# WEB SERVICE BASED INFORMATION RETRIEVAL IN CONVENTIONAL DICOM REPOSITORIES TO SUPPORT TEACHING AND RESEARCH IN RADIOLOGY

Eleni Kaldoudi, Dimosthenis Karaiskakis, John Manavis
School of Medicine, Democritus University of Thrace
6 I. Kaviri Str.
Alexandroupolis
Greece 68100
Tel: +30 2551 030329
E-mail: kaldoudi@med.duth.gr

## Abstract

Although the Digital Imaging and Communications in Medicine (DICOM) protocol is currently the ubiquitous standard for the communication of medical images within the healthcare enterprise, seamless image distribution to research and educational information systems is still hard to achieve, as software developers of such third-party applications have to adapt to the rather complex DICOM protocol. This paper presents a web services environment that allows complex search and dynamic information retrieval in conventional DICOM sources (e.g. PACS systems, stand-alone databases, etc) in order to support educational, research and administrative activities. The proposed solution is based on a web service which acts as a façade for conventional DICOM sources allowing DICOM image data and related information to be transformed into XML documents encapsulated in SOAP messages. At a second level, additional web services collaborate for the dynamic extraction and indexing of all data attributes represented in the DICOM information model, thus allowing for advanced complex searches in the DICOM information space. The purpose of this work is to create easily searchable clinical data repositories (with the ability to give compound structure to both queries and results) that can support research, teaching and data mining for administrative and clinical protocol planning, budget control, etc.

## Introduction

At the typical contemporary radiology department, adherence to the Digital Imaging and Communications in Medicine (DICOM) protocol (DICOM, 2004) is practically always assumed. Moreover, medical image distribution within the healthcare intranet is also based on DICOM. However, seamless image communication to research and educational information systems is still hard to achieve, as software developers of such third-party applications have to go through the rather cumbersome task of adapting the DICOM communication model, implementing the DICOM protocol and keeping up with new changes in the standard. As a result, most research, educational and administrative software tools lack the ability to directly communicate with DICOM sources. In cases where such third-party tools provide an implementation of the DICOM protocol, this mainly addresses the simple task of retrieving DICOM objects and is limited to the standard manipulation of medical images provided by the DICOM protocol which pertains to the clinical task of patient management. As such, the wealth of image related information stored in and communicated with DICOM images is usually concealed and complex data queries and searches are not possible.

Recently, the ubiquitous adoption of Internet and related technologies has made a considerable impact on extending medical image services outside the radiology department. Initially, web-based applications have enabled users to access DICOM image servers through a common web browser. Additionally, eXtensible Markup Language (XML) (W3C, 2004) technologies have been used to represent DICOM data in a generic standard way to promote data interchange amongst disparate healthcare

systems. This paper proposes a new technological approach for distributing DICOM images and related data through commonplace Internet technologies, based on the emerging web services software paradigm. The purpose of this work is to create easily searchable clinical data repositories (with the ability to give compound structure to both queries and results) that can support research, teaching and administrative activities.

## Web Services in Healthcare Intranets

During the past years, the Internet and related technologies have been widely introduced in the radiology department to provide an easy and commonplace way of accessing clinical data and supporting various healthcare processes. In the case of medical imaging, respective work addresses both aspects of image management as supported by the DICOM standard, namely data representation and data communication.

In terms of medical image communication, nowadays most Picture Archiving and Communication Systems (PACS) manufacturers provide a web browser access to medical images, tightly coupled to their PACS implementation. Common to all approaches is the fact that DICOM images that originate from dedicated medical imaging repositories and archiving systems can be accessed by individual users via a commonly available web browser, without any need for special software or hardware. Although an easy and efficient solution for simple image distribution services to individual users, web browser access to DICOM images has some drawbacks: (1) initial DICOM data are usually transformed into a web compatible format, and thus related information included in the header file is not entirely preserved (2) functionality for advanced management of images and related information is not easy to develop as the web browser is a thin client, and (3) access to DICOM images and related information is limited to human users, while other software applications cannot take advantage of this web-based communication. The need for software applications to independently access DICOM data through standard Internet technologies is addressed by the WADO (Web Access to DICOM Persistent Objects) supplement to DICOM standard (DICOM, 2004a). WADO specifies a simple mechanism for accessing a DICOM object (image, waveform, structured report, etc) from web pages or other software applications (e.g. an e-mail system), through the HTTP communication protocol. However, WADO does not support any other conventional DICOM services (e.g. find or store), and assumes that the unique identifiers required for retrieving an object from a DICOM archive are known by other means.

In terms of data representation, the widespread adoption of XML technologies has made a notable step towards easier integration of clinical data in healthcare information systems. Transforming DICOM objects in XML format mainly aims to enhance exchange of medical image data across the healthcare enterprise (Cohen, 2002 & Rassinoux, 2003), especially as other clinical information systems and relevant standardization bodies begin to adopt XML technologies.

Recently, the XML/SOAP web services programming paradigm (W3C, 2004a) has been a catalyst for achieving both data and control integration among applications through commonplace Internet technologies. Web services are loosely defined as self-contained, self-describing, modular applications that can be located and invoked over the Internet. They are based on open Internet standards: build on the HyperText Transfer Protocol (HTTP), they use XML for data presentation while messaging is described in an XML-based messaging protocol, SOAP (Simple Object Access Protocol). Web services describe themselves through a standardized Web Service Description Language (WSDL) document, and can be published to one or more Intranet or Internet repositories for potential users to locate through a standard Universal Description, Discovery and

Integration (UDDI) registry. A whole suite of additional standards are currently being developed to formally address issues such as security, reliability, transactions, etc. In essence, web services are a middleware technology for developing service-oriented architectures. However, it is expected that initially web services are mainly used as wrappers of existing applications serving to interconnect legacy systems without altering their code, as well as to decompose their usually complex functionality and offer it as separate, well-defined targeted services.

The web services paradigm has already gained broad industry support. In the healthcare sector, it has recently been identified as an especially important technology for the future of healthcare delivery and administration (Hashem, 2003, & Malcolm, 2004), and the first implementations are currently emerging. For example, web services technology has been employed to enable integration of biology sequence data banks over the web (Sugawara, 2003), as well as to support a core infrastructure of components and services for the management of data from diverse sources in cancer informatics (Covitz, 2003). Moreover, web services are expected to play a key role in integrating the healthcare enterprise as they are currently being adopted by relevant standardization bodies. For example, HL7 is promoting web services for message exchange in the forthcoming versions of the standard (HL7, 2000), while IHE (Intergating the Healthcare Enterprise initiative) is already employing wed services for the implementation of certain transactions of its integration profiles (IHE, 2004).

Web services have already been introduced in medical image management to create a façade for conventional DICOM image servers, and expose the principal DICOM services of query, retrieve, and store to any other software application over the Internet, using standard XML documents communicated via SOAP messages (Kaldoudi, 2004, & Delistamatis, 2004). This paper builds on this previous work and presents a collaborating web services environment that allows complex search and dynamic information retrieval in conventional DICOM sources. The proposed web services can be engaged by any general purpose application that is not DICOM aware in order to support educational, research and administrative activities that require complex manipulation of DICOM medical images and related data.

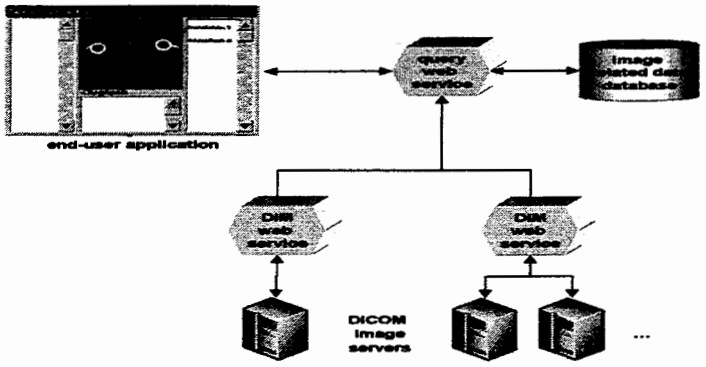
## DICOM Query Management Web Service

The DICOM information model organizes data related to medical images using a logical structure of information entities and their relationship. Basically, the standard specifies the entities involved in radiological operations, such as patients, visits, studies, images, structured reports, other data objects, etc. Each entity is characterized by a collection of attributes, which carry all information related to the particular entity. These collections of various attributes that characterize a study, a series or a data object carry in essence all image related information. They are represented by a tag (which identifies the attribute) and its value and are grouped according to the information entity they describe. Since the primary scope of DICOM is medical image communication amongst different imaging modalities and radiology information systems in order to support the task of patient management, DICOM is not built to be particularly friendly to research and academic demands. So, although the protocol supports basic search functionality, this is limited to exploit only certain attributes and of a specific information entity at a time. As a result it is not possible to formulate complex queries that combine any DICOM attribute from different levels of the information model.

In our work, complex query management in DICOM image servers is achieved though a collaborating web services environment as shown in Figure 1. At a first level, simple DICOM Image Management (DIM) web services act as a wrapper to any conventional DICOM image server, and expose the principal DICOM services of query, retrieve, and

store to any other software application over the Internet, using standard XML documents communicated via SOAP messages (Delistamatis, 2004). This approach combines the advantages of using XML for DICOM data representation with the benefits of employing open ubiquitous web technologies for messaging and communication. Thus, DICOM image data and related information can be discovered, retrieved and maintained in a third party non-DICOM application in its fullness, and through open, standard messaging. At a second level, additional web services collaborate with the DIM web services to provide added-value for search and academic purposes. In particular, this paper presents a web service for the dynamic extraction and indexing of all data tags represented in the DICOM information model, thus allowing for advanced complex searches in the DICOM information space.

Figure 5: Collaborating Web Services for DICOM Query Management.




The DICOM Query Management (DQM) web service retrieves image related data from DICOM image servers to create a special database with indexed DICOM attributes. This database is then available for complex queries to any third party application via SOAP/XML communication. The search functionality of the DQM web service is summarized in the methods exposed:

- *GetAvailableTags*: provides a list of all DICOM attributes that are indexed in the database and can be used for the formulation of a complex query
- *QueryDatabase:* supports formulation of complex queries with any logical combination of available attributes
- *ImportDicomData:* manages the discovery and collection of image related data for newly generated DICOM objects in the designated DICOM image servers (imaging modalities, PACS, related medical information systems, etc.)

The attribute database maintained by the DQM web service is based on the dynamic model shown in Figure 2. It should be noted that it does not follow the DICOM patient-study-series relationship structure, but it is a fully dynamic model where no attribute based tables are used. Each DICOM attribute is represented by an object defined by its tag, tag value, data type, and some references. The model can be dynamically expanded to support new additions to the DICOM protocol (i.e. new attributes) as well as additional user defined attributes (such as comments and task related flags).

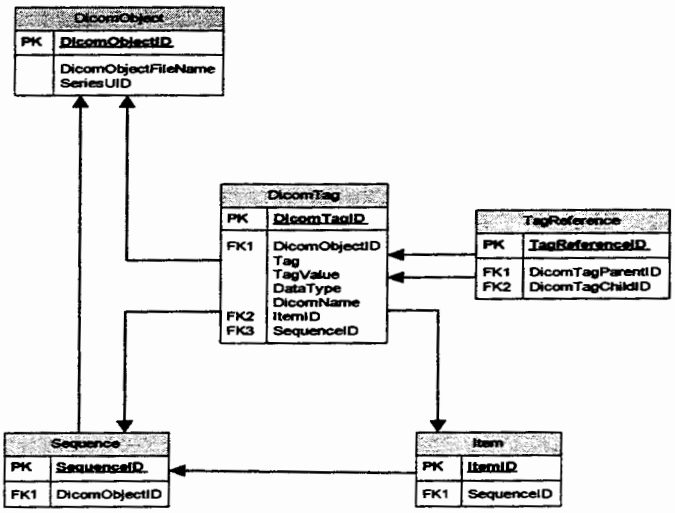
## Implementation Issues

The web services described in this paper have been developed in C# using the MS .Net Framework 1.1 (Microsoft, Redmond, USA) and use the DICOM library DicomObjects 4.1 (Medical Connections, Reynoldston, UK). System requirements at runtime include the MS Internet Information Server >5.x, MS .Net Framework 1.1, and MS SQL Server 2000 Desktop Engine (Microsoft, Redmond, USA). Basic security is achieved through

the Secure Sockets Layer encryption mechanism for data transmission, as provided by the Internet Information Server, and this is enhanced by user authorization process. An asynchronous communication pattern is used to enhance the scalability & reliability of the web service environment and enable long running operations.

The web service methods presented in the previous section are formally described in the corresponding WSDL document. The complete DQM web service WSDL document, as well as detailed documentation and other related information, can be retrieved from: http://iris.med.duth.gr/link/link.aspx?id=107. Using this document, software developers of third party applications can build the appropriate SOAP messages to invoke and consume the proposed web services. It should be noted, that most current Integrated Development Environments (IDE) can generate automatically the required SOAP messages out of the WSDL document.

Figure 2: DICOM Attribute Database Model.




## Discussion

Concluding, we have built collaborating web services to achieve communication and integration of DICOM image sources with other applications using internet standards and commonplace Internet technologies and allow complex search and dynamic information retrieval in conventional DICOM sources (e.g. PACS systems, stand-alone databases, etc) in order to support educational, research and administrative activities. The goal of this work is twofold: (1) to extend the basic functionality of DICOM protocol (to support patient management in the radiology department) so that research and academic tasks can be performed equally well, and (2) to allow DICOM images and related data to be manipulated by third party applications that are not DICOM aware. Examples of such applications, that do not necessarily implement DICOM, could include Internet based medical e-learning environments, medical research support tools, medical expert systems and other healthcare information systems. This work is part of our efforts to seamlessly integrate clinical data into a generic Internet based e-learning environment that will support undergraduate medical education in Democritus University of Thrace, Greece.

## Acknowledgements

This work was partly carried out within the project: "Reforming Undergraduate Education in the School of Medicine, Democritus University of Thrace", which is funded by the

Managing Authority of the Operational Programme for Education and Initial Vocational Training, a joined funding programme of the Greek Ministry of National Education and Religious Affairs and the European Community (European Social Fund and European Regional Development Fund).

## References

Cohen, S., Gilboa, F., & Shani, U. (2002). PACS and electronic health records. In: The Proceedings of SPIE Medical Imaging 2002 (vol. 4685, pp. 288-298). San Diego California.

Covitz, P.A., Hartel, F., Schaefer, C., De Coronado, S., Fragoso, G., Sahni, H., Gustafson, S., Buetow, K.H. (2003). caCORE: a common infrastructure for cancer informatics. Bioinformatics, 19(18), 2404-12.

Delistamatis A., & Kaldoudi, E. (2004). Internet technologies for the distribution of DICOM structured reports within healthcare intranets, Proceedings of the 2nd International Conference on Information Communication Technologies in Health, 334-339.

DICOM: Digital Imaging and Communications in Medicine (2004). DICOM Standard Status. Retrieved May 5, 2004 from http://www.dclunie.com/dicom-status/status.html#BaseStandard2001

DICOM: Digital Imaging and Communications in Medicine and ISO TC215/SC2 Working Groups, (2004a). DICOM supplement 85: Web Access to DICOM persistent Objects (WADO), October 28, 2003. Available from: ftp://medical.nema.org/medical/dicom/supps/sup85_lb.pdf. Retrieved on May 25, 2004

Hashem, A., & Ruggeri, R., (2003). Infrastructural considerations for supporting POC devices, Health Management Technology, 24, 24-27.

HL7 (2000). Application Protocol of Electronic Data Exchange in Healthcare Environments, v2.4. HL7 Inc. Ann Arbor, MI, USA.

IHE: Integrating the Healthcare Enterprise. (2004). Available from: http://www.rsna.org/IHE/mission.shtml. Retrieved on June 10, 2004.

Kaldoudi, E., Vargemezis, V., & Simopoulos, K., (2004). Information and communication technologies in medical undergraduate education, Proceedings of 2nd International Conference on Information Communication Technologies in Health, 129-134.

Malcolm, D.J., (2004). Web services: don't overlook this silver bullet, Health Management Technology 25, 89-90.

Rassinoux, A.M. Lovis, C., Baud R., & Geissbuhler, A. (2003). XML as standard for communicating in a document based electronic patient record: a 3 years experiment, International Journal of Medical Informatics, 70, 109-115.

Sugawara, H., Miyazaki, S. (2003). Biological SOAP servers and web services provided by the public sequence data bank. Nucleic Acids Research, 31, 3836-9.

World Wide Web Consortium (W3C) (2004). Extensible Markup Language (XML) 1.0. Recommendation. Retrieved May 5, 2004 from http://www.w3.org/TR/2004/REC-xml-20040204/

World Wide Web Consortium (W3C) (2004a). Web Services Architecture. W3C Working Group Note. Retrieved May 5, 2004 from http://www.w3.org/TR/2004/NOTE-ws-arch-20040211/

# INEAG

RESEARCH AND TRAINING INSTITUTE OF EAST AEGEAN

# 3rd International Conference On Information Communication Technologies in Health

## "Telemedicine and eHealth in recent years: Meeting the Challenges"

**7-9 July, 2005**

**Samos Island, Greece**

Organizers:
INEAG
University of Athens
EPEAEK
GHIA
University of Missouri
The friedrich Alexander University
of Erlangen - Nuernberg
The UMIT

Under the auspices of:
The Hellenic Ministry of Health and Social Welfare