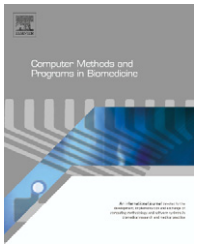
COMPUTER METHODS AND PROGRAMS IN BIOMEDICINE 86 (2007) 255–269

journal homepage: www.intl.elsevierhealth.com/journals/cmpb



A database for medical image management

Esperanza Marcos[a], *César J. Acuña*[a,*], *Belén Vela*[a], *José M. Cavero*[a], *Juan A. Hernández*[b]

[a] *Kybele Research Group, Rey Juan Carlos University, C/Tulipán s/n, 28933 Móstoles, Madrid, Spain*
[b] *GTEBIM Research Group, Rey Juan Carlos University, C/Tulipán s/n, 28933 Móstoles, Madrid, Spain*

ARTICLE INFO

Article history:
Received 8 May 2006
Received in revised form
9 March 2007
Accepted 9 March 2007

Keywords:
Databases
Medical images
Model driven development
XML
DICOM

ABSTRACT

MEDIMAN (Medical Image MANagement) is a web information system (WIS) for medical image management and processing currently used by neuroscientists and clinicians at several medical and research centres in Spain for research and clinical trials. While developing the MEDIMAN database (DB) we encountered several design challenges unlike those arising in traditional DBs. This paper describes the development of MEDIMAN focusing on the database and the use of the database development process proposed in Midas, a model-driven framework for WIS development. Special attention is given to the design decisions made at each stage to address the challenges encountered.

1. Introduction

At present, medical image modalities like positron emission tomography (PET) and functional magnetic resonance imaging (fMRI) make it possible to obtain images of quantitative and qualitative cerebral activity, which neuroscientists and clinicians like neuroradiologists, neurologists, neuropsychiatrists, etc., use in their research. If, for example, they are running clinical trials to evaluate a new drug or pathology, the images acquired must be processed in very different ways for their later correction and statistical analysis.

The volume of data and the complexity of its management mean that researchers do not only need a suitable storage service for the huge number of images they use, but they also need to process them. This involves applying standardized procedures to different sets of images for such purposes as registering, filtering, statistical analysis, viewing and storing in several formats. Moreover it is very useful to represent the different formats they use in one single format to enable cross-application information exchange. Most of the images are acquired in DICOM [1] (digital imaging and communication in medicine), the most widely accepted standard for the interchange of medical images in digital format [2]. Not only is DICOM used as a format but ANALYZE [3] is also considered, which, though a proprietary format, is widely used for fMRI and PET. MEDIMAN [4–6] is a web information system (WIS) for medical image management and processing through the Web involving different hospitals and research institutions, mainly, though not only, in the neuroimaging area. Neuroscientists and clinicians are using it successfully for research and clinical trials in several medical and research institutions such us Ruber International, Fundación Hospital de Alcorcón and the University Hospital of Salamanca, among others. The research currently under way with the images stored in MEDIMAN includes studies of psychic diseases such as schizophrenia and photophobia, and of migraine and other types of pain.

* *Corresponding author*. Tel.: +34 91 488 8140; fax: +34 91 488 8530.
E-mail addresses: esperanza.marcos@urjc.es (E. Marcos), cesar.acuna@urjc.es (C.J. Acuña), belen.vela@urjc.es (B. Vela), josemaria.cavero@urjc.es (J.M. Cavero), juan.tamames@urjc.es (J.A. Hernández).
0169-2607/$ – see front matter © 2007 Elsevier Ireland Ltd. All rights reserved.
doi:10.1016/j.cmpb.2007.03.006

The main goals of MEDIMAN are to offer neuroscience researchers and clinicians easily accessible case history image storage and suitably standardized processing and analysis of medical images. It also stores the results of medical image processing jointly with the original images in a database (DB). Another of MEDIMANs major aims is to facilitate medical information exchange by means of a common representation of the different formats for image file representation.

Today, medical imaging units can store and transmit medical images using picture archiving and communication systems (PACS) and give information about them using part of the function provided for radiological information systems (RIS). Such clinical information systems are oriented to individual patient information management, but not to group management, under different experimental conditions. Management of groups of patients and of the mathematical procedures applied to their data is not possible with existing software tools of this type. Similar tools have been designed for cardiology and cancer multicentre trials [7–10], but with no use of images. Nor do existing tools permit storage of information on clinical trials with images, such as the post-processing procedures applied, algorithm parameters, and in general, information for further qualitative and quantitative analyses. MEDIMAN fills this gap by permitting the storage of the image, post-processing information, and image processing results.

Some existing tools share some of MEDIMANs objectives, e.g. MyPACS [11], Casimage [12], MedPix [13] and MIRC [14], the first three of which aim to provide an online medical image storage service for teaching and research purposes. Nevertheless, the MIRC site allows users to access materials published on several participating sites, i.e., MIRC allows information retrieval from several image repositories. These tools only offer storage facilities but no other standardized image processing service.

Although most of the mentioned proposals use the eXtensible Markup Language (XML), none of them store the XML information in a XML database as is the case of MEDIMAN. Some of them manage the XML files individually as a file system which became a problem when the number of image files grows. The other approaches use OR databases to store the XML data so, as we discuss in Section 6, part of the semantic of the XML file could be lost. It should be pointed out that they do not use XML to encode the information of the DICOM and Analyze files, but they do use it to encode specific teaching files, including metadata about the image files.

During the MEDIMAN DB development process we encountered several design challenges differing from those found in traditional DBs, which we attribute to two causes: the kind of information that MEDIMAN has to handle – structured data, semi-structured information, images, repetitive groups, etc. – and the exchange facilities required for image file information.

The work presented here focuses on the storage service of MEDIMAN, and on its database design and development, with special emphasis on the ways the different challenges were overcome.

Two categories of data are managed by MEDIMAN: firstly, medical image information, i.e. the data of each image file and the results of the image processing; secondly, information on clinical trials, etc. Although the former may be considered as semi-structured data, the information in the latter category is highly structured. The main reasons for differentiating between the two types of data, which constituted one of the main challenges in the design and development of the MEDIMAN database, are explained in detail in Section 2. This study addresses the problem by merging two different database technologies for storing the different types of information: an object-relational (OR) database for the highly structured data and an XML-DB to manage the semi-structured set. Moreover, representing DICOM and Analyze information in XML improves the exchange and integration of medical information in a broader context than the purely medical. There is a range of tools and equipment compliant with the DICOM and Analyze formats which can read and process image data in them, but the exchange capability among them is limited. However, by using XML, which is a standard exchange language, not only in the medical environment but also in almost any discipline, we ensure the sharing of medical image data by means of any tool or equipment able to read XML data.

Multimedia data usually requires special technology for optimal storage, access, indexing and retrieval. Relational, object-oriented and OR DBs have been used (sometimes extended with appropriate characteristics such as types, objects, query languages) for storing this kind or data [15,16]. Several database management systems (DBMSs) support multimedia data types (Informix Dynamic and DB2 Universal DataBase of IBM, Oracle 10g, CA-JASMINE, Sybase, etc.). Of these, we chose Oracle 10g because, in addition to supporting multimedia data types, it integrates XML and OR technologies very well.

In the development of the MEDIMAN database we applied Midas [17], a model-driven framework for the development of WIS with architecture based on the model-driven architecture (MDA) proposed by the Object Management Group (OMG) [18,19]. This model-driven approach has several advantages, e.g. maintenance and migration are easier (owing to the development of different models at different levels of abstraction), as is the semi-automatic transformation between models by means of a set of mapping rules, which facilitates the creation of development tools. This study will show the usage of these models and Midas techniques in relation to WIS database development.

There have been several attempts to address the database development process, such as [20,21] for OR databases, and initiatives such as [22–25] to transform unified modeling language (UML) conceptual models into XML schemas. Nevertheless, we have found no other methodological framework for the systematic development of DBs which takes into account both structured (OR) and semi-structured (XML) data jointly, as was necessary for the development of the MEDIMAN database.

This study addresses the development of MEDIMAN from the point of view of its database, using the development process proposed in Midas, with special emphasis on the design decisions made at each stage.

The remainder of the paper is structured as follows. Section 2 describes the background to the project, with mention of related work; Section 3 discusses the design considerations made during the development of MEDIMAN and the methods used to develop its database; Section 4 describes MEDIMAN, detailing its goals, architecture, functionality and the database development process; Section 5 presents a status report of