

[DELOS Home](#)[DELOS Newsletter Front Page](#)

## Contents of Issue 3

[\[PDF version of this issue\]](#)

### Feature Articles:

#### [Heterogeneity in Digital Libraries: Two Sides of the Same Coin](#)

Georgia Koutrika provides an overview of the challenges facing digital library design caused by the diversity both of data resources and users. She describes how the Information Access and Personalization cluster's user surveys are addressing this difficulty.

---

#### [Audio/Visual and Non-traditional Objects: Progress and Futures](#)

George Ioannidis provides us with a detailed view not only of A/V-NTO's current work and progress but also of the sort of new research the cluster is anticipating in the next phase of the Joint Programme of Activities. His article looks at research items such as:

- [video annotation with pictorially enriched ontologies](#),
- [multimedia interfaces for mobile applications](#),
- [description, matching and retrieval of 3D objects by content](#),
- [a variety of processes to handle news telecasts](#) and also
- [content- and context-aware multimedia content retrieval](#).

In addition there is information on the [DEMOS portal for demonstrators and testbeds](#) and material on [demonstrators currently available](#).

---

### Cluster Reports

All issues of the Newsletter will carry [reports from 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:

- [Digital Library Architecture \(DLA\)](#)
- [Information Access and Personalization \(IAP\)](#)
- [Audio/Visual and Non-traditional Objects \(A/V-NTO\)](#)
- [User Interfaces and Visualization \(UIV\)](#)
- [Knowledge Extraction and Semantic Interoperability \(KESI\)](#)
- [Evaluation \(EVAL\)](#)

## **Promoting Knowledge in the Fields of Information Access and Personalization: Past and Upcoming Events**

Georgia Koutrika collates reactions to past events organized by IAP and gives information on events planned for the future.

---

## **DELOS 9th Thematic Workshop: Digital Repositories**

Rachel Heery provides an overview of the 9th Thematic Workshop on Digital Repositories entitled Interoperability and Common Services.

---

## **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.

- [News page](#)

[Return to top](#)

[DELOS Home](#)

[DELOS Newsletter Front Page](#)

Publication Date: June 2005

File last modified: Tuesday, 12-Jul-2005 11:45:41 BST

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



[DELOS Home](#)

[Issue 1 - April 2004](#) ... [Issue 2 - October 2004](#) ... [Issue 3 - June 2005](#)

Welcome to the DELOS Newsletter. The role of the Newsletter is to provide members of the DELOS Network of Excellence with a channel for communicating to each other and the wider digital library community with news and information about their work. The Newsletter's readership, we hope, will be a broad one therefore, ranging from those researchers, developers and other colleagues working within the NoE, to their opposite numbers in other projects as well as information providers, collection managers, users and a host of other interested parties who all share the aims of the DELOS Project in one form or another.

The [current issue](#) carries [feature articles](#) from Cluster 2 (Information Access and Personalization (IAP)) and Cluster 3 (Audio/Visual and Non-traditional Objects (A/V-NTO) ) on aspects of their research. In addition there are [reports](#) from the other clusters in the Network. Each report carries the name and contact details of the leaders of the clusters should you have queries or comments. In addition each issue will carry a [news](#) selection.

I hope you will enjoy [Issue 3](#).  
The Editor

[Return to top](#)

[DELOS Home Page](#)

File last modified: Monday, 11-Jul-2005 15:42:36 BST

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





# Heterogeneity in Digital Libraries: Two Sides of the Same Coin

**Georgia Koutrika** provides us with an overview of the challenges facing digital library design caused by the diversity both of data resources and users and describes how IAP user surveys are addressing this difficulty.

## Introduction

Heterogeneity may be regarded as a benefit to digital libraries, but the truth is, it also represents something of a problem to developers as well. We come across it in a number of forms. Even at a very high level abstraction, we should consider data sources and users as two of the most fundamental constituents of a digital library. Accordingly, two basic types of heterogeneity are evident: data source heterogeneity and user heterogeneity. One cannot design and implement a digital library without considering these issues extremely carefully. As a consequence, their importance to digital library design has instigated the production of three separate surveys within the DELOS context.

## Data Source Heterogeneity

A digital library can be a vast collection of objects stored and maintained by multiple information sources, including databases, image banks, file systems, e-mail systems, the World Wide Web, and others. Therefore, assembling information of relevance on a specific topic involves searching for correct information items emanating from a wide variety of sources.

The issue of data source heterogeneity can represent significant problems when accessing multiple data sources. In effect it is the degree of dissimilarity between the component data sources that determines the amount of difficulty involved in implementing a data integration system. Data sources may differ in many ways. At a lower level, heterogeneity arises out of differing hardware platforms, operating systems, networking protocols and access interfaces. At the higher level, heterogeneity arises out of differences among different programming and data models as well as different perceptions and modelling of the same real world. Moreover, the fact remains that sources are evolutionary, i.e. where at one point they may be included on a system, there also comes a time when they are removed.

Four types of data source heterogeneity have been identified:

- **System heterogeneity:** arising from different hardware platforms and operating systems
- **Syntactic heterogeneity:** caused by discrepancies across the different protocols, encodings and languages used by the information sources (e.g. query languages, browsing interfaces, data formats, communication protocols and so forth)
- **Structural heterogeneity:** encountered among sources using different data models, data structures and schemas
- **Semantic heterogeneity:** produced by semantic conflicts arising from the fact that the meaning of the data can be expressed in different ways, as every metadata scheme defines its own set of data elements or categories for data

Consequently, there is a need to provide users with the capacity to access digital library objects both seamlessly and transparently despite the heterogeneity and dynamism across the various information sources involved. *Interoperable* information sources and services allow users to focus on information use instead of their being

obliged to acquire and combine the required content manually from the different sources.

Syntactic and structural interoperability supports the handling, exchange and combining of data properly, having proper regard to formats, encodings, properties, values, data types and so forth. A *data integration system* is one that provides users with transparent access to a collection of related data sources as if these sources, as a whole, constitute a single data source. The main objective of a data integration system is to facilitate users' attempts to focus on specifying *what* data they want, rather than on describing *how* to obtain it. To achieve this, the system provides an integrated view of the data stored in the underlying data sources. In a data integration system, users are interested mainly in querying the integrated data rather than updating the data through the integrated view. It is therefore something of an understatement to suggest that heterogeneous data sources invariably present designers of data integration systems with a raft of challenging difficulties.

## The Data Integration Services Survey

Given such challenges, the aim of the **Data Integration Services Survey** is thoroughly to describe and compare the different approaches, schemes, frameworks and systems mentioned in the current literature on supporting information integration from *structurally heterogeneous* sources. This is a survey on the following data source description approaches:

- GAV (Global As View)
- LAV (Local As View)
- GLAV (Combining LAV & GAV)
- as well as related query writing algorithms and systems

## Semantic Interoperability

Semantic interoperability, on the other hand, allows users to negotiate and understand the meaning of the metadata items both in the same application domain and between application domains. Semantic interoperability refers to the extent to which different metadata schemes express the same semantics in their categorization. Successful interoperation requires clarity on how the categories of metadata relate to each other across different schemes. To this end, several questions must be answered:

- When do elements have the same meaning?
- When elements are derivatives, subsets, or variations of each other?
- When elements are completely unrelated?

Furthermore, different application domains have established different metadata standards, making the interoperation of applications from different domains a tricky task. The problem becomes even more complicated when a vast body of standards already exists for the same application domain.

## Semantic Interoperability Survey

The state of the art survey on Semantic Interoperability in Digital Libraries focuses on semantic interoperability issues, and in particular on:

- domain-specific metadata models
- models used for audiovisual content description
- models for the description of items in the cultural heritage domain (including the holdings of archives, libraries and museums)
- metadata schema interoperability resolution of semantics using taxonomies, thesauri and ontologies

## User Heterogeneity

On the other hand, Internet access has resulted in digital libraries being increasingly used by diverse communities for a variety of purposes; among these sharing and collaboration have become important social elements. In addition, a user's information-seeking activities are no longer bound, neither geographically nor temporally. Information access can be achieved through a variety of devices from users' offices, homes, hotel rooms or even on the move, at any time of the day or night, seven days a week. As a result, information systems are seeing far greater use. More importantly still, the kind of people doing so now range well beyond librarians or scientists, as

was once the case.

User Heterogeneity is, hence, a significant problem for digital libraries. Users have ever more complex needs and different users have differing requirements. At the same time, users want to achieve their goals with a minimum of cognitive load and as much enjoyment as possible. Furthermore, we must factor in the matter of information overload which fuels the need for more sophisticated and user-centered services which can provide access to the content of digital libraries. Individuals as much as groups of users have to be better supported if they are to capture, structure and share knowledge successfully. Furthermore, in the same context, both formal and informal learning activity requires similar support.

## **Personalization**

An integral step towards these ends lies in building effective profiles of their users. A user profile is an appropriate description of the user, created manually by either the user, or automatically by the system. It is used by the system during its interaction with digital library users in order to anticipate their needs and satisfy them in the best possible way. This is achieved by adapting presentation, content, and services based on a person's task, background, history, device, information needs, location, and so forth, as dictated by the user profile. Digital libraries which fail to meet the personalization requirements posed by their users will ultimately find it difficult to retain their user base or indeed attract new users.

Therefore, this has led to the development of personalization systems which adapt their behaviour to the goals, interests, and other characteristics of their users, either as individuals or as members of particular groups.

Central to all personalization systems is the issue of user profile representation. This provides the means to record the user's preferences and status and so filter the content retrieved, personalize the services offered as well as track user access behaviour and needs. However the construction of user profiles can represent considerable effort which remains largely invisible to the layman.

The aim of the **User Modeling for Personalization in Digital Libraries Survey** is to study user profiling in Information Retrieval and Information Filtering. It describes different user profile representations, such as history-based, vector space model, weighted n-grams, and classifier-based profiles, explicit and implicit methods for user profile acquisition, user context, existing standards and models, and user profile management in major commercial systems and research projects.

User profiles can be used in a variety of ways to individualize user experience which means of course that approaches to personalization also differ. However it has been commonly observed that the largest proportion of research derives from the Information Retrieval community, with that of the Database community next most in evidence, in many cases inspired by Institutional Repositories (IR).

The **Profile Usage for Personalization in Digital Libraries Survey** covers personalization methods proposed in the IR and Database communities. It describes information filtering, continuous queries, recommender systems and personalized search engines.

## **Other Vital Work**

Heterogeneity is by no means the only issue to consider in digital library design. During the first year of work, the IAP cluster has been drafting a set of comprehensive surveys and reports on other key relevant areas of interest to provide broad overviews of existing models and approaches as well as identify problems. These surveys formed the basis for establishing common approaches on information access, information integration and personalization; they were also instrumental in initiating joint research in a number of the aforementioned areas.

Apart from the surveys mentioned above, other surveys already in draft relate to the following topics:

- Information Access Models and Modes
- Metadata in the Context of DL
- Peer-to-Peer Data Management Systems
- Data Annotation and Provenance in Large Scale InformationIntegration Systems

Work carried out on the formulation of these surveys has served to identify major themes in research on both

information access and personalization, as follows:

- Information access: data indexing for complex similarity measures
- Information integration: query processing and routing in P2P architectures
- Personalization: modelling of user preferences and more general contexts

The surveys are available from the Information Access and Personalization cluster website:

<http://delos.di.uoa.gr/transactions.php?type=Reports>

## Author Details

### **Georgia Koutrika**

University of Athens

Email: [koutrika@di.uoa.gr](mailto:koutrika@di.uoa.gr)

Telephone: +30 210 727 5242

Fax: +30 210 727 5214

[Return to top](#)

[DELOS Home](#)

[DELOS Newsletter Front Page](#)

[Delos Newsletter Contents](#)

Publication Date: June 2005

Publication: Delos Newsletter Issue 3

Originating URL: /projects/delos/newsletter/issue3/feature2/intro.html

File last modified: Tuesday, 12-Jul-2005 13:57:24 BST

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



## Audio/Visual and Non-traditional Objects: Progress and Futures

**George Ioannidis** provides us with a detailed view not only of A/V-NTO's current work and progress but also of the sort of new research the cluster is anticipating in the next phase of the Joint Programme of Activities.

### [1. Current Status of the work](#)

#### [2. New research activities anticipated](#)

##### [2.1. Video Annotation with Pictorially Enriched Ontologies](#)

##### [2.2. Multimedia Interfaces for Mobile Applications](#)

##### [2.3. Description, Matching, and Retrieval By Content of 3D Objects](#)

##### [2.4. Automatic, Context-of-Capture Based, Categorization, Structure Detection and Segmentation of News Telecasts](#)

##### [2.5. Content and Context Aware Multimedia Content Retrieval, Delivery and Presentation](#)

##### [2.6. Natural Language and Speech Interfaces to Knowledge Repositories](#)

#### [3. DEMOS portal for demonstrators and testbeds](#)

##### [3.1. DEMOS Content Browser](#)

##### [3.2. DEMOS Content Manager](#)

#### [4. Available demonstrators](#)

##### [4.1 MILOS - Multimedia Content Management System](#)

##### [4.2. 3D Content-Based Retrieval](#)

##### [4.3. VideoBrowse](#)

##### [4.4. UvA Parallel Visual Analysis in TRECVID 2004](#)

##### [4.5. Audio feature extraction with Rhythm Patterns](#)

##### [4.6. TZI Demonstrators for Delos WP3](#)

##### [4.7. The Video Segmentation and Annotation Tool Demonstrator](#)

##### [4.8. The UP-TV Demonstrator](#)

##### [4.9. The Campiello Demonstrator](#)

#### [5. References](#)

##### [5.1 Publications](#)

## 1. Current Status of the work

### Introduction

Over the first 12 months of the project WP3 aimed to develop a common understanding and foundation for the work that has to be done in DELOS in terms of State of the Art Reports, support for Forum and Testbeds, and efforts at understanding the expertise of the partners and their possible cooperation towards the objectives of DELOS as they are described in the Technical Annex.

### Progress on Reports

The reports entitled *State of the Art on Metadata Extraction and State of the Art in Audiovisual Content-Based Retrieval, Information Universal Access & Interaction including Data Models & Languages* have been completed. A preliminary draft of the state of the art report in *Audiovisual Metadata Management* has been produced.

### Portals and Demonstrators



The Delos Collaborative Portal has been released. The portal is intended to foster exchange of ideas and useful information within the DELOS Community. It includes news, a discussion forum, and a calendar. Administrative access to the system has been granted to all partners, in order to allow decentralized content management of the system.

Based on an analysis of the requirements for supporting testbeds and demonstrators, the DEMOS portal for demonstrators and testbeds has been created. The DEMOS portal is described in further detail in [Section 3](#) of the feature. Several demonstrators have already been ingested, some of which are described later in [Section 4](#) of the feature. Some testbeds have also been provided. They include images and video segments from various sources, e.g. soccer and swimming videos with manual ground truthing of events. They will not be described here, but may be accessed through the DEMOS portal.

## **Metadata-related Activity**

For ontology-based metadata definition, a tool named GraphOnto has been implemented. An OWL upper ontology that captures the MPEG7 MDS is utilized. This upper ontology is extended with domain knowledge through appropriate OWL domain Ontologies. The component provides a graphical user interface for interactive ontology browsing and definition of OWL RDF metadata. The component also provides functionality for exporting the metadata into MPEG7-compliant XML documents. A set of transformation rules from OWL RDF metadata to MPEG7 and TV-Anytime compliant metadata completes the tool. In the same context, a study for the integration of the TV-Anytime Metadata model with the SCORM 1.2 Content Aggregation Model has been completed that defines a detailed mapping between the two metadata standards. This mapping allows for the provision of eLearning services on digital TV systems as well as the reuse of TV programs in order to build educational experiences. This is considered as an essential infrastructure for digital libraries of audiovisual content that conforms to the TV-Anytime metadata specifications in order to support eLearning services.

## **MPEG-7-related Work**

An analysis of the applicability of MPEG-7 descriptors to the existing video annotation tools that are based on home-grown XML annotation formats was carried out.

Based on MPEG-7, a modelling language for magazine broadcasts has been specified. It is capable of describing classes of telecasts, instead of specific telecast instances, for automatic segmentation into semantic structural elements.

A Java class framework has been implemented for the modelling of MPEG-7 descriptions (MDS, Video, Audio). These can be stored in an implemented persistence management framework for media descriptors.

## **Other Developments**

An automated image classifier based on SVM techniques has been designed and realized. An automatic region grouping method for improving semantic meaning of features using psychology laws has been developed. The classifier has been integrated in the MILOS Content Management System, which is also available as a demonstrator through the DEMOS portal. It is described in [Section 4.1](#) of the main feature.

For video analysis, annotation, and retrieval, a prototype video content management system, named VCM, has been developed. It is available through the DEMOS demonstrator portal, and is described in [Section 4.6](#).

A multimedia authoring tool has been defined, which supports content-based constraints for personalizing the presentation of multimedia objects according to users' preferences and skill level.

A prototype system was developed to explore the multimedia content of a digital library (images, text, videos, and audio) relating to theatrical works in 19th Century Milan and which supplies a VR (Virtual Reality) interface (namely, a reconstruction of a 19th Century Milanese theatre).

A front-end of a music search engine has been developed, which is accessible through a web browser to allow users to interact using a query-by-example paradigm. Moreover the typical query-by-humming paradigm is also supported. A preliminary version of a component for semi-automatic extraction of song metadata (title, lyrics,

cover) from ID3-tags and by querying via web services has also been created. Methodologies for music indexing and retrieval have been extensively evaluated, based on a data fusion approach, with encouraging initial results.

Preliminary tests on the use of APIs provided by Web-based CD dealers were made to examine the potential of automatic creation of a network of composers/performers with scope for extracting information about their similarities, and reflecting to customers' behaviour.

Feature extraction systems for audio content, named Marsyas and SOMEJB, have been installed and tested. Evaluation measures on a larger sample collection based on audio files have been collected and will subsequently be used to define scenarios for interactive retrieval and evaluation of retrieval performance in different scenarios.

An audio classification framework for the participation in the International Conference on Music Information Retrieval (ISMIR) audio contests in the disciplines of Rhythm Genre and Artist detection, has been implemented. It was awarded winner of the Rhythm Classification Competition, was ranked fourth in the genre classification contest, and was again winner in the "stress-test" performance of the genre classification contest. A corresponding demonstrator is available through the DEMOS portal. It is described in more detail in [Section 4.5](#).

A web crawler, which is based on APIs provided by a major Web Search Engine, has been developed to create a collection of MIDI files automatically, to be used as a testbed for Music Information Retrieval techniques. When launched, the crawler is able to collect and store thousands of MIDI files in a database, partially overcoming the classic problem of lack of test data.

A syllable-based speech recognition engine for English has been developed. A speech recognizer named ISIP was trained with huge amounts of American English broadcast data. Hidden-Markov-Models were used forming context-dependent cross-word-triphone models. The syllable inventory was generated using tools from NIST. The syllable recognition rate is 88.0%. A syllable retrieval system could be implemented with the syllable recognizer, similar to what has been done for German.

## **NIST TRECVID Evaluation**

Delos members participated in the 2004 NIST TRECVID evaluation - the de facto international standard benchmark for content-based video retrieval. Members participated in the feature extraction task, the shot detection task, and the search task. For the latter task the UvA TRECVID Semantic Video Search Engine was developed, showing the effectiveness of the approaches to content-based retrieval by audio-visual libraries, as well as the parallel implementation thereof. The Semantic Video Search Engine is described in the feature, [Section 4.4](#), and is accessible through the DEMOS portal. The shot detection algorithms implemented for TRECVID participation are also available through the portal. They are referred to in [Section 4.6](#).

## **Other Advances**

Several software components have been continuously refined. These include software for 3D objects modelling and retrieval, as well as tools for MPEG-7 manual annotation of videos and real-time automatic video annotation, in particular for soccer video analysis. Further improvements have been done on automatic audio-visual metadata extraction tools.

Advances have been made with the development of a test-bed and demonstrator for the extraction and integration of most of the MPEG-7 standard visual descriptors. The output of the demonstrator is collected in an MPEG-7 stream and testing on the interoperability is being analyzed.

Other work has included the following:

- Improvements of the ISIS/OSIRIS system for easier DL maintenance and deployment were made. An automatic/dynamic process will take care of visual feature extraction within ISIS.
- Issues relating to the computational requirements and parallelization of emerging applications in the field of audio-visual digital libraries have been investigated, as well as issues relating to the automatic detection of semantic concepts in multi-modal video repositories.
- Various music information retrieval frameworks have been set up and music retrieval performance on benchmark datasets has been evaluated.
- A study of a model for the specification of synchronized multimedia presentations and of methods for automatic and semi-automatic presentation generation has been started.

Documents from public forums, relating to DLs and describing technological innovation and available prototypes, are collected. These are in the process of being catalogued and indexed to provide fast access to public knowledge.

## 2. New research activities anticipated

The efforts of the next 18 months will build on the existing infrastructure, experience gained, and cooperation established. In particular, the building up of cooperation among partners and the establishment of common foundations will continue through the use and expansion of the functionality and contents of the Forum and the Testbed infrastructure. The long-term research activities in the Delos Technical Annex for WP3 cite as objectives:

- Metadata Capturing for Audio-Visual Content
- Universal Access and Interactions with Audio-Visual Libraries
- Management of the Audio-visual Content

The new WP3 tasks scheduled all fall within those three objectives. In particular the six tasks listed under WP3, along with some other tasks (listed below) in which WP3 members participate, (and which are partially overlapping with the above WP3 objectives) cover the three main objectives of the cluster as follows:

1. Metadata Capturing for Audio-Visual Content:
  - Video Annotation with Pictorially Enriched Ontologies
  - Automatic, Context of Capture-Based Categorization
  - Structure Detection and Segmentation of News Telecasts, and partially
  - Multimedia Interfaces for Mobile Applications (all WP3)
2. Universal Access and Interactions with Audio-Visual Libraries:
  - Content and Context-Aware Multimedia Content Retrieval, Delivery and Presentation
  - Description, Matching and Retrieval by Content of 3D Objects and
  - Natural Language and Speech Interfaces to Knowledge Repositories (all WP3)
3. Management of the Audio-Visual Content:
  - Advanced Access Structures for Similarity Measures (WP2)
  - Interoperability of e-Learning Applications with Digital Libraries (WP5), and partially
  - Ontology-Driven Interoperability (WP5)

To achieve these goals, the following Tasks are planned.

### 2.1. Video Annotation with Pictorially Enriched Ontologies

To support effective use of video information, and to cater for ever-changing user requirements, tools for accessing video information are essential. Access must be at a semantic level rather than a technical level as the librarian and the user will not connect the two. Semantic indexes must therefore be as rich and complete as possible.

The ultimate goal of this Task is to *automatically extract high-level knowledge from video data, permitting the automatic annotation of videos*. In order to obtain effective annotation (both in the manual and automatic cases), one must rely on a domain-specific ontology defined by domain experts. The ontology is typically defined by means of a set of linguistic terms capable of describing high-level concepts and their relationships. However, it is often difficult to describe appropriately all interesting highlights purely in terms of (a set of) concepts. Particularly in sport videos, while we can use concepts appropriately to describe basic types of highlights, like goal, counterattack, etc., it must be recognized that each one might occur in multiple contexts, each of which will be worthy of its own individual description. Distinguishing subclasses of these occurrences which group together instances that share the same or similar spatio-temporal characteristics will have to be identified.

The linguistic terms of an ontology are too vague to take effective account of the distinguishing features of these subclasses of spatio-temporal events. Therefore, this Task aims at defining methodologies and techniques to describe concepts and their specializations by *augmenting an ontology of linguistic terms with "visual concepts" that represent these instances in a visual form*. The visual concepts should be learned from occurrences of the highlights through analysis of their similarity (in the spatio-temporal domain) and automatically extracted from both raw and edited videos and integrated into the ontology.

The end result is a *pictorially enriched ontology* (PE-Ontology) that fully supports video annotation, allowing classification and annotation of events up to very specialized levels. Visual concepts, once added to the ontology,

will integrate the semantics described through linguistic terms up to a more detailed representation of the context domain. Visual concepts will be defined by means of global features, meaningful spatial segments (such as regions of frames or key-frames) and temporal segments (as highlights or representative shots). The PE-Ontology will thus be both the support for segmentation and annotation and will represent an efficient approach to handling both summarization and effective access to multimedia data guided by semantics, and in accordance with users' interests.

The Task aims to analyze (A) methodological and (B) implementation aspects of the problem and in particular will seek:

- A1: to define *linguistic* ontologies for specific sport domains, namely Soccer and Formula One motor racing; to define the framework to support *pictorially enriched ontology*. Support for graphically linking pictorial annotations to domain ontology concepts will be provided.
- A2: to identify *distinguishing features*, providing quantitative descriptions of visual concepts;
- A3: to define video analysis and pattern recognition solutions to extract visual concepts, perform their clustering so as to extract prototypes (cluster centres), and add these prototypes to the PE-Ontology. In particular we want to analyze both static and spatio-temporal visual concepts, using both raw visual features (embodied in the video's scenes) and visual features of edit effects.
- B: to address *implementation aspects*, covering the design with automatic annotation and summarization engines, the integration into standards, the prototype development for sport digital libraries and the evaluation of related computational aspects.

This task is a continuation of activities started during the previous Joint Programme of Activities (JPA) for the production of a toolkit of algorithms for metadata extraction and test beds, and in particular is a continuation of Tasks 3, 4 and 5 of WP3. (See "Cluster Activities" under [Audio/Visual and Non-traditional Objects](#) in Issue 1.)

## 2.2. Multimedia Interfaces for Mobile Applications

This Task intends to investigate several strictly interrelated sub problems, producing results in the framework of multimedia access for video presentation on mobile devices. This Task will be conducted in cooperation with cluster 4- UIV (User Interfaces and Visualization). The main subjects of investigation will be:

1. Automatic video extraction of meaningful objects and events according to user's interests;
2. User profiling and design of flexible small screen device interface, able to minimize the user interaction and adapt to devices' characteristics;
3. Performance measures and quantitative/qualitative indexes of user experience and satisfaction.

The Task aims to develop a prototype system composed of three subsystems: *Video Annotation*, *Video Summarization*, and *User Interface*. The anticipated field of application is transmission of sports and news video, enhanced by video summaries.

### Video Annotation

Off-line annotation takes place on uncompressed video, producing a more precise annotation, extracting highlights and significant objects/events. Highlights must be represented with appropriate knowledge models based on the a-priori knowledge of the spatial-temporal structure of events and recognized by a model checking engine, based on statistical or model-based classification frameworks. Image processing and analysis is used to extract the salient features of the video such as motion vectors (that quantify the activity), color patterns (that distinguish background zones), lines, corners and shapes (that identify objects). Text appearing on the video can be extracted and recognized. Players' position in the playground can be detected in order to build statistics of the field occupancy.

### Video Summarization

This subsystem manages the construction of video summaries upon user's request. Summaries are obtained dynamically, combining the user request with the annotations obtained from the off-line annotation process.

### User Interface

The User Interface subsystem is in charge of handling the interaction with the user, and is faced with two main

objectives:

- it should nicely fit the device characteristics and the user preferences;
- it should include new interaction and visualization techniques to convey effectively the information produced by the annotation and video summarization systems.

The overall goals of this task will also be accomplished by integrating contributions from WP4 (UIV), which in particular will contribute to the development of the User Interface subsystem.

## **2.3. Description, Matching, and Retrieval By Content of 3D Objects**

The goal of this Task is to develop a system to support structural as well as view-based retrieval of 3D objects by content. In this context, the Task aims to investigate models for the extraction of view-based and structurally based descriptors of 3D objects, models for indexing and similarity matching of structural and view-based descriptors as well as models and metaphors for querying archives of 3D objects. The theoretical investigation of these models will lead to the design and development of a prototype system. In particular, Task activities will address the following issues:

### **Task 1: Content Description**

Models will be investigated and trialled to extract descriptors of 3D object content from multiple viewpoints. These descriptors should capture prominent features of object views so as to enable retrieval by similarity, based on a single photograph of an object taken from a generic viewpoint. Descriptors of object views should also account for the object's visual appearance in terms of colour and texture features. Models designed for the extraction of 3D object structure will also be investigated and tested. To this end, 3D object segmentation techniques will be developed so as to allow decomposition of a 3D object into its structural components. Each component will be described separately so as to enable description and retrieval based on the characteristics of object parts in addition to global object features.

### **Task 2: Indexing and Similarity Matching**

For both descriptors of object views and object structure, a distance measure should be defined to permit computation - on a perceptual basis - of the similarity between a generic 3D object and a template, the latter being represented either as the image of an object from a particular viewpoint or as a compound set of 3D parts. For the definition of this distance measure, specific constraints should be considered in order to allow combination of the similarity matching process with a suitable index structure that provides efficient access to database content.

### **Task 3: Querying and Presentation**

Despite its wide use to support access by content to image libraries, the query by example paradigm, in its original form (pick one item from the archive and retrieve similar items), exhibits certain limitations when applied to libraries of 3D objects. This is particularly true in the context of this Task where retrieval based on an object photograph (image) and retrieval based on object components are addressed. The former requires the definition of models to manage specification of the query through an external image (representing one view of the object of interest). The latter relies on the user's option to select a subset of the structural components of an archived object and use them only to retrieve objects with similar components in a similar arrangement.

This task is a continuation of activities already described under WP3 in the previous JPA, in particular with the development of a toolkit of algorithms for metadata extraction and test beds.

## **2.4. Automatic, Context-of-Capture-Based Categorization, Structure Detection and Segmentation of News Telecasts**

Context in general is a state. The state of the discussion loosely means what has been discussed and understood by both parties in that discussion. It also reflects the specific subject of the discussion at a certain point in time. Therefore context can be organized into abstraction hierarchies. In general we assume that a particular context is characterized by a set of interrelated concepts described in an ontology. A Context-of-Capture (CoC) may be

inferred by the set of words that appear in a discourse. Knowing the CoC of a discourse, we may be able to do a better job of recognizing what is said in that discourse.

These days automatic audiovisual content segmentation is performed in several systems, mainly at the syntactic level. Only a few systems take into account the semantics of audiovisual content. Furthermore, the CoC concept, which represents the context of information captured in an audiovisual segment (e.g. persons, places, events etc.), is either completely ignored or only superficially utilized. In addition, the CoC supports the automatic assignment of the audiovisual segments detected to appropriate thematic categories since the CoC of a segment contains sufficient information for the determination of the correct thematic category. Alongside the recognition of a specific context for segmentation and indexing purposes, we should recognise the importance of the potential of linking all relevant elements in the knowledge base to a context.

The above necessitates generic models for describing CoC and scenarios of context appearance as well as their use in recognition, segmentation and structuring of the knowledge bases so that complex queries can be answered.

The objective of this Task is to develop a demonstrator for automatic categorization, structure detection and segmentation of news telecasts that uses advanced structural models. Segment boundary detection will be assisted by a powerful CoC model to be used by the appropriate context detection and context change evaluation mechanisms. The segmentation/structural metadata will ultimately be exported in MPEG-7 format. A query API and user interface will be provided in order to evaluate the results. In particular, Task activities will address the following issues:

- **Definition of the Context-of-Capture (CoC) model:** A powerful model for the Context-of-Capture (CoC) and of CoC scenarios will be developed, as well as algorithms for using them for identification and knowledge management for the CoC including its use of recognition and inference.
- **Development of CoC recognition mechanisms:** Appropriate algorithms will be developed for CoC recognition, utilizing image, speech and video text processing for audiovisual feature extraction. Simple audiovisual cues (characteristic colour, texture, loudness), extraction of text inserts and higher-level visual features (e.g. faces, indoor/outdoor) will be taken into account. In addition, semantic concepts will be identified using keyword-spotting techniques on the speech signal.
- **Development of mechanisms for CoC-based segmentation and categorization of telecasts:** A rough syntactic segmentation can be obtained using algorithms for both shot detection on the visual signal as well as speaker and speech/music recognition from the audio signal. Thereafter the sudden changes in the CoC (denoting the segment boundaries) will be used to refine the segmentation. The refinement mainly refers to merging the adjacent syntactic segments with very similar CoC.
- **Development of a query API and user interface for evaluation:** A query API and user interface will be provided in order to evaluate the results.

This task continues with the activities already described in WP3 of the previous JPA, and in particular is a continuation of Metadata Capturing for Audio-Visual Content, Management of Audio-Visual Content in Digital Libraries, and Development of Demonstrators and Testbeds.

## 2.5. Content- and Context-aware Multimedia Content Retrieval, Delivery and Presentation

This Task focuses on the *integration of content-based multimedia retrieval in digital libraries* and the *delivery and consumption* of the retrieved multimedia data. It aims to provide users of digital library systems with a solution for intelligent retrieval in large media collections where visualization of the retrieval process results, media transport and presentation of results are based on adaptation to user preferences.

User preferences can be encapsulated in the MPEG-7 Multimedia Description Schemes (MDS) *User Preferences* descriptor. Unfortunately, this descriptor provides only basic information. Hence, this Task will enrich it by CC/PP profiles (based on RDF descriptions and OWL/RDF ontologies). The user profiles defined in the MPEG-7 MDS, although structured, do not allow for the description of user preferences that take semantic entities into account. Thus, the MPEG-7 MDS user profiles should be enriched, utilizing OWL ontologies as well as the constructs provided by CC/PP profiles and MPEG-21.

Another issue which will be addressed by this Task concerns the personalization of the presentation's content-(or semantic)-based flow and duration with respect to the interests and skills of the end-users. Multiple execution flows, with possibly different duration, for the same multimedia presentation will be provided.

In addition the Task aims to deliver multimedia content that is targeted at a specific person and reflects this person's individual context-specific background, interests and knowledge, as well as the heterogeneous infrastructure of end-user devices to which the content is delivered and presented. Therefore, the multimedia content is selected based on the user profile, adapted to the user's context and assembled into a multimedia composition.

The proposed architecture will allow the integration of content-based retrieval, content adaptation and multimedia presentation delivery. It will use:

- The MM4U (multimedia for you) framework (OFFIS) a generic and modular framework that supports multimedia content personalization applications
- The KoMMA framework (TU Wien), which is used for content adaptation. It is designed as a set of Java APIs which are responsible for the handling of metadata, adaptation decision taking, and the adaptation process itself
- The VizIR framework (TU Wien), employed for metadata extraction and modelling, comprises an growing set of Java classes for media access, content-based metadata extraction (e.g. most MPEG-7 descriptors), media annotation (e.g. the entire MPEG-7 MDS), query formulation and user interface design
- The multimedia authoring system developed by UNIMI which supports constraints on personalizing the presentation of multimedia objects according to users' preferences and skill levels
- The components and methodologies developed in the context of the DS-MIRF framework which comprise:
  - (a) a core OWL ontology, which fully covers the MPEG-7 MDS
  - (b) a methodology for the definition of domain-specific ontologies that extend the core ontology, in order fully to describe the concepts of specific application domains and
  - (c) transformation rules, used for the transformation of semantic metadata (formed according to the core ontology and its domain-specific extensions) to MPEG-7 compliant metadata

The interfaces of the components listed above will be harmonized, so as to provide an integrated toolkit for content-based retrieval, content adaptation and multimedia presentation delivery.

The overall goals of this task will also be accomplished by integrating contributions from WP2 (IAP).

This task continues with the activities already described in the previous JPA, namely:

- WP2 (IAP): Common Foundation on Personalization and Development of Prototypes.
- WP3 (A/V-NTO): Universal Access and Interactions with Audio-Visual Libraries, Management of Audio-Visual Content in Digital Libraries, Demonstrators and Testbeds.

## **2.6. Natural Language and Speech Interfaces to Knowledge Repositories**

The objective of this Task is to provide principles, methodologies and software for the automation of the construction of natural language and speech interfaces to knowledge repositories. These interfaces include the capacity to declare and manipulate new knowledge, as well as support for querying, filtering and ontology-driven interaction formulation. We will also provide a specific application demonstrator of natural language and speech interfaces to knowledge repositories.

The overall technical objective is to automate as much as possible the construction of natural language interfaces to knowledge bases. It has been shown that the overhead of developing natural language interfaces to information systems from scratch is a major obstacle for the deployment of such interfaces. In this design we do not specify what the storage structure for the metadata is. The metadata could be stored in a knowledge repository (such as an RDF repository) or they could be stored in relational systems provided that the inference mechanisms that support the knowledge manipulation language have been built on top of them. In addition to the concept (domain) ontologies, the natural language system will also have to accommodate word ontologies (like WordNet) and the interface between the two.

The Task will investigate the theoretical basis of the proposed approach which employs the domain ontologies to find how a user query in natural language can be converted to an (expanded) query in the knowledge manipulation language using the user profile and context, and allowing for the ranking of the results instead of disambiguation dialogues.

A speech recognizer takes as input a vocabulary produced by the natural language interface subsystem that includes words representing the concepts of the domain ontologies and their relationships with the word ontologies. It uses this input to convert the speech input or a user interaction to possible phrases in natural language. The natural language phrase is processed using the user context and profile as described above for disambiguation and ranking of the results from the knowledge base.

The interaction of the Natural Language Interfaces (NLI) sub-system with the knowledge manipulation language will be based on general query templates. In particular, Task activities will address the following issues:

- Development of a generic framework, as well as the theoretical foundations for the proposed approach, as described above.
- Design of an architecture and construction of a prototype system based on this foundation which automates as much as possible the implementation of Natural Language Interfaces to Knowledge Management Systems.
- Use of the system to define a specific ontology that may include higher-level (like MPEG-7) and domain-specific ontologies and the building of a Natural Language Interface for a particular application environment.
- Investigation of the interplay of speech recognition with the natural language interfaces and the ontology/user profile approach in a realistic application device (for example handheld devices).
- Evaluation of the approach using usability engineering principles with a user population and suggestions for possible improvements.

The overall goals of this task will also be accomplished by integrating contributions from WP2 (IAP), WP3 (A/V-NTO) and WP7 (EVAL).

This task continues with the activities already described in the previous Joint Programme of Activities, namely:

- WP2: Information Access and Personalization, Development of Prototypes.
- WP3: Metadata Capturing for Audio-Visual Content; Universal Access and Interactions with Audio-Visual Libraries; Management of Audio-Visual Content in Digital Libraries; Development of Advanced Multimedia Demonstrators and Test Datasets
- WP4: User Interface and Visualization Design; Context Consideration and Exploitation; Systematic Analysis of User Requirements; Development of a User Interface Design Framework.
- WP7: Development of DL Evaluation Methods; Prototype Evaluation Studies; DL Evaluation Testbeds and Toolkits.

### **3. DEMOS portal for demonstrators and testbeds**

DEMOS is an Information System for Demonstrators and Testbeds in Audiovisual and Non-Traditional Objects Digital Libraries. It has been built to maintain and disseminate demonstrators and testbeds proven or likely to be proven of relevant importance to the Audio/Visual Digital Library field. However, the design of the system is general enough to accommodate demonstrators and testbeds addressed to the digital library research community in general.

All information resides in a relational database implemented using an open source RDBMS. The database is divided into three main sections:

1. Demonstrators
2. Testbeds
3. Resources

The above parts of the database provide users with facilities to insert, access, and give comments on demonstrators, testbeds, and other resources useful for the description of demonstrators and testbeds (e.g. scientific publications, technical reports, user manuals, etc.). A search facility is also available allowing users to search for information based on specific parameters or using classification hierarchies that are based on part of the ACM 1998 computing classification system. User feedback is gathered through comments that the end-users insert with respect to specific demonstrators, testbeds or other resources made available by the system.

The users of the system can be categorized into two classes:

- End-users who use the DEMOS Content Browser to retrieve information about demonstrators, testbeds, and other resources



- Content providers who use the DEMOS Content Manager to insert information about demonstrators, testbeds, and other resources

### **3.1. DEMOS Content Browser**

The DEMOS Content Browser provides detailed information about the contents of the Delos WP3 portal database to the digital library community. It has been developed to maintain and support many resource types such as software demonstrations and testbeds, publications, reports, presentations etc. that have been developed or which are being developed by partners in the context of DELOS NoE.

#### **Resource Categories**

Each resource type can be classified into specific categories or classes. The web user can access the full description of each resource by browsing the specific categories of the resource type or by using the search utility of the content browser.

For the resource type "demonstration" the following categories have been specified:

- Digital libraries
- Digital library applications
- Digital library services
- Automated feature extraction tools
- Efficient and effective search tools
- Metadata repository tools
- Document repository tools

Demonstrators can be on-line or off-line. In the first instance a link to the demonstration is provided and can be immediately used. In the second, a link to a file is provided where users can download and install it locally on their workstation.

For the resource type "testbed" the following categories have been specified:

- Image datasets
- Video datasets
- ED graphics datasets
- Audio datasets
- Music datasets
- Multimedia objects datasets
- Multimedia digital library corpora

For the resource types "publications", "reports" and "presentations" the ACM 1998 classification system for Digital Libraries has been adopted.

#### **Search Utility**

Alternatively, the search utility of the content browser enables web users to find a specific resource by adding some keywords and specifying the resources type(s).

### **3.2. DEMOS Content Manager**

The DEMOS Content Manager is a web application that enables content providers (essentially DELOS members) to insert their content about demonstrators and testbeds into the DEMOS database.

The procedure for adding a new resource item in DEMOS occurs in three separate steps:

- Insert the attributes, the category and the keywords of the resource item
- Insert the persons participated in the resource item and their roles
- Determine the related resources items with the resource

#### **Step 1: Insert the attributes, the category and the keywords of the resource**

In the first step the user can insert the main attributes of a resource item: The *Name*, the *Abstract*, and the *URL*. There are also some extra attributes, customizable for each resource that are determined by the content administrator. For example the extra attributes that have been determined for the resource <<Demonstrator>> are: the <<Demonstration>> (the URL to the demo), the <<Release Date>> (the release date of the demonstrator) and the <<Version>> (the version of the demonstrator).

The user must classify the resource item in a *Category* by selecting one category item from a category hierarchy list.

Finally, the user can select a list of keywords that describe the content of the resource item.

## **Step 2: Insert the persons involved in the resource and their roles**

In every resource item there is a group of persons that has participated in some way (as authors or reviewers in a publication, as developers in a software component or a demo etc.). In this step the user can select the persons involved in the resource item and the role of these participants.

Different person roles have been specified by the content manager for each resource type. For example, for the resource type <<Publications>> the person roles specified are the <<Authors>> and the <<Reviewers>>, for the <<Demonstrators>> there are the <<Creators>>, the <<Designers>> and the <<Developers>> etc.

## **Step 3: Determine the related resources**

The content administration manager provides the capability to create relationships between the different resource types. For example, the relationship <<publish to>> describes where a publication has been published and correlates the publication item with a <<journal>>, a <<proceeding>>, or a <<conference>> item, the relationship <<reference>> describes the references of a publication and correlates the publication item with other publication items etc.

In the case of the "Demonstrators" resource type, the content manager has assigned four relationships:

- <<Publish to>>: describes where the demonstrator has been published or presented
- <<Reference>>: describes related publications or articles with the demonstrator
- <<Technical manual>>: describes a technical manual that may contain installation instructions, system architecture description etc.
- <<Testbed>>: describes some testbeds used by the demonstrator

In this final step the user can specify the related resources of the current resource item.

## **4. Available demonstrators**

Different demonstrators have already been ingested and made available through the DELOS WP3 demo portal. Some of these demonstrators are introduced by their creators in the following sections.

### **4.1. MILOS - Multimedia Content Management System**

MILOS (Multimedia dIgital Library for Online Search) is a general purpose software component tailored to support the design and effective implementation of digital library applications. MILOS supports the storage and content-based retrieval of any multimedia documents descriptions of which are provided by using arbitrary metadata models represented in XML.

Digital library applications are document-intensive applications where possibly heterogeneous documents and their metadata have to be managed effectively. We believe that the main functionalities required by DL applications can be embedded in a general purpose Multimedia Content Management System (MCMS), that is a software tool specialized to support applications where documents, embodied in different digital media, and their metadata are handled efficiently.

The minimum requirements of a Multimedia Content Management System are: flexibility in structuring both multimedia documents and their metadata; scalability; and efficiency.

Flexibility is required both at the level of management of basic multimedia documents and at the level of management of their metadata. The flexibility required in representing and accessing metadata can be obtained by adopting XML as standard for specifying any metadata (for example MPEG-7 can be used for multimedia objects, or SCORM (Shareable Content Object Reference Model Initiative) for e-learning objects). Proper regard for scalability and efficiency is essential to the deployment of real systems able to satisfy the operational requirements of a large community of users over a huge amount of multimedia information.

We believe that the basic functionalities of a MCMS are related to the issues of *storage and preservation* of digital documents, their *efficient and effective retrieval*, and their *efficient and effective management*. These functionalities should be guaranteed by appropriate management of documents and related metadata, according to the following prerequisites:

1. capability of managing different documents embodied in different media and stored with different strategies;
2. capability of describing documents by way of arbitrary, and possibly heterogeneous, metadata;
3. capability of providing DL applications with custom/personalised views on the metadata schema actually handled.

We have designed and built MILOS, a MCMS which satisfies the requirements and offers the functionalities discussed in previous section. The MILOS MCMS has been developed by using Web Service technology, which in many cases (e.g. .NET, EJB, CORBA, etc.) already provides very complex support for "standard" operations such as authentication, authorization management, encryption, replication, distribution, load balancing, etc. Therefore we need not elaborate further on these topics, but will concentrate mainly on the aspects discussed above.

MILOS is composed of three main components:

- the Metadata Storage and Retrieval (MSR) component
- the Multi Media Server (MMS) component
- the Repository Metadata Integrator (RMI) component

All these components are implemented as Web Services and interact by using SOAP (Simple Object Access Protocol). The MSR manages the metadata of the DL. It relies on our technology for native XML databases and offers the functionality described at [point 2](#) above. The MMS manages the multimedia documents used by the DL applications. MMS offers the functionality of [point 1](#) above. The RMI implements the service logic of the repository providing developers of DL applications with a uniform and integrated way of accessing MMS and MRS. In addition, it supports the mapping of different metadata schemas as described at [point 3](#) above. All these components were built choosing solutions able to guarantee the requirements of flexibility, scalability, and efficiency.

## Case study applications

### Reuters case study

The Reuters dataset contains text news agencies and the corresponding metadata. There are two types of metadata: Reuters specific metadata including titles, authors, topic categories, and extended Dublin Core metadata.

The Reuters dataset contains 810,000 news agencies (2.6 Gb) where text and metadata are both encoded in XML. We linked the full text index and the automatic topic classifier to the elements containing the body, the title, and the headline of the news. Other value indexes were linked to elements corresponding to frequently searched metadata, such as locations, dates, countries.

### ACM Sigmod Record and DBLP case study

Both the ACM Sigmod (Association for Computing Machinery Special Interest Group on Management of Data) Record dataset and the DBLP (Digital Bibliography & Library Project) dataset [3] consist of metadata corresponding to the description of scientific publications in the computer science domain. The ACM Sigmod record is relatively small. It is composed of 46 XML files (1Mb), while the DBLP dataset is composed of just one large (187Mb) XML file. Their structure is completely different even though they contain information describing similar objects.

We built one DL application which could access both datasets. We made use of MILOS' mapping functionality to ensure requests on the application were correctly translated for the two schemas. We linked a full text index to the elements containing the titles of the articles, and other value indexes to the more frequently searched elements, such as authors, dates, years, etc.

### **ECHO case study**

The ECHO dataset includes historical audio/visual documents and corresponding metadata. ECHO is a significant example of MILOS' ability to support the management of arbitrary metadata schemas. The metadata model adopted in ECHO, based on the IFLA/FRBR model, is rather complex and highly structured. It is used to represent the audio-visual content of the archive and includes, among others:

- the description of videos in English and in the original language
- specific metadata fields such as Title, Producer, year, etc.
- the boundaries of scenes detected (associated with a textual descriptions)
- the audio segmentation (distinguishing among noise, music, speech, etc.)
- the Speech Transcripts
- visual features for supporting similarity search on key-frames

The collection is composed of about 8,000 documents for 50 hours of video described by 43,000 XML files (36 MB). Each scene detected is associated with a JPEG-encoded key frame for a total of 21GB of MPEG-1 and JPEG files. Full text indexes were linked to textual descriptive fields, similarity search indexes were linked to elements containing MPEG-7 image (key frames) features, and other value indexes with frequently searched elements.

Milos Web site: <http://milos.isti.cnr.it/>

## **4.2. 3D Content-Based Retrieval**

### **Objective**

This demonstrator implements some approaches to retrieval of 3D objects based on their visual similarity. Its main goal is to test and compare the retrieval effectiveness of different solutions for 3D object modelling.

### **Research activity**

Activity undertaken during the first year of the project concentrated on defining a test environment in which different 3D retrieval approaches could be compared. Within this work, retrieval by similarity was achieved through using a number of techniques for object description and similarity computation. Currently, the description techniques implemented include 3D moments, curvature histograms and shape functions. Similarity of content descriptors can be evaluated according to six different distance functions: Haussler Mu, Minkowski L1, Kullback-Leibler, Kolmogorov-Smirnov, Jeffrey divergence and  $X^2$  statistics.

In particular:

Curvature histograms are constructed by evaluating the curvature of vertices of the mesh representing the 3D object. Curvature values are discretized into 64 distinct classes.

To evaluate 3D moments of a 3D object defined by a polygonal mesh, a limited set of points  $P_i$  is considered, where the relevance of each point is weighted by the area of the portion of surface associated with the point. To make the representation independent from the actual position of the model, the first order moments  $m_{100}$ ,  $m_{010}$  and  $m_{001}$  are first evaluated, and higher order moments are then evaluated with respect to the first order moments. In our experiments, moments up to the 6-th order have been computed for each model. This aims to attain a sufficient discrimination among different models.

Shape functions are evaluated by computing the histogram of Euclidean distances between all possible vertex pairs on the object mesh. Distances are normalized with regard to the maximum distance between two vertices, and discretized into 64 distinct class values.

## Plans for next activities

The work on retrieval by content of 3D objects is currently proceeding under a project proposal approved for the next 18 months of the JPA (RERE 3D: Description, Matching and Retrieval by Content of 3D Objects). In particular, we are currently investigating a 3D representation capable of supporting the spatial localization of the properties of an object surface. This is expected to improve existing approaches which currently do not consider local properties of mesh vertices.

## Demonstrator

The 3D CBR (Content-Based Retrieval) demonstrator allows users to test out retrieval by visual similarity over an archive of 3D object models. The system is fully developed in Java Technology and accessible through a Web interface available at: <http://delos.dsi.unifi.it:8080/CV/>.

The archive includes four classes of models:

- taken from the web,
- manually authored (with a 3D CAD software),
- high quality versions of models from the De Espona 3D Models Encyclopedia (<http://www.deespona.com>) and
- variations of the previous three classes (obtained through geometric deformation or application of noise, which caused surface points to be moved from their original locations).

Objects in the database cover a variety of classes, including statues, vases, household goods, transport, simple geometric shapes, and many others.

Each database model is represented in VRML (Virtual Reality Modelling Language) format through the IndexedFaceSet data structure.

The system supports retrieval according to the three content descriptors and the six similarity measures previously described. On the left part of the Web interface, three menus allow the user to:

- request a subsample of database models randomly selected;
- specify which content descriptor to use;
- specify which distance to use in order to compute the similarity between content descriptors.

The type of content descriptor and the similarity measure the user has currently selected are shown on the upper part of the interface. The user can query the system by activating the search button available below every model thumbnail. Once the search process is completed, the system presents retrieved items in decreasing order of similarity from top to bottom and from left to right (the most similar model being displayed on the upper left corner of the results panel).

In order to analyse the effect of using different content descriptors or similarity measures, once a search process is completed the user can change the type of content descriptor or similarity measure. In this case, the system automatically performs a new search evaluating the similarity between every database item and the upper left model, using the newly selected content descriptors and similarity measures.

## 4.3. VideoBrowse

### Overview

This is a tool for fast video accessing and browsing. It provides functionalities for fast decoding and playback of MPEG-1 and MPEG-2 compressed streams, without the need of an external codec. It can also handle reverse playback, one frame forward and one frame backward operations.

Two algorithms for automatic shot detection are included in this tool, one directly operating on compressed data, and the other on uncompressed data. The result of shot detection processing is an index written in the MPEG-7 standard. The index is then parsed at the successive accesses to the same video file, and then used to generate a

storyboard by selecting a single keyframe for each shot in the index.

## Shot Detection

Two different algorithms are included in the tool:

1. **Shot detection with MPEG features:** a number of features extracted from the stream are considered, namely the DC coefficients of I-frames, the number of intra-coded macroblocks and the number of forward-backward-bidirectional predicted macroblocks in a GOP (Group of Pictures). Derivatives of these quantities are also considered. Then Linear Discriminant Analysis has been used to calculate the weights of each feature in a linear combination, whose final value is used (with a threshold) to discriminate between GOPs containing and not containing shot changes. This algorithm is extremely fast, but it cannot determine the exact location of the shot change within the GOP.
2. **Shot detection with uncompressed features:** this algorithm is slower than the previous one, but it gives more accurate results. For each couple of frames we calculate the mean R,G,B values, and the maximum difference between the three channels is used as the distance measure. This value is then compared with the median distance calculated on a window of 20 frames centred in the current frame. If the ratio is greater than a manually imposed threshold, the current frame is marked as a shot change. Whenever multiple adjacent frames satisfy this condition, the frame with the maximum difference measure is selected.

Both algorithms depend on the choice of the threshold, which the user can adjust manually. To help in this operation, at the end of the shot detection processing some statistics on the value to threshold are shown in a dialog box. Furthermore, the value in each frame is written to a CSV file.

## Interface

The graphical interface includes the common playback controls such as play, forward one frame, etc., and a trackbar for fast movements in the video stream. When an index is available for the current video, a browsing window allows users to navigate through the representative keyframes, and to start the playback from a specific shot.

## 4.4. UvA Parallel Visual Analysis in TRECVID 2004

### Introduction

The Parallel-Horus framework, developed at the University of Amsterdam, is a unique software architecture that allows non-expert parallel programmers to develop fully sequential multimedia applications for efficient execution on homogeneous Beowulf-type commodity clusters. Previously obtained results for realistic, but relatively small-sized applications have shown the feasibility of the Parallel-Horus approach, with parallel performance consistently being found to be optimal with respect to the abstraction level of message passing programs. Our demonstrator shows the most serious challenge Parallel-Horus has had to deal with so far: the processing of over 184 hours of video included in the 2004 NIST TRECVID evaluation.

### The 2004 NIST TRECVID Evaluation

TREC is a conference series sponsored by the National Institute of Standards and Technology (NIST) with additional support from other U.S. government agencies. The goal is to encourage research in information retrieval by providing a large test collection, uniform scoring procedures, and a forum for organizations interested in comparing their results. An independent evaluation track called TRECVID was established in 2003 devoted to research in automatic segmentation, indexing, and content-based retrieval of digital video streams.

The 2004 NIST TRECVID evaluation defines four main tasks, at least one of which must be completed to participate in the evaluation. The University of Amsterdam participated in TRECVID 2004 by completing the feature extraction task.

This task was defined as follows: Given the 2004 NIST TRECVID video dataset, a common shot boundary reference for this dataset, and a list of feature definitions, participants must return for each feature a list of at most 2000 shots from the dataset, ranked according to the highest probability of detecting the presence of that feature.

The 2004 NIST TRECVID video dataset consisted of over 184 hours of digitized news episodes from ABC and CNN.

In addition, ten feature definitions were given, including 'Bill Clinton', 'beach', 'airplane takeoff', and 'basket scored'.

## Generic Semantic Concept Detection

Our approach to the feature extraction problem is based on the so-called Semantic Value Chain (SVC), a novel method for generic semantic concept detection in multimodal video repositories. The SVC extracts semantic concepts from video based on three consecutive analysis links, i.e. the Content Link, the Style Link, and the Semantic Context Link. The Content Link works on the video data itself, whereas the Style Link and the Semantic Context Link work on higher-level semantic representations.

In the Content Link we view video documents from the data perspective. In general, three modalities can be identified in video documents, i.e. the auditory, textual, and visual modality. In our approach, detectors are first applied to individual modalities. The results are then fused into an integrated Content Link detector. Based on validation experiments the best hypothesis for a single concept serves as the input for the next link.

Our demonstrator shows the processing of the visual modality only, as this is by far the most time-consuming part of the complete system.

## Visual Analysis

The visual modality is analyzed at the image (or video frame) level. After obtaining video data from file, for each 15th video frame visual features are extracted by using Gaussian colour invariant measurements. RGB colour values are decorrelated by transformation to an opponent color system. Then, in succession, acquisition and compression noise are suppressed by Gaussian smoothing. A colour representation consistent with variations in target object size is then obtained by varying the size of the Gaussian filters. Global and local intensity variations are suppressed by normalizing each color value by its intensity, resulting in two chromaticity values per color pixel. Furthermore, rotationally invariant features are obtained by taking Gaussian derivative filters, and combining the responses into two chromatic gradient magnitude measures. These seven features, calculated over three scales, yield a combined 21-dimensional feature vector per pixel.

The obtained invariant feature vector serves as the input for a multi-class Support Vector Machine (SVM) that associates each pixel to one of the predefined regional visual concepts. The SVM labelling results in a weak semantic segmentation of a video frame in terms of regional visual concepts. This result is written out to file in condensed format (i.e.: a histogram) for subsequent processing.

Note that this segmentation of video frames into regional visual concepts at the granularity of a pixel is computationally intensive. This is especially the case if one aims to analyze as many frames as possible.

In our approach the visual analysis of a single video frame requires around 16 seconds on the fastest sequential machine at our disposal. Consequently, when processing two frames per second at a frame rate of 30, the required processing time for the entire TRECVID dataset would be around 250 days. Application of the Parallel-Horus framework, in combination with a distributed set of Beowulf-type commodity clusters significantly reduced this required processing time to less than 60 hours. These performance gains were obtained without any parallelization effort whatsoever, which was an important contributing factor in our top ranking in the TRECVID results.

## 4.5. Audio feature extraction with Rhythm Patterns

Content-based access to audio files, particularly music, requires the development of feature extraction techniques to capture the acoustic characteristics of the signal and so permit the computation of similarity between pieces of music, reflecting the similarities perceived by human listeners.

'Rhythm Patterns' are feature sets derived from content-based analysis of musical data and reflect the rhythmical structure in the musical pieces. Classification of sound into musical genres as well as automatic organization of music archives according to sound similarity are made possible through the psycho-acoustically motivated 'Rhythm Patterns' features.

The feature extraction process for the Rhythm Patterns is composed of two stages. Firstly, the specific loudness

sensation in different frequency bands is computed, by using a Short Time FFT (Fast Fourier Transform). The resulting frequency bands are then grouped into psycho-acoustically motivated critical-bands, applying spreading functions to account for masking effects and successive transformations into the decibel, Phon and Sone scales. This results in a power spectrum that reflects the human sensation of loudness. In the second step, the spectrum is transformed into a time-invariant representation based on the modulation frequency; this is achieved by applying another discrete Fourier transform, resulting in amplitude modulations of the loudness in individual critical bands. These amplitude modulations have different effects on human hearing sensation depending on their frequency, the most significant of which, referred to as the fluctuation strength, is most intense at 4Hz, decreasing towards 15Hz. From that data, reoccurring patterns in the individual critical bands, resembling rhythm, are extracted, which - after applying Gaussian smoothing to diminish small variations - result in a time-invariant, comparable representation of the rhythmic patterns in the individual critical bands. The proposed feature set then serves as a basis for an unsupervised organization task, as well as for machine learning or classification tasks.

This feature set was submitted to the Audio description contest of the International Conference on Music Information Retrieval (ISMIR 2004), winning the rhythm classification track.

## **4.6. TZI Demonstrators for DELOS WP3**

### **Video Content Manager**

The Video Content Manager is a tool for analysis and annotation of digital videos. It has been developed in cooperation with researchers from the cultural sciences and from the arts.

The faculty of cultural sciences at the University of Bremen possesses thousands of hours of digitized videos. These include videos of lectures, digitized telecasts and video works from students. For students and teachers in cultural studies, access to this video footage is often needed for practising, lecture reruns or the creation of new videos. Annotation of the video material is required to provide such access which can be achieved through use of the Video Content Manager.

At the University of the Arts Bremen, a group of art historians interested in the medium of video built a prototype of an international archive of video arts. Currently it is very difficult to gain access to as is additional information on artistic media works, including video arts. The information about these works is scattered across the world and is very difficult to obtain outside conventional channels, such as exhibitions, festivals, or conferences. These problems are tackled by the prototype archive. The Video Content Manager is used to facilitate annotation of the video works and their ingestion into the archive.

Annotating a video with the Video Content Manager is a three-stage process. Firstly, an automatic shot boundary detection algorithm is run on the video. Its results yield a temporal segmentation of the video. For each shot, a key frame is automatically extracted from the video. Such a key frame allows for a quick overview of the content of a shot and is suitable for browsing the video without having to view it as a whole.

In the second step, successive shots that cover the same topic or show the same location are merged together to form what we call a "scene". This has to be done manually. The result of this step is a hierarchical temporal segmentation of the video with three levels of different granularity: shot, scene, and video.

The final step is a textual annotation follows an annotation scheme tailored to the users' needs, but based consistently on Dublin Core [1]. The annotation is guided by the temporal segmentation from the second step and may use the keyframes from the first step for efficiency purposes. Shots are annotated on a more syntactical level (what can currently be seen in the video?). Scenes are annotated on a more semantical level (what is going on, what is the topic?).

The results of the annotation process may be exported as XML for ingestion into a database. The video data itself are not manipulated. The Video Content Manager is available on the demonstrator website of the DELOS cluster 3 (A/V-NTO).

### **Automatic shot boundary detection for TRECVID 2004**

The TRECVID workshop [2] is an annual meeting of users and researchers in the field of content-based video



analysis, retrieval, and digital video libraries. The workshop began life as a "track" of the Text Retrieval Conference (TREC), but became a separate workshop in 2003. Its goal is to provide a forum for evaluation of video retrieval algorithms together with a common collection of videos. In 2003 and 2004, the video material provided consisted mainly of news broadcasts, including sports material, weather forecasts and commercials.

Several tasks are made available to participants: shot boundary detection, high-level feature extraction, and search. TZI- Bremen University has taken part in the high-level feature extraction task (2002) and the shot boundary detection task (2002-2004). The shot boundary detection tool used to produce the results [3] submitted in 2004 is available as a demonstrator on the DELOS WP 3 demonstrator website.

### Automatic realtime text extraction

Telecasts, especially news and magazine broadcasts, are often, if not always, enhanced with text inserts. The information contained in these inserts may cover topics, names of presenters, interviewers or interviewees, news tickers or casts. Automatic recognition of the text displayed can represent a considerable benefit to content-based video search and retrieval.

The automatic recognition of text in text-based documents (OCR) is a well-researched field. However the recognition of text inserts in video often proves more difficult. It includes segmentation of the text from the background, which is usually much more complex than in black and white text-only documents. To simplify the task, detection of those areas of the video containing text can be very useful.

A fast detector for text areas which extracts the locations of text inserts in video, and which tracks these text areas over multiple frames for scrolled text, is provided in the demonstrator section in the DELOS WP 3 portal. It is based on statistics on the visual properties of text inserts and can be run in realtime. The resulting location data may be used as hints for a subsequent video-OCR. Alternatively they can be used alone as an indicator in the discrimination between different video segments, for example between an anchor shot and a credits sequence.

## 4.7. The Video Segmentation and Annotation Tool Demonstrator

"Video Segmentation & Annotation tool" is a full system that supports the segmentation, indexing and annotation of audiovisual content and the creation of segmentation metadata compliant with the TV-Anytime (Standard for digital video content) Segmentation Metadata Model. The system comprises a graphical application where the metadata are created or edited and a relational database where these metadata are stored. More specifically, its functionality includes the creation of video segments and video segment groups according to the TV-Anytime Segmentation Metadata Model as well as the semantic annotation of these segments through the application of domain-specific ontologies and transcription files.

The architecture of the tool follows a multi-tier approach consisting of the following tiers:

1. Application Tier: the graphical application shown to the user.
2. XML-DB middleware tier: a set of software components responsible for the management of TV-Anytime XML documents (hidden tier).
3. Database Tier: a database management system used to store relational data that derived from the TV-Anytime metadata.

The system offers the following functionality:

- **Creation of a new project:** Users of the tool can start the segmentation process of a video programme and create a new project in one of 3 ways:
  - a) search on the database for recorded programmes (this requires a connection with a database server)
  - b) load XML file containing TV-Anytime Segmentation metadata and
  - c) open a new unsegmented video programme. (Note that the segmentation process cannot proceed without starting a new project as almost all of the buttons and fields of the graphical application are disabled.)
- **Play video programme:** One can use the media player frame of the application to play a video file.
- **Create/Edit segments:** Users can create or edit segments of the video programme playing according to the TV-Anytime Segmentation Model by defining or modifying their start and end time, their keyframes, their description (title, synopsis, keywords and related material), their version and their unique segmentID. (For more information about the above definitions see the TV-Anytime Segmentation metadata model.)
- **Create/Edit segment groups:** Users can create or edit segment groups composed of segments previously

created, according to the TV-Anytime Segmentation Model by defining or modifying the segments which contain them, their group interval, their keyframes, their description (title, synopsis, keywords and related material), their group type, their version and their unique segmentGroupID.

- **Save created metadata:** Every metadata created for a video programme (segments and segment groups) are saved in a structure called the "Segment Information Table" found in the main memory. In order to store this information in a permanent way the tool provides two methods:
  - a) store metadata onto a database (where a connection with a database server already exists) and
  - b) export metadata in an XML file compliant with the TV-Anytime Segmentation metadata model.
- **Search within segments and segment groups:** this tool provides three ways in which to browse within the created segment and segment groups:
  - a) Groups & Segments Schema (a synoptic view of the segments and segment groups of the current Segment Information Table) ,
  - b) Segment Information Table Explorer (a more analytical view of the segments and segment groups of the current Segment Information Table) and
  - c) Text-based search for Segments & Groups (search within segments and segment groups of the current Segment Information Table based on the segment or segment group title or keywords).
- **On-line help:** The tool has an on-line help system that describes in an analytical way the use of the application and may be able to provide immediate answers to users' questions.
- **Security and Recovery:** The tool provides a mechanism so as to prevent loss of unsaved metadata in case of system crash. In that case when the application is restarted the user can reload these metadata.
- **Importing and Using Ontologies:** This tool provides the ability to import domain-specific ontologies (the current version supports ontologies that are based on keywords) in order to achieve a more accurate segment annotation by using words from these ontologies.
- **External Java API video programmes of soccer matches:** This tool has been integrated with an external application that supports an existing ontology covering the domain of soccer matches. Users can insert keywords based on a MPEG-7-compliant ontology for soccer matches. The API creates set of phrases according to this ontology which are then used as keywords in the description of segments and segment groups.
- **Advanced Searching in transcription files:** Users can employ a transcription file produced during the segmentation process to find in which part of the video specific phrases are heard. They can browse within the entire transcription file or search for specific words or phrases. By clicking on one of these phrases they can identify the section of video corresponding to the speech identified. The reverse procedure is also supported: users can find which phrases are heard in each video segment while the video is played.

## 4.8. The UP-TV Demonstrator

The UP-TV system is based on the TV-Anytime architecture for digital TV systems and follows the corresponding metadata specifications of audiovisual content and user descriptions. It is directly related to audiovisual digital libraries as it relates to the development of single-user and server systems for the management of audiovisual content compliant with TV-Anytime specifications.

The UP-TV system follows a multi-tier architecture [1]. The lowest tier handles the metadata management. The middleware tier includes all the logic for interfacing the system with the outside world. The application tier enables the exchange of information between the server and heterogeneous clients through different communication links. The core of the system is the metadata management middleware that takes over the storage of the TVAM program and user metadata descriptions as well as providing advanced information access and efficient personalization services. The implementation was based on the following decisions:

The metadata management system should be able to receive and create all kinds of valid XML documents with respect to the TVAM XML Schema.

The database management system should follow the relational model and support the SQL standard as the language for data manipulation and retrieval, so that it can be easily integrated with additional information on the servers, allow concurrent access etc.

The solutions developed include functionality for storing the program metadata onto relational databases, functionality for storing TVAM consumer metadata on databases and functionality for retrieving data from the relational databases and assembling valid TVAM documents or document fragments. Mapping the TVAM XML structure onto relational databases provides efficient mechanisms for matching program and profile metadata as well as user profile adaptation and data mining for viewing histories through the use of the SQL language, thus facilitating the implementation of powerful services for both final users and service providers. The XML-DB middleware (figure 1) is a set of software components responsible for the manipulation of TVAM XML documents and the mapping of TVAM XML Schema to the underlying relational schema. It is supported by a relational

database management system along with the relational database, used to store the data of TVA metadata descriptions.

TVAM-compliant clients use XML documents to communicate with the system. These documents contain data that could be used in conjunction with data from other TVAM XML documents. Document (or document fragment) retrieval is supported by a special purpose Application Programming Interface (API). In this environment the data management software should not rely on XML document modelling solutions (like DOM) but rather on a data binding approach. Data binding offers a much simpler approach to working with XML and supports effective separation between document structure and data modelling.

There are numerous XML data binding products capable of transferring data between XML documents and objects. Design-time binders (which require configuration based on a DTD or an XML Schema before they can be used) are usually more flexible in the mappings that they can support. The overall system architecture assumes a design-time binder. Therefore a configuration process was necessary to create the appropriate classes. The XML data binder considered for the implementation of our system is data-centric. It is capable of fully representing XML documents as objects or objects as XML documents (the serialization of the object tree to XML document is encapsulated in class (un)marshal methods). The data binder uses a SAX-based parser and the corresponding validator can be used to ensure that incoming and outgoing XML documents conform to the TVAM XML schema.

The communication with the relational database management system relies on the use of standard interfaces like JDBC. Standard SQL statements are used to store-retrieve data from the underlying relational database. To do so, the classes created during the data binding configuration process are extended with DB-Insert/Retrieve methods. DBInsert methods use the object tree to create INSERT/UPDATE statements to give persistence to data on the object tree. These methods can also query the database to avoid duplicates of data. DBRetrieve methods retrieve data from the database in order to build object trees that could be used to create TVAM XML documents. The DB-Insert/ Retrieve methods rely on both the class hierarchy created by the data binding configuration process and the relational schema of the underlying database. The relational database is responsible for the storage and retrieval of information that is represented in TVAM XML documents.

In order to support ubiquitous access, special device-specific components of middleware were developed for the UP-TV environment. Java technology was chosen for the application development in hand-held devices since it is adequate for dynamic delivery of content, provides satisfactory user interactivity and ensures cross-platform compatibility. Two components were built, one suitable for cellular phones compatible with the MIDP profile and one for PDAs that support the Personal profile. In order to keep the communication scheme simple and uniform over different devices, we have chosen to use HTTP since it is suitable for the transfer of XML documents and is the network protocol supported by the MIDP libraries. The front-end of the server consists of Java servlets that accept HTTP requests from the clients and embody software adapters that adapt appropriately the information that will be exchanged and the functionality that can be provided, depending on the kind of the device requesting service.

## 4.9. The Campiello Demonstrator

The Campiello system is a system for intelligent tourism information and interaction between visitors (or potential visitors) of cities possessing significant cultural heritage (e.g. Chania and Venice) and their local citizens. This system has been developed through the use of innovative technologies including non-traditional objects like 3D reconstructions of archaeological sites and interactive city maps.

This section describes the architecture of the Campiello PC Interface. By the term 'architecture' we mean the functionality that the interface was designed to offer, the layout of the screens used in the interface and the way the interface is structured, i.e. the different sections used and the navigation between these sections.

The Campiello website as it is right now fully supports the first requirement. It currently supports 4 languages (English, Greek, Italian and French) but their number can be increased arbitrarily without modification to the implementation.

All the text that appears on the interface is read from a database where it is organized in terms of "Interface Contexts" and "Interface Topics". Contexts refer to, as the name implies, discrete contexts, i.e. sections or subsections of the interface. An Interface Context usually corresponds to one screen of the interface. Topics, on the other hand, refer to specific items in a Context, i.e. to elements within a screen.

Using this convention one can describe all elements on each screen with an intuitive, easy to remember name. For example, the title of the Places page is referred to as Context: Places, Topic: Name/Title. The caption for the Search button (which appears on every page) is referred to as Context: Any, Topic: Search.

In the database we have a description in all the available languages for each Context/Topic, so the system can fetch the appropriate text based on the language the user has selected.

Through the notion of a "Step" parameter to the interface labels stored in the database, we can have multiple texts that refer to the same Context/Topic. This was done to support cases where we needed multiple messages that would correspond to the same Context/Topic pair. An example of this are customized error messages that can vary subtly based on some parameter.

It is also possible to provide different text for each language regardless of whether the user enters the Campiello system through a Mac with a Netscape browser or through a PC with Internet Explorer.

We should point out here the difference between the language of the interface and the language of the content of Campiello. The texts for the interface should be provided in all the available languages, but the content of Campiello will not necessarily display in all languages, i.e. if someone has not posted an appropriate translation. As a result of that, changing the language when viewing Campiello content may lead to an error message if the same content is not available in the language selected, whereas this is not the case in respect of text used for the interface.

## 5. References

1. Dublin Core Metadata Initiative: <http://dublincore.org>
2. TRECVID Video Retrieval Evaluation: <http://www-nlpir.nist.gov/projects/trecvid/>
3. Results can be seen here: <http://www-nlpir.nist.gov/projects/tvpubs/tv.pubs.org.html>

### 5.1 Publications

- [1] G. Amato, C. Gennaro, F. Rabitti, P. Savino "Milos: A Multimedia Content Management System". Extended abstract, SEBD 2004, S. Margherita di Pula (CA), Italy, June 21-23, 2004.
- [2] G. Amato, F. Debole, F. Rabitti, P. Savino, and P. Zezula "A Signature-Based Approach for Efficient Relationship Search on XML Data Collections". *XML Database Symposium (XSym 2004)* in Conjunction with VLDB 2004, Toronto, Canada, 29-30 August 2004
- [3] F. J. Seinstra, D. Koelma, and A. D. Bagdanov: "*Finite State Machine-Based Optimization of Data Parallel Regular Domain Problems Applied in Low-Level Image Processing*". IEEE Transactions on Parallel and Distributed Systems, 15(10):865-877, 2004.
- [4] C. G. M. Snoek, M. Worring, and A. G. Hauptmann: "*Detection of TV News Monologues by Style Analysis*". In Proceedings of the IEEE International Conference on Multimedia & Expo (ICME) - Special session on Multi-Modality-Based Media Semantic Analysis, Taipei, Taiwan, 2004.
- [5] N. Sebe, M.S. Lew, T.S. Huang: "*Computer Vision in Human-Computer Interaction*". HCI/ECCV 2004. Lecture Notes in Computer Science, Vol. 3058, Springer-Verlag, ISBN 3-540-22012-7, 2004
- [6] M. Bertini, A. Del Bimbo, A. Prati, R. Cucchiara, "*Semantic Annotation and Transcoding for Sport Videos*". In Proceedings of International Workshop on Image Analysis for Multimedia Interactive Services (WIAMIS), 2004.
- [7] M. Bertini, R. Cucchiara, A. Del Bimbo, A. Prati, "*Content-based Video Adaptation with User's Preference*". In Proceedings of International Conference on Multimedia & Expo (IEEE ICME 2004), 2004.
- [8] M. Bertini, R. Cucchiara, A. Del Bimbo, A. Prati, "*Object-based and Event-based Semantic Video Adaptation*". In Proceedings of International Conference on Pattern Recognition (IAPR-IEEE ICPR 2004), vol. 4, pp. 987-990, 2004.

- [9] M. Bertini, A. Del Bimbo, A. Prati, R. Cucchiara, "*Objects and Events Recognition for Sport Videos Transcoding*". In Proceedings of 2nd International Symposium on Image/Video Communications over fixed and mobile networks (ISIVC), 2004.
- [10] C. Grana, G. Pellacani, S. Seidenari, R. Cucchiara, "*Color Calibration for a Dermatological Video Camera System*". In Proceedings of The 17th International Conference on Pattern Recognition (ICPR 2004), Cambridge, UK, pp. 798-801, August 23-26, 2004.
- [11] R. Cucchiara, C. Grana, G. Tardini, R. Vezzani, "*Probabilistic People Tracking for Occlusion Handling*". In Proceedings of The 17th International Conference on Pattern Recognition (ICPR 2004), Cambridge, UK, pp. 132-135, August 23-26, 2004.
- [12] N. Orio, P. Zanuttigh, G.M. Cortelazzo. "*Content-Based Retrieval of 3D Models Based on Multiple Aspects*". Accepted at IEEE International Workshop on Multimedia Signal Processing, Siena, IT, 29 September 1 October, 2004. In press.
- [13] G. Neve, and N. Orio. "*Indexing and Retrieval of Music Documents through Pattern Analysis and Data Fusion Techniques*". Accepted at International Conference on Music Information Retrieval, Barcelona, ES, 10-14 October, 2004. In press.
- [14] D. Schwarz, N. Orio, and N. Schnell. "*Robust Polyphonic MIDI Score Following with Hidden Markov Models*". Accepted at International Computer Music Conference, Miami, USA, 1-6 November, 2004. In press.
- [15] E. Bertino, E. Ferrari, D. Santi, A. Perego: "*Constraint-based Techniques for Personalized Multimedia Presentation Authoring*". Submitted for publication.
- [16] S. Valtolina, S. Franzoni, P. Mazzoleni, E. Bertino: "*Dissemination of Cultural Heritage Content through Virtual Reality and Multimedia Techniques: a Case Study*". Accepted for publication in IEEE MMM 2005: 11th International Multi-Media Modelling Conference. Melbourne, Australia, 12 - 14 January 2005.
- [17] S. Valtolina, S. Franzoni, E. Bertino, E. Ferrari, P. Mazzoleni: "*A virtual reality tour in an Italian Drama Theatre: A journey between architecture and history during 19th century*". Proceedings of the EVA 2004: Electronic and Imaging & the Visual Arts. London, United Kingdom, 26 - 31 July 2004.
- [18] M. Bertini, R. Cucchiara, A. Del Bimbo, A. Prati: "*Semantic Annotation and Transcoding for Sport Videos*". WIAMIS 2004, Lisbona, April 2004.
- [19] C. Colombo, D. Comanducci, A. Del Bimbo and F. Pernici: "*Accurate automatic localization of surfaces of revolution for self-calibration and metric reconstruction*". In Proceedings IEEE Workshop on Perceptual Organization in Computer Vision (POCV 2004), Washington, DC, USA, June 2004.
- [20] M. Bertini, A. Del Bimbo W. Nunziati: "*Common Visual Cues for Sports Highlights Detection*". Proc. IEEE International Conference on Multimedia & Expo (ICME'04), Taipei, Taiwan, June 27-30 2004.
- [21] S. Berretti, G. D'Amico, A. Del Bimbo: "*Shape Representation by Spatial Partitioning for Content Based Retrieval Applications*". Proc. IEEE International Conference on Multimedia & Expo (ICME'04), Taipei, Taiwan, June 27-30 2004.
- [22] J. Assfalg, G. D'Amico, A. Del Bimbo, P. Pala: "*3D content-based retrieval with spin images*". Proc. IEEE International Conference on Multimedia & Expo (ICME'04), Taipei, Taiwan, June 27-30 2004.
- [23] M. Bertini, R. Cucchiara, A. Del Bimbo, A. Prati: "*Objects and Events Recognition for Sport Videos Transcoding*". ISIVC 2004, Brest, France, July 2004.
- [24] S. Berretti, A. Del Bimbo, P. Pala: "*A Graph Edit Distance Based on Node Merging*". Proc. International Conference on Image and Video Retrieval (CIVR'04), pp.464-472, Dublino, Ireland, July 21-23 2004.
- [25] S. Berretti, A. Del Bimbo: "*Multiresolution Spatial Partitioning for Shape Representation*". Proc. IEEE

International Conference on Pattern Recognition (ICPR'04), vol.II, pp.775-778, Cambridge, United Kingdom, August August 23-26 2004.

[26] J. Assfalg, A. Del Bimbo, P. Pala: "*Spin Images for Retrieval of 3D Objects by Local and Global Similarity*". Proc. IEEE International Conference on Pattern Recognition (ICPR'04) vol.III, pp.906-909, Cambridge, UK, August 23-26, 2004.

[27] M. Bertini, R. Cucchiara, A. Del Bimbo, A. Prati: "*Semantic Video Adaptation based on Automatic Detection of Objects and Events*". Proc. IEEE International Conference on Pattern Recognition (ICPR'04) vol.IV, pp.987-990, Cambridge, UK, August 23-26, 2004.

[28] Jacobs and Th. Hermes and O. Herzog: "*Hybrid Model-based Estimation of Multiple Non-dominant motions*". In: Proceedings of the 26th DAGM Symposium on Pattern Recognition, Tübingen, Germany, 2004.

[29] M. Crucianu, M. Ferecatu and N. Boujemaa: "*Reducing the redundancy in the selection of samples for SVM-based relevance feedback*". Research report INRIA 5258, May 2004.

[30] H. Shao, T. Svoboda, L. Van Gool: "*Distinguished Color/Texture Regions for Wide Baseline Stereo Matching*". Submitted for publication.

[31] Cotsaces, M.A. Gavrielides, and I. Pitas: "*Video Shot Boundary Detection and Condensed Representation: a review*". IEEE Transactions on Circuits and Systems for Video Technology, submitted, September 2004.

[32] M. Frantzi, N. Moumoutzis, S. Christodoulakis: "*A Methodology for the Integration of SCORM with TV-Anytime for Achieving Interoperable Digital TV and e-Learning Applications*". In the Proceedings of the International Conference on Advanced Learning Technologies (ICALT 2004 ), August 2004, Finland

## Author Details

### George Ioannidis

Technologie-Zentrum Informatik (TZI)

University of Bremen

Germany

email: [george.ioannidis@tzi.de](mailto:george.ioannidis@tzi.de)

website <http://www.tzi.de>

[DELOS Home](#)

[DELOS Newsletter Front Page](#)

[Delos Newsletter Contents](#)

Publication Date: June 2005

Publication: Delos Newsletter Issue 3

Originating URL: </projects/delos/newsletter/issue3/feature1/intro.html>

File last modified: Tuesday, 12-Jul-2005 13:57:22 BST

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



## Reports from the DELOS Clusters

Each issue of the DELOS Newsletter carries reports from clusters working within the DELOS Network of Excellence.

- [Digital Library Architecture](#)
- [Information Access and Personalization](#)
- [Audio/Visual and Non-traditional Objects](#)
- [User Interfaces and Visualization](#)
- [Knowledge Extraction and Semantic Interoperability](#)
- [Evaluation](#)

---

### Digital Library Architecture

[Can Türker](#) has provided us with an update on activity by the Digital Library Architecture cluster in terms of workshops, exchange programme work and site development.

Over recent months considerable progress has been made in preparing a survey on service-oriented architectures (SoA), peer-to-peer architectures (P2P), and grid infrastructures, which is one of the central objectives of WP1 within the first 18 months of the DELOS Project. Input from all WP1 partners on applying SoA, P2P, and/or grid technology to Digital Libraries has been collected and integrated. The survey was completed in February 2005. For further information, do contact the [cluster leader](#).

Many WP1 partners were also involved in preparing a revised, extended version of their paper presented at the Sixth Thematic Workshop of DELOS on Digital Library Architectures at S. Margherita di Pula (Cagliari), Italy, 24-25 June 2004. (See [http://ii.uit.at/research/delos\\_website/6thworkshop.html](http://ii.uit.at/research/delos_website/6thworkshop.html)) These papers have now undergone a second reviewing phase. Accepted papers will be published in the workshop post-proceedings and appear as a book in the Lecture Notes in Computer Science (LNCS) series of Springer. Can Türker (ETHZ), Hans-Jörg Schek (ETHZ/UMIT), and Maristella Agosti (U Padua) are in charge of editing and producing this book. For further information, do contact the [cluster leader](#)

Over the same period, joint work between several DELOS partners has intensified within the researcher exchange programme. For instance, MUNI hosted Giuseppe Amato from ISTI-CNR Pisa for two weeks in November 2004 and ETHZ hosted Harald Krottmaier from Graz for three weeks in total from September to November 2004.

The DELOS OpenDLib (<http://www.opendlib.com>) managed by ISTI/CNR has been updated by including documents produced during the DELOS events organized over the autumn. In particular, it now contains material from the Third International Summer School on Digital Library Technologies (ISDL 2004), the DELOS WP7 Workshop on the Evaluation of Digital Libraries, and CLEF 2004.

The Spring 2005 joint WP1/WP2 workshop organized by MPII took place at Dagstuhl Castle, Germany over 29 March - 1 April. This workshop, the 8th in the DELOS series of Thematic Workshops, devoted itself to two critical themes in digital libraries: system architecture and information access. The main workshop objectives were to bring together researchers and practitioners interested in these two areas and their inter-connections, to identify fundamental system services that allow the development and operation of future Digital Libraries, and to explore the main relevant technical directions. More information about this workshop is available under

[http://ii.uit.ac.at/research/delos\\_website/8thworkshop.html](http://ii.uit.ac.at/research/delos_website/8thworkshop.html)

### **Can Türker**

Member  
ETH Zurich,  
Institute of Information Systems,  
CH-8092 Zurich

E-mail: [tuerker@inf.ethz.ch](mailto:tuerker@inf.ethz.ch)

url: <http://www.dbs.ethz.ch>

Telephone: +41 1 6327248

Fax: +41 1 6321172

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

[Return to top](#)

## **Information Access and Personalization**

The Information Access and Personalization cluster enters the second period of the project with a new, exciting research agenda. [Georgia Koutrika](#) outlines for us the new research activities envisaged in the new phase.

## **The New Research Agenda for IAP**

### **Introduction**

The Information Access and Personalization cluster enters the second period of the project with a new, exciting research agenda. This agenda has come as a product of the efforts of the cluster so far as well as a logical continuation of them. For this reason, it comprises a list of vital research topics with a fair distribution of interest and effort among the major areas of the cluster's interests. In addition, it provides a strong incentive for collaboration among researchers with different background and interests in order to solve new problems that require expertise from different fields.

### **IAP's Strategic Goals**

This agenda represents the combined effect of the cluster's past efforts. These have been planned in line with the following strategic goals:

1. **Construction of a common, comprehensive framework** for information access and personalization approaches. Towards this end, a set of comprehensive surveys and reports have been generated. These provide a broad coverage of the general areas of interest to the cluster, and describe problems, existing models and approaches.
2. **Promotion of knowledge** about available practices in the fields of information access and personalization in digital libraries. To further this objective, several information-sharing and dissemination activities have taken place.
3. **Initiation of research** on new information access and personalization models and methodologies. This was made possible through the afore-mentioned activities and the researchers' exchange programme.

All these activities have led to the identification of major research trends and significant obstacles in the fields of information access and personalization as well as the establishment of cooperation between researchers and of several research proposals. In addition, they have provided the proper basis for the new objectives.

### **Phase Two**

During the second phase, the Information Access and Personalization cluster will build on and continue the work performed during the first period and will pursue the following strategic goals:



1. **Substantial support of cooperation** between individual research groups initiated during the first phase of the project.
2. **Research on new models and methodologies** in order to eliminate inefficiencies in existing ones.
3. **Development of toolkits and systems** for purposes of re-use and demonstration of proposed methods and models.

## IAP Research Agenda

For the second phase of DELOS Project, we have drawn three basic lines of research:

- Advanced Information Access Methods
- P2P resource-sharing for Digital Libraries
- User Context Modelling

Along these lines, five new activities have been planned overall. These, along with those activities continuing from the first period of the project's lifetime, constitute the new research agenda of the IAP cluster. In what follows, I will try to give an overview of the aforementioned research directions.

### 1. Advanced Information Access Methods

The importance of this topic is justified by the ever-growing volume of multimedia objects in digital libraries. Added to this is the critical role multimedia plays in the improvement of a user's experience when accessing a Digital Library. Therefore, effective and efficient information access of multimedia content becomes increasingly more important. In particular, the main challenge here is as follows:

- *Query processing for non-alphanumeric data* (e.g. similarity search in large collections of media objects) is often conducted by means of some distance metric that computes a score quantifying the similarity of two objects. Unfortunately, distance metrics fail to answer queries efficiently as soon as the dimensions of the feature vectors exceed a certain "usability threshold".

The main *issues* to be investigated include the following:

- *Development of novel indexing and query processing algorithms*: The goal pursued is indexing and query processing approaches that :
  - (i) efficiently support several types of queries over multimedia, such as nearest neighbour, range, and ranking queries,
  - (ii) operate on any kind of multimedia data, and
  - (iii) are generic.
- *Design of distributed access methods for similarity search*: The goal is to study new distributed storage structures for similarity search of data by exploiting the P2P paradigm.
- *Similarity search approaches in digital library applications exploiting XML encoded metadata*: Emerging metadata standards (e.g. MPEG-7) encode metadata (including audio/visual features) with XML. In this respect, the objective is to investigate techniques for efficient processing of combined exact and similarity match queries on XML-encoded metadata to support complex search capabilities in multimedia digital libraries.

### 2. P2P Resource-sharing for Digital Libraries

The peer-to-peer (P2P) paradigm is an intriguing approach for coping with dynamically evolving federations of loosely coupled digital libraries. The large scale and high dynamics of a P2P system poses very challenging issues as far as information access is concerned:

- A P2P network comprises Digital Library (DL) nodes, which should be *autonomous*, and, ideally, should have *no restrictions* on how to organize their data and what kind of query capabilities to offer.
- DL services should support *decentralized sharing and management of data* through a P2P network of DL nodes. In such a network, a DL node must be able to provide data to other DL nodes and, at the same time, to have access to data of other DL nodes, while retaining its own autonomy by making local decisions.
- In addition to the libraries, *user agents* with powerful personalized tools may participate as peers. User peers should also remain autonomous and participate in the P2P system and collaborate on behalf of user requests at their discretion. In a typical scenario, a user peer would utilize its local profile to compile an information demand and issue queries to the library peers that are most suitable for the given demand.

Such a P2P system promises unlimited scalability, robustness to failures, fluctuation, and load dynamics, and also much reduced vulnerability to attacks and information manipulation.

The main issues that will be investigated include:

- *Query reformulation in P2P DLs:* To answer a query in a P2P setting, queries are reformulated to match the schemas of the data sources. To perform this reformulation task the key issue is mapping rules. Ideally, the assumption of a single global schema should not have to be a prerequisite in a P2P Digital Library. The objective is to investigate the foundations of query reformulation using mappings in P2P DLs, where the query language will be either XML-based or RDF/S-based.
- *Query roaming in P2P DLs:* The 'roaming' facility allows a user to enjoy uninterrupted communication (through a mobile network or through the Internet) from anywhere in the entire network coverage space. During query roaming, planning and execution must be interleaved, since the precise evaluation steps of the computation and the participating peers cannot be foreseen in advance. Furthermore, plans may have to be revised at runtime and contingency choices made, e.g., when a peer involved in the computation suddenly leaves the system. The objective is to investigate routing, planning, and processing of declarative queries in order to support roaming in P2P DLs.
- *Database selection:* While an issue that has been intensively studied in the literature on distributed IR and meta-search engines, the large scale and high dynamics of a P2P system pose greater challenges, including issues such as the dissemination, maintenance, and effective use of statistical summary information about the peers' information content. The objective here is to investigate opportunities for advanced personalization of information selection and query execution based on users' access patterns, query logs, and further user-behaviour information kept in peers.

### 3. User Context Modelling

The *interpretation* and *suitability* of data increasingly depends on *changing conditions*, like for example the current position of the user or the media he is using (laptop, mobile, PDA, etc.), and on user characteristics, e.g. *preferences*. All these parameters constitute the *user's context* which should be taken into account during information access in order better to serve the user. Towards this direction, it is essential that information providers specify the context under which information becomes relevant. Conversely, information users can specify their (context-dependent or context-free) preferences as well as their current context when requesting for data.

Towards the enhancement of the personalization capabilities of DL systems, the following issues will be investigated:

- *Foundations for a context-aware database system:* It is important to investigate how the notion of context can be incorporated into data stored in DLs of relational form. The investigation will be done at two levels: an abstract and an applied one. Initially the focus will be on laying the foundations for a context-aware database management system, and then the task will attempt to define a uniform mechanism for personalized access to Digital Libraries.
- *Management of user preferences:* Towards the same end, enrichment of the DL information organization and retrieval services with preference capabilities for supporting personalized services will be studied. Such services include query personalization (i.e. enhancing a query by incorporating into it user-specific preferences), user notification (notifying a user when an 'event' of interest happens), and document customization (allowing the user to compose documents that will be created by the system).

### Concluding Remarks

The outcomes from the efforts of the Information Access and Personalization cluster so far have served to underline the significance of the research directions selected. Furthermore, the research agenda of the IAP cluster retains two additional and unique features:

- It achieves a fair distribution of interest and work among the major dimensions of the cluster's work, namely Information Access (interaction with a single information provider), Information Integration (interaction with multiple information providers) and Personalization (customization of information and interaction to the user)
- It provides incentives for collaboration among researchers of differing backgrounds and interests so as to solve emerging problems which require expertise from different fields

With this research agenda before us, the second period of the project for the Information Access and Personalization cluster promises to be both challenging and exciting.

## Author Details

Georgia Koutrika  
 University of Athens  
 Email: [koutrika@di.uoa.gr](mailto:koutrika@di.uoa.gr)  
 Tel: +30 210 727 5242  
 Fax: +30 210 727 5214

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

[Return to top](#)

## Audio/Visual and Non-traditional Objects

**George Ioannidis** gives an outline of the cluster's progress and refers us to greater detail and future directions in the [feature of this issue](#).

### Introduction

Over the first 12 months of the project WP3 has aimed to develop a common understanding and foundation for the work that has to be done in DELOS in terms of State of the Art Reports, support for Forum and Testbeds, and efforts at understanding the expertise of the partners and their possible cooperation towards the objectives of DELOS as they are described in the Technical Annex.

### Progress on Reports

The reports entitled *State of the Art on Metadata Extraction* and *State of the Art in Audiovisual Content-Based Retrieval, Information Universal Access & Interaction including Data Models & Languages* have been completed. A preliminary draft of the state of the art report on *Audiovisual Metadata Management* has been produced. For further information, do contact the [cluster leaders](#).

### Portals and Demonstrators

The [Delos Collaborative Portal](#) has been released. The portal is intended to foster exchange of ideas and useful information within the DELOS Community.

The DEMOS portal for demonstrators and testbeds has been created based on an analysis of the requirements for supporting testbeds and demonstrators. The DEMOS portal is described in further detail in [Section 3](#) of the feature. Several demonstrators have already been ingested, some of which are described in [Section 4](#) of the feature. Some testbeds have also been provided. They will not be described here, but may be accessed through the DEMOS portal.

### Metadata-related Activity

For ontology-based metadata definition, a tool named GraphOnto has been implemented. An OWL upper ontology that captures the MPEG7 MDS is utilized. This upper ontology is extended with domain knowledge through appropriate OWL domain Ontologies.

In the same context, a study for the integration of the TV-Anytime Metadata model with the SCORM 1.2 Content Aggregation Model has been completed that defines a detailed mapping between the two metadata standards. This mapping allows for the provision of eLearning services on digital TV systems as well as the reuse of TV programs in order to build educational experiences.

### MPEG-7-related Work

An analysis of the applicability of MPEG-7 descriptors to the existing video annotation tools that are based on home-grown XML annotation formats was carried out. Based on MPEG-7, a modelling language for magazine broadcasts has been specified. It is capable of describing classes of telecasts, instead of specific telecast instances, for automatic segmentation into semantic structural elements.

A Java class framework has been implemented for the modelling of MPEG-7 descriptions (MDS, Video, Audio). These can be stored in an implemented persistence management framework for media descriptors.

## **Other Developments**

An automated image classifier based on SVM techniques has been designed and realized. An automatic region grouping method for improving semantic meaning of features using psychology laws has been developed. The classifier has been integrated in the MILOS Content Management System, which is also available as a demonstrator through the DEMOS portal. It is described in [Section 4.1](#) of the main feature.

For video analysis, annotation, and retrieval, a prototype video content management system, named VCM, has been developed. It is available through the DEMOS demonstrator portal, and is described in [Section 4.6](#)

A multimedia authoring tool has been defined, which supports content-based constraints for personalizing the presentation of multimedia objects according to users' preferences and skill level.

A prototype system was developed to explore the multimedia content of a digital library (images, text, videos, and audio) relating to theatrical works in 19th Century Milan and which supplies a VR (Virtual Reality) interface (namely, a reconstruction of a 19th Century Milanese theatre).

A front-end of a music search engine has been developed, which is accessible through a web browser to allow users to interact using a query-by-example paradigm. Moreover the typical query-by-humming paradigm is also supported. A preliminary version of a component for semi-automatic extraction of song metadata (title, lyrics, cover) from ID3-tags and by querying via web services has also been created. Methodologies for music indexing and retrieval have been extensively evaluated, based on a data fusion approach, with encouraging initial results.

Preliminary tests on the use of APIs provided by Web-based CD dealers were made to examine the potential of automatic creation of a network of composers/performers with scope for extracting information about their similarities, and reflecting to customers' behaviour.

Feature extraction systems for audio content, named Marsyas and SOMEJB, have been installed and tested. Evaluation measures on a larger sample collection based on audio files have been collected and will subsequently be used to define scenarios for interactive retrieval and evaluation of retrieval performance in different scenarios.

An audio classification framework for the participation in the International Conference on Music Information Retrieval (ISMIR) audio contests in the disciplines of Rhythm Genre and Artist detection, has been implemented. It was awarded winner of the Rhythm Classification Competition, was ranked fourth in the genre classification contest, and was again winner in the "stress-test" performance of the genre classification contest. A corresponding demonstrator is available through the DEMOS portal. It is described in more detail in [Section 4.5](#) in the feature.

A web crawler, which is based on APIs provided by a major Web Search Engine, has been developed to create a collection of MIDI files automatically, to be used as a testbed for Music Information Retrieval techniques. When launched, the crawler is able to collect and store thousands of MIDI files in a database, partially overcoming the classic problem of lack of test data.

A syllable-based speech recognition engine for English has been developed. A speech recognizer named ISIP was trained with huge amounts of American English broadcast data. Hidden-Markov-Models were used forming context-dependent cross-word-triphone models. The syllable inventory was generated using tools from NIST. The syllable recognition rate is 88.0%. A syllable retrieval system could be implemented with the syllable recognizer, similar to what has been done for German.

## **NIST TRECVID Evaluation**

Delos members participated in the 2004 NIST TRECVID evaluation - the de facto international standard benchmark for content-based video retrieval. Members participated in the feature extraction task, the shot detection task, and the search task. For the latter task the UvA TRECVID Semantic Video Search Engine was developed, showing the effectiveness of the approaches to content-based retrieval by audio-visual libraries, as well as the parallel implementation thereof. The Semantic Video Search Engine is described in the feature, [Section 4.4](#), and is accessible through the DEMOS portal. The shot detection algorithms implemented for TRECVID participation are also available through the portal. They are referred to in [Section 4.6](#).

## Other Advances

Several software components have been continuously refined. These include software for 3D objects modelling and retrieval, as well as tools for MPEG-7 manual annotation of videos and real-time automatic video annotation, in particular for soccer video analysis. Further improvements have been done on automatic audio-visual metadata extraction tools.

Advances have been made with the development of a test-bed and demonstrator for the extraction and integration of most of the MPEG-7 standard visual descriptors. The output of the demonstrator is collected in an MPEG-7 stream and testing on the interoperability is being analyzed.

Other work has included the following:

- Improvements of the ISIS/OSIRIS system for easier DL maintenance and deployment were made. An automatic/dynamic process will take care of visual feature extraction within ISIS.
- Issues relating to the computational requirements and parallelization of emerging applications in the field of audio-visual digital libraries have been investigated, as well as issues relating to the automatic detection of semantic concepts in multi-modal video repositories.
- Various music information retrieval frameworks have been set up and music retrieval performance on benchmark datasets has been evaluated.
- A study of a model for the specification of synchronized multimedia presentations and of methods for automatic and semi-automatic presentation generation has been started
- Documents from public forums, relating to DLs and describing technological innovation and available prototypes, are collected. These are in the process of being catalogued and indexed to provide fast access to public knowledge.

Readers are referred to the [contents of the feature](#) in this issue to which this summary relates. For further information on the above report, please contact the [cluster leaders](#).

## Author Details

### George Ioannidis

Technologie-Zentrum Informatik (TZI)

University of Bremen

Germany

url <http://www.tzi.de>

email: [george.ioannidis@tzi.de](mailto:george.ioannidis@tzi.de)

DELOS  
Name of Cluster: Network of Excellence  
Audio/Visual and Non-traditional Objects

Name of Leaders: Stavros Christodoulakis

email: [stavros@ced.tuc.gr](mailto:stavros@ced.tuc.gr)  
phone: +30-2821-037399  
fax: +30-2821-037399  
url: <http://www.music.tuc.gr>

Alberto del Bimbo

email: [delbimbo@dsi.unifi.it](mailto:delbimbo@dsi.unifi.it)  
phone: +39-055-479 6262  
fax: +39-055-479 6363  
url: <http://viplab.dsi.unifi.it/>

[Return to top](#)

## User Interfaces and Visualization

**Tiziana Catarci** and **Stephen Kimani** describe the outcomes from a questionnaire-based study and gives her conclusions.

### Results from the Questionnaire-based Study

The User Interfaces and Visualization cluster carried out a questionnaire-based study in order to establish the functional and non-functional requirements of digital libraries as described previously in our [report in issue 1](#). In this issue, we report the digital library requirements based on the results that we obtained from the study. (See the study at <http://delos.dis.uniroma1.it/C2/Deliverables/default.aspx>). In particular, we present:

- an overview of the set-up of the questionnaire-based study
- background information regarding DL users and stakeholders who participated in the study and
- an analysis of the results obtained from the results from the study

### Overview of the set-up of the questionnaire-based study

The cluster adopted an online questionnaires-based approach and composed two separate questionnaires, one for digital library end-users and one for digital library stakeholders. The questionnaires were designed such that they could gather information pertaining to user background and demographics; users' current experience; DL functional requirements; and DL non-functional requirements.

### Demographics and user background

There were 45 library end-users (14 female, 25 male, 6 not specified) who responded to the online questionnaire. Most of the respondents ranged from 20 to 55 years in age and they all came from Europe. It was noted that while many of them worked in the field of computer science, the sample contained very divergent respondent backgrounds (from computer scientists to humanities studies and librarians). A small number of participants reported a considerable degree of disability in one or more of the cognitive, intellectual and visual categories. This sample of 45 DL users was also characterised by multilingualism, a high level of education, considerable experience of computing and the Internet, as well as relatively high experience of DLs.

The questionnaire results indicated that the users frequently accessed a digital library and thus they were generally aware of the weaknesses, advantages and drawbacks of current digital systems. In addition, it was noted that as far as the type of access used for data retrieval was concerned, the vast majority could be typified as public or free access, indicating that most users were not willing to pay a lot for retrieving data and knowledge from a digital archive. Moreover, Web access was by far the most popular medium. Two thirds of the DLs identified in the completed questionnaires supported English as the only language of interaction. Slightly less than a quarter of all the DLs mentioned supported both English and some other local national language. Very few DLs identified by respondents offered multilingual support or just the relevant national language.

### Data analysis regarding functional and non-functional requirements

The cluster analyzed the results in order to determine the needs and requirements concerning both the functionality of digital libraries and other non-functional characteristics related to interaction which would be important for user interface design. Toward this end, high- and low-importance requirements were identified for both DL stakeholders and end-users, in order to provide a basis for the development of a taxonomy of functionalities and interaction characteristics which will inform the design, implementation and evaluation of future digital libraries.

Stakeholders appeared to pay particular attention to functions for locating and organising resources, including functions for creating cross-reference links among similar resources, as well as functions for storing metadata about resources and checking for inconsistencies among the DL resources. Another interesting observation was that all accessibility requirements (i.e. for all kinds of disabilities) occupied a significant position in the list of high-priority requirements of DL stakeholders; whereas usability requirements reached only the list of medium-priority requirements, except the need for 'ease of use' of the DL. DL stakeholders also placed all kinds of functionalities related to the administration and management of registered DL users on the list of high-priority requirements. On the other hand, the group of requirements that appeared to be of lower value to DL stakeholders included most of

the miscellaneous functional and non-functional requirements, as well as usability requirements related to novice users. Furthermore, History facilities and multilingual support also proved to be of relatively low value to these DL stakeholders.

DL end-users, just like stakeholders, paid a lot of attention to all types of DL facilities for locating useful information by subject. Nevertheless, they, in contrast to stakeholders, seemed to pay particular attention to certain miscellaneous non-functional requirements such as system performance, security, privacy, safety, and other ethical requirements. Printing and Up- or Downloading facilities were also assigned significant importance by DL users, followed by general usability requirements and accessibility for people with motor impairments (i.e. mobility and dexterity impairments). On the other hand, personalisation did not appear to assume great importance for DL end-users, and facilities for user-to-user communication and collaboration hardly proved of interest at all.

Overall, it appeared that DL stakeholders were striving for enriched functionality, whereas DL users paid more attention to the perceived behaviour and reliability of a DL.

## Conclusions

Three issues surfaced in the analysis conducted. Firstly, while end-users view DLs as personalised environments where privacy is protected, stakeholders appear to view DLs as more collaborative environments. Secondly, the traditional "paper document" metaphor is still seen as prevailing, which may prove a challenge when it comes to a purely digital environment. Finally, there is a conceptual rift between the end-user and the stakeholder in respect of DL non-functional aspects.

The prioritization of requirements identified in this deliverable has the potential to provide a framework for DL user interface design. Toward this end, at least two future steps are planned. First of all, a further extension of the study. Secondly, an investigation into the DL lifecycle in order to get an insight into how a digital library is expected to evolve as regards the interaction of users and stakeholders, as well as how the different phases of the lifecycle relate to both functional and non-functional requirements.

## Author Details

### **Tiziana Catarci**

Cluster Leader

User Interfaces and Visualization Cluster (UIV)

Università degli Studi di Roma "La Sapienza"

E-mail: [catarci@dis.uniroma1.it](mailto:catarci@dis.uniroma1.it)

url: <http://www.dis.uniroma1.it/~catarci/>

Telephone: +39-06-4991 8331

Fax: +39-06-4991 8331

### **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](mailto:kimani@dis.uniroma1.it)

Telephone: +39-06-49918548

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

[Return to top](#)

## Knowledge Extraction and Semantic Interoperability

Martin Doerr provides us with an overview of a comprehensive report on Semantic Interoperability.

The DELOS WP5 cluster on Knowledge Extraction and Semantic Interoperability is finishing a comprehensive report on Semantic Interoperability in Digital Library Systems [1].

The Internet and more particularly the Web have been instrumental in making widely accessible a vast range of digital resources. However, the current state of affairs is such that the task of pulling together relevant information involves searching for individual bits and pieces of information gleaned from a range of sources and services and manually assembling them into a whole. This task becomes increasingly intractable with the rapid rate at which resources are becoming available online.

Interoperability is therefore a major issue that affects all types of digital information systems, but has gained prominence with the widespread adoption of the Web. As far as digital libraries are concerned, interoperability is becoming a paramount issue as the Internet unites digital library systems of differing types, run by separate organisations which are geographically distributed all over the world. Federated digital library systems, in the form of co-operating autonomous systems, are emerging in a bid to make distributed collections of heterogeneous resources appear as a single, virtually integrated collection.

The report defines interoperability very broadly as enabling any form of inter-system communication, or the ability of a system to make use of data from a previously unforeseen source. Interoperability in general is concerned with the capability of different information systems to communicate. This communication may take various forms such as the transfer, exchange, transformation, mediation, migration or integration of information.

Semantic interoperability ("SI") is characterised by the capability of different information systems to communicate information consistent with the intended meaning of the encoded information (as intended by the creators or maintainers of the information system). It involves:

- the processing of the shared information so that it is consistent with the intended meaning
- the encoding of queries and presentation of information so that it conforms with the intended meaning regardless of the source of information

The issue is addressed from the following perspectives:

- Definition and theoretical aspects of semantic interoperability in the light of the information life-cycle in digital libraries
- Enabling factors and technologies to enhance semantic interoperability, as well as relevant actual methods and processes in use or under development
- Implications for a research agenda

SI issues are analyzed from a practical point of view for the following extended list of **information life cycle** elements that reveals the extraordinary relevance of SI in all aspects of Digital Libraries:

1. Creation, modification
2. Publication
3. Acquisition, selection, storage, system and collection building
4. Cataloguing (metadata, identification/naming, registration), indexing, knowledge organisation, knowledge representation, modelling
5. Integration, brokering, linking, syntactic and semantic interoperability engineering
6. Mediation (user interfaces, personalisation, reference, recommendation, transfer etc.)
7. Access, search and discovery
8. Use, shared application/collaboration, scholarly communication, annotation, evaluation, reuse, work environments
9. Maintenance
10. Archiving and preservation

From a theoretical point of view, the report distinguishes SI at three levels of abstraction:

1. **Data structures**, be it for metadata, content data, collection management data, service description data.



2. **Categorical data**, i.e. data that refer to universals, such as classification, typologies and general subjects.
3. **Factual data**, i.e. data that refer to particulars, such as people, items, places.

It shows in the sequence how these levels relate to different problems, methods and systems to achieve SI. Arguments are made that interoperability is always achieved by a reasonable combination of adhering to common standards and providing methods for dynamic interpretation of non-standardized contents. The above levels of abstraction greatly differ in the scale of concepts or data to be integrated. Consequently, standards are more easily promoted for the upper level, whereas the lower levels have to be addressed more by automated, dynamic methods of integration. The report also tries to bridge some gaps between the emerging different terminology of the libraries and the computer science communities for the same concepts.

The analysis of **enabling factors** and technologies to enhance SI begins with the role of foundational and core ontologies. They are not only perceived as a means to improve contents and consistency of terminological systems, but particularly as a means to assist mediation between different data structures and the transition zone of data structures and upper-level terminology.

A central role is played by Knowledge Organisation Systems (KOS) and their use in networks (NKOS), which deserve a particular classification due to the large variety in size and sophistication of intellectual analysis. KOS represent the shared agreement on concepts (categorical data) and important factual data, such as place names, very important people etc. Particular methods to enhance semantic interoperability are KOS transformation, correlation, mapping and others, but also questions of availability and rights are addressed.

An analysis of the role of architecture and infrastructures connects to how communication protocols and central services can support the global communication on standards and shared concepts, starting with metadata registries at level one to gazetteer services at level three.

In particular when discussing implementation strategies of integrated services, standardization, mediation and data warehousing are frequently controversially discussed, each one as the best solution. The report sees these techniques as alternatives with different application characteristics. Therefore a particular chapter is devoted to the pros and cons of these approaches so that designers may have better decision criteria at hand for their specific application.

Finally, some implications for a research agenda are discussed. At least some important areas for further R&D are identified:

Methodologies and tools for schema matching, mapping, and semantic data transformation, including graphical visualization methods to assist domain experts to formulate equivalences following their conceptualization as well as automated tools proposing schema matching to the expert.

Future issues for Thesaurus and KOS protocols include possible provision of more complex services, such as semantic expansion (beyond basic broader and narrower expansion), more advanced natural language functionality for identifying controlled terminology in free text (documents or query), cross-mapping provision (important for semantic interoperability) and possible data-dependent filters such as the number of postings associated with a concept.

The vision of employing imprecise semantic equivalences between multiple KOS (as "switching languages" etc.) requires a revision of query languages and engines in order to control dynamically the respective information loss.

Overall, methods and services are sought that lead to a convergence of global resources to higher states of semantic consistency, against the diverging forces of information isolation, update and local innovation.

## References

1. Available draft: Patel M., Koch T., Doerr M., Tsinaraki C., DELOS Deliverable D5.3.1: *Semantic Interoperability in Digital Library Systems*, February 2005

## Author Details

**Martin Doerr**

Principal Researcher  
 Institute of Computer Science  
 The Foundation for Research and Technology - Hellas (FORTH)  
 Vassilika Vouton  
 P.O.Box 1385  
 GR 711 10 Heraklion, Crete  
 Greece

email: [martinATics.forth.gr](mailto:martinATics.forth.gr)

Tel: +30 2810 391625

Fax: +30 2810 391638

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

[Return to top](#)

## Evaluation

[Sarantos Kapidakis](#) provides a summary of activity across the differing tasks for the Evaluation cluster.

### Task 1: Evaluation Forum

The Evaluation forum website (<http://dlib.ionio.gr/wp7/>) that brings together DL developers and evaluators was harmonized with the DELOS website and guidelines and further content was added. The forum is represented by two distinct virtual spaces, where the communication and collaboration of WP7 members takes place.

The first virtual space consists of the Evaluation cluster website which hosts collections of existing evaluation approaches and testbeds. The collection of existing evaluation approaches (<http://dlib.ionio.gr/wp7/literature.html>), as expressed in the form of a list of publications, related to the evaluation of digital libraries, is available in two bibliographic formats (Harvard and BibTex), which allows for easy inclusion in publications. The collection of existing testbeds and toolkits (<http://dlib.ionio.gr/wp7/testbeds.html>) operates as a linking point to testbed collections for the evaluation of digital libraries or to the results of other research projects. In order to ensure the holistic inspection of the research area and awareness of previous work, these results are carefully selected and reflect diverse forms of research, methodologies, measurements and metrics.

Finally the WP7 website is an area for publishing information to other research communities or to the public. Visitors can read about the aims of WP7, the work already completed, the partners and the events that are organized by the cluster.

The second virtual space implements the discussion forum, which serves as an area of communication among the members of the WP. A list of threads reflecting the general interests of the WP and the specific Tasks enables WP members to communicate in a centralized way.

In addition, a satellite website supporting the WP7 workshop on the evaluation of digital libraries was created (<http://dlib.ionio.gr/wp7/workshop2004.html>) which now contains the presentations from the workshop and the electronic proceedings.

### Task 2: Evaluation Models and Methods

In our Workshop on DL evaluation, which took place in Padova over 4-5 October, 5 keynote speakers covered major aspects of DL evaluation:

- Tefko Saracevic: Evaluation of Digital Libraries: An Overview
- Ann Blandford: Understanding Users' Experiences: Evaluation of Digital Libraries

- Christine L. Borgman: Evaluating the Uses of Digital Libraries
- Edward A. Fox: Towards a Quality Model for Digital Libraries
- Derek Law: Content Evaluation

In addition, 3 representatives from DL-related Integrated projects currently funded by the EU described their plans and expectations concerning DL evaluation. The workshop gave a very good analysis of the state of the art on digital library evaluation. During the final panel session major issues for further research on DL evaluation were discussed. The presentations from the workshop together with the electronic proceedings are online at [http://dlib.ionio.gr/wp7/workshop2004\\_program.html](http://dlib.ionio.gr/wp7/workshop2004_program.html).

### **Task 3: INEX**

In 2004, INEX consisted of three tracks: ad-hoc retrieval, iTrack focusing with interactive retrieval, and the heterogeneous track dealing with heterogeneous collections. Retrieval effectiveness and efficiency are the evaluation criteria currently considered. Appropriate evaluation measures that consider structural relationships between different answers are still to be developed, as well as usage-oriented measures for the interactive track.

#### **Ad-hoc Retrieval Track**

For the ad-hoc retrieval track, the following steps were performed: The creation and selection of topics, the submissions of runs, the pooling of runs, and relevance assessments by the participating groups. Computation of results is currently underway.

#### **Interactive Track**

In the interactive track, a base system has been developed and tested, topics were selected based on the ad hoc topics, questionnaires were created and guidelines written. Each of the 10 participants ran the base system with at least 8 users and submitted their results and their interaction logs along with the completed questionnaires. In addition, some participants ran their own interactive system with additional users. Evaluation of results and analysis of the interaction logs is underway.

#### **Heterogeneous Track**

For the heterogeneous track, six subcollections from different DLs were gathered. Then topic selection guidelines were created before topics were selected from the ad-hoc track; furthermore new topics were created. After topics had been distributed and submission guidelines had been formulated, participating groups submitted their runs.

#### **INEX Annual Workshop**

The INEX annual workshop took place in Schloss Dagstuhl on December 6-8. Working notes for this workshop are available at <http://inex.is.informatik.uni-duisburg.de:2004/pdf/INEX2004PreProceedings.pdf>. The final proceedings will appear in the Springer series 'Lecture Notes in computer science'. At the workshop, participants presented their work, results were discussed and the evaluation campaign for 2005 was planned.

### **Task4: CLEF**

For the 2004 evaluation campaign, 6 different evaluation tracks were defined to assess different aspects of the MLIR (Multilingual Information Retrieval) paradigm:

- Ad Hoc (monolingual/bilingual/multilingual free text retrieval)
- GIRT(domain/specific retrieval with controlled vocabulary)
- iCLEF (interactive retrieval)
- QA@CLEF (question answering)
- ImageCLEF (image retrieval) and
- CL-SDR (spoken document retrieval)

CLEF activity included: setting up and managing the CLEF 2004 website; preparing and distributing Calls for Participation; preparing and extending the data collection (a Portuguese newspaper collection for 1994/95 has been added to the existing multilingual comparable corpus in nine European languages); preparing and distributing topics for the Ad Hoc and GIRT tracks; receiving and analysing the results. In addition, DELOS supported the

overall coordination of the other six tracks - managed on a voluntary basis by research groups with expertise in the areas covered - and the organisation of the CLEF2004 workshop held in Bath, UK, 15-17 September (immediately following ECDL2004).

Of the 64 registered, 55 groups submitted results:

- Europe 38
- North America 13
- Asia 4

This represented an increase on the 42 groups in CLEF2003. 96 people attended the workshop. 15 European and 2 North American research groups collaborated in the organization under the overall coordination of ISTI-CNR, supported by DELOS. The test collections (consisting of data, queries and relevance assessments for a number of tasks) were expanded. The main multilingual comparable document collection now contains nearly 2 million news documents in ten European languages; new collections were added for cross-language image retrieval.

There was a shift of focus from textual document retrieval to information extraction and multimedia retrieval over languages. Evaluation methodologies were tested for the two new tracks (cross-language question answering and cross-language retrieval in image collections via a combination of text- and content-based methods). The aim has been to stimulate research towards next-generation CLIR (Cross-Language Information Retrieval) systems. A CLEF Steering Committee meeting was held in Bath on 15 September 2004.

## 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](mailto:sarantos@ionio.gr)  
Tel: +30 26610 87413  
Fax: +30 26610 87436

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

[Return to top](#)

[DELOS Home](#)

[DELOS Newsletter Front Page](#)

[Delos Newsletter Contents](#)

Publication Date: June 2005

Publication: Delos Newsletter Issue 3

Originating URL: </projects/delos/newsletter/issue3/cluster-reports/intro.html>

File last modified: Tuesday, 12-Jul-2005 11:17:48 BST

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



# Promoting Knowledge in the Fields of Information Access and Personalization: Past and Upcoming Events

**Georgia Koutrika** collates reactions to past events organized by IAP and gives information on events planned for the coming months.

Promotion of knowledge about available and new practices in the fields of information access and personalization in digital libraries is a long-term goal of the cluster activities. Towards this end, information sharing, and dissemination activities have taken place in the past, and more are planned in for the near future.

## Past Events

### 3rd International Summer School on Digital Library Technologies (ISDL 2004) Pisa, Italy, 6-10 September 2004

The 3rd International Summer School on Digital Library Technologies (ISDL 2004), focused on 'User-Centered Design of Digital Libraries' in collaboration with User-Interface and visualization cluster. The school was directed towards members of the research community in the wide sense, i.e. graduate students and young researchers and professionals involved in R&D in digital library-related areas.

30 people took part. These represented mostly the academic computer science community, but with participants coming from the industry and the user communities interested in digital library technologies (libraries, archives and museums). Seven lecturers from various user environments analyzed a broad set of requirements for digital library systems and discussed the state of the art in several user-related themes such as user interfaces, information visualization, personalization, and recommendation. One of the participants, Jehad Najjar of Katholieke Universiteit, Leuven, gives us his personal view of the event:

*"I really enjoyed the summer school! I'm not doing my research in the field of digital libraries, but in the field of learning objects repositories, which is a special sort of digital libraries. I'm using empirical analysis to evaluate the actual use made of learning objects and metadata in learning object repositories. This summer school was a wonderful opportunity for me to:*

- *learn more about digital libraries from experts who lectured in the school.*
- *meet key persons from the field of digital libraries and human computer interaction.*
- *have interesting discussions and get useful feedback from colleagues in the field of digital libraries.*

*It was a great opportunity for me to meet Prof. Alan Dix (Lancaster University, UK). I enjoyed his interactive lectures on user interaction design and information visualization as well. Also, I enjoyed the discussions we had with him during the coffee breaks and at the dinner we had in the city of Pisa. I also appreciated the chance to meet Prof. Yannis Ioannidis (University of Athens, Greece) who gave a very useful lecture on personalized systems behaviour, which is quite relevant to my research. Meeting Prof. Barry Smyth (University College Dublin, Ireland) was also a very welcome opportunity. He presented a number of interesting personalization applications being*

*developed by his research group. I particularly liked the I-Spy personalized search engine <http://ispy.ucd.ie/fc?actionId=index>*

*The School social events were perfectly organized! Thanks to the organizers from ISTI-CNR (Pisa, Italy), DELOS group and to the Italian students whom I met there! This year I'm presenting part of my research at the ECDL 2005 conference in Vienna. I hope to have make further contacts among people from the community of digital libraries there!"*

## **Future Digital Library Management Systems Workshop Schloss Dagstuhl, Germany, 29 March - 1 April 2005**

The workshop entitled Future Digital Library Management Systems was a collaborative effort between the Digital library Architecture and Information Access and Personalization clusters. This workshop, the 8th in the DELOS series of Thematic Workshops, was devoted to two critical themes: Information Access and System Architecture. The General Chair, Gerhard Weikum, and the two Programme Co-Chairs, Yannis Ioannidis and Hans-Jfgrg Schek had compiled an interesting programme with a keynote (Peter Buneman), three invited talks (Stefan Gradmann, Keith Jeffery, and Michalis Vazirgiannis) and four regular sessions.

With respect to system architecture, the topics of primary interest were Peer-to-Peer (P2P) Data Management, Grid Middleware (Grid), and Service-oriented Architecture (SoA). Within respect to information access, among the key issues of interest were organization of diverse types of information within an individual source, efficient and effective search and support of new trends in user interfaces for future systems.

The objective of the workshop was to bring together researchers and practitioners interested in these two areas and how they relate to one another. Principal aims were to identify fundamental system services that allow the development and operation of future digital libraries, and to explore the main directions in which technical development was likely to go. One of the participants, Harald Krottmaier of IICM, Graz University of Technology, gives us his personal view of the event:

*"The first day started with two sessions where both architectural issues and semantics and context in the field of digital libraries were discussed. We heard eight presentations from DELOS partners and they presented actual scientific results. On the second day another regular session took place related to search and indexing. Four papers were presented. Thereafter we discussed working group topics. Because of the bad weather the planned (and, as we heard traditional) hiking tour didn't take place and we continued our discussions. Another search and indexing session happened on the third day with four presentations. On the final day the results of our working group discussions were presented.*

*Some 50 participants took part in the workshop. Dagstuhl as a venue is fantastic: the largest computer-science related library in Germany is located there where there is 24/7 open access to the library. Some clever touches (e.g. randomised seat arrangement) made it possible to communicate intensively with other workshop participants as compared with traditional conferences. Summarising it was a very successful workshop with many discussions and interesting presentations."*

The programme and technical details of the two events can be found as follows:

DELOS Summer School <http://www.delos.info/eventlist/isdl3.html>

Dagstuhl workshop [http://ii.uit.at/research/delos\\_website/8thworkshop.html](http://ii.uit.at/research/delos_website/8thworkshop.html)

## **Upcoming Events**

The workshop entitled Personalization in Digital Libraries, originally planned for the summer of 2005 has been postponed for the summer of 2006. The reason for the change is to avoid conflict with the Personalized Information Access workshop that has been independently initiated as part of the User Modelling Conference, in Edinburgh, UK, 24-25 July 2005. This event will now be co-sponsored and co-organized by DELOS, hence maximizing the breadth of the relevant networking activities.

## **PIA 2005 - Workshop on New Technologies for Personalized**

## Information Access, in Edinburgh, UK 24-25 July 2005

As the amount of information available, especially on the Web and in modern digital libraries, is increasing at a tremendous rate, the roles of user modelling and personalized information access are also increasing in significance. Equipped with user modelling tools capable of comprehending specific user information needs, new retrieval tools will be able to filter out irrelevant information more effectively, rank information in the most suitable way, compare the contents of different documents, personalize information presentation and to tailor man-machine interaction more satisfactorily. A range of new technologies for personalized information access is responding to these new challenges within all information access paradigms - from classic "ad-hoc" information retrieval to information filtering, browsing, and visualization.

The goals of this workshop are to intensify the exchange of innovative ideas between the different research communities involved, to provide an overview of current activities in the area of personalized information access, and to point out connections between them. The workshop will focus in particular on researchers working on ontologies, computational linguistics, user modelling and profiling, user-adaptive interfaces, digital libraries in their various combined forms.

This workshop is collocated with the International Conference on User Modeling 2005 and is extremely interesting and relevant to the field of digital libraries. Therefore, in addition to the above, it is a welcome opportunity for DELOS colleagues active in the area of personalization to exchange ideas with the user modelling community and attend the UM conference.

### Acknowledgements

My thanks to Jehad Najjar, Computer Science Department, Katholieke Universiteit Leuven, Belgium, Harald Krottmaier, Institute for Information Systems and Computer Media (IICM), Faculty of Computer Science, Graz, University of Technology, Austria and Yannis Ioannidis, leader of the Information Access and Personalization cluster, for their contributions to this report.

### Author Details

**Georgia Koutrika**

University of Athens

Email: [koutrika@di.uoa.gr](mailto:koutrika@di.uoa.gr)

Telephone: +30 210 727 5242

Fax: +30 210 727 5214

[Return to top](#)

[DELOS Home](#)

[DELOS Newsletter Front Page](#)

[Delos Newsletter Contents](#)

Publication Date: June 2005

Publication: Delos Newsletter Issue 3

Originating URL: </projects/delos/newsletter/issue3/iap-events/intro.html>

File last modified: Monday, 11-Jul-2005 15:09:40 BST

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



# DELOS 9th Thematic Workshop: Digital Repositories Crete, May 2005

**Rachel Heery** provides an overview of the 9th Thematic Workshop on Digital Repositories entitled Interoperability and Common Services.

## Introduction

The 9th DELOS Network of Excellence thematic workshop on Digital Repositories: Interoperability and Common Services took place over 11-13 May, 2005 hosted by FORTH near Heraklion on the Greek island of Crete. This workshop was organised jointly by the Semantic Interoperability (KESI) and Preservation (PRESERV) clusters of the DELOS Project and sought to provide different communities with the opportunity to learn about the latest research developments in this field. The workshop website with links to papers and presentations from the [programme](#) is available at:

<http://www.ukoln.ac.uk/events/delos-rep-workshop/>



**Panorama of Heraklion, Crete.**  
(Photo courtesy of Martin Doerr, FORTH)

## Workshop Themes

The workshop included eleven submitted papers focusing in particular on the role of repositories within e-learning and e-research and related digital library services. The papers examined the role of such repositories as providers of both preservation and access services.

In an opening keynote presentation, Sandy Payette, Co-director of the Fedora Project spoke on Beyond Storage: Rethinking the role of repositories in scholarly communication that suggesting we stand away from our traditional perspective of digital repositories as document-centric storage mechanisms and consider their role within a broader service-oriented architecture that enables integration of data, content, and services. Sandy's vision is that the current evolutionary development of the scholarly communication process might be revolutionised by a more process-oriented approach. Implementing repositories as flexible compositions of services based on a rich resource object model, would facilitate tasks such as content repurposing, and information reuse.

In a second invited presentation, Leona Carpenter, acting JISC Programme Manager, gave an overview of repositories which are being investigated and implemented as part of an of an integrated infrastructure for education within the JISC Information Environment. Two funding programmes within the UK are particularly focusing on repositories: one is supporting digital preservation and asset management in institutions and the second is a new Digital Repositories Programme aiming to embed repositories within the everyday information landscape of Higher Education.



There were presentations across a number of themes including: Repository Workflow and Design where both the optimisation of metadata workflows in a distributed environment and information design for cultural documentation were examined. In the theme Institutional Repositories experiences of building an open access institutional repository in a UK scientific research organisation were shared. Delegates also learnt of DEPTAL as a framework for institutional repositories as well as the significance of SHERPA DP as a means of establishing a persistent preservation environment for institutional repositories. In another paper on the theme of preservation, delegates also heard of the role reliability modelling had to play in the field of long-term digital preservation.

Finally the theme of Providing Services also offered presentations on the use of RDF aggregation and querying to extend the Open Journal System software to provide additional information about authors, searching heterogeneous e-learning resources and two shorter presentations on the OAI service provider, METALIS, and the digital object lifecycle in the context of the dLibra Digital Library Framework.

In the course of the three days the conference not only covered specific work undertaken by a wide variety of research groups but also included breakout groups that looked at a range of issues and challenges confronting the community with respect to digital libraries. The issues included for example how repository research fits within the wider digital library research agenda and with the development of institutional management systems. There was also discussion on the potential for research collaboration within the context of digital repositories and in particular whether national funding of repository research and development limited the scope for sharing experience and best practice.

There was some consideration, as one might expect, of what exactly was understood by the term repositories and how they fit in relation to other services. Were they typically components of other services? Discussion moved to identifying common services for a repository reference model.

The workshop concluded with a review from Liz Lyon, Director of UKOLN, of the challenges facing research, in particular considering how repositories fit into the knowledge cycle of scholarly communication from research to learning. Discussion initiated at the workshop is being continued by means of an online forum mailing list.

## Author Details

### **Rachel Heery**

Assistant Director, Research and Development

UKOLN

University of Bath

Bath

BA2 7AY

UK

E-mail: [r.heery@ukoln.ac.uk](mailto:r.heery@ukoln.ac.uk)

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

Telephone: +44 (0) 1225 386724

Fax: +44 (0) 1225 386838

[Return to top](#)

[DELOS Home](#)

[DELOS Newsletter Front Page](#)

[Delos Newsletter Contents](#)

Publication Date: June 2005

Publication: Delos Newsletter Issue 3

Originating URL: </projects/delos/newsletter/issue3/heraklion/intro.html>

File last modified: Thursday, 07-Jul-2005 10:13:10 BST

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



---

## 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.

- [DELOS Welcomes Proposal to Create European Digital Library](#)
- [Presentations Available from Joint Workshop on Electronic Publishing](#)
- [DELOS Digital Repositories: Heraklion Workshop Presentations Now Available](#)
- [Programme Overview for ECDL2005 Reveals Topics for the Conference](#)

[Return to top](#)

---

### DELOS Welcomes Proposal to Create European Digital Library

A recent letter of proposal to create a European Digital Library has been welcomed by the DELOS Network of Excellence. The letter to the current European Council and Commission presidents from the French, Italian, Spanish, German, Polish and Hungarian heads of state proposes making works held in European libraries available online in a European digital library. Giving his reaction to the proposal, Costantino Thanos, scientific coordinator of the Network, remarked, "That all citizens, anywhere, anytime, should have access to Internet-connected digital devices to search all of human knowledge, regardless of barriers of time, place, culture or language has been the vision of DELOS since its inception. DELOS believes that, in the near future, networked virtual libraries will enable anyone from their home, school or office to access the knowledge contained in the digital collections created by traditional libraries, museums, archives, universities, governmental agencies, specialised organizations, and individuals around the world."

Costantino went on to elaborate on the potential of the next-generation digital libraries DELOS has in its sights: "These new libraries will offer digital versions of traditional library, museum and archive holdings including text, video, sound and images, 2D and 3D objects. But they will also provide powerful new technological capabilities that enable users to refine their requests, analyse the results, access collections in other languages, share resources, and work collaboratively. No matter where the digital information resides physically, sophisticated search software can find it and present it to the user on demand. A European Digital Library is a step towards making this vision a reality."

[Return to top](#)

---

### Presentations Available from Joint Workshop on Electronic Publishing

A reminder to delegates and all readers that the presentations given by speakers at the Joint Workshop on Electronic Publishing organised by DELOS, SVEP and ScieCom at Lund University in April 2005 are now available.

They can be accessed via the [Workshop Programme page](#). With some 90 participants mostly from Scandinavia but also from Germany, the Netherlands, Switzerland and beyond, the workshop covered a number of themes including Integrating Advanced Features into Institutional Repository Platforms, e-Journal Publication Software, e-Publishing Policies, Author Support and Long-term Access and Preservation.

[Return to top](#)

---

## DELOS Digital Repositories: Heraklion Workshop Presentations Now Available



**Panorama of Heraklion, Crete.**  
**(Photo courtesy of Martin Doerr, FORTH)**

The 9th DELOS Network of Excellence thematic workshop on *Digital Repositories: Interoperability and Common Services* took place over 11-13 May, 2005 in Heraklion on the Greek island of Crete. This workshop was organised jointly by the Semantic Interoperability (KESI) and Preservation (PRESERV) clusters of the DELOS Project and sought to provide different communities with the opportunity to learn about the latest research developments in this field.

We have received a [report on the workshop](#) from Rachel Heery and presentations and papers are available from the workshop website available at:

<http://www.ukoln.ac.uk/events/delos-rep-workshop/>

[Return to top](#)

---

## Programme Overview for ECDL2005 Reveals Topics for the Conference

It is now possible to see the tutorials, papers and posters envisaged for the European Conference on Digital Libraries to be held in Vienna over 18-23 September 2005 and organised by TU Wien.

The whole conference is prefaced by a series of [tutorials](#) which this year will be concentrating on an Overview and Formal Framework of Digital Libraries, Context-enhanced Digital Library Services, Thesauri and Ontologies in Digital Libraries and Building Digital Library Collections with Greenstone.

The main body of the conference takes place over 19-21 September and offers delegates, in addition to keynote speeches and panels, two tracks to choose from on each day. Over the 3 main days this means some 20 papers per track will be available to each delegate attending the whole programme. Consequently delegates need to choose between competing themes:

- Digital Library Models and Architectures - or - Multimedia Digital Libraries
- XML - or - Building Digital Libraries
- User Studies (1) - or - Digital Preservation
- User Studies (2) - or - Metadata
- Digital Libraries and e-Learning - or - Text Classification in Digital Libraries
- Searching - or - Text Digital Libraries

On the first day of the main conference delegates are invited to attend two panel sessions. The first will consider [Does eScience need digital libraries?](#) and will promote discussion on the role that digital libraries can and should

play in the emerging eScience computational infrastructure. The panel will explore how the eScience and digital libraries research communities might work together to support the conduct of eScience and move towards a shared research agenda. The second will consider the question [Digital Libraries over the GRID: Heaven or Hell?](#) There has been considerable attention given to computation in the Grid environment and yet the literature offers a good deal less on information and service management of the kind required by digital libraries. Consequently the aims of this second panel are to identify the opportunities and problems that derive from the development of the Grid and examine whether its benefits outweigh the disadvantages.

There will be a series of workshops on a range of themes including Digital Libraries in Health Care, Knowledge Extraction and Deployment for Digital Libraries and Repositories, Web Archiving and Digital Preservation, Cross-Language Evaluation Forum (CLEF) as well as the NKOS Workshop on Mapping Knowledge Organisation Systems. The [workshops](#) will begin directly after the close of the main conference.

Added to which the conference will be offering delegates the opportunity to view in excess of 40 [posters and demonstrations](#) at time of writing. These will range from overviews of projects and software to directions and outcomes of research activity in areas across the board from metadata, ontologies, thesauri, models, policy and issues such as digital preservation to impact evaluation, cost estimation and beyond.

[Return to top](#)

[DELOS Home](#)

[DELOS Newsletter Front Page](#)

[Delos Newsletter Contents](#)

Publication Date: June 2005

Publication: Delos Newsletter Issue 3

Originating URL: </projects/delos/newsletter/issue3/news/intro.html>

File last modified: Thursday, 07-Jul-2005 10:36:12 BST

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