

# A medical knowledge management system based on expert tagging (MKMST)

A. Baigorri, J. Villadangos, J. J. Astrain & A. Córdoba  
*Mathematics and Computers Engineering Department,  
Universidad Pública de Navarra, Spain*

## Abstract

This paper introduces a prototype of a medical knowledge management system (MKMST) for visual expertise in medical diagnostic reasoning and medical education. The system proposed here, developed following medical expert (radiologists) recommendations, is based on medical image tagging and annotation. Medical images stored in picture archiving and communications systems (PACS) are accessed by medical experts to diagnose pathologies by means of a tagging system and the generation of a taxonomy. The system stores the annotations in a relational database management system and interacts via DICOM standard and HL7 with PACS and the health records. The knowledge base consists of the translation from tacit knowledge to explicit knowledge.

*Keywords: medical knowledge, tagging system, medical diagnostic imaging, expert annotation, UMLS.*

## 1 Introduction

Healthcare organizations are not able to fully manage the knowledge that they possess and KM tools allows them to acquire potentially useful knowledge and to make it available to those who can use it at a time and place that is appropriate. The use of knowledge management systems (KMS) in healthcare organizations is growing notably despite healthcare is a captive market where medical technology manufacturers do not encourage the interoperation among equipment from different manufacturers and proprietary formats are widespread. Furthermore, healthcare managers are aware that knowledge management offers organizations many tools, techniques and strategies to apply to their existing business processes.



Healthcare is a domain which offers a unique opportunity to analyze extremely large, rich and complex data sets of information concerning medical aspects, but also administrative and management ones. The increasing pressures to both reduce costs and improve the quality of healthcare, to reduce medical errors, to improve the training of professionals and bring current research developments into practice are some of the reasons for the introduction of KMSs.

The four key steps involved in KM (creating/generating, representing/storing, accessing/using and disseminating/transferring knowledge) are associated with numerous complications due to the confidentiality of many clinical data. Medical Knowledge requires residents/fellows to demonstrate knowledge of established and evolving biomedical, clinical, epidemiological, and social-behavioral sciences, as well as the application of this knowledge to patient. Every day, new patients give physicians new inputs to change, test, validate or consolidate their knowledge. Collecting this knowledge is a very important start point to later provide on with reliable second-diagnosis procedures, better residents instruction and identification of future disease trends in any early stage. The main challenge is to collect high-quality medical knowledge in a comprehensive, useful, and accurate way. Medical knowledge must be updated whenever new medical findings become available without disturbing physicians from their daily activities. The volume of medical knowledge doubles itself every 17 years. There is also the issue of the various formats in which information exists as well as the diverse disparate medical information sources. Taking these aspects into account, it is not surprising that knowledge management is attracting so much attention as described by Stroetmann and Aisenbrey [1].

Tacit knowledge relevance in medical diagnosis and its implantation on medical education is addressed by Heiberg [2], which explains the importance that educators and trainees are aware of the elements of the tacit dimension and include this awareness in their educational planning, as suggested by Leron and Hazzan [3]. The role of tacit knowledge in medical diagnostic reasoning is discussed in Henry [4], Goldman [5] and Tsoukas [6]. Henry [4] argues that evidence-based medicine presupposes an inaccurate and deficient view of medical knowledge and tacit knowledge may remedy this deficiency. As opposite, Goldman [5] exposes that clinical judgment is a fully explicit process compared to the relatively neglected view that tacit knowledge plays a substantial role in the clinician's mental operations. Tsoukas [6] argues that tacit and explicit knowledge are two sides of the same coin since even the most explicit kind of knowledge is presupposed by tacit knowledge. A recent study concerning tacit and explicit knowledge in public health in Canada (Kothari *et al.* [7]), provides insights into how tacit knowledge comes into play during the program planning process in public health. These results will be useful in guiding the development of future knowledge translation strategies for public health organizations and decision makers.

Crowley *et al.* [8] argue that analytical reasoning is very important in the visual specialties of pathology, radiology and dermatology. This is because these specialists are based on their own experience. Consequently, it is extremely



important that novices of visual specialties are given enough time to be able to build up an ever-growing knowledge base.

Some proposals focus on knowledge-based community that connects hospitals, clinics, pharmacies, physicians, and customers for sharing knowledge, reducing administrative costs, and improving the quality of care as described in Metaxiotis [9]. Furthermore, mobile electronic medical record and web applications are used to provide additional tools which help physicians provide safe, cost-effective, better-quality care, allowing online access to scientific medical information, review articles, graphical presentations of disease workflows, tele-medicine... Clinical knowledge management (CKM) techniques and tools are addressed in Sittig *et al.* [10]. All those tools are implemented using web technology. Proprietary CKM tools, as those of Siemens are addressed in Stroetmann and Aisenbrey [1]. A web portal that integrates information from PatientsLikeMe, WebMD, and PubMed is introduced in Chun and MacKellar [11], although such tools, perhaps because of their complexity, are not well accepted by the medical community (Sittig *et al.* [10]). Other tools closer to that proposed in this article are addressed in Rubin *et al.* [12], Bo Hu *et al.* [13] and Möller *et al.* [14]. All of them use ontologies for the annotation and retrieval of medical images in order to provide semantics to those annotations. In our case, these tools have been qualified by the radiologists with whom we have collaborated as "overly complex and not too much reliable in retrieving the information from the medical images and unsuitable for the formation of novices". They underline the relevance that the radiologist manages himself the available semantic information of tags himself instead the system.

This paper presents a prototype of a knowledge management system (MKMST), which allows both visual expertise in medical diagnostic reasoning and medical education using tagging on medical images annotated by medical experts. In our case, we have collaborated with a group of radiologist from the *Complejo Hospitalario de Navarra*. These experts look for an efficient, simple and easy to use MKMST where radiodiagnostic tacit knowledge may be collected and accessed by both experts and novices.

The rest of the paper is organized as follows: Section 2 is devoted to describe the architecture of the system, while the tool that manages the system (MKMTool) is described in depth in Section 3; finally, conclusions, acknowledgements and bibliographical references end the paper.

## 2 MKMST architecture

The system comprises a hardware controlled by a tool (MKMTool) that manages the system's knowledge. The hardware consists of one or more picture archiving and communications system (PACS) servers, one or more database servers where annotations are stored, one or more UMLS servers and a set of workstations connected by a network. The PACS is the information repository which provides storage and convenient access to both medical images (from multiple modalities: X-ray, computed tomography, magnetic resonance imaging...) and clinical reports. Database servers manage and store the



annotations of those images and reports. The annotation information stored concerns users, taxonomies, tags and resources. The Unified Medical Language System (UMLS) [15], maintained by the US National Library of Medicine, is a compendium of many controlled vocabularies in the biomedical sciences which provides a mapping structure among these vocabularies allowing the translation among different terminology systems. Some users consider it as an ontology of biomedical concepts. The UMLS semantic servers store information about the terms of the annotations. MKMTool connects to these servers using web services and workstations run instances of MKMTool. Workstations interact with the PACS following the DICOM standard. Figure 1 illustrates the architecture of MKMST.

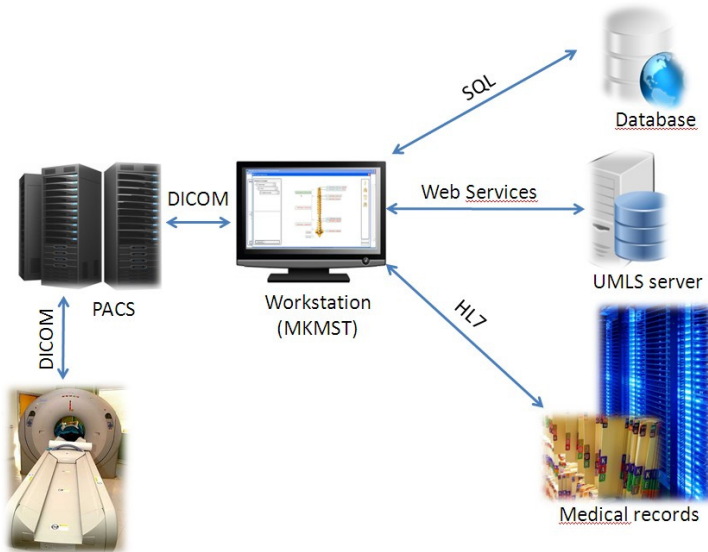


Figure 1: System architecture.

The medical knowledge contained in diagnostic medical imaging may be adequately managed with the aid of expert tagging. Physicians tag medical images stored in PACS in order to diagnose pathologies and injuries. These annotations, which are tuples  $\langle \text{user}, \text{tag}, \text{resource} \rangle$ , performed over image resources allow the automatic classification of medical resources (computed tomography or magnetic resonance imaging, blood tests...) facilitating the relationship of medical cases according to diagnosed diseases, the annotations made on diagnostic tests and the medical history.

As medical images are obtained, they are achieved in the data storage server and then radiologists start the diagnostic process. They annotate the images including the pathologies detected and those other aspects they consider relevant. Such annotations, performed using tags, are stored in the Database handled by a relational database management system or a set of them. MKMTool stores in its database, all the metadata provided by the service Averroes [16]. This service

provides the educational services required for learning all the pathologies of the human body. The radiologist starts selecting the anatomical structures of the different human body systems in order to facilitate and automatic the tagging procedure. As physicians annotate the resources in their workstations, the system will be creating a taxonomy which will guide the annotation process. Annotations are saved into the Database together with the identifier of the medical history and the image reference. Therefore, it is added content and knowledge to medical images. In order to cope with this knowledge, MKMTool integrates an algorithm for the automatic classification of resources, described in Córdoba [17] and denoted SAM, which classifies resources according to their annotations. SAM performs its work in background without disturbing the physician work. At this point medical knowledge derived from expert tagging is classified, and then it may be retrieved, according to the medical taxonomy and also according to the automatic semantic classifier. The Unified Medical Language System (UMLS) allows the expansion of terms, since this system represents numerous relationships among concepts. Consequently, new tags are added to improve taxonomy.

All tagging produced previously will serve to make studies and/or research about pathologies observed due to knowledge stored in the database, which is classified and related to the assistance of an expert system as SAM or manually following the taxonomy provided by UMLS. As the tool is used by medical professionals, the knowledge will be increased with each new annotation, and the system will be enriched permanently. Also, this tool can be used in any medical field and improve the training of new medical professionals as well.

This tool is devoted to facilitate the daily tasks of physicians, but also to improve the training of new medical professionals. Tacit knowledge management is based on medical image tagging (the process which achieves the knowledge) and on the exploitation of the information by searching, retrieving and relating this information. Physician's experience is collected and shared in a collaborative way. Figure 2 shows the information management performed by MKMTool.

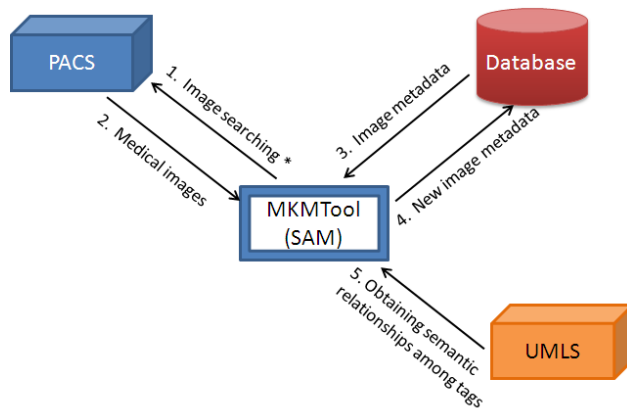


Figure 2: MKMTool, basic operation.

Medical images hosted into the storage servers of the PACS are accessed by MKMTool using the DICOM standard. The same occurs with the medical records, which are accessed via HL7. The information concerning the annotations performed by the experts is stored on a relational database management system and accessed using the Standard Query Language (SQL).

Image annotation requires the storage of information concerning the resource (medical image in such case), the tag and the user which performs the annotation. Tags may be obtained from the UMLS server through a web service and experts may also define their own tags. The resulting taxonomy is stored into the relational database.

A relevant aspect of the annotations is the geographical location of the tags. Radiologists mark the areas to annotate and situate the tags in relevant places of the image. This location is taken into account preserving a relative coordinate system for each image and resolution which is also achieved into the relational database.

Metadata provided by the service Averroes is also achieved into the database, allowing a reduction in the communication latency of the system.

### 3 MKMTool description

Medical experience is reflected during the diagnostic process. A type of mark in a radiological image, to physicians represents information. For example, their experience said that the mark has an irregular shape and is located in a certain position so, there is a high probability that this mark is a fat accumulation. Physicians can tag this area with several labels indicate location, pathology, diagnostic etc. In conclusion, tag reflects physician's opinion, and hence, their experience is reflected. The system allows the access to several physicians to value the opinion of their colleagues or add their own tags as well.

Initially, MKMTool authenticates the user by its login and password and then MKMTool retrieves of the metadata provided by the Averroes service in order to allow the physician to select the body system to annotate (circulatory, respiratory, bone...). Figure 3 depicts the main window obtained by the physician.

Then, the physician will access the PACS to retrieve those resources (images and reports) that want to annotate and/or query. Resource search and retrieval may be performed by the medical history of the patient, by pathologies (considering different granularity levels) and tags; or a combination of them. Once images are obtained, the physician starts (or continues) its annotation work. The annotation implies the storage of tuples  $\langle \text{user, resource, tag} \rangle$  where the user is well known due to the authenticated login, the resource is the image or report analyzed and the tag is the term, comment or note introduced by the physician for such resource. A resource may be tagged, or not, by many experts, an expert may tag many resources, and many tags can be used/defined by experts. At this point, MKMTool connects to the UMLS servers in order to retrieve the semantic information concerning the terms that the expert may use in their annotations. The metathesaurus forms the base of the UMLS and comprises over 1 million biomedical concepts and 5 million concept names, all of which stem from the

over 100 incorporated controlled vocabularies and classification systems as MeSH, SNOMED CT and others. Each concept in the metathesaurus is assigned one or more *semantic types* or categories, which are linked with one another through *semantic relationships*. The semantic network is a catalogue of these semantic types and relationships. This is a rather broad classification; there are 135 semantic types and 54 relationships in total. UMLS provides facilities for natural language processing, so MKMTool includes this information in its left panel.

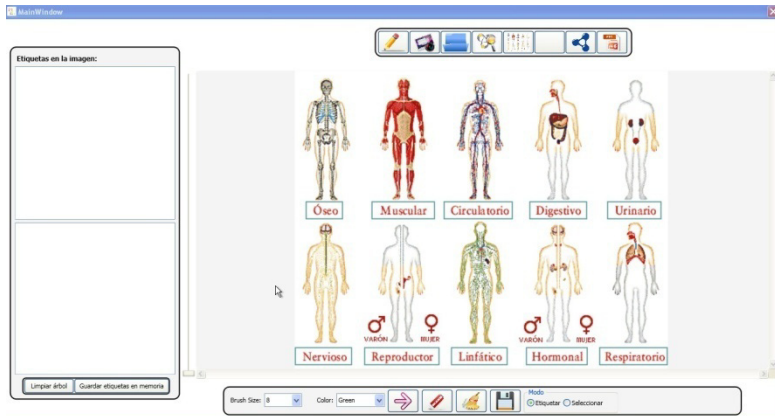


Figure 3: MKMTool, main window (in Spanish).

Annotations, as well their location coordinates in the resource annotated, are achieved into the database. Resources and annotations are stored in different places (PACS' storage server, and database respectively) and may be linked when desired. The tags used are showed in the left panel of MKMTool as taxonomies.

Our system allows selecting different elements for the diagnostic: image localization (system) and pathology, with a simple click of mouse. If all of them were encoded, physicians could make whole process by mouse clicks, but as design requirement, physicians want to get the chance to write down free text to determinate de tag that identifies the final diagnostic.

The main window consists of: a) a main panel, consisting of figures representing the different body systems; b) a left panel where taxonomy and tags are showed; and c) two toolbars. As focusing on one part of the figures of the main panel, other figures relating to the first appear in the window. All of them are tagged (UMLS) and automatically generate a taxonomy in the left panel. Once selected a given tag, MKMTool retrieves the images related to this tag from the PAC. Figure 4 illustrates the selection of the cervical vertebrae tag. Then, MKMTool catalogs all the images, classified by pathology. The top toolbar has some buttons which allow the definition of closed surfaces into the image, the retrieval of a new image, the search of new images, the export of the medical report and so on. The bottom toolbar manages the zoom of the image and the brush size, color and shape of the annotation.

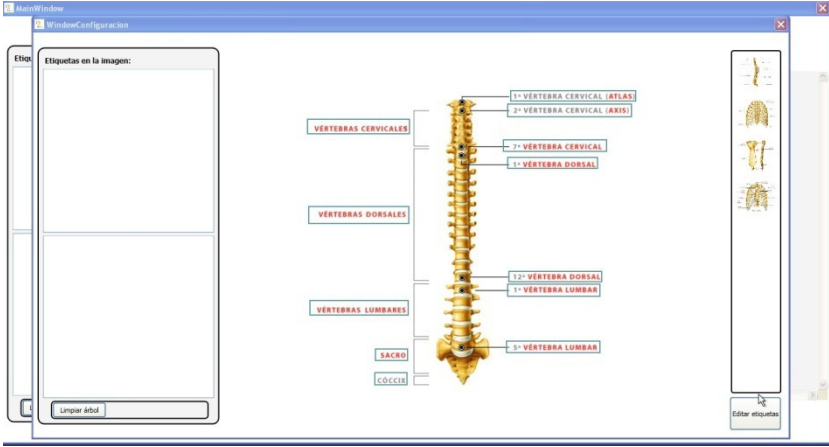


Figure 4: Location/system selection (in Spanish).

Figure 5, shows the annotation process, where physicians can paint over the image, add tags and comments and save/exchange diagnostics.

MKMTTool has a complete editing system, allowing reuse annotations in many parts of medical imaging. Physicians can tag first pathology or location. There is not restriction in the order to do things. In addition, an image can be tagged with different pathologies for different locations. For example, X-ray film can give information about thickening of the hearth. That it belongs to circulatory system and a rib fracture belongs to skeletal system.

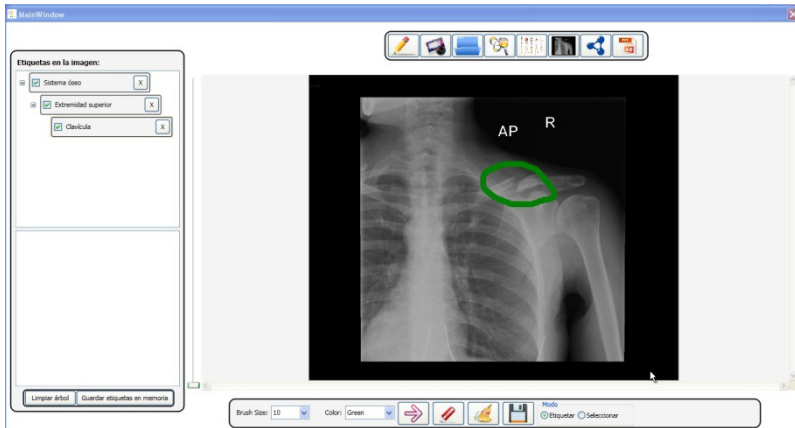


Figure 5: Pathology selection and final results of the tagging process (in Spanish).

MMKTool allows advanced search, regardless of direct search. One can search by medical history, general pathology and specific pathology. In addition, one can search by tags, indicating the number and depth he/she wants. Figure 6



shows the Advanced Search window. It may be noted that the injury tag column has a value of 4. Once the information is located, the images can be displayed, and their tagging too.

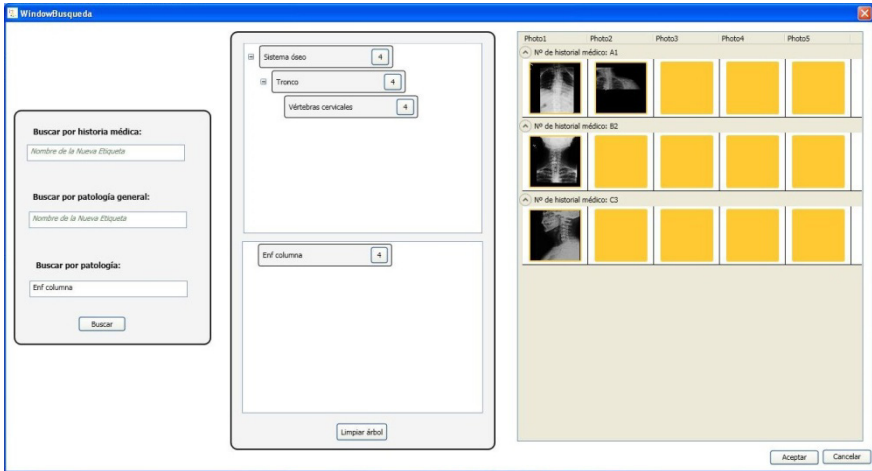


Figure 6: Results of the advanced searching process (in Spanish).

Finally, Figure 7 shows the relationships established between the terms *pulmonary emphysema* and *Swyer-Meyer syndrome*, obtained by MKMTool after a search over the UMLS servers.

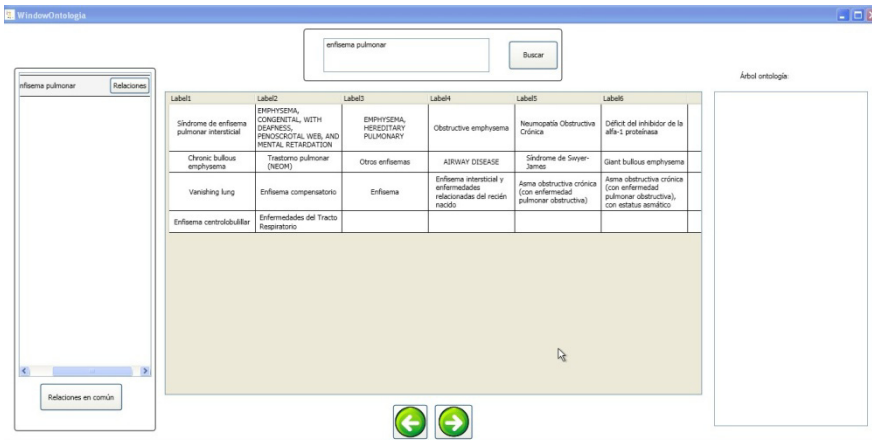


Figure 7: Related terms according to UMLS for the search pulmonary emphysema and Swyer-Meyer syndrome (in Spanish).

## 4 Conclusions

This paper has introduced a prototype of a medical knowledge management system (MKMST) for visual expertise in medical diagnostic reasoning and medical education. MKMST has been designed for radiologists and it is efficient and easy to use, meeting the objectives of its designers. Medical user access the system to diagnose the pathologies from medical images obtained from PACS, annotating the images and contributing to enrich the clinical knowledge of the community.

Medical experience is reflected during the diagnostic process and knowledge is stored and retrieved in a semantic way by means of a tagging system.

At the moment it is in pre-production on a medical service, estimating its production launch in the coming months.

## Acknowledgements

The authors wish to acknowledge the collaboration in the design, implementation and use of the proposed tool to doctors Jorge Olier and Natalia Álvarez de Eulate from the Radiology Service of the Complejo Hospitalario de Navarra.

This work has been supported by the Spanish Government: TIN2010-17170, TIN2011-28347-CO1-02, TIN2011-28347-CO2-02 and INNPACTO RailTrace IPT-370000-2010-036.

## References

- [1] Stroetmann, B. and Aisenbrey, A. “Medical Knowledge Management in Healthcare Industry”, World Academy of Science, Engineering and Technology: 64:557-562, 2012.
- [2] Heiberg Engel, P. J., “Tacit knowledge and visual expertise in medical diagnostic reasoning: implications for medical education”, *Medical Teach.*, 30:e184–e188, 2008.
- [3] Leron U, Hazzan O. 2006. The rationality debate: Application of cognitive psychology to mathematics education. *Educ Stud Math* 62:105–126. Ark *et al.* 2007.
- [4] Henry, S.G., “Recognizing tacit knowledge in Medical Epistemology”. *Theor. Med. Bioeth.*, 27:87–213, 2006.
- [5] Goldman, G.M., “The tacit dimension of clinical judgment”, *Yale J Biol Med*, 63:47–61, 1990.
- [6] Tsoukas, H., “Do we really understand tacit knowledge?”, Presented to Knowledge Economy and Society Seminar, LSE Department of Information Systems, 14 June 2002.
- [7] Kothari, A., Rudman, D., Dobbins, M., Rouse, M., Sibbald, S. and Edwards, N., “The use of tacit and explicit knowledge in public health: a qualitative study”, *Implementation Science*, 7:20, 2012.
- [8] Crowley, R.S. *et al.*, “Development of visual diagnostic expertise in pathology: An information-processing study”, *JAMIA*, 10:39–51, 2003.



- [9] Metaxiotis, K., "Healthcare Knowledge Management", *Encyclopaedia of Knowledge Management*, 2006, pp. 204-210.
- [10] Sittig, D.F., Wright, A., Simonaitis, L., Carpenter, J.D., Allen, G.O., Doebbeling, B.N., Sirajuddin, A.M., Ash, J.S. and Middleton, B., "The state of the art in clinical knowledge management: An inventory of tools and techniques", *International Journal of Medical Informatics*, 79(1): 44-57, January 2010.
- [11] Chun, S. and MacKellar, B., "Social health data integration using semantic Web", *Proceedings of the 27th Annual ACM Symposium on Applied Computing, SAC' 12*, 2012, Trento, Italy, 392-397.
- [12] Rubin, D.L., Mongkolwat, P., Kleper, V., Supekar, K. and Channin, D.S., "Medical Imaging on the Semantic Web: Annotation and Image Markup", *Journal IEEE Intelligent Systems*, 24(1):57-65, 2009.
- [13] Bo Hu; Dasmahapatra, S.; Lewis, P. and Shadbolt, N., "Ontology-based medical image annotation with description logics", *Proceedings of the 15th IEEE International Conference on Tools with Artificial Intelligence*, pp. 77-82, 3-5 Nov. 2003.
- [14] Möller, M., Regel, S. and Sintek, M., "RadSem: Semantic Annotation and Retrieval for Medical Images", *Proceedings of the 6th European Semantic Web Conference on The Semantic Web: Research and Applications, ESWC 2009*, Heraklion, Crete, Greece, 21-35, 2009.
- [15] Unified Medical Language System<sup>®</sup> (UMLS<sup>®</sup>), U.S. National Library of Medicine. <http://www.nlm.nih.gov/research/umls/>
- [16] Averroes. Regional Ministry of Education of the Junta of Andalucía. <http://www.juntadeandalucia.es/averroes/averroes/impe/web/portadaPresentacion?idMenu=mCPre>.
- [17] Córdoba, A., Astrain, J.J., Villadangos, J. and Echarte, F. "A self-adapted method for the categorization of social resources", *Expert Systems with Applications*, Available online 27 December 2012, ISSN 0957-4174, 10.1016/j.eswa.2012.12.075.