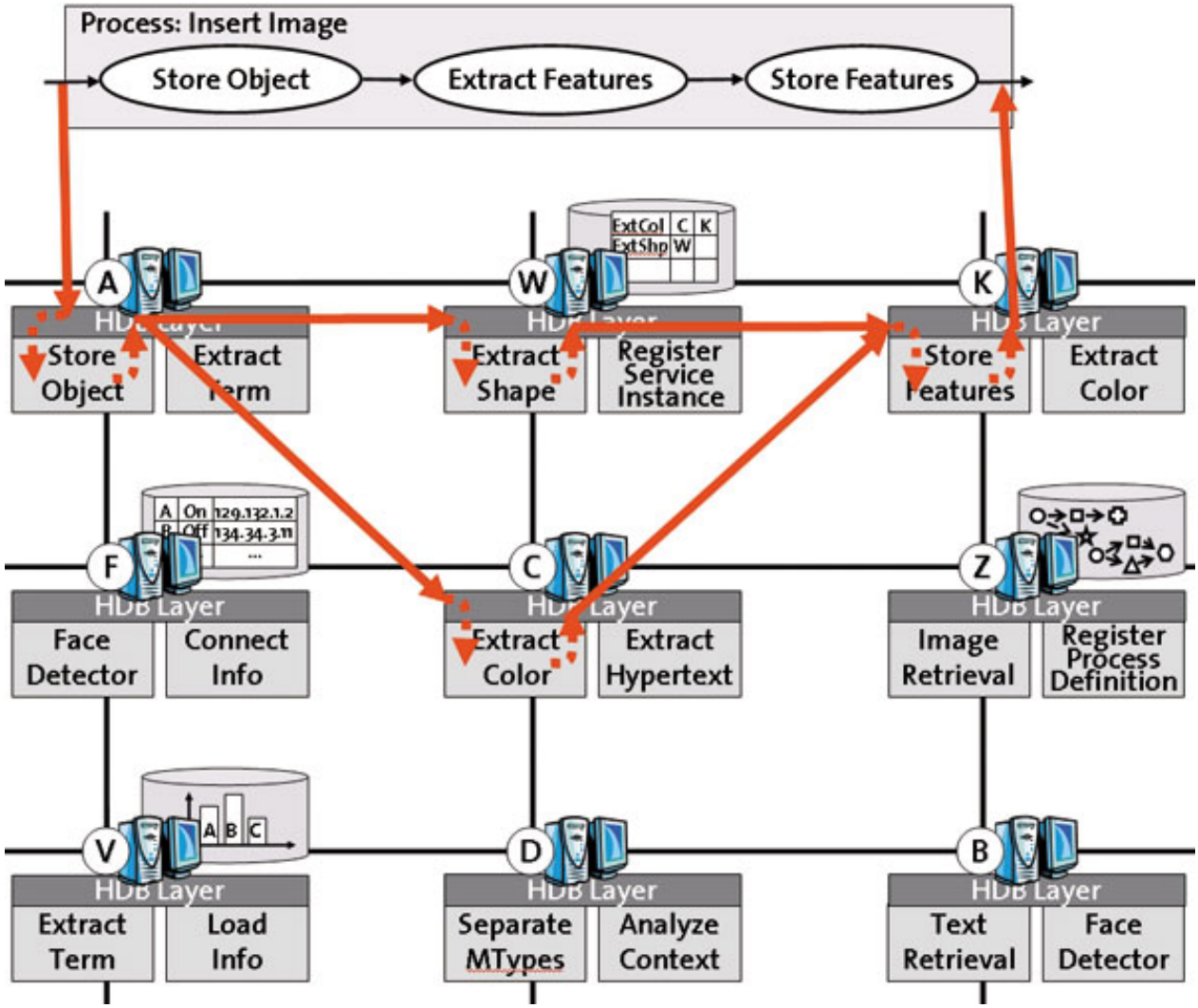


**Figure 1: Digital library architecture based on a hyperdatabase infrastructure**

### Research Relevant to WP1

Our research on hyperdatabases is strongly related to the objectives of WP1. With the implementation of our OSIRIS hyperdatabase prototype, we hope to show how a synthesis of concepts and techniques from database systems, process management systems, service-oriented architectures, grid and peer-to-peer computing can work together. The results obtained from many experiments with OSIRIS show the considerable scalability potential of hyperdatabase technology. Moreover, OSIRIS demonstrated its usefulness for advanced search and co-ordination of multimedia documents. While in the past we have developed sophisticated metadata replication and peer-to-peer process execution techniques in the hyperdatabase framework, current research focuses on concepts and mechanisms for the transactional execution of processes which did not involve any dedicated grid components. Such transactional executions will even be provided in the context of mobile grid peers and grid partitioning.

### Main Contributions to WP1

As project management lead for this work package, the ETH database research group co-ordinates the tasks in WP1. Furthermore, it contributes in particular to the digital library architecture which supports a peer-to-peer execution of composite services under transactional guarantees and in the presence of dynamically changing combinations of the grid peers. Using the experience gained from the implementation and investigation of the OSIRIS hyperdatabase prototype, the ETH database research group contributes to the WP1 survey on service-oriented architectures, peer-to-peer systems and grid infrastructures. This is done in close collaboration with the DELOS partner UMIT. Furthermore, the ETH database research group leads the survey on synchronization techniques in E-health applications. Currently, the following members of the ETH database group are involved in WP1 activities: Hans-Jörg Schek, Sören Balko, Michael Mlivoncic, Christoph Schuler, Hao Shao, and Can Türker.

## ISTI-CNR:

# Meeting the Demands of Virtual Organizations

## Introduction to the Group

The ISTI-CNR group belongs to the Multimedia Networked Information Access Laboratory, one of the fourteen research laboratories of the Istituto di Scienza e Tecnologia dell'Informazione, A. Faedo- CNR, located in Pisa, Italy. Currently, this laboratory, lead by Dr. Costantino Thanos, comprises thirty members including both permanent and temporary staff. This group has been conducting research on Digital Libraries (DLs) since 1996. This research has mainly focused on the design of generic digital library systems able to satisfy the needs of many different application frameworks. Particular attention has been dedicated to the introduction of innovative services on digital objects and to the study of flexible DL architectures capable of supporting different requirements in terms of handled content, functionality, policies, distribution, availability, etc.

## Vision of a Future DLA

Our research has always been driven by the aim of satisfying concrete user requirements. Recently, we have been facing radical new user expectations in respect of digital libraries. A large part of the demand for DLs comes from "virtual organizations", i.e., organizations composed by remotely distributed individuals, not usually highly skilled in computer science, who often work together for a limited period to achieve a common goal. These individuals need DLs to support temporary activities such as projects, exhibitions, courses, etc. These users demand easy, cheap and quick DL development models. In order to satisfy this demand we have started to explore the introduction of generic, customizable and highly dynamic software environments capable of providing, in addition to the functionality of a library, all the necessary management functions required to maintain the library, e.g. to preserve the content and services, and to guarantee the quality of the entire DL service, e.g. to support availability, performance and scalability. These environments must also support the sharing of resources in order to maximize re-use and to decrease the costs of creating a DL. Note that in the currently emerging context, the notion of sharing is not confined to "content", as has been the case until now. It also spans applications, software platforms, storage and computing elements. Sharing these resources must then, of necessity, be highly controlled. Most resource providers will only usually open access to their resources when the technology is sufficiently mature to guarantee that the resources shared are only used according to the policies established by their owners.

The creation of these new DL environments calls for appropriate architectural frameworks. We are currently working on the definition of an innovative distributed service-oriented architectural framework composed of three main elements:

1. A technical infrastructure, which provides all the necessary functionalities for supporting basic capabilities, like dynamic allocation and sharing of resources, transparent distribution, security, interoperability, quality of service, activation and de-activation of DLs, etc
2. A set of services that implement the typical digital library functionality
3. A number of application specific services, possibly supplied by third parties, which provide access to shared repositories of content and application specific tools by following the standard rules imposed by the technical infrastructure

In principle this architectural framework has the capacity to support multiple dynamically created customised virtual views of the underlying resources. A specific DL can thus be defined as one capable of providing such views and can therefore be created and destroyed dynamically on demand with limited effort in a short period of time.

## Research Relevant to WP1

Over recent years our research has mainly focused on the design of a distributed, dynamically configurable, service-oriented architecture for Digital Library Management Systems (DLMSs). As an outcome of this research we have designed and implemented one of these systems, OpenDLib (http://www.opendlib.com), which is now fully operational. OpenDLib has been used for building a number of DLs:

- eLibrary:http://elibrary.isti.cnr.it
- ARTE:http://odl-server1.isti.cnr.it/
- DELOS:http://delos-dl.isti.cnr.it:8007

The development of a system with this kind of architecture required greater implementation effort than that of a

centralized one since a number of services devoted to the co-ordination, management and optimal allocation of the different service instances also had to be provided. However, the experience acquired so far in building the different DLs has validated our choice as we have been able to satisfy quite easily a large number of application-specific requirements that could not otherwise have been met.

By exploiting the outcomes of this activity, we are now working on the definition of a more general architectural framework as described above. We are confident that the new architectural approaches that have emerged since we started the design of the OpenDLib system, e.g. Web services, P2P, Grids, now provide features that simplify the implementation of a distributed DLMS architecture and offer a number of new opportunities to implement novel user functionalities and enhance the quality of the overall system. In particular, we are now exploring, together with other DELOS partners, the use of the Web service paradigm on a Grid infrastructure as a basis for building a DL environment with the desired characteristics.

### Main Contributions to WP1

We contribute to the WP1 activities by bringing to this workpackage the results of our past experience on building and experimenting a DLMS with a distributed service-oriented architecture and by reporting our new research on Grid-enabled DL environments. In particular, we are leading the preparation of a survey on current aspects and tools of the Grid technology that potentially can be exploited for the construction of DL environments. We are also working on the identification of the gap between the functionality provided by some of the best known Grid middleware and the functionality required by a technical infrastructure capable of supporting the on-demand creation of transient DLs. The ISTI-CNR researchers participating in this specific work package activity are: Henri Avancini, Leonardo Candela, Donatella Castelli, Pasquale Pagano and Manuele Simi.

## FhG/IPSI: Grid-based Infrastructures the Basis of Future Virtual Digital Libraries

### Introduction to the Group

The Fraunhofer-Gesellschaft zur Förderung der angewandten Forschung e.V. (FhG), the leading organization of institutes of applied research and development in Europe, is a link between science and industry, i.e. between research and the application of its results. It was founded in Munich in 1949 as a registered non-profit association. The Fraunhofer-Gesellschaft is an autonomous body with a decentralized organizational structure, which currently maintains 58 research institutes and a patent office in locations throughout Germany. A staff of approximately 13000, the majority of whom are qualified scientists and engineers, works with an annual research budget of about one thousand million (1,000,000,000) Euros.

The Fraunhofer Institute IPSI (Integrated Publication and Information System Institute) focuses its research and development work on software applications for co-operative work, publication and information, innovation support, and lifelong learning in real and virtual environments. Our research areas comprise knowledge management and e-commerce, systems for individual or group learning, security in media and document management, digital libraries and e-Science, information systems, database-supported publication tools, distributed publication environments for the common maintenance of extensive data, and services for mobile communication. These activities also cover the fields of planning and installing modern working environments, i.e., building elements and furniture equipped with high-quality information technology.

### Vision of a Future DLA

The IPSI group sees the future role of a digital library and of digital library services as an important part of the information and knowledge environments which support E-Science and other innovative processes. According to IPSI's understanding, the digital library is currently undergoing a transition from a statically integrated system to a dynamic federation of services. This transition is inspired by new trends in technology which include developments in technologies like Web services and Grid infrastructures as well as by the success of new paradigms like Peer-to-Peer Networking and Service-oriented Architectures. The transition is driven by DL "market" needs. This includes a requirement for

- a better and adaptive tailoring of the content and service offer of a DL to the needs of the relevant

community as well as to the current service and content offer, and
- a more systematic exploitation of existing resources like information collections, metadata collections, services, and computational resources

Such new decentralized and service-oriented architectures for digital libraries make the library functionality available in a more cost-effective and tailored way and thus open up new application areas for digital libraries.

In essence, the creation of virtual digital libraries on the basis of Grid-based infrastructures, support for the integration of metadata, personalization services, semantic annotation and the on-demand availability of information collections and extraction services will make digital libraries more useful and attractive to a wider clientele.

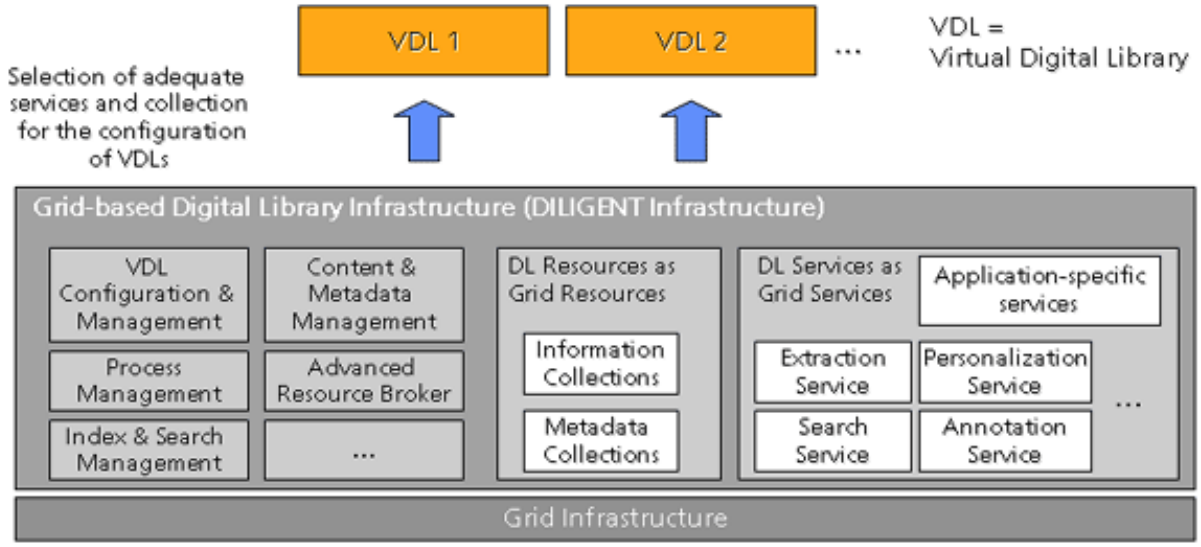


**Figure 2: The Infrastructure to Support Virtual Digital Libraries**

### Research Relevant to WP1

IPSI is part of two key European projects related to the WP1 objectives: Bricks and DILIGENT. Both projects are related to the next generation of digital library infrastructures. In these projects the IPSI group develops concepts and services for the architecture, metadata and content management, metadata integration, personalization in distributed architectures, manual and automatic annotation, distributed data management in the Grid and decentralized data management as well as content and services security.

### Main Contributions to WP1

In WP1 IPSI participated in the discussion on potential architectures for the future digital library with a special focus on XML-based decentralized metadata management, personalization approaches that also work in heterogeneous dynamic environments and the systematic support and exploitation of annotations in such environments. This work was summarized in a publication for the WP1 Workshop in June 2004.

## Max-Planck-Institut für Informatik Saarbrücken: Achieving Service-quality Collaborative Searching in a Peer-to-Peer Environment

### Introduction to the Group

The research group Databases and Information Systems is headed by Prof. Dr. Gerhard Weikum and located at the Max-Planck-Institut für Informatik in Saarbrücken, Germany. The group is being built up and currently consists of about 10 researchers. The overriding long-term goal is to develop rigorous service-quality guarantees for Internet-

based information systems, comprising provably correct behaviour, predictably acceptable response times, very high availability with failure masking, and satisfactory result quality for various kinds of information search.

The current focus is on the following topics:

- Automatic organization and effective and efficient search of semi-structured data in intranets, digital libraries, federations of deep-web portals and scientific data repositories
- Self-organization and self-optimization of large peer-to-peer systems, with applications in the areas of workflow management and search engines

### Vision of a Future DLA

We are addressing the problem of collaborative search across a large number of digital libraries and query routing strategies in a peer-to-peer (P2P) environment. Both digital libraries and users are equally regarded as peers and, thus, as part of the P2P network. Our system provides a versatile platform for a scalable search engine combining local index structures of autonomous peers with a global directory based on a distributed hash table (DHT) as an overlay network.

### Research Relevant to WP1

The peer-to-peer (P2P) approach, which has become popular in the context of file-sharing systems such as Gnutella or KaZaA, permits the handling of huge amounts of data in a distributed way. In such a system, all peers are equal and all of the functionality is shared among all peers so that there is no single point of failure and the load is balanced across a large number of peers. These characteristics offer potential benefits for building a powerful search engine in terms of scalability, resilience to failures, and high dynamics. In addition, a P2P search engine can potentially benefit from the intellectual input of a large user community; for example, prior usage statistics, personal bookmarks or implicit feedback derived from user logs and click streams.

Our framework combines closely studied search strategies with new aspects of P2P routing strategies. In our field of digital libraries, a peer can either be a library itself or a user that wants to benefit from the huge amount of data in the network. Each peer is a priori autonomous and has its own local search engine with a crawler and a corresponding local index. Peers share their local indexes (or specific fragments of local indexes) by posting the meta-information into the P2P network, thus effectively forming a large global, but completely decentralized directory. In our approach, this directory is maintained as a distributed hash table (DHT). A query posed by a user is first executed on the user's own peer, but can be forwarded to other peers for better result quality. Collaborative search strategies use the global directory to identify peers that are most likely to hold relevant results. The query is then forwarded to an appropriately selected subset of these peers, and the local results obtained from there are merged by the query initiator.

### Main Contributions to WP1

Our prototype system is described in:

- Bender M., Michel S., Zimmer Ch., Weikum G.: Leveraging the Power of Peer-to-Peer Systems for Collaborative Web Search. Digital Library Architectures. Preproceedings of the 6th Thematic Workshop of the EU Network of Excellence DELOS.
- Bender M., Michel S., Zimmer Ch., Weikum G.: Bookmark-driven Query Routing in Peer-to-Peer Web Search. Proceedings of the SIGIR Workshop on Peer-to-Peer Information Retrieval.

Our current and future work covers the auto-generation of web services (for deep-web sources), automatic query mapping and P2P exploiting "collaborative intellectual input" (bookmarks, ontologies, query logs). We are also working on personalized ontologies based on long-term relevance feedback. An architecture of a scalable, self-organizing DL federation with intelligent and efficient searching represents, we feel, a valuable contribution to the DELOS Network of Excellence.

## Masarykova universita v Brne: Working Towards a Uniform Computational Framework

### Introduction to the Group

The group consists of one professor, three associate professors, and five PhD students. The main research focus of the group is on advanced data processing methods. We are mainly interested in indexing techniques for the new types of data, for example multimedia data, as well as the new forms of data, such as the XML format.

### Vision of a Future DLA

In our vision, the future architecture of digital libraries should respond to the ever-growing need for the data processing software systems that abstract over programmable infrastructure, which while provided locally, is distributed on a world-wide scale. We could visualize this architecture as a global computer that consists of the networked integration of individual computing units with potentially different computing, storage and networking capabilities. Each unit exposes a common interface that is co-ordinated so as to provide a global paradigm. As such, the global computer provides an abstraction that co-ordinates a potentially wide range of units, with the advantage of providing a uniform computational framework for particular sets of digital library applications.

### Research Relevant to WP1

Recently, our effort has focused on developing a distributed storage structure for similarity searching in metric spaces that would scale up with constant or moderately increasing search times. In this respect, our proposal, called the Distributed Generalized Hyperplane Tree (GHT*), can be seen as a Scalable and Distributed Data Structure (SDDS) which uses the P2P paradigm for communication in a Grid-like computing infrastructure. We can achieve the desired effect for arbitrary metric data by linearly increasing the number of network nodes (whole computers), where each of them can act as a client and some of them can also operate as servers. A client inserts metric objects and issues queries, but there is not a specific (centralized) node to be accessed for all (insertion or search) operations. At the same time, insertion of an object, even the one causing a node split, does not require immediate update propagation to all network nodes. A certain degree of data replication can be tolerated. Each server provides some storage space for objects and also has a capacity to compute distances between pairs of objects. A server can send objects to other peer servers and can also allocate a new server.

The parallel search time for the similarity range and nearest neighbour queries in the GHT* becomes practically constant for arbitrary data volumes - the larger the dataset the greater the potential for inter-query parallelism. The GHT* has no hot spots - all clients and servers use as precise an addressing scheme as possible and they all incrementally learn from mis-addressing. Finally, updates are performed locally and a node splitting should never require the sending of multiple messages to many clients or servers.

### Main Contributions to WP1

A member of the group acted as a PC member of the Sixth International Workshop on Digital Library Architectures. We have co-operated closely with the IST-CNR Pisa. As a result of this co-operation, we have published the following papers:

- BATKO, M., GENNARO, C., ZEZULA, P.: A P2P System for Searching in Metric Spaces. In Proceedings of the Twelfth Italian Symposium on Advanced Database Systems. Cagliari : LITHOSgrafiche - Cagliari, 2004. pp. 410-417.
- BATKO, M., GENNARO, C., SAVINO, P. ZEZULA, P.: Scalable Similarity Search in Metric Spaces. In Pre-proceedings of the Sixth Thematic Workshop of the EU Network of Excellence DELOS. Cagliari : Edizioni Progetto Padova, 2004. pp. 213-224.
- BATKO, M., GENNARO, C., ZEZULA, P.: A Scalable Nearest Neighbor Search in P2P Systems. Accepted for the Second International Workshop On Databases, Information Systems and Peer-to-Peer Computing, August 2004, Toronto, Canada.

The following group members are involved in WP1 activities: Pavel Zezula, Michal Batko and Vlastislav Dohnal.

## OFFIS:
## Super Peer Networks: Providing Scalability and Autonomy

### Introduction to the Group

Two of the five research divisions of OFFIS are involved in research on digital libraries. These are the division on

Business Information and Knowledge Management and the division on Multimedia and Internet Information Services.

## Vision of a Future DLA

The described network in Figure 1 represents a first step in order to capitalize on the advantages of peer-to-peer technology for digital libraries. For personal or project reference libraries most of the upcoming traffic will remain within sub-areas of the network where co-workers co-operate intensely. For specialized collections which focus on special topics or special media type queries can be routed directly to selected collections or even library experts without flooding the entire network. Precision and query performance can hence be improved. Additionally, a self-organization of collections and libraries is possible. Scalability and administrative autonomy are also ensured.

## Research Relevant to WP1

Our research focuses on super peer networks. The figure depicts a hierarchical super peer network for digital libraries. Users are able to search for artefacts and offer artefacts independently. They are therefore supplied with person peers. On the next organizational level, the artefacts are grouped within collections managed by collection peers. Collection peers offer functionality relating to collection organization as for example the provision of a common classification scheme. A digital library can combine a number of different collections and is associated with a digital library peer. A digital library peer supports the integration of different collections, for example by offering merging services for different classification schemes. Furthermore, it manages access to the digital library artefacts, for example by ensuring a certain mode of payment. Person peers and collection peers can also exist independently of a superordinate peer and offer artefacts autonomously.

Peers are organized in disjoint clusters. Super peers route the messages along clusters to the destination cluster. Within the clusters the messages move through the hierarchical structure of the peers. The hierarchical peer structures in combination with their super peers form a hierarchical super peer network. The super peers hold a common metadata index of available artefacts which are distributed over the different organizational units or peer types, respectively. They are able to answer simple queries. Detailed queries additionally pass through the hierarchical structure of the peers. The exchange of the artefacts located takes place directly from peer to peer.

Super peer networks have some advantages over pure peer-to-peer networks. They combine the efficiency of the centralized client-server model with the autonomy, load balancing, and robustness of distributed search. They also take advantage of the heterogeneity of capabilities across peers.

The most important benefits of the approach discussed in this paper are scalability and administrative autonomy. A super peer can route messages independently within its cluster. Similarly, digital library and collection peers can route the messages to subordinated peers using their own strategy. Queries to selected organizational units do not flood the entire network but can be routed directly.

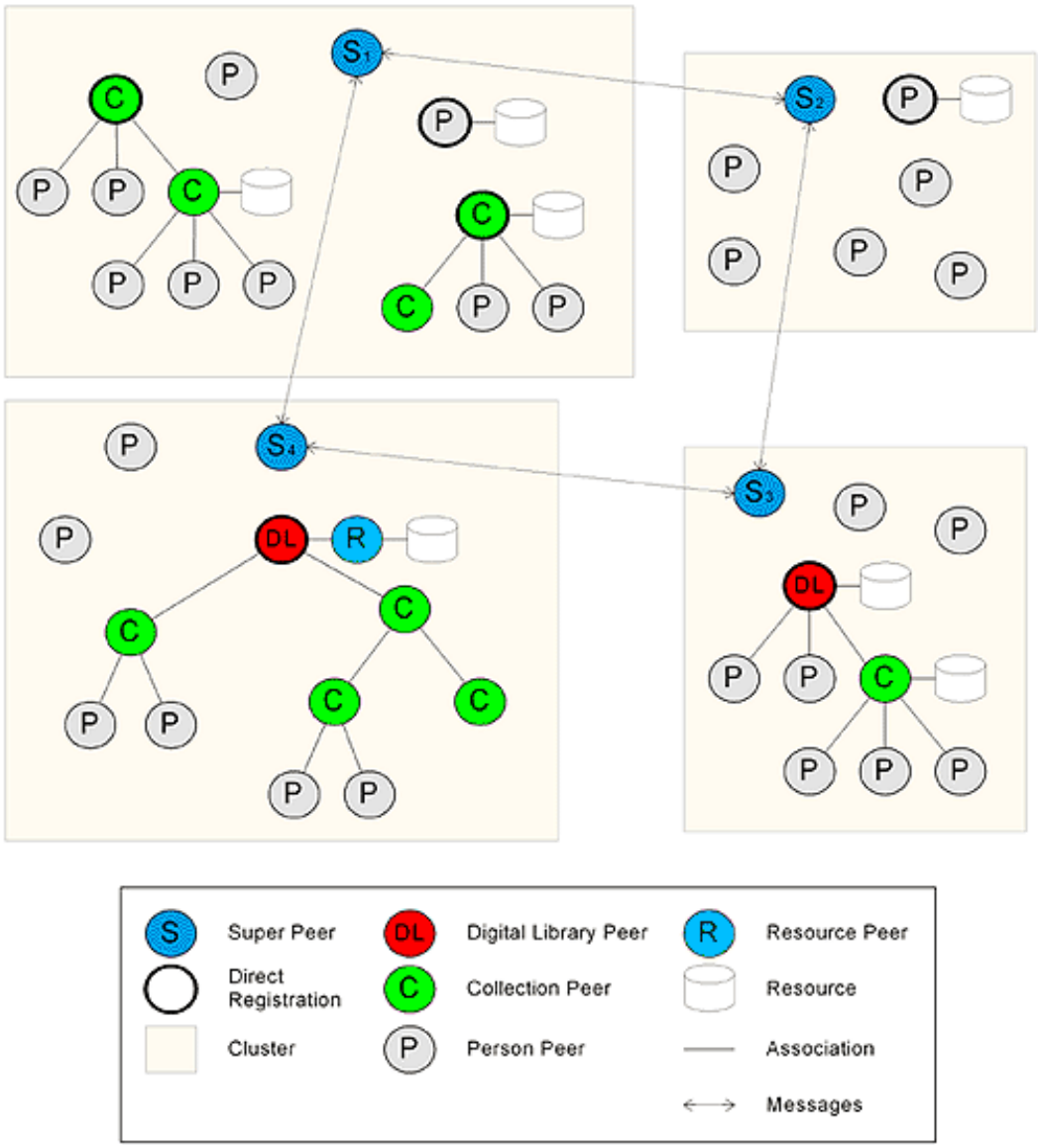


**Figure 3: A hierarchical super-peer network for distributed artefacts**

The hierarchical super peer network supports the flexibility and self-organization of widely distributed, loosely coupled and autonomous digital library systems. The architecture allows for searching over collections of arbitrary artefacts as for example traditional documents, on-line books, digital images, and videos, which is a basic service requirement for digital libraries. Beyond this, the network also enables library users to store, administer and classify their own artefacts. Therefore, it supports scenarios like the construction of personal or group reference libraries and collaborative authoring.

In the new Probado Project, tools for locating, storing, and releasing non-textual, multimedia documents automatically will be developed. Current digital libraries do not support such documents adequately because they usually assume documents to be purely textual in content.

## Main Contributions to WP1

A research visit of a research assistant from OFFIS at the UMIT Innsbruck took place over 20 September - 1 October 2004. The purpose of the visit was to exchange knowledge about peer-to-peer architectures and to investigate the use of structured/hierarchical super-peer networks [1] [3] within the medical sector in order to solve the availability problem for distributed patient records. A patient record can be regarded as a specific kind of digital library artefact. These artefacts are typically distributed over several institutions and yet may not be regarded as available anywhere and everywhere to the doctor who needs them. A joint paper [2] describes a first approach regarding the use of super-peer networks to solve the described problems. To gain a deeper insight, however, additional research is needed.

**Related Publications:**


1. Bischofs, Ludger; Hasselbring, Wilhelm; Schlegelmilch, Jürgen; Steffens, Ulrike: A Hierarchical Super Peer Network for Distributed artefacts. In: Pre-proceedings of the Sixth Thematic Workshop of the EU Network of Excellence DELOS. S. Margherita di Pula (Cagliari), Italy, 2004, pp. 105-114
2. Bischofs, Ludger; Hasselbring, Wilhelm; Niemann, Heiko; Schuldt, Heiko; Wurz, Manfred: Verteilte Architekturen zur intra- und inter-institutionellen Integration von Patientendaten. In: Tagungsband der 49. Jahrestagung der Deutschen Gesellschaft für Medizinische Informatik, Biometrie und Epidemiologie (GMDS 2004) (2004), September
3. Bischofs, Ludger; Hasselbring, Wilhelm: A Hierarchical Super Peer Network for Distributed Software Development. In: Proceedings of the Workshop on Cooperative Support for Distributed Software Engineering Processes (CSSE 2004). Linz, Austria, 2004, September


# Technical University of Crete: Focusing on Service-oriented Architectures

## Introduction to the Group

The Laboratory of Distributed Multimedia Information Systems and Applications of the Department of Electronic and Computer Engineering (ECE) of the Technical University of Crete (TUC/MUSIC) is a centre of research, development and education in the technological fields of Information Systems and their applications in the Information and Knowledge Society. In particular the Laboratory operates in the fields of Technologies, Architectures and Web Services Systems over the Internet, Databases and Knowledge Bases, Information Retrieval, Digital Libraries, Geographic Information Systems, Human Computer Interaction, Multimedia Management Systems, Digital TV Systems and Applications in e-Commerce, Tourism, Culture and e-Learning. TUC/MUSIC has been active since 1990 and it has participated in more than 40 European and national projects as well as in many Excellence Networks of the European Union.

The director of MUSIC/TUC is Prof. Stavros Christodoulakis. The research staff consists of six permanent staff members and about twenty postgraduate (Masters and PhD) students.

## Vision of a Future DLA

Architectures could be Peer-to-Peer or pure Grid (with centralized co-ordination). The emphasis in the Peer-to-Peer approach is more on service provision by independent organizations and the composition of the independent services offered by DLs within a given network. These DLs may involve other smaller DLs (by SMEs) or even personal DLs. The emphasis in grid architectures with centralized knowledge of resources (CPU power, disks, network, etc) is more on dynamic workload balancing for large DLs or DLs which manage very large multimedia objects requiring real-time interaction and synchronization.

## Research Relevant to WP1

Our current research focuses on Peer environments and service-oriented architectures. Service description according to MDA standards, service differentiation, service search and service synthesis are the current research topics being pursued.

In the future we are also planning to expand our research activities in the area of Grid architectures. Relevant research we have done in the past includes parallel streaming of audiovisual data from servers to support communities of users.

## Main Contributions to WP1

Preliminary work done in the architectures for peer-to-peer computing and service-oriented environments was published in the cluster workshop in Sardinia. The work continues.

# UKOLN: Open Standards Key to the Creation of Long-Term Viable Applications and Resources

## Introduction to the Group

UKOLN, based at the University of Bath, is a UK centre of expertise in digital information management, providing advice, support and services to the library, information, education and cultural heritage communities. UKOLN seeks to use its expertise to influence policy and inform best practice, to promote community-building and consensus-making by actively raising awareness, to advance knowledge through research and development, to build innovative systems and services based on Web technologies and to act as an agent for knowledge transfer.

The organisation is sub-divided into 4 main areas of activity, formally entitled Policy and Advice, Research and Development, Distributed Systems and Services and Resources and Administration. The four groups offer a complementary set of skills adding up to a wide-ranging coverage across all the related disciplines of Digital Library management.

UKOLN is currently active in, among others, the new Digital Curation Centre (DCC) where it is one of the four key partners in this important new initiative, maintenance of the Resource Discovery Network (RDN), the JISC Information Environment and in the provision of technical support and consultancy to large-scale nationwide projects. UKOLN holds JISC focus posts in Collection Description, QA (Quality Assurance), Web (including the UK HE seat on the W3 consortium) and Interoperability and is instrumental on many library and e-learning initiatives.

UKOLN is principally funded though the UK's JISC and MLA organisations. Further information is available on our website http://www.ukoln.ac.uk/.

## Vision of a Future DLA

UKOLN shares the vision of a digital library offering open access at any time, from any place, to all stored digital resources.

To achieve this, UKOLN is committed to the development and uptake of open standards and to oversee the application of these standards in new architectures and network middleware. UKOLN see the support and use of open standards as a key to creating long-term, viable applications and resources.

UKOLN understands that a future DLA must be built on functional, useful and agreed international metadata standards. UKOLN will be integral in the development and acceptance of well structured and globally accepted metadata definitions and schemas and will support the advancement and uptake of new and existing standards upon which the future DLA will be built.

UKOLN recognises that a future DLA will require new technologies to support distributed searching across multiple targets. UKOLN will continue its research and development work in this area through its existing commitments within the JISC Information Environment and will work with its partners to bring about the creation of exciting new developments in the application of cross-searching and data harvesting to the emerging peer-to-peer and GRID environments.

## Research Relevant to WP1

UKOLN has a long history of achievement in the field of metadata for description and interchange. Currently, UKOLN is significantly involved in the DCMI (Dublin Core Metadata Initiative) with Andy Powell sitting on the usage committee and Pete Johnston working in the collection description area. Peter Dowdell, Monica Duke and Greg Tourte are currently engaged in implementation work integrating schemas into existing and new applications such as the RDN (Resource Discovery Network), EnrichUK and the JISC IESR (Information Environment Schema Registry). UKOLN will continue to use its expertise to develop schemas further, to propose and direct new research into the area and to use its reach to publicise developments in the field.

It also continues with the development of distributed information systems incorporating XML-based record sharing and harvesting, news dissemination through RSS, distributed searching and web services.

UKOLN has an important role as a promoter in the development and maintenance of open technical standards and good practice. UKOLN has been the custodian of the NOF-digitise technical standards and guidelines and is a partner in the new Digital Curation Centre, contributing to the establishment of a suitable standards framework in

this important new nationwide initiative. Brian Kelly holds the JISC Web Focus post and is extensively involved in the promulgation and dissemination of open standards and good practice across the UK Higher and Further Education sectors.

UKOLN continues to maintain its extensive publication and event schedule as part of a key role in keeping its community informed and up to date with developments in the digital library sphere. UKOLN continues to publish the Web magazine 'Ariadne' and has recently successfully held ECDL 2004 at Bath, bringing together the principal players on the DL stage.

### Main Contributions to WP1

UKOLN will hold a workshop during 2005 to initiate a discussion into standards frameworks suitable for underpinning the architecture of a DLA. This workshop will provide an open forum that will lead to the identification of all protocols and standards that will be relevant to the Future DLA. Furthermore, we will encourage the introduction of new research and development in this area, especially in the rapidly-evolving fields of P2P and GRID where we expect DELOS as a whole will be instrumental in the advancement and codification of new technical standards to support these new technologies.

The outcomes from this workshop will be firstly a healthy discussion and opening-up of the issues that confront us in the establishment of a DLA. UKOLN will lead and marshal these discussions, and will subsequently publish a draft document that will be our first guiding reference to the agreed standards and protocols upon which the DLA will be built.

## UMIT: University for Health Sciences, Medical Informatics and Technology: Designing an Infrastructure for Highly Networked Information in a Pervasive Computing Environment

### Introduction to the Group

The University for Health Sciences, Informatics and Technology (UMIT) is located in Innsbruck, Tyrol, Austria and was founded in 2001. UMIT's role is to explore the potential and practical applications of information and communication technologies, to contribute to high-quality, efficient health care which satisfactorily serves both the individual and society, and to contribute to progress in medical and health sciences research. Two research groups at UMIT actively participate in DELOS: the Institute of Information Systems (IIS) http://ii.umit.at/, headed by Prof. H.-J. Schek, and the unit for Software and Information Engineering (ISE) http://ise.umit.at/, headed by Prof. H. Schuldt.

IIS activities address the infrastructure for highly networked information in a pervasive computing environment - the "infrastructure of the information space". The vision of the group is a new infrastructure called "hyperdatabase" which is particularly appropriate to e-Heath applications. In short, a hyperdatabase can be characterized as a synthesis of database technology, peer-to-peer computing, and a grid infrastructure. The activities of ISE support the vision of building reliable and dependable process-based systems, i.e., systems their users can count on. In most domains, and especially in digital libraries, specialized and well-engineered applications and databases are already in place and offer dedicated services providing access to information. In addition to process support and dependability, the infrastructure has to support automatic adaptation to changing environments. This requires the combination of aspects of workflow management, transactional process support, and grid infrastructures.

For more information about UMIT, see http://www.umit.at.

### Vision of a Future DLA

A Digital Library does not exclusively focus on the management of static data, information, and knowledge that can be accessed anywhere from any place, but has increasingly also to consider information that is dynamically modified and/or continuously generated. Examples of such continuously generated information are sensor data streams as they have to be processed and stored in an eHealth Digital Library for health monitoring applications, i.e., in applications where elderly people or patients with chronic ailments are equipped with wearable devices and

non-intrusive sensors to monitor and record their condition permanently. Therefore, the more general consideration of a Digital Library as Dynamic Ubiquitous Knowledge Environment (DUKE) must also have consequences for future Digital Library Architectures. Future Digital Libraries have to support a service-oriented architecture in order to make use of existing services and to combine services of the peers which are hosting services and information resources. Moreover, in order to achieve a high degree of scalability, (i.e., to scale with the number of resources, services, service providers and applications in the information space), aspects of grid computing are needed.

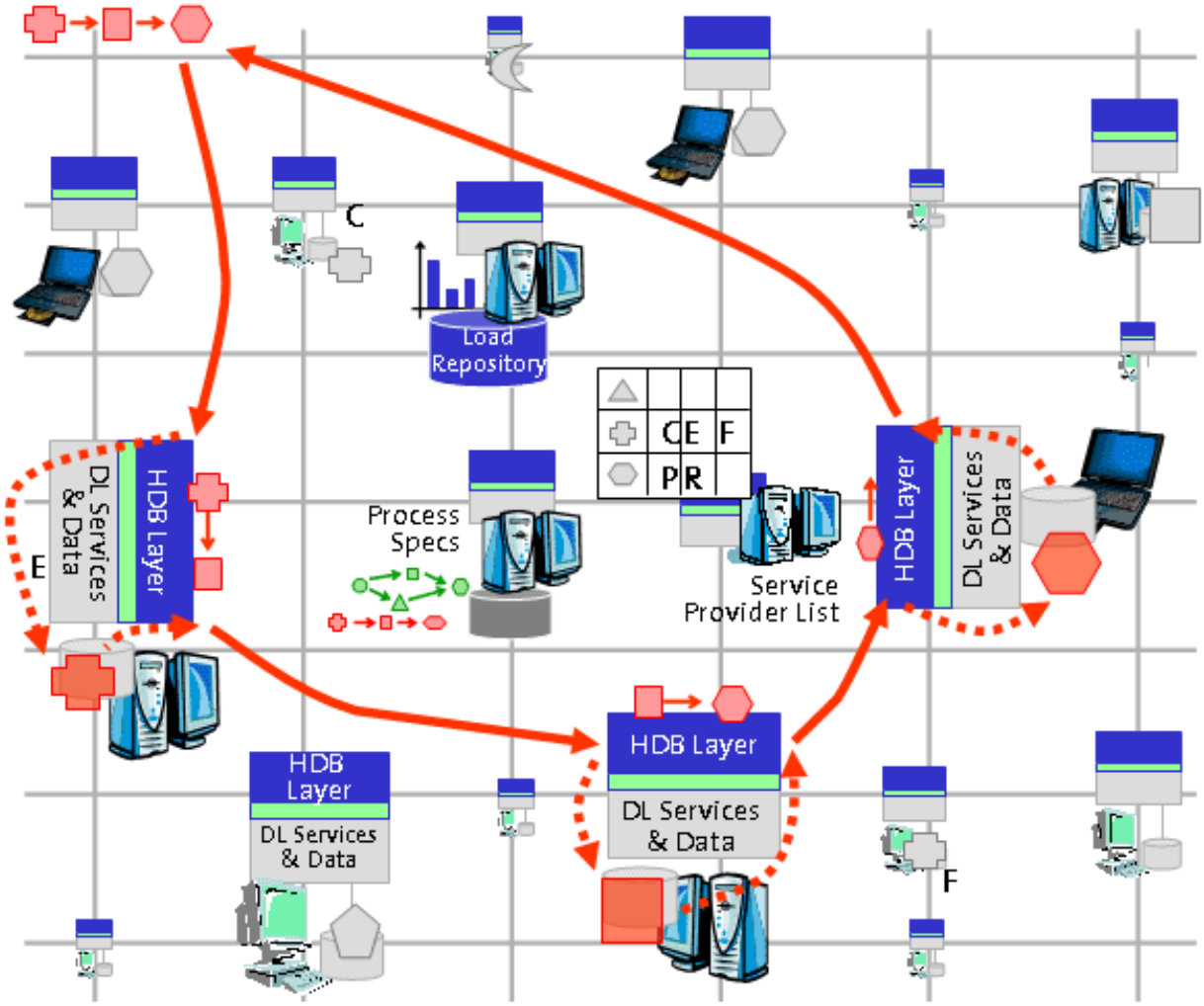


**Figure 4: Future DLA based on hyperdatabase technology for composite services**

In our vision, the future DLA follows a hyperdatabase architecture (see Figure 1) which supports composite services and processes. A small hyperdatabase layer is installed on each peer in the information space. By applying clever replication management, meta information on:

1. the applications (composite services and processes) to be executed,
2. the providers of other services, and
3. their load

is distributed among these hyperdatabase layers.

This allows for decentralized, peer-to-peer execution. Another important feature of future DLs is the potential to define sophisticated failure handling strategies within composite services and to be able to validate their correctness and derive quality characteristics during the design stage. This is particularly important for composite services and processes in eHealth DLs where correct functioning is critical.

## Research Relevant to WP1

Research by IIS and ISE at UMIT focuses on the above-mentioned vision to build a highly dependable, scalable, and adaptable infrastructure for process-based applications and its application in health care. In several research projects, particular aspects of this infrastructure are being considered. One project seeks to combine features of a grid infrastructure with hyperdatabase functionality [1]. While the hyperdatabase controls the execution of composite services and processes, the additional grid features make it possible to split up the invocation of a service dynamically into a set of calls that can be issued in parallel. By taking into account the providers and services that are currently available, this allows us to distribute complex service requests dynamically between several providers and to make efficient use of their resources. This is particularly important when different medical databases and repositories need to be queried with given time constraints. Another focus of current research is on the identification, definition, provision, and combination of building blocks that support efficient searching over multimedia patient records [2]. While patient information is not usually stored in a centralized system but rather remains under the control of the treating physician or hospital, the capacity to query these resources efficiently is essential in order to build a virtual electronic health record of a patient. In the area of health monitoring, we have extended a hyperdatabase prototype which supports the combination of stream operators and which controls the execution of stream processes, i.e., processes that are continuously fed with different sensor data streams. A major result is a highly reliable hyperdatabase infrastructure that combines stream operators and (web) services within the same application [3], [4].

### Main Contributions to WP1

A major contribution to WP1 is the architecture of an infrastructure for the execution of composite services in a peer-to-peer style following the notion of a Digital Library as Dynamic Ubiquitous Knowledge Environment (DUKE). This work is being addressed in close collaboration with another DELOS WP1 partner [5], the database research group of ETH Zürich. The experience gained with this infrastructure and its application in health care will also contribute to the WP1 survey on service-oriented architectures, peer-to-peer systems and grid infrastructures.

In a co-operative effort with OFFIS (Prof. W. Hasselbring's group), we are addressing the problem of availability in medical patient records by applying clever replication management over a peer-to-peer network. Preliminary results of this co-operation [6] will be reinforced by extending such collaboration, especially by means of exchanging PhD students. UMIT is co-ordinating a survey on service-oriented architectures and in particular on their potential for the construction of next-generation DL environments.

#### Related Publications:


1. M. Wurz, G. Brettlecker, H. Schuldt: Data Stream Management and Digital Library Processes on Top of a Hyperdatabase and Grid Infrastructure. In: Pre-Proceedings of the 6th Thematic Workshop of the EU Network of Excellence DELOS: Digital Library Architectures - Peer-to-Peer, Grid, and Service-Orientation (DLA 2004), pages 37-48, Cagliari, Italy, June 2004, Edizioni Progetto Padova.
2. M. Springmann, H-J. Schek, H. Schuldt: Kombination von Bausteinen zur ähnlichkeitsbasierten Suche in elektronischen Multimedia-Patientenakten. To appear in: Tagungsband der 49. Jahrestagung der Deutschen Gesellschaft für Medizinische Informatik, Biometrie und Epidemiologie (GMDS 2004), Innsbruck, Austria, September 2004. In German.
3. G. Brettlecker, H. Schuldt, R. Schatz: Hyperdatabases for Peer-to-Peer Data Stream Processing. In: Proceedings of the 2nd International Conference on Web Services (ICWS'2004), pages 358-366, San Diego, CA, USA, July 2004, IEEE Computer Society.
4. G. Brettlecker, H.-J. Schek, H. Schuldt: Information Management Infrastructure for Telemonitoring in Healthcare. To appear in: Tagungsband der 49. Jahrestagung der Deutschen Gesellschaft für Medizinische Informatik, Biometrie und Epidemiologie (GMDS 2004), Innsbruck, Austria, September 2004.
5. C. Schuler, R. Weber, H. Schuldt, H.-J. Schek: Scalable Peer-to-Peer Process Management - The OSIRIS Approach. In: Proceedings of the 2nd International Conference on Web Services (ICWS'2004), pages 26-34, San Diego, CA, USA, July 2004, IEEE Computer Society.
6. L. Bischofs, W. Hasselbring, H. Niemann, H. Schuldt, M. Wurz: Verteilte Architekturen zur intra- und inter-institutionellen Integration von Patientendaten. To appear in: Tagungsband der 49. Jahrestagung der Deutschen Gesellschaft für Medizinische Informatik, Biometrie und Epidemiologie (GMDS 2004), Innsbruck, Austria, September 2004. In German.


## University of Athens: Information Access Methods: Respecting the Autonomy of Differing Digital Libraries

### Introduction to the Group

The University of Athens (UoA) group participating in DELOS is led by Professor Yannis Ioannidis and has twelve members, all coming from the Department of Informatics and Telecommunications. The group's research focuses on several aspects of Digital Libraries, including Digital Library Architectures (DLAs). Specifically, it studies information access in distributed DL architectures, especially those based in the P2P or Grid paradigm, and in particular on query personalization and distributed information access. Advanced distributed information access methods are critical, as the forthcoming large-scale DL networks will have to provide prompt responses to end-user queries and traditional approaches do not scale up. Another focus of the group is related to query personalization, which enables DLs to provide end-users exactly the data they need by automatically tailoring the behaviour of the searching facilities to the users' preferences.

## Vision of a Future DLA

As the volume of available information increases, the size of future Digital Libraries is most likely to lead to the adoption of large-scale distributed architectures, such as those developed for federated databases or those based on the GRID or the P2P paradigm. Independent of any progress made in the hardware available, distributed architectures are the only solution to scalability problems; additionally, they provide single points of (homogeneous) access to information held by multiple institutions.

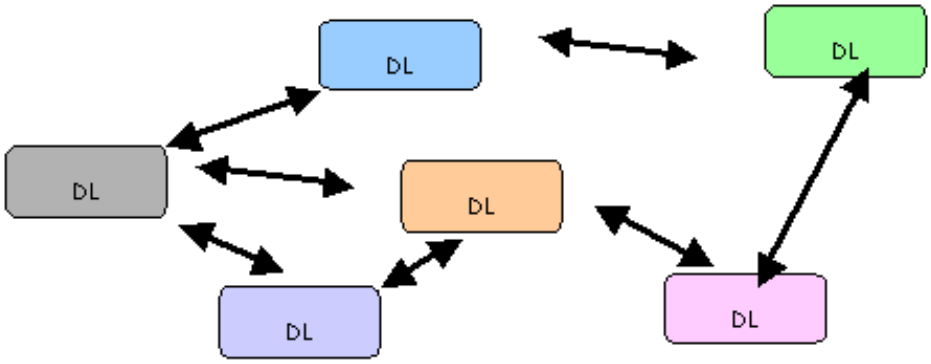


**Figure 5: Interacting Digital Libraries**

An important prerequisite to the success of distributed DLAs is respecting the autonomy of participating DL systems. Search and browse access methods should be developed for an environment where participating heterogeneous nodes or DLs designate the exact hardware resources and information content (data and metadata) that are available to inquiring nodes, while the latter make no a priori assumptions on the capabilities of the former. Each individual DL may be of a co-operative inclination to provide information for the benefit of all or may be in the competitive business of selling information. Either way, DLs should be allowed to interact freely with each other in unspecified patterns, each time making decisions based on the needs of the current query or processing request.

Incidentally, availability of large-scale distributed DL systems will increase the volume of information that can be browsed and searched. Taming the potential information overload will be helped considerably by the adoption of effective personalization techniques.

## Research Relevant to WP1

Within WP1, UoA has focused its research on developing information access methods suitable for distributed DLAs in the direction mentioned above, i.e., where DL autonomy is respected at all times. Such an environment poses significant challenges to query processing and optimization, as it results in lack of knowledge about any particular node with respect to the information it can produce and its characteristics, e.g., cost of production or quality of produced results. Potential inter-node competition also creates difficulties, as it results in potentially inconsistent behaviour of the nodes at different times. UoA envisages query processing and optimization of a form that resembles a commodity-trading negotiations framework, where in this case, information is the object being traded between independent DLs. In a recent paper (EDBT Conference, 2004), such a framework has been demonstrated, respecting node autonomy, being able to handle node heterogeneity, and offering effective query processing.

## Main Contributions to WP1

The UoA members involved in DELOS WP1 activities are primarily Professor Yannis Ioannidis and PhD Candidate Fragkiskos Pentaris. The team has participated in the 6th DELOS Thematic Workshop in Sardinia (June 2004), where a paper was presented on the overall framework for autonomous query optimization and execution in a distributed DL system. The paper recognized queries and query answers as commodities, modelled query optimization as a trading negotiation process, and outlined several aspects of this approach which required further investigation.

# Università degli Studi di Milano: Providing Ubiquitous Access through Personalized and Content-aware Presentation

## Introduction to the Group

The Università degli Studi di Milano (UNIMI) is one of the largest Italian Universities with 60,000 students enrolled, and several schools (Mathematical, Physical, and Natural Sciences, Humanities, Laws, Political Sciences, Medicine, Pharmacy, Agriculture). The UNIMI team involved in DELOS consists of people from the Database & Security Group (DB&Sec)and a research group from the Department of Computer Science and Communication (DICo).

The DB&Sec group consists of 8 members and 10 PhD students. Research is focused on the following areas:

- Security & Privacy
- Multimedia Systems
- Object-Oriented Systems,
- Temporal Systems
- GIS and Spatial Data
- Innovative IT Applications in the area of humanities

The DB&SEC group is affiliated with the Center for Education and Research in Information Assurance and Security (CERIAS) of Purdue University, Indiana, USA.

DICo has about 35 faculty members and supports 4 different degrees (Computer Science, Digital Communication, Informatics for Telecommunications, Science and Technology for Musical Communication), 2 master degrees, and a PhD programme. Its main research areas are concerned with:

- Databases and Security
- Computer Networks
- CSCW and Software Engineering
- Multimedia Systems

## Vision of a Future DLA

Future DLA should take advantage of current advances in system architectures and networks, resulting in architectural paradigms like grid-computing systems, wireless grid systems and P2P which provide unlimited computing and storage capabilities. Emphasis should be on providing ubiquitous access to DLA by, however, supporting a personalized and context-aware content presentation. Collaborative learning and discovery processes should also be supported. Finally, models and mechanisms for security, privacy and IPR should be part of any such solution.

## Research Relevant to WP1

Digital library-related activities at the UNIMI currently include the investigation of techniques supporting query formulation and data presentation from virtual reality (VR) environments. Thus, VR and database techniques are being integrated. Such research is carried out in the framework of the DHX Project (project IST-2001-33476). In addition, techniques and tools are being developed to generate multimedia presentations automatically, based on constraint languages. A third area of activity concerns security issues for database systems and advanced data management systems. Among the various research directions pursued, the most relevant are:

- the Author-X Project - investigating security, integrity, completeness, secure third-party publishing and

secure dissemination strategies for XML data
- the DLAM Project - investigating access control techniques for textual digital libraries
- the EUFORBIA Project (project IAP 26505) - investigating filtering techniques for Web documents and pages

### Main Contributions to WP1

We have started to develop a discretionary access control system to resources in a grid architecture. The system is based on the XACML (eXtensible Access Control Markup Language) standard and exploits the notion of virtual community to group together grid nodes adopting the same policies. We have developed scheduling algorithms which when given a request for accessing resources by a computation allow the system to determine the available resources by taking into account the access control policies. We are currently developing a preliminary prototype of our system in order to compare it with the performance of the Condor system. Finally, we are investigating an approach that would allow users, submitting computation on a grid, to specify security requirements. This approach is tailored to computations organized according to workflow systems.

#### Related publications:

E. Bertino, P. Mazzoleni, B. Crispo, S. Sivasubramanian, E. Ferrari, "Towards Supporting Fine-Grained Access Control for Grid Resources" to appear in Proceedings of IEEE 10th International Workshop on Trends in Distributed Computing Systems - FTDCS 2004, Suzhou, China, May 26-28, 2004.

E. Bertino, B. Crispo, P. Mazzoleni, "Support Multi-Dimensional Trustworthiness for Grid Workflow" submitted to the DELOS Workshop on "Digital Library Architectures: Peer-to-Peer, Grid, and Service-Orientation", 2004.

## Università degli Studi di Padova: Flexibility of Implementation to Reflect Differing Architectural Paradigms

### Introduction to the Group

The Information Management Systems (IMS) Research Group (http://www.dei.unipd.it/) is one of the research groups within the Department of Information Engineering of the University of Padua, Italy (http://www.unipd.it). The IMS group led by Professor Maristella Agosti has a strong programme of research, based on both theory and experiment, aimed at producing new multimedia information management and retrieval tools. The group, which began work more than fifteen years ago, currently numbers ten members all from the Department of Information Engineering but with teaching commitments in different faculties of the University.

The group has a good research history in the area of information retrieval, database management and digital libraries research. Present interests include aspects of:

- Design, modelling and implementation of advanced digital library systems that need to support novel information management and access services;
- Multimedia document and information management, and specifically Music Information Retrieval (MIR), which is an emerging research area that focuses on the content-based retrieval of musical documents against musical queries, where both documents and queries may be in acoustic or notated form;
- Web and document information retrieval, such as link analysis, automatic text categorization, and information retrieval evaluation.

The group plays a role in both the national and international research communities. In the past the group was active in the network of excellence IDOMENEUS, and the working group Mira, as well as projects such as JUKEBOX and EUROIEMASTER.

Members of the group have also been extensively involved in organising conferences, workshops and summer schools in the area of digital libraries and information retrieval. Some relevant examples are:

- The European Summer School in Information Retrieval (ESSIR) launched for the first time in 1990 by members of the IMS research group and which has now reached its fifth incarnation
- The First DELOS International Summer School on Digital Library Technologies (ISDL 2001)
- The European Conference on Digital Libraries of 2002 in Rome, Italy

Recently the IMS research group has co-operated in the organization of SPIRE 2004, the Eleventh Symposium on String Processing and Information Retrieval, which took place in the Department as part of Dialogues 2004, and two DELOS workshops, the Sixth Thematic Workshop of the EU Network of Excellence DELOS on Digital Library Architectures, in Sardinia in June 2004, and the Workshop on the Evaluation of Digital Libraries, again held in the Department in October 2004.

## Vision of a Future DLA

The architecture of a digital library allows and supports the functionalities that the digital library systems designed on it are going to make available to the different categories of users. This means that the architecture strongly influences the effective capabilities and functionalities of digital library systems that can be based on it. The architecture of future digital libraries needs to be designed and implemented in a way that easily supports the evolution of information access and management services that are offered to end-users. This means that an architecture of this type needs to cope with different user requirements and must allow the addition of new functionalities to a digital library based on it without the need to re-design the underlying architecture.

Another important aspect is the flexibility of such an architecture which must permit its implementation according to different architectural paradigms, such as Web Services (WS) or Peer-to-Peer (P2P).

## Research Relevant to WP1

In the past the IMS research group has participated in relevant European and national projects contributing to the design of advanced architectures for digital archives and libraries. Most relevant projects were:

- the EEC project JUKEBOX (LIB-JUKEBOX/4-1049) (1993);
- the IDOMENEUS IRIDES Project for the design and development of a prototype for the electronic publishing, storage, distribution and retrieval of a scientific journal over the Internet (1995);
- the feasibility study for the project of a digital library for the Venetian music of the eighteenth century to be developed by "Discoteca di Stato" (Rome), National Library "Marciana" (Venice), and University National Library of Turin (1997-1999).

The most relevant projects currently under development and to which the IMS research group is contributing are the ECD and the IPSA projects, briefly introduced below.

The group participates in Action 1 of the ECD (Enhanced Content Delivery) Project, a national research project supported by both the Italian National Research Council (CNR) and the Italian University (MIUR). The project is addressing the development of methodological tools and technologies for the delivery of enhanced contents to end-users. The main objective of the IMS group's research in this project is the design and development of a prototype for an Annotation Service (AS) for Digital Libraries. The service is going to deal with the different and relevant aspects of annotations, such as creation, management, access and retrieval of both manual and automatically created annotations.

The IPSA Project, launched at the University of Padua in 2002, is relevant to DELOS WP1 activities since it aims to design and construct a digital library of drawings and illustrations of historical documents, where the digital library in this instance is to serve researchers of the art and history of scientific illustration. Having carried out an analysis of user requirements and developed a methodology for accessing a digital herbal, the development of a prototype system, also named IPSA, is now under way. IPSA is a Web application that is based on a three-tier architecture. The system is based on software distributed under open source licence; the application has been developed on a Debian GNU/Linux platform.

### Related Publications (in chronological order):


- M. Agosti, F. Crestani, M. Melucci (1995). Automatic Construction of Hypermedia for Information Retrieval. ACM Multimedia Systems, (1995), 3, 15-24.
- M. Agosti, F. Crestani, M. Melucci (1998). Electronic Publishing, Storage, Dissemination and Retrieval of a Scientific Journal through the Web. Proc. of IEEE Forum on Research and Technology Advances in Digital Libraries (ADL'98), Los Alamitos, CA, USA, IEEE Computer Society, 1998, pp. 137-146.
- M. Agosti, F. Crivellari, M. Melucci (1998). Evaluation methods to improve information content extraction from the Web. In: F. Sadri (Ed), Proc. of First Int. Workshop on Web Information and Management (WIDM'98), Bethesda, Maryland, USA, 1998, pp.25-28.

- M. Agosti, F. Bombi, M. Melucci, A. Mian (1998). Towards a digital library for the Venetian music of the eighteenth century (Abstract). Proc. of Third International Conference on Digital Resources in the Humanities (DRH98), Glasgow, Scotland, Sept 1998, pp. 75-77.
- M. Agosti, F. Crivellari, M. Melucci (1999). The Effectiveness of Meta-data and other Content Descriptive Data in Web Information Retrieval. Proc. of Third IEEE Meta-Data Conference (META-DATA '99), Bethesda, Maryland, USA, 1999.
- M. Agosti, F. Bombi, M. Melucci, G.A. Mian (2000). Towards a digital library for the Venetian music of the eighteeenth century, in: M. Deegan, J. Anderson, H. Short (Eds). DRH98: Selected Papers from Digital Resources for the Humanities 1998.


DELOS Home | DELOS Newsletter Front Page | Delos Newsletter Contents

Publication Date: October 2004
Publication: Delos Newsletter Issue 2
Originating URL: /newsletter/issue2/feature1/intro.html
File last modified: Friday, 29-Oct-2004 14:54:06 CEST

The Delos Newsletter is published by the DELOS Network of Excellence and is edited by Richard Waller of UKOLN, University of Bath, UK.

Newsletter
Issue 2 : October 2004



# Reports from the DELOS Clusters

Each issue of the DELOS Newsletter carries a report from the clusters working within the DELOS Network of Excellence. In this issue clusters are reporting events and meetings which have taken place over the summer and autumn of 2004.



---

## Digital Library Architecture

In addition to introducing The Work and Vision of Work Package 1, Can Türker has provided us with information on the main event in WP1's summer calendar, the Sixth Thematic Workshop on Digital Library Architectures.

## Sixth Thematic Workshop on Digital Library Architectures

This summer period was dominated by the Sixth Thematic Workshop of the EU Network of Excellence DELOS on Digital Library Architectures, which was held at S. Margherita di Pula, Cagliari, Italy, over 24-25 June 2004. This first workshop in DELOS FP6 was co-organized by the WP1 members Maristella Agosti (UNIPD), Hans-Jörg Schek (ETHZ/UMIT), and Can Türker (ETHZ). The workshop was co-located with the 12th Italian Symposium on Advanced Database Systems (SEBD 2004) facilitating the exchange of research results and information at European and Italian level between the digital library and database communities. The local organization of the workshop was carried out jointly by the DELOS Network of Excellence and the Department of Information Engineering of the University of Padua, Italy.

The objective of the workshop was to bring together researchers interested in the architecture and related basic services which form the basis of building and operating DLs to identify ongoing research directions. In particular, the impact of newer computing paradigms, such as peer-to-peer, grid, and service-oriented computing, on future DL architectures was the focus of the workshop. Many lively discussions arose around the opportunities and challenges of applying these technologies for digital library architectures.
We had more than forty participants, twenty regular talks and three invited speakers. Nearly all WP1 cluster members were represented at the workshop by one or more attendees.

The papers accepted for this workshop were published in the following pre-proceedings:

Maristella Agosti, Hans-Jörg Schek, Can Türker (eds): Digital Library Architectures: Peer-to-Peer, Grid, and Service-Orientation, Pre-Proceedings of the Sixth Thematic Workshop of the EU Network of Excellence DELOS, S. Margherita di Pula, Cagliari, Italy, 24-25 June 2004. Edizioni Progetta Padua, 2004. ISBN 88-87331-60-X.

Springer has recently agreed to publish post-proceedings of this workshop in the LNCS series. The papers of the

pre-proceedings will undergo a second reviewing phase. Selected papers will then move into these post-proceedings in an extended, revised form. The post-proceedings will also include papers from the workshop's invited speakers.

For further details about this workshop, please refer to http://www.dbs.ethz.ch/delos/6thworkshop.html. There you will also find the pre-proceedings mentioned above ready for download in PDF format.

## Author Details

**Can Türker**
Member
ETH Zurich,
Institute of Information Systems,
CH-8092 Zurich

url: http://www.dbs.ethz.ch
E-mail: tuerker@inf.ethz.ch
Telephone: +41 1 6327248
Fax: +41 1 6321172

| | | |
|---|---|---|
| DELOS | Network of Excellence | email: schek@inf.ethz.ch |
| Name of Cluster: | Digital Library Architecture | phone: +30 210 727-5224 |
| Name of Leader: | Hans-Jörg Schek | fax: +30 210 727-5214<br>url: http://www.dbs.ethz.ch/ |

Return to top

## Information Access and Personalization

**Harald Krottmaier**, currently participating in the DELOS researcher exchange programme at ETHZ, provides an overview of current activity in a subtask of the Information Access and Personalization workpackage concerning metadata.

## Introduction

To reach the objectives of the DELOS WP2 cluster, several tasks and subtasks have already been defined. We at the IICM (Institute for Information Systems and Computer Media) are working on one of these tasks. This short report will outline the study on metadata models currently being undertaken. The study will promote the knowledge of available technologies in the context of metadata and digital libraries.

## Metadata is...

Within our community the short-hand expression is generally well known: that metadata is data about data. You may find this statement several thousand times on the Web as a definition of metadata. There are however many different models available, and depending on the application, slightly different implementations exist.

MARC, Dublin Core, MPEG-7, RDF etc. are just some of the currently available models used in the field of digital libraries. One goal of the study is to list the standards and link to the responsible institutions. It should also give a comprehensive overview of these models and point to existing applications using the models. Since this is 'old hat' in the field of digital libraries, we do not intend to replicate existing material, but instead point to definitions available, for example, at UKOLN (e.g. Metadata, BIBLINK), W3C and other research bodies.

Metadata in the field of services offered by digital libraries is, on the other hand, relatively new. It is necessary to describe services and put these descriptions into repositories. A short exploration of currently available services will be provided as an impetus to increase interest in this subject. To furnish one instance: we are going to examine in detail OSIRIS (http://www.osiris.ethz.ch/), an Open Services Infrastructure for Reliable and Integrated process Support.

The technologies used in OSIRIS and other systems will be explored and described. WSDL (Web Service Description Language), UDDI (Universal Description, Discovery, and Integration), and DAML (DARPA Agent Markup Language), are just a few of the items to be covered in the study. Historically important developments, such as RPC (Remote Procedure Calls) will also be listed. Additionally, initiatives have been established to promote sharing and reuse of metadata. A short overview in this respect those will also be provided.

Questions and current answers regarding service registries will complete the study.

## Implementation

To make it possible for all members of DELOS to participate in the creation of the study and to make it possible to reuse this content in lectures and other dissemination events, we are going to use a WIKI-system.

Currently we are exploring several implementations of WIKI-systems (JSPWiki, MediaWiki to mention just two). If such a system is deployed, we will distribute the URL and invite you to participate actively and contribute to the survey.

## Author Details

**Harald Krottmaier**
Institute for Information Systems and Computer Media (IICM)
Faculty of Computer Science
University of Technology
Graz Austria

url: http://www.informatik.tugraz.at/
E-mail: hkrott@iicm.edu
Telephone: +43-316-873-5631

| | | |
|---|---|---|
| DELOS | Network of Excellence | email: yannis@di.uoa.gr |
| Name of Cluster: | Information Access and Personalization | phone: +30 210 727-5224 |
| Name of Leader: | Yannis Ioannidis | fax: +30 210 727-5214<br>url: http://www.di.uoa.gr/~yannis |

Return to top

## Audio/Visual and Non-traditional Objects

**Nektarios Moumoutzis** scopes the work and describes the current activity of WP3 - Audio/Visual and Non-traditional Objects.

## Cluster Organization

The cluster is co-ordinated by Professor Stavros Christodoulakis of the Technical University of Crete, Greece and Professor Alberto Del Bimbo of the University of Florence, Italy. The cluster comprises seventeen organizations, (see the list of the organizations and contact details)

## Objective

The major objective of the Audiovisual and Non-Traditional Objects Cluster of DELOS II NoE (http://delos.music.tuc.gr) is to focus on research activities related to the management of the lifecycle of Audiovisual and Non-Traditional Objects including:

- Capturing of digital objects in order to integrate them into digital libraries.
- Analysis of the content of these objects emphasizing metadata extraction and the correlation between them.
- Management of audiovisual metadata and content including the use of databases, indexing mechanisms, and support for versioning.

- User interactions exploiting the semantics of content for efficient retrieval and dissemination.
- Transformation of the digital objects and synthesis of new objects to support services such as multi-channel delivery, personalized browsing and repurposing (e.g. creation of educational experiences).
- Advanced applications

## Scope of the Work

The scope of the work of the cluster is two-fold: The first part is to develop common foundations for:

1. **Metadata capturing for audiovisual content**: In particular, the cluster is investigating the state of the art of automatic extraction of audiovisual metadata and produces new models and experiments with prototype systems for automatic and semi-automatic extraction models in domain-specific and context-specific audiovisual application environments.
2. **Universal access and interactions with audiovisual libraries**: The cluster is looking into the state of the art of audiovisual information access and interaction exploring models based on the integration of existing frameworks of multimedia content (e.g. MPEG-7, TV_Anytime) and of domain-specific extensions of those models in order to enhance the effectiveness of retrieval and provide value-added services. Moreover it is also exploring models of user interaction based on the concept of context and context-based retrieval in audiovisual digital library applications. The cluster is also studying models and interfaces for advanced applications of audiovisual digital libraries. It intends to develop new solutions and conduct experiments with prototype systems for processing, semantic transcoding, interaction and presentation of audiovisual content for different delivery media and interaction devices so as to adapt the users' preferences regarding content to the specific characteristics of the delivery media and the users' terminals.
3. **Management of audiovisual content**:The cluster is also looking at state-of-the-art database models and data structures appropriate to the storage, retrieval and dissemination of audiovisual metadata and will conduct experiments with prototype systems. It is also investigating state-of-the-art models for profiling and stereotyping users of the most important multimedia applications and current dissemination strategies.

The second part of the scope of work within the cluster is to support common research activities in specific areas identified by the partners, taking into account the current trends and the future directions in the field emphasizing the development of innovative solutions for emerging digital libraries which are integrating Audiovisual and Non-Traditional Objects.

## Cluster-supported Research Activities

Research activities identified by the technical annex and currently being investigated include:

1. Research on the development of automatic annotation algorithms at the semantic level, taking into account both higher ontologies supported by multimedia content standards and domain Ontologies addressing the needs of specific application domains such as sports, education, cultural heritage, etc.
2. Research on universal access and interaction with audiovisual libraries including the development of interaction models that capture issues related to the use of digital objects by final users; context models that represent contextual information related to user intentions, physical location of the user, history of interactions etc.; multimodal interfaces integrating new means of user interactions including natural language and speech; models for interoperability of applications of non-traditional objects following diverse standards or Ontologies.
3. Research on the management of audiovisual content in digital libraries including semantic retrieval languages, models for segmentation and summarization as well as models for multimedia databases and multimedia P2P computing.
4. Research on techniques for 3D objects modeling and retrieval.
5. Research on the integration of semantic video annotation and transcoding so as to obtain a semantic transcoding system; extension to real time applications.
6. Joint activity on the development of advanced multimedia demonstrators and test-beds in order to provide a reference point for the development of innovative tools and mechanisms for digital library services by ensuring a common evaluation framework.

## Cluster Activities So Far

During the eight months from the official start of the project, an effective management structure for the cluster has been established which began with the cluster kick-off meeting in January 2004. The meeting also nominated a Steering Committee for the cluster. In addition, the skill sets of cluster members have been identified in detail and mapped to the activities of the cluster.

Cluster workshops have been scheduled for the dissemination of cluster results and for the discussion and development of new proposals, solutions and achievements for audiovisual digital libraries. An internal workshop will be held in Chania, Crete, during November 2004.

Cluster 3 - Audio/Visual and Non-Traditional Objects - is organizing a joint workshop with cluster 4 - User Interfaces and Visualization - entitled "Seventh International Workshop of the EU Network of Excellence DELOS on AUDIO-VISUAL CONTENT AND INFORMATION VISUALIZATION IN DIGITAL LIBRARIES (AVIVDiLib'05)". Detailed information about this workshop can be found at http://delos.dis.uniroma1.it/Workshops/default.aspx

To facilitate communication between the cluster members as well as the dissemination of cluster information to the other Delos II NoE members and to the outside world, a site was established at the address http://delos.music.tuc.gr. The site is being updated on a daily basis with new material and news related to the activities of the cluster.

The co-ordination with the other relevant Delos II NoE clusters is a matter of major importance. To support closer co-ordination and collaboration, liaison partners have been nominated as follows:

- **Prof. Hans Schek** with cluster 1 - Digital library architecture
- **Prof. Stavros Christodoulakis** with cluster 2 - Information access and personalization - and with cluster 5 - Knowledge extraction and semantic interoperability
- **Prof. Alberto Del Bimbo** with cluster 4 - User interfaces and visualization

A major activity since the beginning of the project has been the development of State of the Art reports in the major axis of interest of the cluster. These are:

- State of the Art report of Metadata Extraction for images, videos, 3D objects, audio, domain-specific ontologies and representation frameworks
- State of the Art report of Audiovisual Content-based Retrieval, Information Universal Access & Interaction including Datamodels & Languages
- State of the Art report of Audiovisual Metadata Management

The State of the Art reports are also intended to map the expertise and recent research of cluster members in the international research framework.

An important task under way is the development of a website offering the demonstrators and test datasets to be used during the project. Demonstrators can be on-line or off-line. With the former, a link to the demonstrator is given ro provideimmediate use. With the latter, case a link is provided to downloadable software which users can then install on their machines. Different types of test datasets will be chosen to be made available for evaluation purposes, including datasets for images, videos, 3D graphics, audio, music, etc.

## Plans for the Next Months

The cluster is currently finalizing the State of the Art reports including previous research by the cluster in the field. The reports will include a section on the future directions that will guide the work of the cluster.

In order to schedule the work in the next 18-month phase of the project and take decisions regarding the funding of the participating partners, the cluster co-ordinators have issued a request to the cluster members for joint research proposals . The proposals will be evaluated on the basis of their interest and will serve to identify the work to be done over the next 18 months. The next 18-month plan will also take account of the Technical Annex, the State of the Art reports and the proposals submitted.

## Author Details

**Nektarios Moumoutzis**
Project Co-ordinator & Researcher
Laboratory of Distributed Multimedia Information Systems & Applications
Technical University of Crete (MUSIC/TUC)
Chania
Crete

Greece

url: http://www.music.tuc.gr/~nektar
E-mail: nektar@ced.tuc.gr
Telephone: +30-821-64803

| | | |
|---|---|---|
| DELOS | Network of Excellence | |
| Name of Cluster: | Audio/Visual and Non-traditional Objects | |
| Name of Leaders: | Stavros Christodoulakis | email: stavros@ced.tuc.gr<br>phone: +30-2821-037399<br>fax: +30-2821-037399<br>url: http://www.music.tuc.gr |
| | Alberto del Bimbo | email: delbimbo@dsi.unifi.it<br>phone: +39-055-479 6262<br>fax: +39-055-479 6363<br>url: http://viplab.dsi.unifi.it/ |

Return to top

# User Interfaces and Visualization

**Stephen Kimani** reports on a major item of activity by his cluster based on four case studies on digital libraries.

# Case Study Analysis

The DELOS WP4 User Interfaces and Visualization cluster has carried out an analysis of a number of case studies of digital libraries. The analysis was headed up by Risoe National Library which is responsible for Task 1, the provision of an empirical basis.

The case studies analyzed were:

- LAURIN http://laurin.uibk.ac.at/ - Libraries and Archives Collecting Newspaper Clippings (analysis by Roma1)
- SCHOLNET http://www.ercim.org/scholnet/ - a test bed for retrieval and annotation of software concepts for development of digital libraries (analysis by Fraunhofer-IPSI, prime contractor)
- COLLATE http://www.collate.de/ - a Collaboratory for Annotation, Indexing and Retrieval of Digitized Historical Archive Material (analysis by Fraunhofer-IPSI, prime contractor)
- i-dove http://i-dove.ics.forth.gr/ - Interactive Guidelines Support Tool for the Development of Virtual Environments, developed in the context of the VIEW Project by Forth-ICS (and analysis by Forth-ICS, Project Partner)

This analysis specifically concentrated on the following overall aspects of digital libraries: domain(s), users, tasks/services, methodology (empirical perspective)as well as input for a digital library taxonomy.

## Domains

As far as the domains are concerned, they are divided into two categories: work domains and collection content domains. The work domain refers to the primary territory of work at which the digital library is targeted. The collection content domain refers to the primary types of knowledge that are accessible in the digital library. COLLATE and LAURIN address the work domains of archives and libraries, i-dove covers the work domain of development of virtual environments, and SCHOLNET addresses the work domain of development of digital library environments. With regard to the collection content domains, COLLATE addresses film knowledge, LAURIN covers newspaper content, i-dove covers ergonomic virtual environment knowledge, and SCHOLNET addresses development knowledge.

## Users

The DELOS WP4 technical annex identifies three classes of digital library users: experts in knowledge mediation, experts in knowledge content and end-users. In COLLATE, experts in knowledge mediation are film archive staff;

experts in knowledge content are film archive staff, film researchers, and university students; whereas end-users are almost any type of user (from experts to 'lay' users). In LAURIN, experts in knowledge mediation are librarians in public libraries; experts in knowledge content are professionals (e.g., journalists and researchers); whereas end-users are public library users. In i-dove, experts in knowledge mediation are knowledge creators/reviewers (e.g. moderators); experts in knowledge content are managers and producers of sources of knowledge; whereas end-users are developers of industrial virtual environment applications. In SCHOLNET, experts in knowledge content and end-users are digital library developers and managers; it is not evident however who the experts in knowledge mediation are.

### Tasks and Services

Each of the four digital library cases was analyzed in terms of tasks and services. In COLLATE, the primary tasks are integration of knowledge through indexing; cataloguing and annotation; and access to knowledge. LAURIN also provides access to knowledge and supports the integration of knowledge through thesaurus-building. In i-dove, the primary tasks are development and management of distributed knowledge; access to knowledge; and knowledge exploitation for the creation of virtual environments. SCHOLNET supports the management of distributed digital library knowledge and provides access to knowledge.

### Methodology

As for the methodology, the empirical aspect of the digital library and user interface development were analyzed. In particular, the kind of data gathered and used to define user requirements, design and evaluation, how the data are gathered and analyzed, and the kinds of supporting technology (guidelines/available technology) that are used, etc. The analysis presents an overview of the main digital library characteristics that are derived from the dimensions: user domain and knowledge domain. As far as the user domain is concerned, the most common unit of analysis is individual users, viewed as representative of particular professions or stakeholders. Study methods vary from participant observation, interviews and questionnaires to testbeds and interactive prototyping. Dialogue between users and developers appears to be important in analysis of user needs. As for the knowledge domain, the gathering of knowledge content has two phases (which may iterate): creating an initial knowledge content base, and maintaining/revising the knowledge content base. The four digital library efforts have approached the creation of an initial knowledge base in quite different ways.

### Digital Library Taxonomy

With regard to the input for a digital library taxonomy (in a thesaural tree-structure format), the analysis produced a tree structure with the following top terms: domain, user, task, services. Task and service categories cover the so-called “functional requirements”. Among the case studies examined, the following general non-functional requirements were identified: scalability, platform independence, reliability, strength, flexibility, adaptability, customizability, and usability. All case studies have featured the importance of reliability, flexibility, and usability. However, none of the digital libraries analyzed has been designed in a manner so as to support accessibility.

### Conclusions

The case studies analyzed presented significant differences in terms of domains, users, and tasks but they all were developed following user-focused approaches. Iterations of design solutions and evaluations involving users were features identified in most of the reported DL development lifecycles. However, the analysis did not provide insights sufficient to identify definitively the various phases in the DL usage lifecycle in general. The digital library lifecycle phases that just one of the case studies supports include: the acquisition process, the clipping archive management (indexing, thesaurus), the access to the selected clippings through the queries, data retrieval access, maintenance of the database and future exploitation. Further studies are needed to arrive at a general characterization of such lifecycle phases.

## Author Details

**Stephen Kimani**
University of Rome "La Sapienza"
DIS, Piano 2, Stanza 233
Via Salaria 113
00198 Rome

Italy

url: http://www.dis.uniroma1.it/~kimani/
E-mail: kimani@dis.uniroma1.it
Telephone: +39-06-49918548

| | | |
|---|---|---|
| DELOS | Network of Excellence | email: catarci@dis.uniroma1.it |
| Name of Cluster: | User Interfaces and Visualization | phone: +39-06-4991 8331 |
| Name of Leader: | Tiziana Catarci | fax: +39-06-4991 8331<br>url: http://www.dis.uniroma1.it/~catarci/ |

Return to top

# Knowledge Extraction and Semantic Interoperability

**Michael Day** and **Manjula Patel** report on the WP5 KESI Workshop on Semantic Interoperability in Digital Libraries, held at the University of Bath, 17 September 2004.

**Liz Lyon**, Director of UKOLN and co-ordinator of the DELOS research cluster on Knowledge Extraction and Semantic Interoperability (WP5), welcomed delegates to the University of Bath for the cluster's first workshop on the 17 September 2004. The objectives of the workshop were as follows:

- To raise awareness of cluster activities to the wider community;
- To initiate discussion of key issues in the fields of knowledge representation and semantic interoperability;
- To share current research experience with interested parties and cluster members.

The workshop consisted of three presentations by members who had recently joined the cluster together with a description of DELOS activities on semantic interoperability.

**Colleagues from WP5 who attended the KESI Workshop:
(back row, left to right) Traugott Koch, Lund University; Les Carr, University of Southampton; (middle row, left to right) Jane Hunter, University of Queensland; Michael Day, UKOLN; David Alsmeyer, BT; Doug Tudhope, University of Glamorgan; Martin Doerr, ICS-FORTH; (front row, left to right) Manjula Patel, UKOLN; Liz Lyon, UKOLN; Koraljka Golub, Lund University.**

The first speaker was **Jane Hunter** of the Distributed Systems Technology Centre (DSTC), Brisbane, Australia, who introduced tools and approaches for enhancing semantic interoperability based on research being undertaken at DSTC. Key topics in this area included the description of Web-based resources and services, query mediation between different metadata formats, metadata lifecycles, and provenance. She introduced the MAENAD (Multimedia Access across Enterprises, Networks And Domains) Project, which is concerned with the development of tools that can facilitate the discovery, use and management of multimedia regardless of domain. Key components include common data models, ontologies and metadata schemas. The project also has an interest in metadata generation and capture, e.g. through the statistical generation of e-science workflows or the semantic indexing of multimedia. Other important issues include presentation interfaces and visualisation, annotation, and

preservation tools.

Within DSTC, they have been using a version of the ABC ontology [1] refined with reference to the CIDOC Conceptual Reference Model (CRM) (http://cidoc.ics.forth.gr/) to harmonise a number of domain specific ontologies [2]. Jane then spoke in more detail about a joint project with the Smithsonian Museum of the American Indian (http://www.nmai.si.edu/) concerning the management of indigenous knowledge [3]. This had developed software tools that could be used by members of indigenous communities to annotate and control access to digitised cultural resources. Annotation tools allow the members of indigenous communities to document the meaning and significance of resources from their own perspectives. For some resources, the question of access could be quite complex, e.g. it could be dependent on the time of year, the tribe or gender of the user. In the interface, access constraints are indicated by logos and it is possible to generate different views of the same database based on different access conditions. The implementation was based on Vannotea, an application for the collaborative indexing and annotation of video, also developed at DSTC [4].

In terms of forthcoming activity, Jane spoke very briefly about a DSTC project concerned with the assimilation of the different kinds of data used to help optimise Hydrogen fuel cell design and efficiency [5].

The next presentation was by **David Alsmeyer** of British Telecommunications plc, who introduced the SEKT (Semantically Enabled Knowledge Technologies) Project (http://www.sekt.semanticweb.org/), an EUR 12.5m project funded as part of the European Union's Sixth Framework Programme. He began his presentation by invoking what he called the Eastbourne Test; i.e. making sure that Semantic Web technologies addressed the real questions that people wanted answering. He felt that one key issue was the 'bottleneck' of creating annotation or metadata. Solutions might include the use of human language technologies, e.g. for the automatic extraction of metadata from articles or to support on-the-fly metadata creation. SEKT were investigating some of these issues through three case studies, covering the requirements of newly appointed judges in Spain, IT consultants in Germany and a corporate digital library in the UK.

David noted the transformation of BT's own library into a digital library where the majority of services are delivered online directly to engineers at their desks. The aim ought to be to apply semantic thinking in the library context. Examples might be personalised services, e.g. to document search terms used to build up a framework of users' interests or mapping between thesauri and personal profiles to help find colleagues with the same interests.

**Doug Tudhope** of the Hypermedia Research Unit at the University of Glamorgan, UK then spoke on "Recent developments from the perspective of networked knowledge organisation systems and services". He started by emphasising that Glamorgan's approach mainly concentrated on doing better with existing thesauri and using these to build towards the Semantic Web, the development of terminology services or semantic terminology services. He thought that one thing that the DELOS cluster could do would be to revisit the useful taxonomy of knowledge organisation systems devised by Hodge [6].

Doug then introduced the FACET Project, which had experimented with integrating thesauri into user interfaces, e.g. for semantic query expansion and disambiguation [7]. The project had developed a Web demonstrator (http://www.comp.glam.ac.uk/~FACET/webdemo/) using the Art and Architecture Thesaurus (AAT) and the collections of the UK National Museum of Science and Industry to show how thesauri could be used for query expansion in a realistic application. A live demonstration showed how some of the features of the FACET system worked.

Looking to the future, important work was proceeding on revisions to the BSI's thesaurus standards (BS 8723) and the definition of a Resource Description Framework (RDF) vocabulary for describing thesaurus data as part of the SWAD-Europe thesaurus activity. Doug finally noted some key differences between ontologies and thesauri - one reflecting scientific precision, the other cost-effectiveness - but suggested that there should be some convergence.

**Manjula Patel** of UKOLN then gave an overview of key issues for semantic interoperability from the perspective of the cluster. DELOS Knowledge Extraction and Semantic Interoperability (WP5) was preparing a state-of-the art report on semantic interoperability in digital library systems, identifying why it was important and noting key gaps. She started by introducing the authors of the report and their affiliations (Manjula Patel, UKOLN; Traugott Koch, Netlab, Lund University; Martin Doerr, ICS FORTH; Chrisa Tsinaraki, Technical University of Greece). An outline structure of the report in terms of a table of contents was distributed to the delegates for reference.

The current state of the report is such that all sections that have been allocated to authors have been completed to

first draft stage. The report needs some work with regard to improving consistency and coherence between the various sections. A major aim of the report is to integrate views from overlapping communities working in the area of semantic interoperability, these include: Semantic Web, artificial intelligence, knowledge representation, ontology, library and information science and computer science. The types of issue that the report is trying to address include:

- Why is Semantic Interoperability (SI) important in Digital Library systems (DLs) and how can it be used in DLs?
- An analysis of different types or levels of SI
- A clarification of the relationship between syntactic and semantic interoperability
- Describing relevant methodologies, prerequisites, standards and tools
- How can SI in DLs be enhanced?

In examining such issues, the authors developed a broad outline of the report:

- Overview
- Introduction and definition of SI
- Importance of SI in DLs
- Theoretical Considerations
- Prerequisites to enhancing SI
- Methods and processes to enhance SI in DLs
- SI in DL Services

Definitions of interoperability, syntactic interoperability and semantic interoperability were then presented noting that SI is very much about matching concepts as a basis. SI has been identified as being of primary importance in DL research by the NSF Post Digital Libraries Futures Workshop: Wave of the Future [8]. The overall goal of SI is to support complex and advanced, context-sensitive query processing over heterogeneous information resources. The report examines several areas in which SI is important in DLs, these include: improving the precision of search, enabling advanced search, facilitating reasoning over document collections and knowledge bases, integration of heterogeneous resources, and its relevance in the information life-cycle management process. The report also investigates some theoretical issues such as clarification and selection of relevant terminology, standardisation and interpretation and the differing levels of SI in DL environments. It notes that information structure, language and identifiable semantics are prerequisites to SI, as is consensus building and standardisation.

Other important areas include the role of foundational and core ontologies, knowledge organisation systems (KOS), syntactic interoperability and encoding systems and the role of semantic registries, tools and architectures. Rights issues are also relevant in the integration and reuse of information. The final sections of the report on enhancing SI in DLs and SI in DL services are still being addressed, although an outline of areas to be covered appears in the table of contents. Manjula welcomed participants to comment on the structure and content of the report and to identify any obvious gaps.

A lively discussion ensued with regard to the issues raised in the report and the presentation. On closing the workshop, Liz Lyon thanked participants for their contributions and summarised the key issues as being: harmonisation, integration, bridging gaps, migration, integration of heterogeneous information and ontologies.

## References


1. Lagoze, C., & Hunter, J. "The ABC ontology and model". Journal of Digital Information, 2(2), 2001, article no. 77. Retrieved from: http://jodi.ecs.soton.ac.uk/Articles/v02/i02/Lagoze/
2. Hunter, J. "Enhancing the semantic interoperability of multimedia through a core ontology." *IEEE Transactions on Circuits and Systems for Video Technology*, 13(1), 2003. 49-58. Retrieved from: http://archive.dstc.edu.au/RDU/staff/jane-hunter/events/paper.html
3. Hunter, J., Koopman, B., & Sledge, J. "Software tools for indigenous knowledge management." Museums and the Web 2003, Charlotte, North Carolina, USA, 20-22 March 2003. Retrieved from: http://www.archimuse.com/mw2003/papers/hunter/hunter.html
4. Hunter, J., Schroeter, R., Koopman, B., & Henderson, M. "Using the Semantic Grid to build bridges between museums and indigenous communities." GGF11 Semantic Grid Applications Workshop, Honolulu, Hawaii, USA, 10 June 2004. Retrieved from: http://www.semanticgrid.org/GGF/ggf11/ggf11-semgrid-proceedings.pdf
5. Hunter, J., Drennan, J., & Little, S. "Realising the Hydrogen economy through Semantic Web technologies." IEEE *Intelligent Systems*, January/February 2004, pp. 2-9 Retrieved from:

http://maenad.dstc.edu.au/papers/2004/ieee-is.pdf
6. Hodge, G. *Systems of knowledge organization for digital libraries: beyond traditional authority files.* Washington, D.C.: Council on Library and Information Resources, 2000. Retrieved from: http://www.clir.org/pubs/abstract/pub91abst.html
7. Binding, C., & Tudhope, D. (2004). "KOS at your service: programmatic access to knowledge organisation systems." *Journal of Digital Information*, 4(4), 2004, article no. 265. Retrieved from: http://jodi.ecs.soton.ac.uk/Articles/v04/i04/Binding/
8. NSF Post Digital Libraries Futures Workshop: Wave of the Future, June 2003, http://www.sis.pitt.edu/%7Edlwkshop/

## Author Details

**Michael Day**
UKOLN
University of Bath
E-mail: m.day@ukoln.ac.uk
url: http://www.ukoln.ac.uk/

**Manjula Patel**
UKOLN
University of Bath
E-mail: M.Patel@ukoln.ac.uk
url: http://www.ukoln.ac.uk/

| | | |
|---|---|---|
| DELOS<br>Name of Cluster:<br><br>Name of Leader: | Network of Excellence<br>Knowledge Extraction and Semantic Interoperability<br>Elizabeth Lyon | email: e.lyon@ukoln.ac.uk<br>phone: +44 1225 386580<br>fax: +44 1225 386838<br>url: http://www.ukoln.ac.uk |

Return to top

## Evaluation

**Sarantos Kapidakis** provides a summary of activity for the Evaluation cluster and describes the objectives and proposed content on the Work Package's forthcoming workshop.

## Meetings

After our Kick-off meeting in Duisburg, Germany on 23 Jan 2004, the next Work Package meeting took place in Corfu, over 3-4 May 2004 in which we planned the WP activities mentioned as well as our workshop.

## WEB

The Evaluation - WP7 website http://dlib.ionio.gr/wp7 is working and shows the aims and the objectives of the Work Package, the WP7 partners, related events, our publications, bibliography and useful links. It also includes a discussion forum to enable communication on evaluation issues. The design of the website is an iterative process and the collection of its content is a continuous WP activity.

## INEX

INEX has seen intense activity during the last few months. We looked at the issue of accessing the INEX testbed. DELOS members can get access to the documents and queries, but not the relevance assessments. Approximately 50 groups are now registered to use it this year and there were 4 evaluation tracks. We have selected the CO topics for INEX 2004. 40 queries were selected this year. Many topics need modifications or corrections to be usable for the various tracks including the ad hoc retrieval task. Topic Assignment has been finalized and two topics are assigned to each participant keeping in view the topic authors. The deadline is 8 October 2004.

## CLEF

After the detailed design of evaluation tracks, approximately 60 groups registered for one or more of the 8 evaluation tracks. 50 topics are prepared for mono- / bi- / multilingual tracks in 13 languages and 25 topics for Structured Scientific Data in 3 languages. Finally, CLEF is organizing a 3-day Workshop, co-located with the ECDL2004 conference, at Bath University, UK (15-17 September 2004), where the results will be presented.

## Evaluation Workshop

The DELOS Workshop on the Evaluation of Digital Libraries http://dlib.ionio.gr/wp7/events.html is going to be held at the Department of Information Engineering (http://www.dei.unipd.it/), University of Padua, Via Gradenigo 6/a, 35131 Padova, Italy over 4-5 October 2004.

The objective of the workshop is to formulate a new model for DL evaluation and the results and recommendations emerging from the workshop will provide input for the production of a White paper (which will be ready by mid-December 2004). The workshop will cover a discussion on all components of a DL. We have invited representatives from each DELOS WP, and also a few speakers to cover different areas including some people from FP6 DL projects. Participation in the workshop will be by invitation only.

The Workshop is co-located with the DIALOGUES conferences (http://thepadovadialogues.dei.unipd.it) Padova Dialogues 2004 comprises three conferences:

- SPIRE 2004 (String Processing and Information Retrieval) http://thepadovadialogues.dei.unipd.it/spire04
- ALT 2004 (Algorithmic Learning Theory) http://www.tcs.uni-luebeck.de/pages/alt04.jhtml
- DS 2004 (Discovery Science) http://www.slab.dnj.ynu.ac.jp/DS04

and satellite meetings for exchange and integration in the modelling, design and implementation of advanced tools for the representation, encoding, storage, search, retrieval and discovery of information and knowledge.

The workshop will focus on the definition of a workable evaluation framework. Specifically, this workshop will try to find answers to the following questions:

- Why are DLs evaluated?
- What are standard approaches for DL evaluation?
- Which aspects/components are evaluated in these approaches?
    - collections, users, usage, systems
    - content, community, services, technology
    - societies, scenarios, spaces, structures, streams
- Which evaluation criteria are used in these approaches?
- What are the strengths and weaknesses of these approaches?
- What should be done (in the future) in DL evaluation?

In addition, two general issues are to be addressed:

- What is an appropriate conceptual DL model to which evaluation could refer?
- Are there conceptual frameworks for evaluation?

## Author Details

**Sarantos Kapidakis**
Laboratory on Digital Libraries and Electronic Publishing
Archive and Library Sciences Department
Ionian University
Platia Eleftherias, Palea Anaktora,
Corfu 49100, Greece

email: sarantos@ionio.gr
Tel: +30 26610 87413
Fax: +30 26610 87436

| | | |
|---|---|---|
| DELOS | Network of Excellence | email: fuhr@uni-duisburg.de |
| Name of Cluster: | Evaluation | phone: +49-203-379-2524 |
| Name of Leader: | Norbert Fuhr | fax: +49-203-379-2549 |
| | | url: http://is.informatik.uni-duisburg.de |


Return to top

DELOS Home | DELOS Newsletter Front Page | Delos Newsletter Contents

Publication Date: October 2004

Publication: Delos Newsletter Issue 2

Originating URL: /newsletter/issue2/cluster-reports/intro.html

File last modified: Wednesday, 03-Nov-2004 17:40:30 CET

The Delos Newsletter is published by the DELOS Network of Excellence and is edited by Richard Waller of UKOLN, University of Bath, UK.

Newsletter
Issue 2 : October 2004



# DELOS Information Day, Bath 15 September 2004

**Michael Day** provides an overview of the meeting held by DELOS partners to brief interested ECDL2004 delegates on the aims and issues for the DELOS Network of Excellence.

## Introduction

Following the closing session of the 8th European Conference on Digital Libraries (ECDL 2004), delegates met at the University of Bath in the United Kingdom for an information day on the DELOS Network of Excellence on Digital Libraries [1]. The DELOS Network is a project funded by the European Commission under the Information Society Technologies (IST) priority [2] of the European Union's Sixth Framework Programme (FP6) [3].

Delegates were welcomed by Vittore Casarosa of the Institute of Information Science and Technologies of the Italian National Research Council (CNR-ISTI) and leader of the DELOS Network's dissemination effort. The first presenter was Claude Poliart, the project officer from the European Commission, who noted the growing importance of the information society in the twenty-first century. His expectation was that the DELOS network would become a catalyst for uniting the global digital library community as well as a forum to exchange good practice and different perspectives. He predicted that the socio-economic impact of DELOS and other digital library programmes would increase over the next decades.

Costantino Thanos of CNR-ISTI, the scientific co-ordinator of the DELOS network, then gave a brief overview of the project's organisation, objectives and joint programme of activities. Its main objective was to define a joint programme of activities for next-generation European digital library research. The project is organised into thematic clusters, each with its own co-ordinator.

## Corvara Brainstorming Meeting

Thanos was followed by a report by Yannis Ioannidis of the University of Athens of a DELOS brainstorming meeting held in Corvara, Italy on the 8-9 July 2004. The meeting had concerned the future of digital libraries and its primary goal was to develop a research agenda. It first divided into parallel sessions of three competing research agendas, each group reporting back its findings which at a level of abstraction all showed some commonality. There were 25 invited participants, who all presented their own vision of digital libraries. The main conclusions were that digital libraries had to become more user-centred, that digital libraries should not just be passive repositories but required more active collaboration and communication tools, and that there was a need for more generic digital library management systems. It was felt that in the future no one should be building from scratch; instead the generic digital library management systems should provide a basis for the specialised modules required for each system. There was also a feeling that the term 'digital libraries' did not really reflect the subject as it has now developed [4]. Specific recommendations included the need for reference models with defined roles and for digital libraries to be integrated into larger environments, e.g. health, science, government, learning, etc. Research was required in two separate areas. Firstly, user-centred research, taking into account the different roles users play, the need for good interfaces, personalisation, etc. Secondly, research was needed into the system layer, the bits and bytes of digital information, middleware, generic and pragmatic semantics. From this was developed a list of research topics. This reflected what Costantino referred to as the two layers of a conceptual framework.

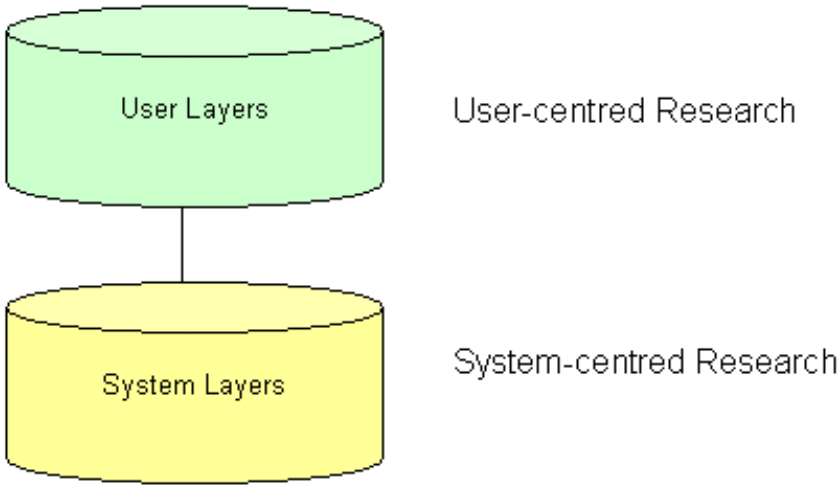


**Figure 1: Conceptual Framework for Research**

Finally, the brainstorming meeting investigated the possibility of developing a phrase that could be used instead of 'digital libraries.' Combinations of various terms were considered, including the adjectives 'pervasive,' 'ambient' or 'collaborative' and the entities 'garden,' 'factory' and 'architecture,' but the phase 'dynamic ubiquitous knowledge environments' was said to be the favoured choice of the meeting.

## Presentations from Clusters

The information day then turned to a more detailed description of each DELOS cluster, all of which addressed their objectives and task plans, while highlighting progress to date.

First, Seamus Ross of the University of Glasgow (UK) introduced the challenges being addressed by the preservation cluster (WP6) and some of its current activities. These included the authenticity and reliability of digital objects, their ingest and retrieval, and technology obsolescence. Other problems related to the identification of adequate preservation metadata, and the legal and organisational contexts of preservation. In thinking about cluster activities for the next phase of the DELOS network, Ross talked about the re-engineering of processes and technologies, including ingest and appraisal. Current cluster activities included work on the development of a digital preservation testbed environment, the design and deployment of digital repositories, file formats and the representation and documentation of function and behaviour. Ross noted the urgent need to integrate preservation requirements with digital library design. He concluded by announcing that there would be a DELOS summer school on digital preservation in the south of France over 5-11 June 2005.

Sören Balko of the Swiss Federal Institute of Technology (ETH) Zurich (Switzerland) then introduced the work of the cluster on digital library architectures (WP1). The objectives of the cluster were to evaluate peer-to-peer, grid middleware and service oriented architecture. Progress to date had included the creation of a cluster Web site [5] and a thematic workshop held in Cagliari, Sardinia in June 2004. Other planned work would include a survey and comparison of the service architecture, peer-to-peer and Grid approaches to digital libraries. The proceedings of the Cagliari workshop had already been published on the cluster Web site and negotiations were underway with a publisher for a book on current trends in digital library architectures. Other future activities include a joint cluster workshop with WP2, to be held in Dagstuhl, Germany over 29 March - 1 April 2005.

Yannis Ioannidis then returned to speak about the cluster on information access and personalisation (WP2). The objectives of the cluster had been further informed by the Corvara meeting, but key topics included access and integration and generic, user-centred approaches to personalisation. The cluster is also co-operating with the architecture clusters on the Dagstuhl workshop, and was planning a further workshop to be held in Greece. Surveys had already started on access and interaction models and metadata. Other surveys would cover integration and interaction management schemes, integrated metadata, data provenance, and personalisation. The cluster had also collaborated with WP4 on the DELOS summer school on 'User-centred design of digital libraries', held in Pisa, Italy over 6-10 September 2004.

Stavros Christodoulakis of the Technical University of Crete (Greece) introduced the work of the cluster on audiovisual and non-traditional objects (WP3). The cluster had initiated research activities on the lifecycle of non-traditional and audiovisual objects, including capture, analysis, metadata extraction, normalisation, management, semantic interaction, retrieval languages and algorithms. Christodoulakis said that there was a need to develop common foundations, e.g. on metadata capture and automatic annotation, universal access, management of multimedia, semantic retrieval, models of segmentation and summarisation. He was of the view that all the issues mentioned could not be viewed in isolation and that advanced applications such as eScience and eLearning would also be relevant to the cluster's work, as well as the issue of interoperability between repositories. An internal workshop was planned for Crete in November 2004 and a joint cluster workshop with WP4 in Cortona, Italy, May 2005. The cluster was also preparing state-of-the-art reports on metadata extraction, content-based retrieval, interaction, and data management.

After a short break, Tiziana Catarci of the University of Rome 'La Sapienza' (Italy) spoke about the activities being undertaken by the cluster on user interfaces and visualisation (WP4). The cluster was mainly addressing the subject of user needs, specifically how to collect and analyse user requirements from end users and other stakeholders. Work was also being undertaken on a theoretical framework for the design of new interfaces, the contexts digital library lifecycles, functional and non-functional requirements, visualisation and standards for usability and accessibility. Initial deliverables had included the cluster Web site [6] and a draft report on functional and non-functional requirements for digital libraries. Planned future events include a workshop on audiovisual content and visualisation to be held in Cortona on the 4-6 May 2005 and a special session on accessing digital libraries at the International Conference on Human-Computer Interaction (HCII 2005), to be held in Las Vegas, USA over 22-27 July 2005.

Liz Lyon of UKOLN (UK) introduced the cluster on knowledge extraction and semantic interoperability (WP5) by highlighting aspects of the constantly evolving landscape of the cluster. Examples were recent reports recommending the deployment of institutional repositories, the need for knowledge extraction in data-intensive e-science contexts, and the growing importance of the Semantic Web. WP5 had a major focus on fostering integration and there had been an emphasis on inviting interested non-funded partners to join the cluster. There was also an awareness of the need to integrate with other DELOS clusters, e.g. with WP6 on repositories. The cluster was already undertaking a survey of semantic interoperability and a workshop on this topic was planned for the following day in Bath.

Claus-Peter Klas of the University of Duisburg-Essen (Germany) then introduced the work of the cluster on evaluation (WP7). He started by stressing the importance of communication between evaluation experts and digital library researchers. An evaluation forum had been set up to facilitate this communication [7] and a workshop was planned for 4-5 October 2004 in Padova, Italy, to which selected external experts had been invited. The cluster was also involved in two major evaluation activities. First, the Initiative for the Evaluation of XML Retrieval (INEX), which is testing the retrieval effectiveness, efficiency and usability of XML documents based on a corpus of 12,000 articles from IEEE Computer Society publications. INEX 2004 [8] currently had 57 participants and a workshop was planned for Dagstuhl in December 2004. Secondly, the cluster was involved in the Cross-Language Evaluation Forum (CLEF), which is concerned with the evaluation of multilingual retrieval systems [9]. It was noted that the CLEF 2004 workshop was simultaneously taking place in a neighbouring building at the University of Bath [10].

Vittore Casarosa then introduced the dissemination and spreading of excellence cluster (WP8). He started by noting that FP6 placed more emphasis on integration and that dissemination could be used to support this. Cluster activities included a network Web site (later to become a portal) and an electronic newsletter that focuses on cluster activities. Past and future events include cross-cluster thematic workshops (e.g., Padova, Dagstuhl and Cortona), workshops hosted with other user communities (e.g. ICA Congress 2004, FIAT-IFTA), regular brainstorming meetings (e.g., Corvara), and national awareness events (e.g. those held in Rome and Lund, June 2004). There had been a recent summer school on user-centred design (Pisa, September 2004) with another on preservation planned for next year (June 2005). Casarosa concluded by noting that there would be a call for proposals to host future ECDL conferences (from 2006) and that DELOS had funding available for a researcher exchange programme (requests to be sent to the cluster co-ordinators) where the host organisation was a DELOS member.

## Acknowledgements

My thanks to Manjula Patel and Richard Waller of UKOLN for their assistance in the compilation of this report.

## References


1. DELOS Network of Excellence: http://www.delos.info/
2. Information Society Technologies priority: http://www.cordis.lu/ist/
3. Sixth Framework Programme: http://fp6.cordis.lu/fp6/home.cfm
4. Yannis Ioannidis, *The DELOS report: digital libraries at a crossroads*, DELOS Network of Excellence, September 2003.
5. DELOS digital library architecture (WP1): http://www.dbs.ethz.ch/delos/
6. DELOS WP4 - user interfaces and personalisation: http://delos.dis.uniroma1.it/
7. DELOS WP7 - evaluation workpackage: http://dlib.ionio.gr/wp7/
8. Initiative for the Evaluation of XML Retrieval 2004: http://inex.is.informatik.uni-duisburg.de:2004/
9. Cross Language Evaluation Forum: http://clef.iei.pi.cnr.it:2002/
10. CLEF 2004 workshop: http://clef.iei.pi.cnr.it:2002/2004/2004work.html


## Author Details


Michael Day
UKOLN
University of Bath
E-mail: m.day@ukoln.ac.uk
url: http://www.ukoln.ac.uk/


Return to top

DELOS Home | DELOS Newsletter Front Page | Delos Newsletter Contents

Publication Date: October 2004
Publication: Delos Newsletter Issue 2
Originating URL: /newsletter/issue2/infoday/intro.html
File last modified: Friday, 29-Oct-2004 16:31:56 CEST

The Delos Newsletter is published by the DELOS Network of Excellence
and is edited by Richard Waller of UKOLN, University of Bath, UK.

Newsletter
Issue 2 : October 2004



# Latest News from DELOS

Each issue of the Newsletter will carry the most recent news items from the DELOS website. The full listing will grow over time.



---

## Presentations now available from "Workshop on the Evaluation of Digital Libraries", 4-5 October 2004

Presentations from the DELOS Workshop on the Evaluation of Digital Libraries are now freely available. The "Workshop on the Evaluation of Digital Libraries", took place over 4-5 October 2004 in the Department of Information Engineering,University of Padova, Italy. Given the lack of appropriate evaluation methodologies in the area of digital library evaluation the workshop took as its aim a survey on the state of the art in DL evaluation, and sought to identify major issues for further research in this area.

Return to top

---

## Workshop on Audio-Visual Content and Information Visualization in Digital Libraries, May 2005

Workshop on Audio-Visual Content and Information Visualization in Digital Libraries, May 2005
The Delos research cluster 4 "User-Interface and visualization" announces the "Seventh International Workshop of the EU Network of Excellence DELOS on Audio-Visual Content and Information Visualization in Digital Libraries" (AVIVDiLib'05), 4-6 May 2005, in Cortona, Italy. A Call for Papers is available at the workshop web site. The workshop aims at providing a forum to present the latest research results, new technology development and new applications in the areas of multimedia content and information visualization in digital libraries.

This workshop, promoted by the DELOS Network of Excellence (http://www.delos.info/), jointly organized by Audio/Visual and Non-traditional Objects (WP3) and User Interfaces and Visualization (WP4), aims to provide a forum to present the latest research results, new developments in technology and new applications in the areas of multimedia content and information visualization in digital libraries.

The following non-restrictive list is given to illustrate items of particular interest for this conference:

- Interface Design
- Visual Filtering
- Visual Mining
- Metadata extraction

- Efficient data base structures
- Content-based retrieval and browsing
- Personalization support
- Audio, sound digital libraries
- Image, Video, 3D data digital libraries
- Multimedia digital libraries
- Large Data and High Dimensional Visualization
- Visualization in Non-traditional Devices
- Universal Accessibility

Return to top

---

## Presentations from "Semantic Approaches in Digital Libraries" workshop freely available

Presentations from "Semantic Approaches in Digital Libraries" workshop freely available
All presentations from the Delos Regional Awareness Event "Between Knowledge Organization and Semantic Web: Semantic Approaches in Digital Libraries", 23 June 2004 in Lund, Sweden, are now online.
They include:

- Boris Lauser: From thesauri to rich ontology: The AGROVOC case
- Boris Lauser: FAO Ontology Portal Prototypes: Fishery
- Boris Lauser: Building a rich ontology from AGROVOC
- Anders Ardö and Koraljka Golub: Automatic Subject Classification and Topic Specific Search Engines - Research at KnowLib
- Douglas Tudhope: Semantic Terminology Services: experiences from the FACET Project
- Alistair Miles: SKOS: Simple Knowledge Organisation with the Semantic Web
- Frida Sandgren: Markup of Educational Content
- Knut Hegna: Exploiting the classification scheme and a controlled vocabulary in a library catalogue

---

## On-line Questionnaire for Collecting User Requirements

One of the main objectives of Cluster 4 (User Interfaces and Visualization) is to collect and analyze user requirements in order to relate them to the different research perspectives and technical implementation options on a digital library. To this end, two integrated questionnaires (for end-users and stakeholders) have been set up for on-line collection and processing of data. Digital library end-users and stakeholders are warmly invited to complete them.

The questionnaires can be found at http://www.dis.uniroma1.it/~delos/questionnaires.

Return to top

DELOS Home | DELOS Newsletter Front Page | Delos Newsletter Contents

Publication Date: October 2004
Publication: Delos Newsletter Issue 2
Originating URL: /newsletter/issue2/news/intro.html
File last modified: Wednesday, 27-Oct-2004 21:03:10 CEST

The Delos Newsletter is published by the DELOS Network of Excellence
and is edited by Richard Waller of UKOLN, University of Bath, UK.