

Web Portal for Genomic and Epidemiologic Medical Data

Mónica Miguélez Rico

University of Coruña, Spain

Julián Dorado de la Calle

University of Coruña, Spain

Nieves Pedreira Souto

University of Coruña, Spain

Alejandro Pazos Sierra

University of Coruña, Spain

Fernando Martín Sánchez

University of Coruña, Spain

INTRODUCTION

Medical data and digital imaging for medical diagnosis currently represent a very important research area in computer science. The generation of medical information is continuously increasing. More specifically, genomic (molecular and histological) data and images have become key points for diagnosis. The specific processing these data require is more and more requested.

This article describes a Web portal based on the most common current standards. This platform is not only able to integrate the medical information available at several sources, but also to provide tools for the analysis of the integrated data, to use them for the study of any pathology. It will provide a common access point to share data and analysis techniques (or applications) between different groups that are currently working in several fields of health area.

BACKGROUND

Nowadays, several studies are being carried on with regard to the different levels of information about health (population, disease, patient, organ, tissue, molecule, and gene) but none of them integrates the information. The biomedical computer science must play an important role at the integration of these viewpoints and their data.

From a classical viewpoint, computer science in public health has been able to confront and solve problems at different population levels; has effectively managed levels of diseases and patients and lastly; has developed tools for image management and analysis to be used in non-invasive

techniques for tissue or organ study. The source of knowledge regarding molecular and genetic levels is greater every day. One of the fields were developing new applications is Genomic Epidemiology, which performs population studies about the impact of genetic human variability on health and disease. Another field, Pharmagenetic, considers the differential genetic aspects among people (e.g., SNPs profiles) when developing new medicines and analyzing its influence after the administration of a medicine.

HUGE NET (from Office of Genetics and Disease Prevention (USA)) is an example of this kind of application. Briefly, it is a communication network that allows sharing epidemiological information about Human Genome.

PharmGKB program (from Stanford University) is used in nine universities and medical centers, which investigate pharmacogenetics. The program makes a knowledge base possible with genomic data, laboratory phenotypes, clinic informations, etc.

However, these examples solve just partial aspects of the aim, but not the complete problem. Nowadays, there are not examples of integrated information systems to cover this kind of study completely. The development of such a system will facilitate the studies about complex diseases.

Digital imaging for medical diagnosis is currently one of the most relevant research areas. Since the discovery of the x-ray in 1895, the techniques for acquisition of medical images have evolved to images in digital format.

Every manufacturer used to design its own image storage format, therefore the development of applications should be specific for every device. Therefore, it makes it impossible to transfer information between different machines. A standard named DICOM (Digital Imaging and Communications in

Medicine) was published (Bidgood & Horii, 1996; Clunie, 2005; Nema, 2005) as a solution for these problems. DICOM unifies imaging storage criteria for their transmission among heterogeneous equipment by a common procedure, which is open and public.

Another problem related to medical imaging is its accurate management, mainly due to the great volume to store. This way, the picture archiving and communication system (PACS) (Huang, 2004) makes the achievement of an imaging service that might integrate images and clinical information without films or paper documents possible.

The PACS DICOM duet, combined with Web technology provides the specialist with the possibility of gaining access to images and their related information from place, using the legally required security mechanisms (BOE, 15/1999, BOE, 994/1999, Garfinkel & Spafford, 2001).

The existing health databases and Web portals are heterogeneous and physically dispersed. These DBs may be relational, as PACS DICOM, public, as NCBI (NCBI, 2005), or HapMap (HapMap, 2005), etc. Therefore, there also exists a great variety of software for data processing. There are some development platforms for Windows and Linux in different programming languages as Java or C, several commercial tools for image management like Quantity One from Bio-Rad Laboratories (Bio-Rad, 2005), LabImage (LabImage, 2005), Phoretix 1D developed by Nonlinear Dynamics (Nonlinear, 2005), or Label Cell Counter Software create in the Image Management Laboratory, Otolaryngology Department, Rochester University Medical Center.

These are potent tools, which cover the requirements of this kind of image, although not always in an automatic way. Besides, they are a commercial software, so it is not possible to add new functionalities and, in most cases, they can deal just with a specific type of images.

The previous circumstances disturb not only the access to information, but also the processing of data and image. For instance, to perform the study of any disease, the first

step will consist on locating the different DBs containing the desirable information. Secondly, it is necessary to generate the appropriate queries to the DBs in a specific language and with a specific structure. Finally, obtained data must be adapted to every program wanted for the analysis. The process is, consequently, a tough task.

It would be desirable to have systems able to store, relate, manage, and visualize all the data and the information coming from several studies, and process it as a homogeneous dataset instead of multiple and separate sources.

These systems should be developed ready for their integration in a Web environment; which would facilitate independence of place and time and user personalization. In addition, the easy use of this environment decreases the learning time.

PROPOSED SOLUTION: WEB PORTAL

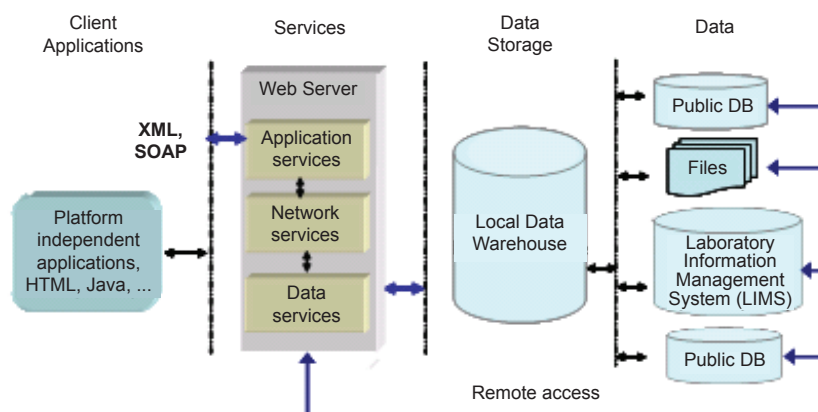
The proposed solution lies in a Web portal for managing and accessing heterogeneous information stored at several repositories. It also provides the different users with a tool repository for data processing.

The system can work in two different ways. First, it is able to generate specific applications for a given pathology. These applications are Web interfaces for retrieving information from several data repositories. Thus, the user can visualize relevant information and images by means of processing algorithms adjusted to his or her needs.

Besides, the system has a services layer where advanced users can include data he or she wants to analyze. This layer also provides the user with Web services (Colin, 2005, Sun, 2005) for accessing information from any application.

The developed system, which fulfills the previous requirements, is a four level platform: user applications, services, data storage, and data source levels are represented in Figure 1.

Figure 1. Platform architecture



5 more pages are available in the full version of this document, which may be purchased using the "Add to Cart" button on the publisher's webpage: www.igi-global.com/chapter/web-portal-genomic-epidemiologic-medical/18021

Related Content

Eidsvoll 1814: Creating Educational Historical Reconstructions in 3D Collaborative Virtual Environments

Ekaterina Prasolova-Forland and Ole Ørjan Hov (2011). *International Journal of Web Portals* (pp. 1-14). www.irma-international.org/article/eidsvoll-1814-creating-educational-historical/60246

Effective Feature Selection Using Hybrid GA-EHO for Classifying Big Data SIoT

Iyapparaja Mand Deva Arul S (2020). *International Journal of Web Portals* (pp. 12-25). www.irma-international.org/article/effective-feature-selection-using-hybrid-ga-eho-for-classifying-big-data-siot/245742

Research Essay: Improving Our Approach to Internet and SOA Projects

Neil Richardson (2010). *International Journal of Web Portals* (pp. 52-56). www.irma-international.org/article/research-essay-improving-our-approach/49567

Learning Geography with the G-Portal Digital Library

Dion Hoe-Lian Goh (2007). *Encyclopedia of Portal Technologies and Applications* (pp. 547-553). www.irma-international.org/chapter/learning-geography-portal-digital-library/17927

Semantic Portals

Brooke Abrahams (2007). *Encyclopedia of Portal Technologies and Applications* (pp. 887-893). www.irma-international.org/chapter/semantic-portals/17981