

# Aplicación Web para la Gestión de una Base de Datos Pública en Mamografía Digital: MamODB

C. Mata Miquel, X. Lladó Bardera, R. Martí Marly, J. Freixenet Bosch, A. Oliver Malagelada, J. Martí Bonmatí

VICOROB, Grupo de Visión per Computador y Robótica

Departamento de Arquitectura y Tecnología de Computadores (ATC), Universitat de Girona

{cmata,llado,marly,jordif,aoliver,joanm}@eia.udg.edu

## Resumen

*Cada vez son más los hospitales que disponen de sistemas computarizados de adquisición y visualización de imágenes digitales, con las ventajas que ello supone cuanto a acceso a la información, capacidad de diagnóstico y aprendizaje. Sin embargo, el volumen ingente de datos requiere de nuevas herramientas para su almacenaje, gestión y recuperación. En este trabajo se propone un modelo de estructura basado en tecnología web como herramienta de ayuda al diagnóstico de Cáncer de Mama. La estructura propuesta se basa en la administración de imágenes y estudios mamográficos con el objetivo de ser un referente en la comunidad científica. Su arquitectura, metodología y aplicación en formato web se presentan en este trabajo así como conclusiones y trabajos futuros.*

## 1. Introducción

La imagen médica desempeña actualmente un papel muy importante en el estudio y comprensión de los exámenes de diagnóstico del paciente, lo que supone un gran incremento en la gestión, almacenaje y recuperación de imágenes digitales y diagnósticos. En este sentido, el desarrollo mediante aplicaciones web [1][2] de arquitecturas y sistemas de archivo de imágenes y comunicaciones PACS [3] supone una mejora de los actuales sistemas informáticos de soporte al diagnóstico. Sin embargo, numerosos centros hospitalarios que disponen de dichos sistemas han restringido su acceso a la red de área local, con lo que se limita su uso a estaciones de trabajo dentro del departamento o del mismo centro hospitalario.

En este sentido, el uso de aplicaciones Web orientadas al ámbito médico se ha incrementado notablemente durante los últimos años, y se ha convertido en un nuevo medio de comunicación para facilitar el acceso a la información tanto local como remota [4]. Los rápidos avances en tecnología informática y el potencial de Internet permiten el desarrollo de software en línea y la recuperación de datos de manera eficaz, fiable y económica, con lo cual desaparecen algunas de las razones que desaconsejaban el uso de aplicaciones web.

Por otra parte, la reciente inclusión del formato de archivos XML (eXtended Markup Language) [5] facilita la compatibilidad entre sistemas, ya que es el lenguaje que garantiza que los datos estructurados sean uniformes

e independientes de aplicaciones y fabricantes, asegurando el intercambio de información sin que se ocasionen problemas de tipo “contenido”. Los trabajos [6][7] son ejemplos de aplicaciones web en el ámbito médico con el uso de XML para anotaciones.

La motivación de este trabajo surge de la necesidad de proporcionar una herramienta a la comunidad médica y científica para la gestión de una base de datos de imágenes mamográficas digitales y sus diagnósticos asociados (MamODB), con el objetivo de facilitar su uso remoto a través de una aplicación web que incluya todas las funcionalidades de uso (autenticación, seguridad, consulta, etc.) asociadas a este tipo de aplicaciones.

## 2. Arquitectura del sistema

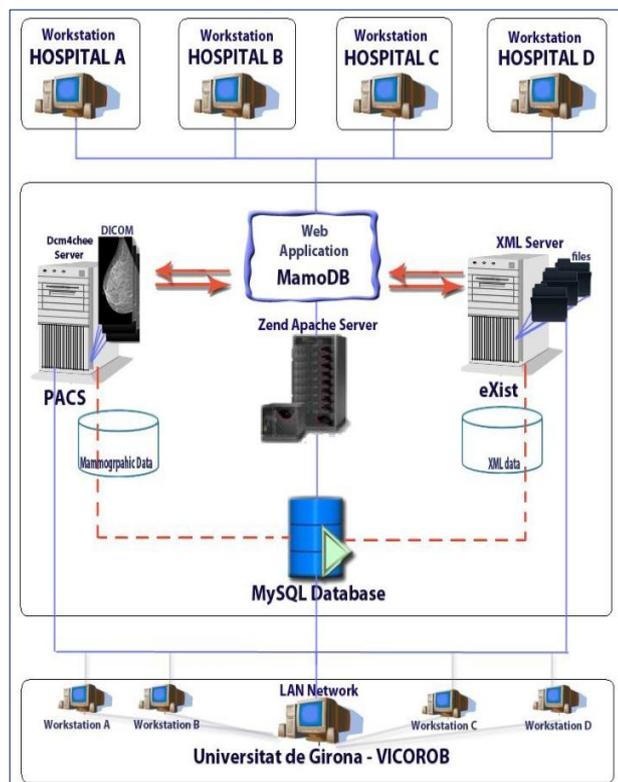
A partir de las características que debe poseer la aplicación MamODB, se ha diseñado un modelo de arquitectura, descrita en la *Figura 1*, con el propósito de facilitar a los usuarios el proceso de recuperación de la información, que se realiza de forma rápida y sencilla. Ello ha supuesto la integración de un sistema PACS que almacena los archivos DICOM procedentes de los diferentes hospitales con una nueva base de datos que contiene los ficheros de anotación XML, además del desarrollo de la propia aplicación web.

### 2.1. Diseño e implementación

La aplicación MamODB está alojada dentro del servidor principal Zend Apache, configurado asimismo como gestor de base de datos MySQL. Es importante remarcar que el sistema PACS crea de forma automática una base de datos propia dentro de MySQL (*pacddb*), donde se almacena la información referente al acceso y almacenaje de las imágenes mamográficas. Así, existe una relación directa entre el servidor Zend Apache con la base de datos de imágenes.

La aplicación MamODB también está relacionada con un gestor eXist, responsable de la base de datos XML, incorporando un propio servidor para su acceso y gestión mediante interface web. Se ha optado por la creación de una base de datos de anotaciones con objeto de preservar la integridad de los ficheros DICOM originales, y para asegurar su total compatibilidad al proceder de sistemas de adquisición diferentes. Todos los archivos XML que se

almacenan en eXist tienen una relación con las imágenes mamográficas a que hacen referencia, ya que se trata de los archivos que contienen las anotaciones de diagnóstico realizadas por los radiólogos. Por esta razón, se puede ver en el modelo de arquitectura de la *Figura 1* una relación directa entre MySQL y eXist. Finalmente, se ha definido el identificador o campo clave de los ficheros con sus respectivas imágenes mamográficas, de modo que la base de datos y los ficheros estén correctamente relacionados.



*Figura 1. Diseño de la arquitectura de la MamODB*

En cuanto al acceso a MamODB, y dado que pretende dar servicio tanto a especialistas médicos presentes en el propio centro hospitalario como a profesionales externos, se ha optado por el diseño de una aplicación web. Uno de los elementos clave en este tipo de aplicaciones es la seguridad de los datos e informaciones que viajan por la red, más aún cuando se trata de datos de carácter confidencial relacionados con exploraciones médicas. En este sentido, los tres módulos en que se ha estructurado la arquitectura del sistema (Zend Server, PACS y eXist-db) disponen de medidas de seguridad internas para evitar vulnerabilidades durante la transferencia de información, de forma que además de usar un control de identificación de usuario de claves privadas, los sistemas disponen de algoritmos que protegen y encriptan la información en su transferencia.

A continuación se explicará de forma resumida las características técnicas y funcionalidades de cada uno de los tres módulos de la arquitectura del sistema.

## 2.2. Zend Server

Zend Server [8] es un administrador para servidores de aplicaciones web que incluye servidor Apache y base de datos MySQL. Es una herramienta dotada de una fácil instalación, configuración y gestión de todas las funcionalidades internas tanto del servidor como de la base de datos. La aplicación web ha sido programada con lenguaje PHP siendo una alternativa a las tecnologías de Microsoft ASP, ASP.NET (que utiliza C# VB.NET como lenguajes) entre otras. Su creación y desarrollo está asociado a sistemas libres con licencia GNU y por ello muchos sistemas IDE incluyen la instalación de PHP en su software. Una de ellas es Zend Server, usado en este trabajo. Finalmente, permite administrar extensiones, actualizaciones y mejoras de su sitio web de forma totalmente gratuita.

## 2.3. PACS

Dado que la aplicación está orientada a la gestión de una base de datos de imágenes mamográficas digitales, la incorporación de un sistema de tipo PACS (Picture Archiving Communicating System) era imprescindible. Se ha usado para almacenar las imágenes mamográficas en formato digital (DICOM) procedentes de los diferentes centros hospitalarios que colaboran en el proyecto, y para su gestión a través de la red. Se compone de un servidor central que almacena una base de datos (*pacsdb*, comentada anteriormente) y de interfaces gráficas de entorno web como medio de comunicación. En cuanto a su instalación, se ha utilizado el sistema PACS *dcm4chee* [9], compuesto por una colección de aplicaciones de código abierto que instala un servidor interno conjuntamente con la interface gráfica para su gestión.

## 2.4. eXist-db: Base de datos nativa XML

eXist-db [10] es un sistema de gestión de bases de datos libre y de código abierto que almacena archivos XML de acuerdo a un modelo de datos XML. A diferencia de la mayoría de los sistemas de gestión de base de datos relacionales, eXist-db soporta distintos lenguajes de consultas XML, y más concretamente el lenguaje XQuery usado en este proyecto. Además, algunas de sus características más destacadas son la indexación de documentos, el soporte para la actualización de datos, y la capacidad de obtener los resultados de las consultas en formato XML. Finalmente, el acceso a eXist se realiza mediante una interface gráfica en entorno web, que facilita la gestión y comunicación con la aplicación web MamODB.

## 3. Aplicación Web MamODB

Esta sección describe de forma detallada y resumida las funcionalidades más importantes de la aplicación MamODB. Se ha estructurado en cuatro subsecciones: identificación y seguridad, panel de control y accesibilidad, gestión y consulta de la MamODB, y administración del PACS y eXist-db.

### 3.1. Identificación y seguridad

El acceso a MamODB por parte de los usuarios se garantiza mediante un nombre de usuario y contraseña

únicos, generados por el administrador del sistema. El sistema incorpora los mecanismos habituales de autenticación y de uso fraudulento de identidad.

Se han creado tres perfiles diferentes de usuario, de acuerdo con el acceso a diferentes funcionalidades del sistema. El primer nivel corresponde al usuario estándar, quien está facultado para realizar consultas sobre los diferentes casos almacenados en MamODB (en forma de imagen y diagnóstico). Básicamente, corresponde a un perfil de especialista médico o científico que utiliza la aplicación MamODB únicamente como consulta (telediagnóstico, formación, etc.). El segundo nivel de usuario incorpora, además de las anteriores funcionalidades de consulta, el acceso a PACS y eXist con el objetivo de modificar el contenido de la base de datos mamográfica, añadiendo (o modificando) casos clínicos (imágenes y diagnóstico). Fundamentalmente, corresponde a un perfil de especialista médico colaborador en la construcción de la base de datos mamográfica, y que tiene acceso al sistema de forma local o remota. El tercer perfil de usuario corresponde al de administrador del sistema, que añade a todas las anteriores funcionalidades la gestión de cuentas de usuario. Finalmente, es importante remarcar que la aplicación MamODB utiliza como encriptación de seguridad los algoritmos propios para cada componente, y se plantea como trabajo futuro el desarrollo de nuevos algoritmos de codificación interna que incrementen los mecanismos de seguridad y autenticación.

### 3.2. Panel de control y accesibilidad

Una vez los usuarios se han identificado correctamente, aparece el panel de control como página de inicio de la MamODB. Dependiendo de la tipología de usuario que accede al sistema, el panel de control muestra las diferentes posibilidades.

El panel de control ha sido diseñado de forma ergonómica, para facilitar al usuario un acceso de forma intuitiva y sencilla a todas las funcionalidades de la aplicación. Esta página de inicio está diseñada en forma de mapa web, pudiendo acceder a todas las funcionalidades del sistema en un solo clic. En la parte superior aparece un menú de barras desplegable común para todas las interfaces de la MamODB. Del mismo modo, existe una barra lateral extensible que el usuario puede hacer visible según sus preferencias. Su función consiste en facilitar un acceso directo a aquellas funcionalidades más usadas por los usuarios.

Además, el panel de control contiene una opción para cambiar los datos del perfil de usuario: datos personales, imagen de perfil y cambio de identificación del sistema (nombre de usuario y contraseña).

### 3.3. Gestión y consulta de MamODB

La petición de consultas a MamODB se realiza mediante un formulario, y a partir de los criterios de selección y/o de búsqueda el sistema proporciona un listado con los casos clínicos almacenados (incluyendo imágenes y diagnósticos). La *Figura 2* muestra un ejemplo de formulario de solicitud de consulta, donde se aprecian los

diferentes criterios de búsqueda que hasta el momento el sistema incorpora.

Además de un listado exhaustivo de los casos clínicos contenidos en la base de datos, el formulario de búsqueda permite realizar peticiones de acceso a casos clínicos de acuerdo a criterios médicos específicos: identificador de casos, diagnóstico, clasificación BIRADS [11], presencia de patologías o lesiones concretas (microcalcificaciones, distorsiones estructurales, masas, etc.)



*Figura 2. Ejemplo de un prototipo de formulario de la aplicación MamODB*

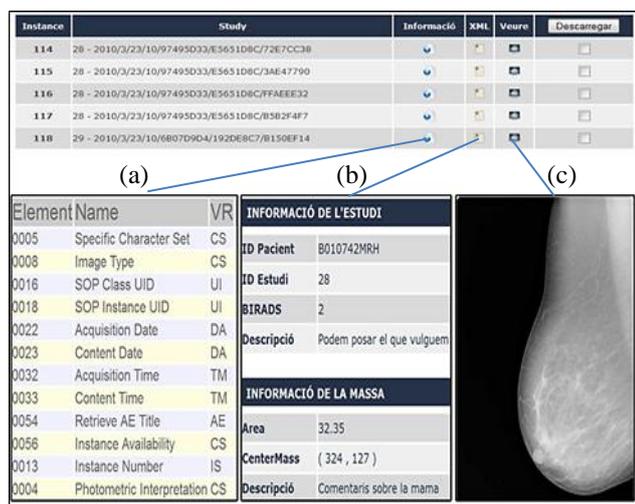
Además, el usuario puede seleccionar qué información referente al diagnóstico desea obtener (identificador de caso, descripción, área, tipo de patología, coordenadas de masa,...). En resumen, permite una búsqueda tan exhaustiva o personalizada como el usuario demandante crea necesario, de acuerdo con sus necesidades. En función del volumen de registros obtenidos en el listado de resultados, la aplicación usa paginación automática para agrupar los registros de una forma más visual.

Para cada uno de los registros obtenidos, el sistema permite realizar tres tipos de acciones: consultar la información referente a la imagen DICOM almacenada en el servidor de PACS, consultar su diagnóstico asociado en formato de fichero XML, y visualizar la imagen mamográfica. Un ejemplo de obtención de resultados está representado en la *Figura 3*, donde se aprecia en la parte superior el listado completo de resultados para una consulta concreta, mientras que la parte inferior muestra el resultado de las tres acciones posibles para un caso clínico concreto (a): información referente a la adquisición; (b): anotaciones diagnósticas asociadas al caso; (c): imagen mamográfica).

Una característica importante del sistema es que las acciones de consulta sobre los registros obtenidos pueden realizarse de forma visual en pantalla, o bien descargar localmente al centro de trabajo del usuario los archivos seleccionados. Para cada registro obtenido en el listado se puede seleccionar de forma manual, seleccionando individualmente las imágenes deseadas a descargar, o usando una selección automática. Una vez seleccionadas, los archivos DICOM conjuntamente con sus archivos XML asociados pasan a formar parte del archivo en formato ZIP "mamODB.zip", que es descargado por el usuario.

### 3.4. Administración: PACS y eXist-db

Se trata de una funcionalidad accesible únicamente para los usuarios de segundo y tercer nivel, con el objetivo de modificar el contenido de la base de datos. El acceso se realiza mediante interfaces gráficas propias de PACS y eXist, y permite no únicamente incorporar nuevos casos a la base de datos mamográfica, sino también tareas de mantenimiento de los casos existentes. En consecuencia, para acceder a dichas interfaces se requiere también una identificación adicional que debe ser facilitada por el administrador del sistema.



The screenshot shows a web-based interface for MamODB. At the top, there is a table listing study instances with columns for Instance, Study, Informació, XML, Veure, and Descarregar. Below this, a detailed view of a study instance is shown, including a table of element names and values, and a thumbnail image of a mammogram. Labels (a), (b), and (c) point to specific parts of the interface.

Instance	Study	Informació	XML	Veure	Descarregar
114	28 - 2010/3/23/10/97495033/E5651D8C/72E7CC38				
115	28 - 2010/3/23/10/97495033/E5651D8C/3AE47790				
116	28 - 2010/3/23/10/97495033/E5651D8C/FFAEE32				
117	28 - 2010/3/23/10/97495033/E5651D8C/B382F4F7				
118	29 - 2010/3/23/10/6807D904/192DE8C7/8150EF14				

Element Name	VR	INFORMACIÓ DE L'ESTUDI
0005	Specific Character Set	CS
0008	Image Type	CS
0016	SOP Class UID	UI
0018	SOP Instance UID	UI
0022	Acquisition Date	DA
0023	Content Date	DA
0032	Acquisition Time	TM
0033	Content Time	TM
0054	Retrieve AE Title	AE
0056	Instance Availability	CS
0013	Instance Number	IS
0004	Photometric Interpretation	CS

INFORMACIÓ DE L'ESTUDI	
ID Pacient	B010742MRH
ID Estudi	28
BIRADS	2
Descripció	Podem posar el que vulguem

INFORMACIÓ DE LA MASSA	
Area	32.35
CenterMass	( 324 , 127 )
Descripció	Comentaris sobre la mama

Figura 3. Ejemplo de un listado de resultados de la MamODB

## 4. Conclusiones

En este artículo se ha presentado una propuesta de arquitectura para implantar y mantener un sistema de gestión de imágenes mamográficas y de diagnósticos asociados a la aplicación MamODB. Este sistema ofrece la posibilidad de gestionar y centralizar una base de datos pública de imágenes digitales para la comunidad médica y científica, y su acceso seguro mediante aplicación web.

Aunque en la actualidad existen bases de datos públicas de imágenes mamográficas, algunas de ellas no han sido adquiridas digitalmente, sino que provienen de adquisiciones analógicas y la posterior digitalización de las películas [12-13]. En el caso de imágenes mamográficas, dicho proceso comporta una inherente pérdida de calidad, dependencia del sistema digitalizador, pérdida de la información de adquisición y la no estandarización de formatos de imagen. La arquitectura del sistema propuesto permite administrar y centralizar una base de datos pública de imágenes mamográficas adquiridas digitalmente. Además, usando la aplicación web MamODB, se tiene acceso a la información tanto de forma local como remota.

Como trabajos futuros para mejorar MamODB, se plantea definir un formato único de ficheros XML para el colectivo médico, correspondiente a las anotaciones mamográficas de sus imágenes asociadas. Asimismo, definir de igual modo formularios de búsqueda personalizados que aseguren resultados óptimos y se

adecuen a las necesidades del usuario. Finalmente, se podrían investigar nuevas mejoras en cuanto a la seguridad y transferencia de datos como el uso de nuevos algoritmos para mejorar el encriptado de datos y ofrecer un sistema más seguro y robusto.

## Agradecimientos

Los trabajos realizados en este artículo han sido parcialmente financiados a través del Ministerio de Ciencia e Innovación (proyecto TIN2007-60553) y la Generalitat de Catalunya (proyecto 2008SALUT00029).

## Referencias

- [1] Munch H, Engelman U. Web-based distribution of radiological images from PACS to EPR. *International Congress Series*, 873-879 pp 2003.
- [2] Lin Y, Feng D, Cai TW. 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 (ISSN: 1445-1336)
- [3] Choplin RH, Boehme JM, Maynard CD, Picture archiving and communication systems: an overview. *Radiographics* January vol 12, 127-129 pp 1992 (ISSN: 0271-5333)
- [4] 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
- [5] Extensible Markup Language (XML). <http://www.w3.org/XML/>
- [6] 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)
- [7] Kim J, Feng D, Cai T. A Web-based medical image data processing and management system, *ACM International Conference Proceeding Series*, vol 9, 89- 91 pp, 2000 (ISSN:1445-1336)
- [8] Zend Server Company. <http://www.zend.com/en/products/server/>
- [9] Dcm4chee Open Source Clinical Image and Object Management. <http://www.dcm4chee.org/>
- [10] eXist-db Open Source Native XML Database. <http://exist.sourceforge.net/>
- [11] 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)
- [12] Heath M, Bowyer KW, Kopans D, Moore R, Kegelmeyer P, The Digital Database for Screening Mammography, *Proceedings of the 5th International Workshop on Digital Mammography*, Toronto, Canada, 2000 (ISBN: 1-930524-00-5)
- [13] Heath M, Bowyer KW, Kopans D, Kegelmeyer P, Moore R, Chang K, MunishKumaran S, Digital Database For Screening Mammography. *Proceedings of the 4th International Workshop on Digital Mammography*, 457-460 pp, 1998