

MAMODB: A WEB-BASED TOOL FOR TRAINING RADIOLOGISTS IN THE DIAGNOSIS OF DIGITAL MAMMOGRAPHY

J. Martí, C. Mata, X. Lladó, A. Oliver

Institute of Informatics and Applications, University of Girona (SPAIN)
joanm@eia.udg.edu, cmata@eia.udg.edu, llado@eia.udg.edu, aoliver@eia.udg.edu

Abstract

The massive growth in applications of radiological imaging and image-guided treatments has met a worldwide shortage of trained radiologists. Therefore, several computer-based radiology training environments, which combine traditional learning opportunities with advanced e-learning platforms, have recently been developed with the goal of supporting the acquisition of radiological expertise. However, the quantity and quality of clinical experience arises as the main issue to ensure the success for such learning environments when training radiologists to interpret medical images.

In this sense, MamoDB is a research web-based tool which has been customized in order to incorporate learning capabilities for training and guidance of newly employed or resident radiologists. MamoDB allows panels of experts to collaborate at different hospitals and research centres by means of integrating a Picture Archiving and Communication System (PACS) to store the Digital Imaging and Communications in Medicine (DICOM) files as well as a database with the eXtended Markup Language (XML) files containing the experts' annotations for each clinical case.

Keywords: MamoDB, e-learning, e-health, mammography, digital imaging, DICOM, radiologist, web-based tool, PACS, XML, and system architecture.

1 INTRODUCTION

Nowadays, medical imaging plays a fundamental role in achieving correct diagnosis for many diseases. However, it involves changes concerning the management, storing and retrieving of digital images. Thus, in order to improve the current imaging systems, it is necessary to develop architectures and medical databases, as well as PACS [1] systems. Although there exists a number of public mammographic databases, none of them has been digitally acquired and the images come from analogical acquisition and subsequent digitization of the films [2][3]. For breast imaging, this process involves an inherent loss of quality, depending on the digitizer system, loss of acquisition information and no standardization of image formats.

Development and publication of a database of digital mammography (Full Field Digital mammography, FFDM) with a high number of representative cases is essential, because the purpose is to become a landmark in the scientific community for the evaluation of aided systems of breast cancer diagnosis.

Besides, systems based on web applications [4-6] are particularly interesting due to recent advances in computer technology and in communications systems and protocols, which have allowed the development of online software and data recovery in a reliable and inexpensive way [7]. Therefore, the use of Web Services applied to medical domain has experienced an important growth during the last years, and has become a communications system in itself. The main advantages of web-applications are obvious: easy accessibility, users can connect from anywhere in the world, time-saving and efficient mechanism to find special information which can be saved directly to one's computer without the tiring tasks of searching in libraries [8], providing high levels of security system, downloading medical annotations using query forms and, can be used as a high level of interactivity among health professionals.

A number of research works, published during the last decade and conducted with this aim, can be found in the literature. However, there are still major problems to solve. The main problem is the access and information sharing between different centres, and the interaction between hospitals and physicians to assist medical diagnosis. In this sense, the goal of our project is to develop a web-application system using a digital mammographic database shared in several hospitals in Catalonia (Hospital Universitari Dr. Josep Trueta, Corporació Sanitària Parc Taulí, Clínica Girona, Hospital Vall d'Hebron).

Moreover, the introduction of the XML (eXtended Markup Language) file format [9] allows information exchange regardless of the platform, thus, introducing a standard format to include diagnosis on the images [10]. Furthermore, we propose to use MamODB as a part of the problem-based learning paradigm, which in fact, is an innovative and challenging approach to medical education. It is a new way of using clinical material to help students to learn, and it requires the medical expert to use facilitating and supporting skills rather than didactic, directive ones. Such paradigm emphasizes the application of knowledge and skills to the solution of problems instead of just recalling the facts. Therefore, stated that the training process for the radiologists should include a large number of cases with multiple benign and malignant findings, MamODB allows testing and training a radiologist's ability to:

- Detect mammographic findings.
- Properly characterize mammographic findings according to their number, location, shape, size, margins, density, and distribution.
- Issue a diagnosis according to the detection, correct identification and characterization a mammographic finding.

The purpose of this work is to provide a useful tool to the medical and scientific communities in order to manage a mammographic image database including their associated diagnoses (MamODB), featuring the advantages and functionalities of a web service (i.e. authentication, security and data retrieval).

1.1 MamODB as an integrative tool for e-learning-health development

Technology in health allows to record, measure, monitor, manage and provide guidance to the patient, together with the opportunity to offer care services through the Internet in real time. E-learning health systems have enabled to maximize the collection, delivery and communication of medical information, clinical messaging, interaction with doctors and medical records [11].

In this context, MamODB is an e-learning health tool designed for students, assistant radiologists, and research users who must possess the following set of skills:

- recognize when there is the need for information request;
- identify the need to control the information overload;
- evaluate and organize the information, and have ability to present results;
- be efficient using information to address the problem;

Moreover, the role of the doctor is very important because the user must spend time to teach students to use the application. In the same way, students must be receptive and willing to learn. If this effort is not mutual the educational action will fail.

MamODB offers the possibility to develop a learning process that is adaptive for students. Accordingly, some of the advantages of the MamODB as an e-learning tool are given below:

- Provides students with a wide range of information;
- Facilitates information and content update;
- Promotes *just in time just for me* training;
- Provides training and collaborative group;
- Facilitates student autonomy;
- Enhances interactivity with information, with the teacher and among students.
- This project involves collaboration of different hospitals in Catalonia with the same goal: learning and support to diagnosis.

2 SYSTEM ARCHITECTURE

MamODB has been designed as a web-based application within the Zend Apache server [12] which is also used as a MySQL database server for data storage. The application links both the PACS and the XML servers which store clinical cases (as DICOM and annotation files). All data transmissions

between users and the web server are encrypted in order to ensure complete confidentiality of the data. Currently, the Secure Socket Layer (SSL) cryptographic protocol is used. Queries to the MamODB are performed through a web form and provide a user with a list of clinical cases according to the query parameters, including the image and the annotation files.

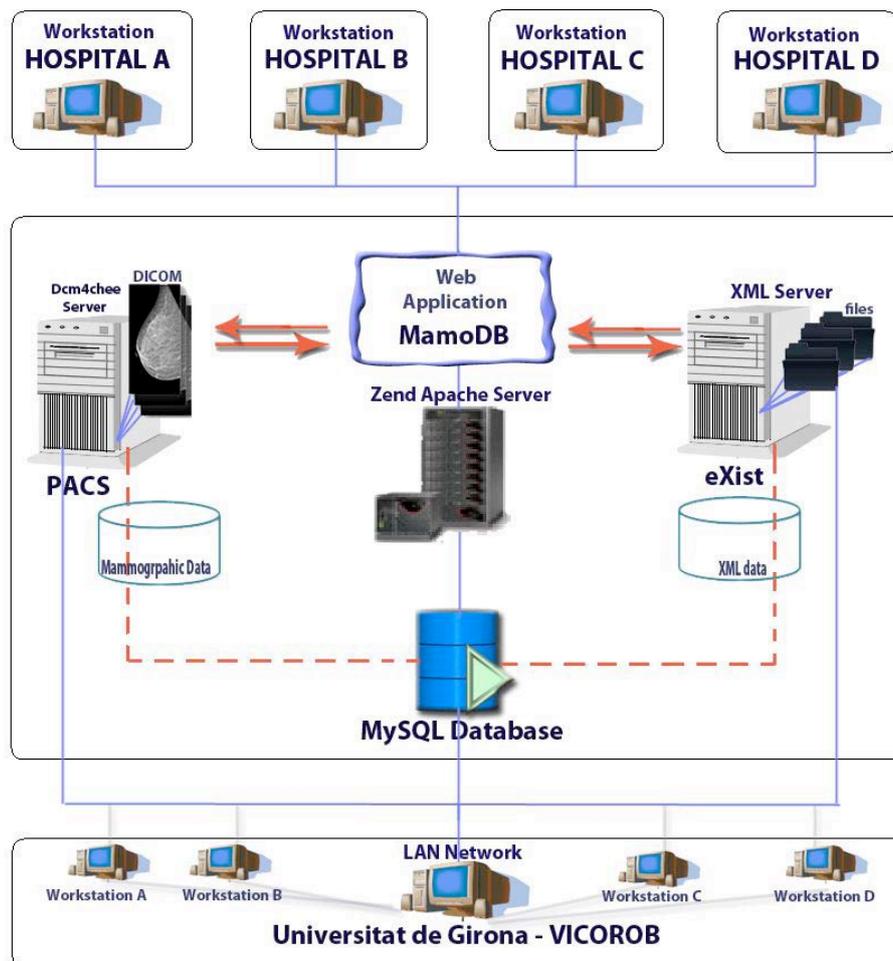


Figure 1. Architecture model of the MamODB

Figure 1 shows the system architecture of the MamODB application. It is important to emphasize that the application is stored in the Zend Apache server. This server is configured with a MySQL database manager. In addition, once the PACS is installed, a database (pacddb) is also created in MySQL and therefore directly related to Zend Apache server. Thus, when a mammographic case is added to the PACS server, the image is stored inside the server but the information is saved inside the MySQL database.

The MamODB application is also related to the eXist manager [13], responsible for the XML database, incorporating its own server for access and management via web interface. We opted for the creation of the XML annotation files in the database in order to preserve the integrity of the original DICOM files and to ensure their full support if they come from different acquisition systems. All XML files, that are stored in the eXist database are related to their corresponding mammographic images as these are the files containing the records of diagnosis made by radiologists. For this reason, it can be seen in the architecture model of Figure 1 that there is a direct relationship between the MySQL database and the eXist manager. Finally, we have defined the key field identifier or files with their breast imaging, so that the database and the files are correctly matched.

In order to facilitate the access to the MamODB application, and with the idea to serve both the medical experts present at the hospital itself and the outside professionals, we decided to design a web application. One of the key elements in this application is the data and information security across the network, especially when it comes to confidential data related to medical examinations. In this sense, the three modules shown in Figure 1, that constitute our system architecture (Zend Server,

PACS and eXist), have internal security measures to avoid vulnerabilities from the transfer of information, so that besides using a control User ID private keys, the algorithms have systems that protect and encrypt the information during its transmission.

Finally, the access system has been designed in order to facilitate both local and external [14]. The users should log in correctly using a unique username and password independently if the connection is via intranet or external.

3 WEB-APPLICATION: MAMODB

We propose the use of the web-application MamODB as a part of the problem-based learning paradigm. This section describes the most important features of the MamODB application. It has been divided into three different subsections: identification and security, control panel and accessibility, and management and training query form.

3.1 Identification and security

Firstly, the secure access user to the MamODB application is ensured by a username and password generated by the system administrator. The system incorporates mechanisms of authentication and protection against fraudulent use of identity.

The system has a three-level security setting for user profiles. The 1st level (Standard User) allows users to submit queries on the various medical cases stored in MamODB (images and diagnosis). This level corresponds to a profile of a student or auxiliary medical personnel, who uses the application for remote diagnostics, training, etc. The 2nd security levels, in addition to the permissions of the 1st level, allows access to PACS and eXist for addition and modification of the mammographic database content and clinical cases (images and diagnosis). Essentially, it corresponds to a profile of advanced student or medical specialist partner in the construction of the mammographic database, who needs access to the system locally or remotely. The 3rd security level serves for system administrators, adding user account management to the list of permissions.

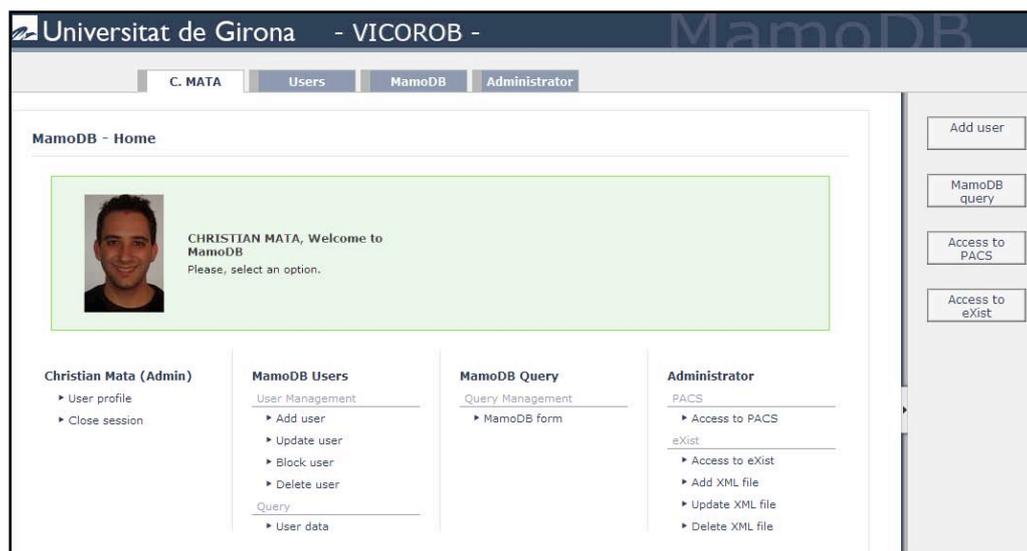


Figure 2. Control panel of the MamODB

3.2 Control panel and accessibility

Once a user is identified correctly in the system, the home page of MamODB as the control panel for his profile appears in the browser. Depending on the type of user accessing the system, the control panel shows the different possibilities.

Figure 2 shows the control panel, which has been designed to take into account UI ergonomics and simplify access to the options of the system. This home page is designed as a web map; the user can access all the functions of the system in a single click. There are two blocks of access elements: a drop-down menu bar and an expandable sidebar, which the user can make invisible. The latter's role is to provide direct access to the most used features. Moreover, the control panel contains a personal information section where a user can edit such information as profile image, username and password.

3.3 MamODB functionality

The most important part associated with the learning capabilities for training and guidance of newly employed or students, is shown in this section. Therefore, the management and training form, finding results, and characterization analysis of the mammographic findings are presented.

3.3.1 Management and training query form

This e-learning platform uses queries to the MamODB through a web form and provides a user with a list of clinical cases according to the query parameters, including both the image and the annotation files.

The first step required from the student for this training process is the use of the query MamODB form. In order to obtain desirable results the user should select appropriate criteria, as directed by the expert or medical staff. Their work is to find mammographic studies according to the criteria used by the user. The search is performed in all the XML files stored in the eXist database and the list of results is shown in the web interface.

Figure 3 shows an example of a standard query form where a user can define the type of query (upper area) and select the information to be shown in the result (bottom area). This training criterion is based on the query form with the purpose to detect a mammographic finding according to its their number, location, shape, size, margins, density, and distribution.

Figure 3. Example of a MamODB query form

3.3.2 Finding results

In addition to a comprehensive list of clinical cases contained in the database, the query forms allow to search according to some specific medical criteria: reported diagnosis, BIRADS classification [15], presence of abnormalities (micro-calcifications, structural distortions, masses, etc.). Depending on the volume of records obtained from the results list, the application uses automatic paging group records in a more visual way as can be seen in Figure 4.

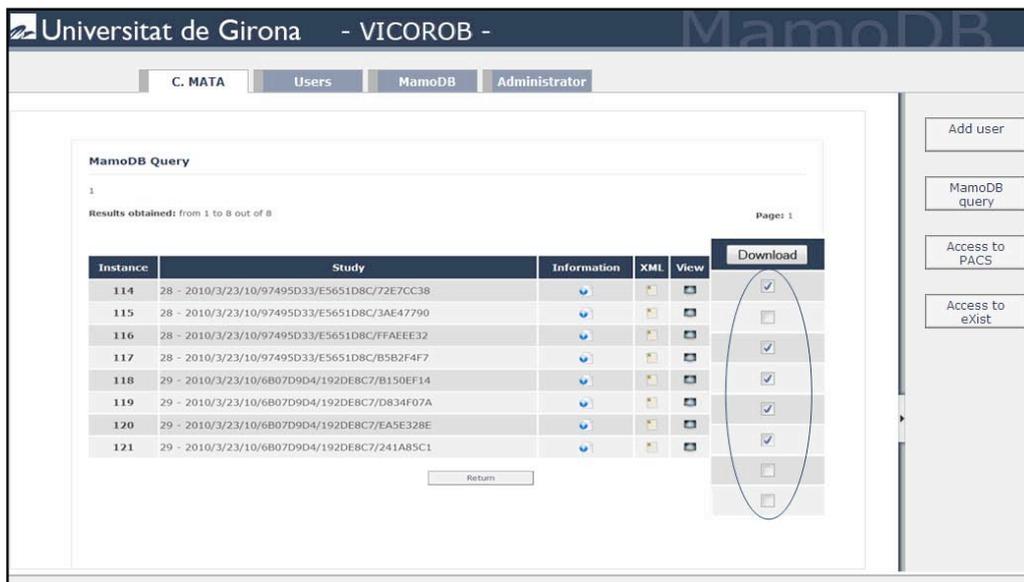


Figure 4. Results obtained after the training query form

For each of the records obtained, the system allows three actions: see information on the DICOM image stored in the PACS server, check the diagnosis associated XML file formats, and display the mammographic image. Figure 4 shows the result of three possible actions for a particular clinical case (a): information concerning the acquisition, (b) diagnostic annotations associated with the case, (c) mammographic image.

3.3.3 Characterization analysis of the mammographic findings

Three different actions are available once the records have been obtained:

- Display the information contained in the DICOM file (Figure 5).
- Get the diagnosis linked to the case, stored in a XML file (Figure 6).
- Display the digitized images contained in the DICOM file (Figure 7).

When the list of results is obtained, the user can get information about all studies and compare them in order to perform a classification and selection of the best results according to the preferences selected in the query form. This is the educational part of this e-learning tool and it is useful to understand its methodology. For each case a student can display acquisition information contained in the DICOM file, read annotations about the study included in the XML file, and display the images included in each study.

Group	Element Name	VR	Length	VM	Value	
0008	0005	Specific Character Set	CS	10	1	ISO_IR 100
0008	0008	Image Type	CS	22	4	DERIVED/PRIMARY/RIGHT
0008	0016	SOP Class UID	UI	28	1	1.2.840.10008.5.1.4.1.1.1.2
0008	0018	SOP Instance UID	UI	56	1	1.3.12.2.1107.5.12.7.1036.30000005022316255773400000031
0008	0022	Acquisition Date	DA	8	1	20050223
0008	0023	Content Date	DA	8	1	20050223
0008	0032	Acquisition Time	TM	14	1	093533.343000
0008	0033	Content Time	TM	14	1	093554.937000
0008	0054	Retrieve AE Title	AE	8	1	DCM4CHEE
0008	0056	Instance Availability	CS	6	1	ONLINE
0020	0013	Instance Number	IS	2	1	1
0028	0004	Photometric Interpretation	CS	12	1	MONOCHROME2
0028	0010	Rows	US	2	1	3328
0028	0011	Columns	US	2	1	2560
0028	0100	Bits Allocated	US	2	1	16
0043	0010		LO	15	1	dcm4che/archive
0043	1013		OB	8	1	0'0'0'0'0'0'0'0'5

Figure 5. Example of the information contained in the DICOM file.

XML FILE	
Image	1
Path	2011/29/11/16/78E4924D/C868A2DE/1FE12C2E
User	Christian Mata
Data	16/03/2011

STUDY INFORMATION	
ID Patient	148644
ID Study	1
BIRADS	2
Description	Description of mammographic image

MASS INFORMATION	
Area	32.35
CenterMass	(324, 127)
Description	Overlay detected in a mass centered

Figure 6. Example of the information contained in the XML annotation file.

Figure 5 shows an example of the DICOM image provided by PACS. Sometimes, it is useful to know the details about the acquisition system machine in order to classify cases. On the other hand, in Figure 6 a user can obtain information from the XML files. As can be seen, this information is divided into different parts. Firstly, the information about the doctor involved in the annotation is shown followed by the study (ID, BIRADS classification and description) and the mass information. It is important to make a remark that the information about the patient is anonymous, and only the Patient ID is shown as it has no relation with the patient's identity. The main objective for users is to learn and to discuss the cases which will be useful for the doctors to do and support mammographic diagnosis.

Finally, in Figure 7 two examples of DICOM images associated with the selected study are displayed. It also shows overlays marked by doctors and their individual annotations. This information is very important because users should compare and classify studies in order to later download all the cases useful for the learning purposes.

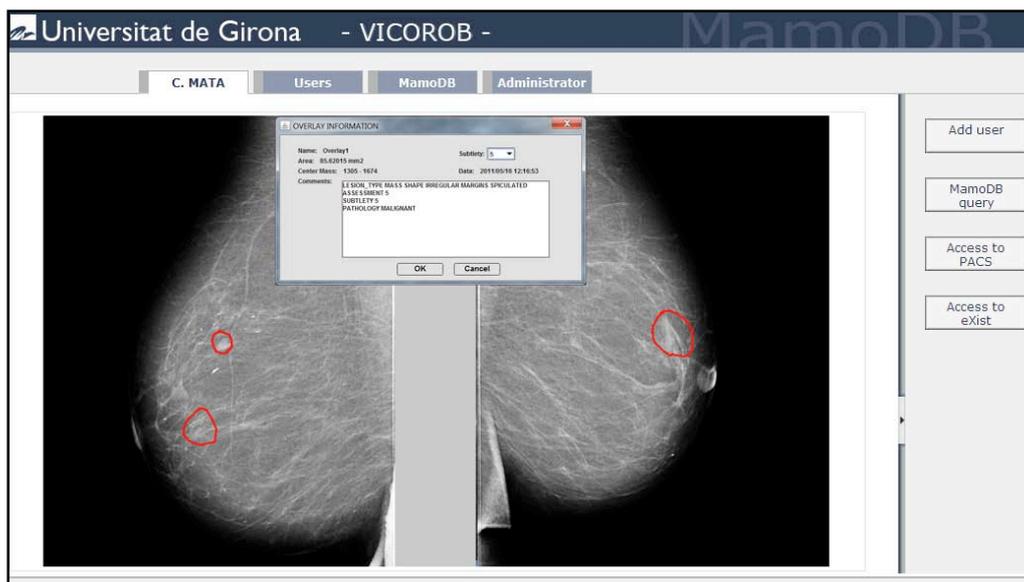


Figure 7. Display the DICOM images with their associated overlays. It is also possible to obtain the description about each overlay.

3.3.4 Downloading selected mammographic studies

The last step when students study all mammographic cases is to download all the causes useful for them. An important feature of the system is that the records can be visualized on the screen or downloaded locally to the workplace selected by the user. The system offers the possibility to download a single record or use an automatic selection. Once selected, the DICOM files together with their associated XML files become part of the ZIP file which is downloaded by the user. As shown in Figure 4, a student can select the cases that are of interest to the learning process and start the download process immediately.

4 CONCLUSIONS AND FUTURE WORKS

Digital medical imaging technologies have become beneficial in modern medical practices and health care systems, providing powerful tools for diagnosis, treatment, and surgery. The volume of medical images being generated has grown rapidly due to an increase in the number of clinical exams performed in digital form and to the large range of available modalities. Therefore, the demand for online medical imaging systems that allow visualization and processing has increased significantly.

Although nowadays public databases of mammographic images are available, many of them have been acquired digitally by analogic acquisition and subsequent digitization of the films. For breast imaging, this process involves an inherent loss of quality, depending on the digitizer system, loss of information acquisition and non-standardization of image formats. The proposed system architecture to manage and centralize a public database of digitally acquired mammographic images. Furthermore, using the Web application MammoDB, a user can have access to information both locally and remotely.

MammoDB has been designed as a web-based application within the Zend Apache server which is also used as a MySQL database server for data storage. The application links both the PACS and the XML servers which store the clinical cases (as DICOM and annotation files). Database queries are performed through a web form and provide the user with a list of clinical cases according to the query parameters, including both the image and the annotation files.

As future work to improve MammoDB, we propose to define a single XML file format for the medical community, which would correspond to the annotations of the images associated with mammography. Moreover, we plan to define custom search forms that ensure optimal results and are tailored to the needs of the user. Finally, further improvements could be investigated in terms of security and data transfer and the use of new algorithms to improve data encryption and provide a more secure and robust system.

REFERENCES

- [1] R.H. Choplin, J.M. Boehme, C.D. Maynard, Picture archiving and communication systems: an overview, *Radiographics* January vol 12, 127-129 pp, 1992.
- [2] M. Heath, K.W. Bowyer, D. Kopans, R. Moore , P. Kegelmeyer, The Digital Database for Screening Mammography, *Proceedings of the 5th International Workshop on Digital Mammography*, 2000
- [3] M. Heath, K.W. Bowyer, D. Kopans, R. Moore , P. Kegelmeyer, R. Moore, K. Chang, S. Munish Kumaran, Digital Database For Screening Mammography, *Proceedings of the 4th International Workshop on Digital Mammography*, 457-460 pp, 1998.
- [4] H. Munch, U. Engelman, Web-based distribution of radiological images from PACS to EPR, *International Congress Series*, 873-879 pp, 2003.
- [5] Y. Lin, D. Feng, T.W. Cai, A Web-based Collaborative System for Medical Image Analysis and Diagnosis, *Conference on Research and Practice in Information Technology*, vol 2, 93-95 pp, 2001.
- [6] J. Zhang, J. Sun, J.N. Stahl, PACS and Web-based image distribution and display, *Computerized Medical Imaging and Graphics*, vol 27 (2-3) 197-206 pp, 2003.
- [7] J. Kim, D. Feng, T. Cai, A Web-based medical image data processing and management system, *ACM International Conference Proceeding Series*, vol 9, 89- 91 pp, 2000.
- [8] P. Wunderbaldinger, W. Schima, K. Turetschek, T. H. Helbich, A. A. Bankier and C. J. Herold. World Wide Web and Internet: applications for radiologists. Volume 9, Number 6, 1170-1182, 1998.
- [9] Extensible Markup Language (XML).<http://www.w3.org/XML/>
- [10] Chung L, Hsu T, Deng C, Woei C, Chia H. A Web-based Solution for Viewing Large-Sized Microscopic Images. *Journal Digital Imaging*, vol 22, 275-285 pp 2009 (ISSN: 1618-727X)
- [11] J.M. Monguet, B. Blanco, E. Huerta, E. Castelán, M. Sampieri, M. Ferruzca, J. Fernández, H. Bascuñana. E-Therapy-Learning. *Proceedings of EDULEARN 2010 Conference*.
- [12] Zend server company official webpage: <http://www.zend.com/en/products/server/>
- [13] eXist official web page: <http://www.exist-db.org/>
- [14] Bolaños M, Vidal A, Navarro C, Valderrama JC, Aleixandre R. Usabilidad: concepto y aplicaciones en las páginas web médicas, *Papeles Médicos*, vol 16 (1), 14-21 2007
- [15] Berg WA, Campassi C, Langenberg P, Sexton NJ. Breast Imaging Reporting and Data System: inter-and intraobserver variability in feature and final assessment. *AJR Am J Roentgenol* vol 174, 1769-77 pp, 2000 (ISSN: 0361-803X)