The Emerald Research Register for this journal is available at
www.emeraldinsight.com/researchregister

The current issue and full text archive of this journal is available at
www.emeraldinsight.com/0264-0473.htm

Development of a medical digital library managing multiple collections

Mara Nikolaidou
Libraries Computer Centre, University of Athens, Athens, Greece

Dimosthenis Anagnostopoulos
Harokopio University of Athens, Athens, Greece, and

Michael Hatzopoulos
Department of Informatics and Telecommunications, University of Athens, Athens, Greece

Abstract

Purpose – Aims to present the authors' efforts towards the development of a digital library environment supporting research at the Medical School of Athens University, Greece.

Design/methodology/approach – The digital library facilitates access to medical material produced by laboratories for both research and educational purposes. As the material produced varies (regarding its type and structure) and the search requirements imposed by potential users differ, each laboratory develops its own collection. All collections must be bilingual, supporting both Greek and English. Extended requirements were imposed regarding the services offered by the digital library environment, due to the following reasons: end-users actively participate in the cataloguing workflow; cataloguers should be able to create and manage multiple collections in a simplified manner; and different search requirements must be supported for different user groups. To formulate and then deal with these requirements, the authors introduced the term "dynamic collection management" denoting automated collection definition and unified collection management within an integrated digital library environment. Digital library components providing the desired functionality and the interaction between them are described. System performance, especially during collection search, and bilingual support are also explored.

Findings – Finds that Athens Medical School Digital Library facilitates access to medical material to researchers and students for both research and educational purposes.

Originality/value – The paper provides useful information on a digital library environment which supports research.

Keywords Collections management, Digital libraries, Medical information systems, Greece

Paper type Case study

1. Introduction

Athens Medical School (AMS) is the largest medical research institution in Greece, where educational activities are combined with everyday practice at the University Hospital. Athens Medical School Digital Library (AMS DL) has more than 700 faculty

The authors would like to express their gratitude to Professor Dr C. Kittas, head of the Medical School, and Associate Professor P. Marinos, researcher in the Laboratory of Histology, for his extensive help during system design. This work is dedicated to the memory of D. Anastasiou, who actively participated in this project before his tragic death.

The Electronic Library
Vol. 23 No. 2, 2005
pp. 221-236

0264-0473
DOI 10.1108/02640470510592933

members and researchers, more than 3,500 undergraduate students and more 800 postgraduate students. A total number of 12 medical laboratories operate in three University Hospitals and the Medical School and produce a large amount of material, mainly medical images and, in some cases, videos, in digital format. This material should be exploited for both research and educational purposes. Digital library systems facilitate the management of large collections of digital material and resources providing advanced accessing capabilities (Borgman, 2002). Medical applications benefit considerably from this technology, as they require organizing and maintaining large amounts of images and videos. In this paper, we present the digital library system, henceforth named AMS DL, built to support the digital collections of AMS laboratories. The system was developed as a joint project between the Libraries Computer Centre and the Department of Informatics and, among other things, provides web access capabilities to researchers and students on all medical material types.

Typical medical image archive systems, such as the ones presented in Suh *et al.* (2002) and Bristol University (2000) facilitate access to medical material, mainly images, using web-based multi-tiered architectures. All material added in such systems is part of a single collection. All images are described by a core metadata set, while no specific research or educational characteristics are maintained. In the case of AMS the required functionality was considerably extended, as:

- AMS DL must provide primal material (images or videos) to teaching staff for presentation in lectures and allow students to search for material relevant to a specific subject and familiarise with it. Thus, the material added in the library must be characterised by educational and specialized research properties useful for specific medical areas. Metadata management becomes more complicated, as combinations of different metadata schemes, such as Dublin Core (Dublin Core Metadata Initiative, n.d.) or IEEE Learning Object Metadata (IEEE, 2002), may be used to describe collections.
- The digital library should support more than one collection. Collections must follow the specific orientation of individual laboratories. Thus, specific features must be provided for individual collections, determining the type and structure of the digital material. Although a core metadata element set can be identified, the metadata used to describe collections differ. To facilitate multiple collection search, the definition of crosswalks between metadata schemes should be supported (Yu *et al.*, 2003).
- The number and nature of supported collections are not predefined. Thus, designing an overall digital library architecture, we need to consider that dynamically creating and administering collections is required. The librarian administrating AMS DL should be able to easily create and manage independent collections using a graphical interface with no programming effort.
- Interaction with a new collection in the Digital Library should be seamlessly supported. Thus, existing services should be functional for all collections without additional programming effort.
- The nature of AMS DL is not archival, as a high degree of interaction with the system is supported for both researchers working in laboratories and librarians working in the library. In the aforementioned medical archive systems (Suh *et al.*,

2002; Bristol University, 2000), material is gathered and processed by a specialised unit before added to the system. AMS DL environment supports a more complex workflow, as research material is added directly by the researcher, while he/she also participates in metadata creation. Furthermore, this material may be used to create presentations and on-line tutorials, which are stored in the Digital Library.

- All features provided need to be bilingual. Both Greek and English languages are supported in material characterisation and the user interface.

To deal with the extended requirements imposed in a large-scale digital library, we have introduced the term "dynamic collection management" and implemented the respective features. This term is used to denote automated collection definition and unified collection management within an integrated digital library environment. Dynamic collection management facilitates:

- Implementation of the provided digital library services under common guidelines in an open environment, where all services operate parametrically supporting any collection.
- Integration of heterogeneous collections in terms of material structure and metadata support, as crosswalks between collection descriptions may be defined.
- Simplification of collection management, as no programming skills are necessary to create and administer new collections.

AMS DL was build based on a modular architecture similar to the one proposed by Open Archive Initiative (Lagoze and Van de Sompel, 2001). Digital material is stored in repositories providing the basic capabilities for storing and accessing data, while additional functionality is implemented as independent software modules called services. Each service provides a specific functionality and accesses other services and repositories to accomplish its tasks. Within AMS DL framework, dynamic collection management is facilitated through Collection Management Service. Regarding other services offered by AMS DL, special attention was given to the implementation of the supported workflow applications, as researchers actively participate in the cataloguing workflow. Bilingual support affects digital library service implementation and performance. Thus, the efficient performance of the system, especially during collection search, was exploited.

The rest of the paper is organised as follows: in section 2, we present the services supported by AMS DL environment, focusing on collection management and workflow services. Desired functionality is described and implementation requirements are identified. Collection definition process and metadata representation issues are addressed in section 3. Histological Collection definition is discussed as an example to identify the advantages offered by dynamic collection management. AMS DL architecture and corresponding components are described in section 4, along with the implementation of supported services. Conclusions reside in section 5.

2. Supported services

As laboratory requirements differ significantly, it was decided to develop a different collection for each laboratory. This decision extends to medical metadata: while most metadata, such as producer, description and format, are common across collections, collection-specific metadata also need to be supported (Darmoni *et al.*, 2001). Collection administration is performed by a librarian of the Central Library of Health Sciences through the Collection Management Service, analytically discussed in the corresponding section. Since one of the main objectives of the system is to simplify the collection definition process and most collections are partially described by the same metadata set, it is important to allow the librarian to define collections based on existing collection descriptions.

Laboratory staff are responsible for selecting the material to be added in the library. The selection process is performed according to criteria related to specific medical file fields. Thus, selected material needs to be characterised only by content-based properties, which are meaningful in medical research: each image/video is accompanied by a description indicating its significance, written by the researcher in both Greek and English. While the researcher adds new material in AMS DL, cataloguers in the Central Library of Health Science are responsible for processing the material (e.g. creating different analysis images or videos) and filling metadata fields. This process is accomplished in cooperation with the researcher produced the material. The corresponding workflow is facilitated by the Medical Object Processing and Cataloguing Service. Since one of the main purposes of AMS DL is to support teaching activities, laboratory staff should easily use AMS DL material to edit presentations or online tutorials corresponding to a specific lecture. The corresponding workflow is facilitated by the Create Presentation/Tutorial Service. Presentations/tutorials are also stored within AMS DL. Both workflow services are described in the following corresponding section.

Access is granted only to staff and students through web-based applications. Different access privileges are granted according to user profile. According to his/her user group, each user may have access to a different quality format of the same material and may be able to download it or not. Users may search collections using simple or complicated criteria based on metadata fields or digital material content in both Greek and English. Search results are presented as XML pages.

Collection Management Service

In an integrated digital library environment, heterogeneous collections – in terms of structure and purpose – need to be supported. Such issues have been exploited in the Greenstone system (Witten *et al.*, 2001), where the Collector application facilitates collection structure definition. This is accomplished by defining the structure of collection material and the related metadata using a collection directory. Aggregating diverse collection-specific requirements and facilitating access to collections through a common access point enables the unified management of all digital material and promotes interoperability, as has been identified in Arms *et al.* (2002). Arms also introduced the collection dictionary concept to facilitate access to heterogeneous collections supporting OAI PMH (Van de Sompel and Lagoze, 2001). We introduced the term dynamic collection management (i.e. automated collection definition and management within an integrated digital library environment) to integrate the functionality supported by collection dictionary concept in both approaches.

Dynamic collection management allows the integration of heterogeneous collections stored in different systems and their unified administration through a common access point. The creation and management of collections should be accomplished with no programming effort, while existing services should be functional for all collections without any additional programming. Dynamic collection management is facilitated by the Collection Management Service which incorporates capabilities similar to those provided by the Gatherer environment (Bainbridge *et al.*, 2003) and offers extended ones, such as the definition of collections based on existing ones and the handling of different kinds of relationships between collections. In this way, collection definition is simplified. Integration of external collections is also supported. Collection Dictionary and Collection Management Service implementation is independent of the digital library systems used to store collections. To accomplish dynamic collection management, the proposed collection dictionary needs to provide the following capabilities:

- defining collections;
- extending/modifying collection definitions;
- defining relationships between collections (common metadata fields, sub-collection definition);
- accessing collections by a common access point; and
- integrating collections supported by different implementation environments, independently of digital object storing and searching mechanisms.

When a collection is defined, its respective collection description is added in collection dictionary. Collection description consists of three parts: collection properties, object structure and object metadata.

The term “digital object” denotes digital material stored in the Digital Library. Digital objects are usually compound objects/documents consisting of parts of different medium type (e.g. text, image, sound, video), which are indexed by different tools. Any object may also be part of another object. The term “object structure” denotes the skeleton used for the construction of any digital object of a specific collection, regardless of whether object parts are mandatory, thus assuming that all objects belonging in a specific collection have the same structure.

Digital objects in a specific collection are characterized by a common metadata set (both at object and part level). Metadata may be general, i.e. common in all collections, and collection specific. For both general and collection specific ones, we maintain four categories of metadata (Besser, 2002; Niu, 2002):

(1) descriptive, used to describe material;
(2) technical, related to object/part type/format (e.g. image quality properties) and storing properties;
(3) rights, used for access control; and
(4) educational, related to educational categorization (e.g. corresponding course or lecture).

The metadata schemes used may be standard ones (e.g. DC, IEEE LOM), a variation of it or even a local one. Implementation properties, e.g. whether a field is bilingual, field labels in both Greek and English, multi-valued or mandatory, are also included in metadata field description along with the value type of each field. Restricted value lists are also supported for specific fields.

Collection properties involve structural information, relationships between different collections and access information. Access information consist of the system where the collection is stored and the access protocol used. Collection search options and limitations are defined by indicating searchable metadata fields and digital object fields and possible combination restrictions. Relationships between collections are also defined (e.g. sub-collections). Each sub-collection may be described by different metadata schemes. The relationships between them must be explored. Collection interoperability rules may also be defined, corresponding to the definition of crosswalks between the metadata characterizing each collection (Lightle and Ridgway, 2003). Crosswalks are usually partial, as they concern a metadata subset, and are useful in common collection search.

Collection descriptions can be derived from existing ones through extending the object structure and metadata model, i.e. as a descendant of an existing collection description, while additional object parts and metadata fields can be defined. This feature provides flexibility and simplifies collection definition process. A critical aspect of dynamic collection management is the dynamic adjustments of supported services to operate using any collection taken into account its specific structure and properties. Such a feature enhances significantly the incorporation of new collections and the modification of existing ones, as far as both programming effort and time are concerned.

Supported workflow services

Two workflow services are supported: Medical Object Processing and Cataloguing and THE Create Presentation/Tutorial Service as shown in Figure 1. The Medical Object Processing and Cataloguing Service facilitates the co-operation of laboratory researchers and cataloguers to develop collections.

Primal medical material is produced by laboratory equipment in digital format. It mainly consists of TIFF images and, seldom, of videos of various, high-resolution formats. Laboratory staff select the material to be added in AMS DL, compose its description in Greek and English and add it to the proper collection. A cataloguer in the Central Library of Health Science processes the primal image or video. He/she produces different analysis images or videos to fulfil the needs of different user groups and watermarks them. Primal material is accessed only by the researcher who produced it. While doing so, the cataloguer might ask the researcher to review some of the produced images/videos. The researcher will suggest discarding or reproducing this material. Using the produced different analysis versions of the primal image/video and its description, the cataloguer constructs a medical object based on the object structure of the corresponding collection, adds it in AMS DL and catalogues it according to the supported metadata scheme. As the metadata characterising each collection is strongly related to each specific medical area, cataloguers were not in position to accomplish this task efficiently. Thus, metadata characterization is partially done by the researcher. The cataloguer is responsible for determining whether cataloguing is complete and publishing the medical object. The corresponding workflow is shown in

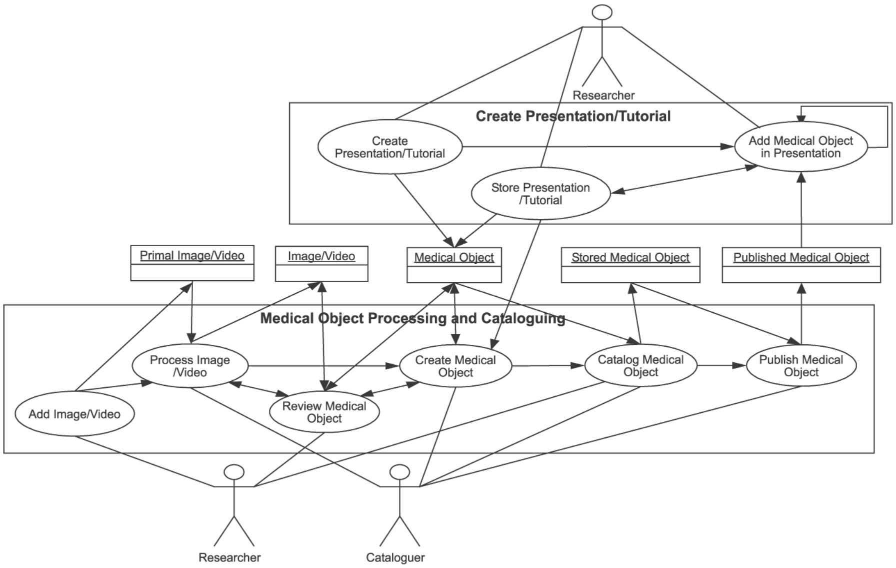

Figure 1.
AMS DL supported workflows

Figure 1 using UML notation. Each medical object entering the workflow is characterised by its state. Only objects characterised as published are actually available for viewing.

Laboratory staff use the material stored in the Library to create presentations and online tutorials for specific lectures. Presentations/tutorials are treated as composite medical objects and stored in AMS DL. The Create Presentation/Tutorial workflow service supports the editing of simple presentations consisting of sequential presentation pages. Each page either presents a medical object or contains comments in text format. The researcher may store the presentation/tutorial as draft and complete its editing at a later time. When the presentation/tutorial medical object is characterized as completed, the cataloguer fills the corresponding metadata fields and publishes it as a medical object. Metadata fields are filled in cooperation with the researcher, who is responsible for editing most of them, especially the educational ones.

Collection search service

Collections search must support the following essential features:

- Simple search on specific metadata fields (e.g. find all images depicting tissues from a human liver).
- Combined search on the same or multiple metadata fields (e.g. find all images depicting tissues from a human liver or lone, or find all images depicting tissues from a human lone having cancer created by a researcher whose name starts with "SMIT" within the last two years).
- Combined search on multiple metadata and data fields (e.g. find all images related with liver cancer).

Multiple collection search must be supported to facilitate searching sub-collections of a specific collection or even related collections. Limitations regarding the material type (e.g. tutorials) and educational characteristics (e.g. course) must also be supported.

Collection Search is accessed by the end-users of AMS DL, thus ensuring its performance is of great importance. Metadata storage and indexing policy and free text search options used to index digital objects text fields have great influence on search performance. The way bilingualism is supported in metadata and digital object parts also affects performance.

3. AMS collections definitions

As all AMS collections include specific material type and a common metadata set was identified, we decided to define a generic Medical Collection and use it as a prototype for the creation of laboratory-specific collections. Each specific collection is defined as an ancestor of Medical Collection, thus only additional features or restrictions need to be defined. This contributes to the simplification of laboratory-specific collections. In the following, we present the Medical Collection description using features facilitated by dynamic collection management and the definition of a laboratory-specific collection, namely Histological Collection, to indicate the advantages offered.

Medical Collection definition

Medical Collection consists of three sub-collections: medical image collection, which includes compound objects consisting of different analysis images (medical image objects), medical video collection, which includes compound objects consisting of different analysis videos (medical video objects) and presentation collection, which consists of presentations edited by researcher using the material included in the two aforementioned collections (presentation objects) (Figure 2). The corresponding collection description must be defined for each sub-collection. The collection descriptions for both medical image and medical video collections are similar. The following parts are included in the medical image/video objects:

- *Original image/video*. The original image/video produced in the laboratory. It is of high quality. It cannot be efficiently transferred over the web and its copyright should be strongly protected. Thus, access to it is restricted.
- *Derivative image/video*. This is produced from the original image/video, usually in JPEG/MPEG format, to enable access through the web.
- *Thumbnail image.* This is shown in Collection Search.
- *Description*. This is in both Greek and English.

The original image and the description are produced by the researcher, while all other formats are produced by the cataloguer during image processing.

Metadata kept for both collections are similar, except for those concerning technical issues – part level metadata. The metadata scheme introduced to describe descriptive and technical metadata of the medical image/video objects is based on Dublin Core (Dublin Core Metadata Initiative, n.d.), although it also supports customisation for medical material. Dublin Core is a widely adopted scheme for medical images archives (Bristol University, 2000) and health care applications (Sakai, 2001; Davenport Robertson *et al.*, 2001). Educational metadata are described using an extension of IEEE

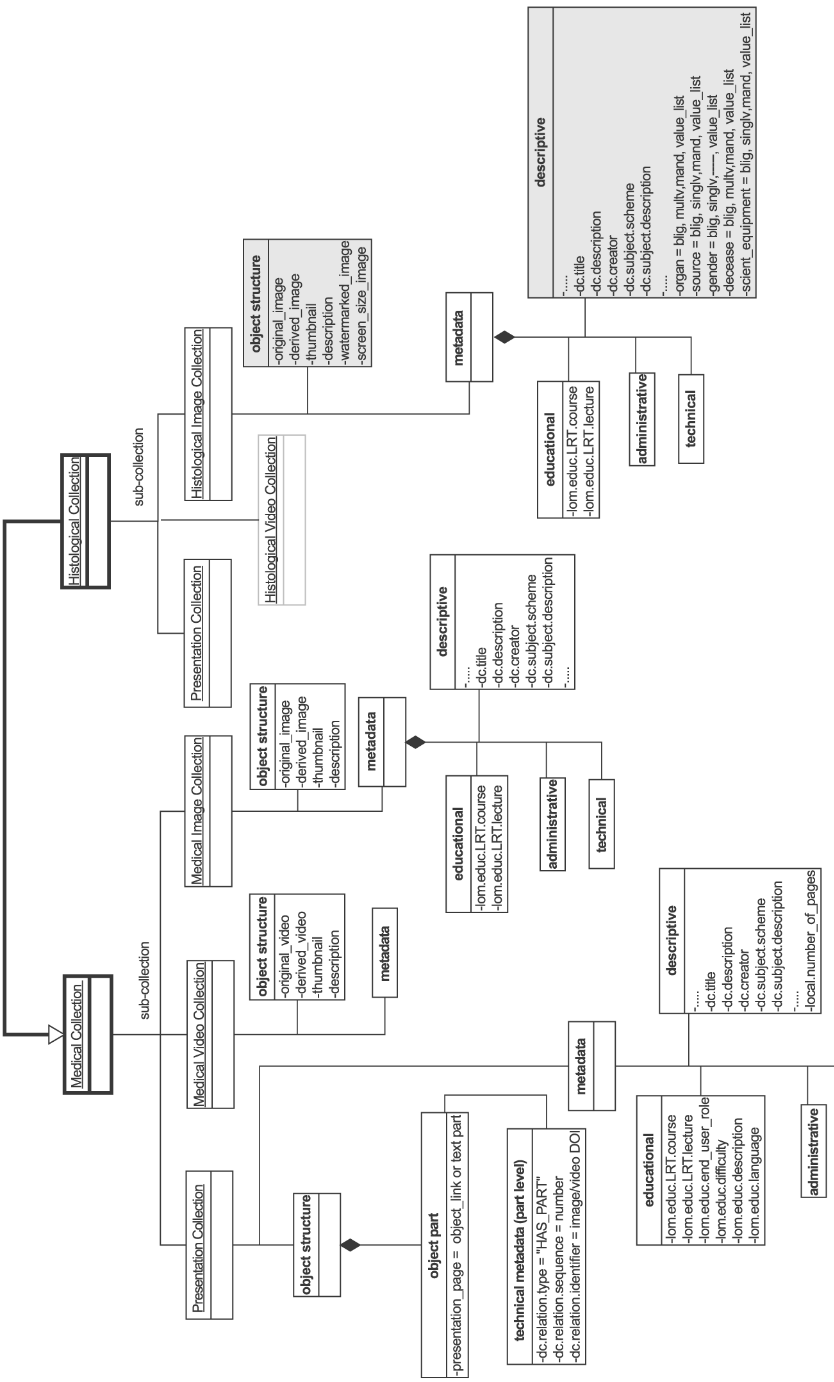

Figure 2.
Medical and Histological
collection descriptions

LOM (IEEE, 2002). Since IEEE LOM is extensively used to describe presentation objects, two fields corresponding to related course and lecture are used in medical image/video description to promote interoperability. Rights metadata are described using local fields.

Presentation objects consist of multiple parts, one for each presentation page. Each part either contains links to video/image objects or is a text part (Figure 2). Dublin Core metadata scheme is used for descriptive and technical metadata. Regarding descriptive metadata, fewer fields were used, while local fields such as Number_of_pages, indicating presentation pages, are also included. Technical metadata associated with this part contains information for its structure. Dublin Core relation field (has_part type) was used for the description of presentation page part. The relation field is extended with the sequence subfield indicating the position of the specific page within the presentation. Educational metadata are more extensive and described according to IEEE LOM educational part.

Collection properties are similar for all sub-collections. Access information indicates that all of them are implemented in IBM Content Management Platform (IBM Corporation, 2000), used for storing all AMS collections. Collection search options and limitations were defined to facilitate the functionality described in the previous paragraph. Simple and combined search is allowed using all metadata fields indicated as searchable and description object part of medical image/video objects. No combination restrictions were defined. Two limitations were defined related to course and lecture fields of the educational metadata. Since the same metadata schemes were used for each metadata category, no crosswalks needed to be defined.

For the Medical Collection only collection properties were defined. Access information was not filled, since the collection has no objects directly related to it. The Collection Search Options definition corresponds to the union of the search options of its sub-collections. No additional restrictions were defined. The Medical Collection is practically empty, while all other collections are easily defined as its descendants by adding collection-specific metadata fields and extending the properties of medical image/video objects. We discuss the definition of the Histological Collection as an example (Figure 2).

Histological collection definition

The Histological Collection was defined to satisfy the needs of the Laboratory of Histology. Its description was derived from Medical Collection Description. Only additional properties and restrictions were defined. There were no alterations in the Presentation Collection description. As the laboratory only produces images, the Medical Video sub-collection remains idle (its usage was deactivated). The Medical Image Collection description was extended, as two new parts were added in the Medical Image Object structure: Watermarked Image, which is produced from the derivative image and is watermarked using the symbols of the University and the corresponding Laboratory; and Screen Size Image, a medium-quality image produced from the derivative image to be easy included in presentations. Collection-specific metadata fields were added in the descriptive metadata. These fields are considered as local, as they are useful only when searching the specific collection. Some of them are: the organs presented in the image, the source of the tissue presented in the image (human, animal) or the decease affecting the tissue. All these fields are searchable. Since the Medical Video sub-collection was de-activated, corresponding search options or limitations are automatically removed from Histological Collection Search Options.

The definition of seven additional collections corresponding to different laboratories was performed in the same way. For each collection, only additional features have to be described. This contributes to the simplification of the overall process. Since most of the laboratories either produce images or videos, the de-activation of specific features of the original collection description proved to be very useful.

Medical object implementation issues

Collection repositories were developed using IBM Content Manager platform. Supporting bilingualism in most metadata fields and restricted metadata value-lists increased system complexity. To support both exact and approximate search in combined multi-valued metadata fields, capabilities of a relational database, such as the one facilitating AMS DL operation, provide a rather poor performance. Thus, database search is applied for exact numerical and date metadata field search, while string exact and approximate search is performed using the free text search capabilities provided. Metadata information is stored within both the underlying database, as field manipulation and presentation capabilities reduce programming effort, as well as a tag-structured text part in the corresponding medical digital object, to improve search performance. The hybrid metadata maintenance scheme provides the performance required by the Collection Search Service.

Medical Object data and metadata internal representation using IBM Content Manager constructs (IBM Corporation, 2000) are presented in Figure 3. Metadata information is stored within the Metadata text part. The Metadata part is indexed using Text Search Server. Different tags are used to support Greek and English languages, while all properties of a specific metadata field, e.g. DC.subject, are included within one tag. For example, the structure of subject field is the following:

<Language_Mark> <DC.Subject_Mark> SUBJECT_VALUE <DC.SubjectType_ Mark> TYPE_OF_SUBJECT <EndOfLine>

All metadata information and text object parts, as Description, are stored in both Greek and English to ensure search performance. Thus, bilingual support adds complexity in object creation. During AMS implementation, we also resolved issues involving full support of Greek language and the restrictions of IBM Text Search engine where, although very efficient, problems were encountered in proximity search options.

4. AMS DL components

AMS DL components are presented in Figure 4. A collection dictionary is used to maintain collection-related information. User-related and access control information is included in the dictionary. Collection repositories facilitate storing and searching digital objects. The Collection Management Service is responsible for managing collection dictionary. It also acts as a "middleware" between collection repositories used to store collections and the end-user services provided by AMS DL integrated environment. The Collection Management Service enables unified access to all collections and transparent implementation of the proposed services. The system is

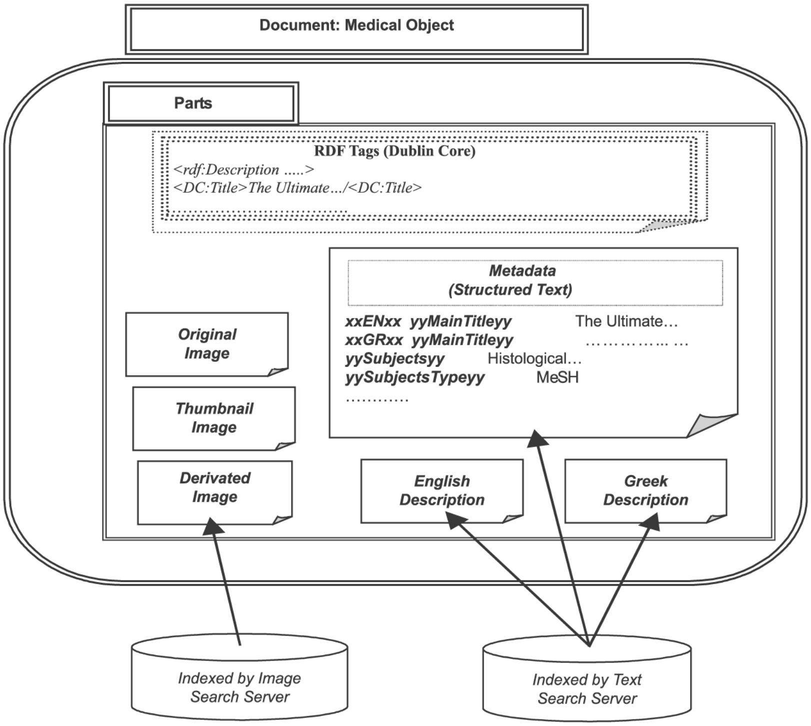

Figure 3.
Medical object internal representation

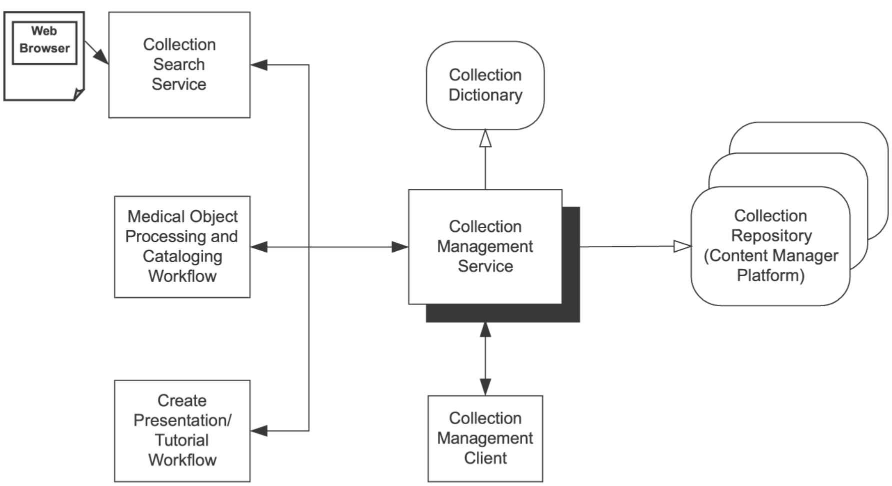

Figure 4.
AMS DL components

implemented using Java and IBM Content Manager platform (IBM Corporation, 2000).

Collection Repository

The Collection Repository consists of Content Manager components. Content Manager is a middleware platform providing tools for storing, searching and managing digital content. Sets of APIs are supported to enable access through standard programming interfaces, and image processing capabilities (e.g. watermarking). The main servers forming the middleware platform are: Library Server, maintaining the data dictionary; and Object Server, facilitating storing and retrieving digital data as objects. Library Server facilitates the maintenance of metadata, DOIs and object access information using an RDBMS (DB2 or Oracle). It also maintains system information regarding object structure. The Content Manager platform was chosen to implement medical collection repositories, since it efficiently supports large volumes of data and easily expand to distributed object servers to ensure search performance.

Collection Management

The Collection Management service provides elementary services for adding/deleting digital material, initiating collection search and forwarding search results. It is a multithreaded environment implemented using component programming. The service consists of two main modules. Repository Access module is responsible for interacting with Content Manager using predefined APIs and is activated upon request whenever there is a need to store or retrieve data or metadata information to/from the Content Manager, but it cannot facilitate structuring and administrating a digital collection. This functionality is included in Collection Manager module providing services to external clients. Thus, clients do not interact with Content Manager platform and consequently have no knowledge of its existence. This ensures system modularity and extendibility and enables supporting different data and metadata models at data storage and data management levels.

End-user services

The Collection Management client is a graphical environment facilitating user and collection management. Medical Object Processing and Cataloguing workflow application facilitates processing and cataloguing workflow and collection access based on user privileges. Create Presentation/Tutorial workflow application facilitates researchers to create presentation. All these applications are implemented as Java clients. They are modular, while modules can be easily implemented as applets. Servlets are used to implement the web-based Collection Search service.

End-user services are implemented under common guidelines in an open environment. They operate parametrically to support any collection. In all cases, application logic remains the same regardless of the collection process. Thus, to achieve unified collection handling, dynamic interface creation is supported based on information stored within Collection Dictionary. The same application logic is applied for all collections, while screens presented to users are dynamically formed based on collection description. Bilingual support for each collection is accomplished in the same fashion, since collection description stored within the Collection Dictionary supports both languages. In a static approach, a separate application instance must be automatically created for each supported collection. As code generation is required, the main disadvantage of the static approach is that collection creation implementation is significantly complex and collection creation cannot be supported in real time. This is the main reason we chose the dynamic approach, although it is a bit slower.

5. Conclusions

AMS DL facilitates access to medical material to researchers and students for both research and educational purposes. We dealt with two strong requirements during system development: organisation and administration of dynamically created collections, and support of advanced workflow capabilities. Medical material can be added in the library directly by researchers, and may be used to create presentation or online tutorials that also are stored in the library.

Dynamic collection management provided automated collection definition and unified collection management within the AMS DL integrated environment. We proposed and implemented a collection dictionary with enhanced features and the corresponding collection management service. Collections are described in detail, concerning their structure, supported metadata and their relationships. Definition and administration of composite objects, such as the ones belonging in Presentation Collection, proved straightforward and efficient. Derived collection definition feature enables the description of new collections by extending the description of existing ones. Simplification was thus achieved, as metadata field definition is time-consuming, especially when dealing with bilingual fields, where labels, predefined values and extended value options must be described in all supported languages. The Collection Management Service, responsible for managing the collection dictionary, facilitates access to collections through a common access point. Both unified management of all collections and interoperability are thus promoted, as the provided services are parametrically implemented independently of the Repository implementation platform.

AMS DL is currently under testing. Processing and cataloguing images are time-consuming tasks: cataloguers receive new material in bursts; it takes approximately two days for a new object to be published, while a cataloguer processes and catalogues less than six objects per day. Integrating Medical Object creation in researcher's daily work was also not easy. In the Laboratory of Histology, for example, selection of material takes place while the researcher is examining tissue samples. Owing to every-day workload, it was not feasible to add this material directly to the Histological Collection. Thus, it was decided to store images during tissue examination, while review and characterisation of images are performed on a weekly basis.

References

Arms, W.Y. *et al.*, (2002), "A spectrum in interoperability", *D-Lib Magazine*, Vol. 8 No. 1, available at: www.dlib.org.dlib/january02/01arms.html

Bainbridge, D., Thompson, J. and Witten, I.H. (2003), "Assembling and enriching digital library collections", *Proceedings of Third Joint ACM/IEEE Conference on Digital Library (JCDL 2003)*, ACM Computer Press, New York, NY, pp. 323-34.

Besser, H. (2002), "The next stage: moving from isolated collections to interoperable digital libraries", *First Monday On-line Journal*, Vol. 7 No. 6, available at: http://firstmoday.org/issues/issue7_6/besser/index.html

Borgman, C.L. (2002), "Challenges in building digital libraries for the 21st century", *Proceedings of IACDL 2002*, LNCS 2555, Springer Verlag, New York, NY, pp. 1-13.

Bristol University (2000), "Technical report: Bristol Biomedical Image Archive case study", available at: www.brisbio.ac.uk (accessed June 2002).

Darmoni, S.J., Leroy, J.P., Baudic, F., Douyère, M., Piot, J. and Thirion, B. (2001), "CISMeF: a structured health resource guide", *Methods Information Medicine*, Vol. 39 No. 1, pp. 30-5.

Davenport Robertson, W., Leadem, E.M., Dube, J. and Greenberg, J. (2001), "Design and implementation of the National Institute of Environmental Health Sciences Dublin Core Metadata Schema", *Proceedings of International Conference on Dublin Core and Metadata Applications 2001*, National Institute of Informatics, Tokyo, pp. 193-9.

Dublin Core Metadata Initiative (n.d.), "Dublin Core metadata element set, version 1.1: reference description", available at: www.dublincore.org/documents/dces (accessed September 2003).

IBM Corporation (2000), *Content Manager Documentation – Planning and Installation Guide*, IBM Corporation, White Plains, NY.

IEEE (2002), "IEEE P1484.12.1/D6.4 draft standard for learning object metadata", available at: http://ltsc.ieee.org/doc/wg12/LOM_WD6_4.pdf (accessed September 2003).

Lagoze, C. and Van de Sompel, H. (2001), "The Open Archive Initiative: building a low barrier interoperability framework", *Proceedings of the First Joint ACM/IEEE Conference on Digital Library (JCDL 2001)*, ACM Computer Press, New York, NY, pp. 54-62.

Lightle, K.S. and Ridgway, J.S. (2003), "Generation of XML records across multiple metadata standards", *D-Lib Magazine*, Vol. 9 No. 9, available at: www.dlib.org.dlib/september03/lightle/09lightle.html

Niu, J. (2002), "A metadata framework developed at the Tsinghua University Library to aid in the preservation of digital resources", *D-Lib Magazine*, Vol. 8 No. 11, available at: www.dlib.org.dlib/november02/ 01niu.html

Sakai, Y. (2001), "Metadata for evidence based medicine resources", *Proceedings of International Conference on Dublin Core and Metadata Applications 2001*, National Institute of Informatics, Tokyo, pp. 81-5.

Suh, E.B., Wang, S.A., Cheung, H., Tangiral, P. and Martino, R.L. (2002), "A web-based medical image archive system", *Proceedings of the Medical Imaging 2002, International Society for Optical Engineering*, available http://cmag.cit.nih.gov/file/SPIE_MI_Paper.pdf (accessed June 2002).

Yu, S.C., Lu, K.Y. and Chen, R.S. (2003), "Metadata management system: design and implementation", *The Electronic Library*, Vol. 21 No. 2, pp. 154-64.

Van de Sompel, H. and Lagoze, C. (2001), *The Open Archive Initiative Protocol for Metadata Harvesting*, Open Archive Initiative, available at: www.openarchives.org/OAI_protocol/openarchives protocol.html

Witten, I.H., Bainbridge, D. and Boddie, S.F. (2001), "Greenstone: open-source digital library software with end-user collection building", *On-line Information Review*, Vol. 25 No. 5, pp. 288-98.

Further reading

Day, M. (1999), "Metadata for images: emerging practice and standards", *Proceedings of the Second UK Conference on Image Retrieval, Newcastle*, available at: www.ukoln.ac.uk/metadata/presentations/ cir99/paper.html

Duval, E., Hodgins, W., Sutton, S. and Weibel, S. (2002), "Metadata principles and practicalities", *D-Lib Magazine*, Vol. 8 No. 3, available at: www.dlib.org.dlib/april02/weibel/04weibel.html (accessed June 2002).

(Mara Nikolaidou currently serves as the Director of Libraries Computer Center in the University of Athens. She received a BSc and PhD degree in Computer Science from the Department of Informatics and Telecommunications of University of Athens. She has published more than 40 papers in international journals and conference proceedings regarding digital libraries, distributed systems, workflow systems and web-based technology. She can be contacted at: Mara Nikolaidou, Libraries Computer Centre, University of Athens, Panepistimiopolis, 15771 Athens, Greece. E-mail: mara@di.uoa.gr

Dimosthenis Anagnostopoulos is an Associate Professor of Information Technology at Harokopio University of Athens. He holds a BSc and PhD degree in Computer Science from Department of Informatics and Telecommunications of University of Athens. He has published more than 50 papers in well-known international magazines and conference proceedings, in the scientific areas of simulation, distributed systems and web-based systems. He can be contacted at: Dimosthenis Anagnostopoulos, Harokopio University of Athens, 70 El. Venizelou Str., 17671, Athens, Greece. E-mail: dimosthe@hua.gr

Michael Hatzopoulos is a Professor in the Department of Informatics and Telecommunications at the University of Athens. He holds a BSc in Mathematics from the University of Athens (1971) and a MSc and PhD in Computer Science from Loughborough University of Technology (1972 and 1974 respectively). He participated in a large number of research projects and has published more than 70 papers in international journals and conference proceedings. He can be contacted at: Michael Hatzopoulos, Department of Informatics and Telecommunications, University of Athens, Panepistimiopolis, 15771 Athens, Greece. E-mail: mike@di.uoa.gr)