Knowledge Management in Medicine
A Framework to Organize, Browse and Retrieve Medical Data

Marios Pitikakis[1]and Imon Banerjee[2]
[1] *Softeco Sismat S.r.l., Via De Marini 1, Genova, Italy*
[2]*CNR IMATI-Genova, Via De Marini 6, Genova, Italy*

Keywords: Knowledge Management, Bio-medical Ontology, Medical Imaging, Computer-Aided Diagnosis.

Abstract: This paper outlines a knowledge-based approach to construct semantically enriched components that can be integrated to form an environment for assisting and supporting medical professionals. We aim at capturing and utilizing invaluable expert knowledge, formalized within ontologies, to improve different kinds of medical diagnosis, monitoring and treatment. Our work is focused on supporting the knowledge management task, targeting the early stage diagnosis of musculoskeletal diseases of the human knee articulation, but is general enough to support similar knowledge management tasks for a wide range of clinical decision-making, research, teaching and learning activities.

1 INTRODUCTION

Knowledge is a very important resource for preserving valuable information, solving problems and creating core competences. Managing this knowledge has become an important research issue and a wide range of technologies for both academic and real world applications have been developed.

In the medicine and health care domain, there are many individual applications and tools that rarely share common semantics. Queries related to medical data (images, clinical records, treatments plans etc.) are not arbitrary; they are based on specific semantics of anatomy, physiology and diseases.

Medical databases have a wealth of digital resources and formalized knowledge can help to organize them in an efficient way in order to support searching, browsing and retrieval tasks, as well as to assist in the diagnosis and follow-up practices. Ontologies can provide the essential glue to ensure semantic consistency of data and knowledge sharing by the different actors in complex medical scenarios.

In this work, we focus on developing an ontology-based knowledge management framework and a set of tools and services for computing different kinds of diagnostic measurements which can support the sharing, access, retrieval and integration of various pieces of information related to Musculoskeletal Diseases (MSD) and other disorders of the human knee region. To this end, our research activities take into account different scales (i.e. molecular, cellular, tissue, organ and behavior scales) and modalities in a Computer-Aided Diagnosis (CAD) context. Combining ontologies with CAD systems could improve the segmentation and analysis processes as well as the follow-up and treatment by applying generic knowledge to highly patient specific data (Catalano et. al., 2012). On the other hand, diagnostic measurements (e.g. cartilage thickness map, bone/organ volume) from standalone applications can create meaningful semantic annotations/associations and contribute to the knowledge formulation.

This paper aims at presenting a knowledge-based approach targeted at constructing semantically enriched components that can be integrated to form an environment that can assist and support medical professionals. Section 2 contains the motivation behind our efforts and reviews related work. Section 3 shortly describes the conceptualization of the domain and the purpose of the knowledge base. Section 4 gives an overview of some preliminary results in the form of standalone and web-based applications. Finally, in Section 5 we present our concluding remarks, the proposed future work, and the envisaged research road-mapping.

Pitikakis M. and Banerjee I..
Knowledge Management in Medicine - A Framework to Organize, Browse and Retrieve Medical Data.
DOI: 10.5220/0004867603740380
In *Proceedings of the International Conference on Health Informatics* (HEALTHINF-2014), pages 374-380
ISBN: 978-989-758-010-9

2 BACKGROUND AND RELATED WORK

Computer-Aided Diagnosis (CAD) is one of the major research subjects in medical informatics and diagnostic radiology. CAD is a well established concept where physicians use the computer output in a complementary way to support their final diagnosis (Doi, 2007).

The need to improve the quality of health care has led to a strong demand for CAD systems that can provide accurate, repeatable and objective feature measurements which can be used by physicians for diagnostic and follow-up purposes. Current CAD approaches concentrate on algorithmic improvements, and mostly fail to include the implicit knowledge of medical applications.

Modern diagnosis depends on the patient's health condition acquired by different modalities. Most of the times it becomes difficult to integrate all these acquired data to obtain the final diagnosis, resulting in a strong manual effort by experts of different domains to combine and interpret these data. Therefore, one of the biggest challenges is to design a common platform for searching, browsing, accessing and combining these data (preferably in an automatic/semi-automatic way) for better diagnosis and follow-up results. This kind of platform could also offer the prospect of integration, or at least interconnection, among doctors and medical professionals and enable them to navigate more easily through the data and directly gather all the relevant information available.

Ontologies up to now were mainly used to provide a common vocabulary in different domains. Especially bio-ontologies are quite popular for providing taxonomies and supporting different knowledge management tasks. The use of ontologies in medicine started with focus on the representation and (re-)organization of medical terminologies, for example FMA (Foundation Model of Anatomy) (Rosse & Mejino, 2003), ICD (International Classification Diseases, online), SNOMED (Systematized Nomenclature of Medicine) (Spackman, 2000) etc. These reference ontologies are used to provide a common vocabulary between the medical experts for the establishment of a shared understanding of concepts used.

In addition, the advancement of semantic web technologies contributed to the widespread usage of ontologies and made possible to extract structural, functional and morphological information from heterogeneous medical data residing in different bio-medical ontologies. This fact can potentially support computational frameworks for clinical decisions e.g. OntoQuest (Chen et. al., 2006), or study/analyze the human anatomy and the functional behavior of organs in a more interactive way e.g. MyCorporisFabrica (MyCF) (Palombi et. al., 2009).

Furthermore, linking clinical knowledge with the geometry extracted from the patient record is likely to open new pathways for clinical analysis. Extracting anatomically and functionally significant regions from medical imagery is another challenging and essential task. In the process of image segmentation, it is highly beneficial to attach semantic information to the segmented parts, addressing not only standardization, but also machine-readability (due to the formalized representation), advanced browsing and searching.

Some efforts on semantic tagging can be seen in the area of generic human body modelling for teaching and training purposes, as well as patient specific modelling for studying the patient condition. The Zygote Body browser (Zygote browser, online), Voxel man (Voxel-man, online) and BioDigital Human (BioDigital Human, online) represent some of the recent work on generic human body modelling combined with semantic knowledge, while 3D anatomical human (Magnenat-Thalmann et. al., 2007) and MyCF browser (MyCF browser, online) focus on patient specific models.

In medicine, this knowledge is useful to drive automated analysis to support diagnosis, therapy planning, surgery and legal medicine. Only a few initiatives have taken on the use of geometric data derived from acquired images and canonical anatomic knowledge e.g. the Virtual Soldier project (Virtual soldier project, online) by the U.S. Defence Advanced Research Projects agency. The use of semantic technologies to the creation of expert systems applied to the medical diagnostic process is presented in (Rodríguez-González et. al., 2012), where a knowledge base containing findings (signs and symptoms) and diagnostic tests was developed.

Formalizing and retrieving images and 3D data from different medical acquisition devices is not a trivial task, since their characterization depends on morphological attributes as well as on other semantic attributes. In this context, one of the main results the AIM@SHAPE Network of Excellence (AIM@SHAPE, 2006, online) was the formalization and sharing of knowledge related to 3D digital shapes and their applications. The scientific community involved in AIM@SHAPE brought a significant contribution to the development of ontologies for 3D applications by proposing a conceptualization of a *shape*, meant as a generalization of a dataset, in terms of geometrical, structural and semantic aspects, complemented by the knowledge related to the application domain in which the shape is used. This resulted in the design of two kinds of ontologies, namely the *Common* Ontologies (for both Shapes and Tools) and three *Domain* Ontologies. The Common Shape Ontology (CSO) (Vasilakis et. al., 2010) was conceived to capture knowledge related to the set of geometric, structural and topological data that define a shape.

The Digital Shape Workbench (DSW) (Pitikakis et. al., 2012) was one of the main results of AIM@SHAPE, which was later upgraded to the Virtual Visualization Services (VVS) of VISIONAIR (Attene et. al., 2013). VVS is a framework based on Semantic Web technologies for managing, storing, reasoning and searching the semantic content. It integrates resources and knowledge, providing functionalities for inserting resources, managing the resource metadata and related ontologies, advanced searching, browsing and downloading of resources.

3 SEARCHING AND BROWSING THE KNOWLEDGE BASE

In a complex medical scenario where multiple agents co-operate in order to allow continuity of care, formalized knowledge can help to organize the different resources e.g. acquired data, exams, anatomical information, patient history and links/relations between them.

Our work focuses on MSD and related disorders of the human knee articulation. The main motivation of this work is to develop a knowledge management and ontology framework, which can facilitate the sharing, access to, refinement and integration of various pieces of information pertaining to the MSD domain. A knowledge-driven framework provides access and search functionalities to concepts, patient-related data, and information related to musculoskeletal pathologies, which are properly addressed via a shared conceptualization.

For the formalization of the MSD domain, our obvious choice was to utilize ontologies. We did not start the ontology design, and the corresponding conceptualization, from scratch, but we capitalized on what other initiatives already built concerning biomedical aspects of the domain. After analyzing existing work and deciding what can be reused, we planned our design process that included the integration of existing ontologies or parts of them. We choose the versatility of the middle-out approach for ontology design, since we are guided by usage scenarios provided by the medical experts. Therefore we started from the actual data and modelled the necessary concepts.

In our case, the diverse nature of medical data is one of the main challenges for the formalization and browsing of the patient records. We use ontologies as a tool to link different scales and perspectives (e.g. anatomical, cellular, behavioral etc.) and to provide an abstract layer for structuring the knowledge and the data to support efficient retrieval.

Another challenge for a successful multi-scale diagnosis (where the data are acquired by different acquisition sessions and modalities e.g. MicroCT, MRI etc.), is to model these various kinds of data representations according to different user's perspectives which could be quite diverse.

One of our main goals is to optimize the visualization, search and browsing of the knowledge base for the patient specific or general purpose data that could assist in the clinical decision making process. This will be supported by the development of a computational framework that will take into account the available morphological, structural, and patient-specific information. There are two distinct but complementary approaches to develop such a framework: through standalone applications which can support the demanding requirements of computing resources and through a web-based interface for sharing, browsing and searching the data remotely.

4 PRELIMINARY RESULTS

In this section we present some work in progress concerning the development of both standalone and web-based applications for browsing and searching the knowledge base, as well as annotating segmented MRI scans and 3D models to facilitate the discovery of heterogeneous medical data.

4.1 Standalone Applications

As mentioned in Section 3, our standalone application provides a way to access the multi-modal medical knowledge through the defined ontology. The main objective of this application is to satisfy various end-user perspectives (such as Radiologist, Orthopedist, General practitioner, Tissue engineer) for the early stage diagnosis of MSD.

In Figure 1, we provide a layered architectural view of our proposed platform that supports the sharing and browsing of multi-modal, multi-scale data. The knowledge management layer creates an abstraction over the relational databases which formalizes the structure of the medical data stored and provides a way to access and update the data.

The system can support the retrieval of the multi-scale data from the knowledge base by the use of appropriate SPARQL queries, according to the usage scenarios and competency questions formulated from the requirement analysis phase. Some example competency questions could be: *Find all the data related to **Femoral cartilages** of a **patient Mr.Brown** who has a **Musculoskeletal disease.** I am interested specifically in studying **osteoarthritis**. Visualize the latest acquired **3D model** of the femoral cartilage of Mr.Brown. What is the **thickness** of the **cartilage**? What was the **thickness** of the **cartilage** one year ago? Show me the latest **Motion capture data (MoCap)** of the patient. Show me all the **cells** contained in the **Femoral Cartilage**.*

Answers to the above competency questions can be derived directly from the knowledge base and, in addition, the platform can provide efficient ways to explore this knowledge further to provide a mapping between the data and the derived knowledge.

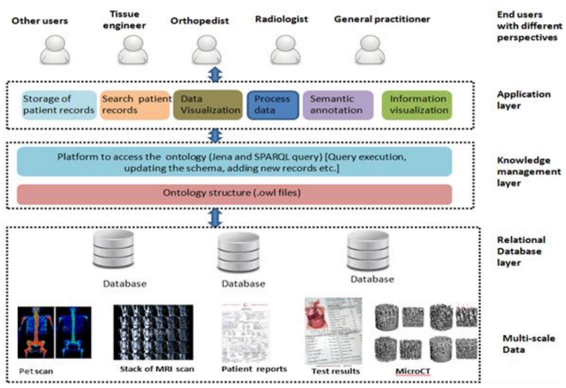

Figure 1: Proposed system architecture for accessing multi-modal medical data using ontologies.

We have developed a user interface using the Jena framework (Jena2 ontology API, online) which provides a way to visualize and navigate through the structured knowledge described by the ontology. It can also support the browsing and searching of patient data and visualization of the selected 3D data (meshes and volumes).

The implemented prototype depicted on Figure 2 serves as a framework to ease the annotation pipeline, by coupling manual or automatic segmentation with automatic computation of diagnostic measurements which may be applied to the segmented 3D parts (such as volume, area etc.); also, we foresee to develop methods for semi-automatically or automatically add/modify the semantic annotation according to the analysis of inter-linked data of different scales, by exploiting the power of multi-scale inference to aid medical doctors in their data analysis processes and follow-up studies. This is our first attempt to provide a way to link the 3D geometry with knowledge by attaching semantic information to 3D data.

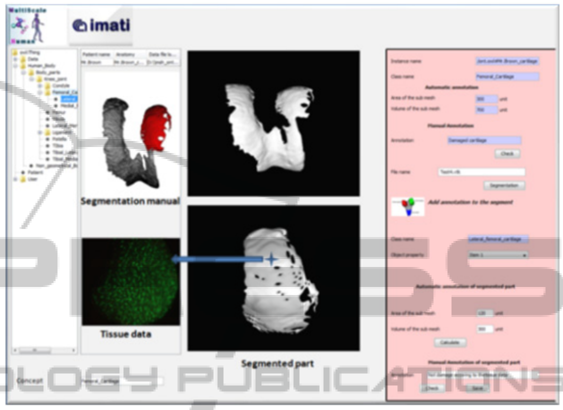

Figure 2: Interface to access and annotate the patient data along with formalized knowledge.

Another developed standalone application addresses a different kind of knowledge regarding the measurement of the femoral cartilage thickness. Accurate and precise assessments of cartilage thickness are important for addressing a number of clinical questions for the prevention, treatment and progression of osteoarthritis. For example, the mechanical loading during walking has been shown to influence the progression of osteoarthritis at the knee as well as the outcome of treatment (Koo et. al., 2005).

3D models of the femoral cartilage were created from segmented magnetic resonance images, which offer the potential of quantifying the cartilage morphology with better accuracy than two-dimensional plane images, and the weight bearing regions of the cartilage that sustain contact during walking were identified. The separation of weight bearing regions from the non-weight bearing regions of the knee joint is an important condition for the study of osteoarthritis.

Using this tool, the cartilage thickness over each region can be calculated and displayed as a color map (see Figure 3). Focusing on the weight bearing regions, which are usually of the greatest clinical interest, this tool can assist in the progression monitoring and affect the treatment outcome. The outputs of this tool (3D models and measurements) are used in the knowledge management system.

4.2 Web-based Browsing, Sharing and Searching using the Knowledge Base

We intent to further exploit semantics and knowledge management in a web based environment, to support the integration of processes within the medical investigation and the visualization pipeline, utilizing scientific visualization using WebGL for displaying 3D models and medical data on the user's web browser.

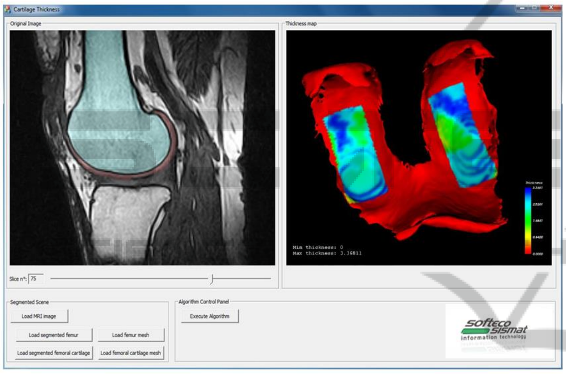

Figure 3: Tool for measuring the femoral cartilage thickness near the weight bearing regions.

The conceptualization of medical content can be the basis for more intelligent representations, with intrinsic meaning (including contextual and domain knowledge) and covering different user perspectives. This way we will be able to identify, collect and link all correlated information, making them accessible to physicians during diagnostic and follow-up processes.

Our goal is to create a web-based platform for data exchange and knowledge sharing. The ontology-driven knowledge management system will support the organization and browsing of all different kinds of stored medical content and information, as well as different ways of searching for these resources.

Three main search mechanisms will be provided: (a) a keyword search for browsing and discovering resources, (b) an advanced (semantic) search which will utilize the knowledge base (e.g. SPARQL queries combined with an OWL reasoner) and (c) a geometric search mechanism, which will be based on different kinds of similarity measures for shape matching. Our aim is to offer an integrated way to explore and visualize the data and support combined search modalities to improve the retrieval effectiveness.

Figure 4: Web-based data browsing and visualization.

We are planning to provide web interfaces for information filtering/refinement and knowledge visualization, as well as guided user interaction for uploading, searching and navigation purposes. Some preliminary work is shown in Figure 4.

5 DISCUSSION AND FUTURE WORK

Nowadays CAD systems are already included in medical imaging modalities such as digital mammography, CT and MRI. Radiologists use this type of CAD systems mainly for consulting purposes before making their final decision thus reducing the overall analysis time and manual efforts.

In order to assist in the diagnosis process, it would be possible to search for and retrieve relevant cases with a known pathology and compare the therapy used in the past, which could increase the physician's confidence in his/her decision. Of course, this would require a storage system that could host and logically organize a large number of cases, and a reliable methodology and definition of appropriate similarity measures. An intelligent knowledge management platform could be able to handle all of the above activities and interactions.

In addition, semantic search is essential for connecting and exploiting this information. Ontology management tools could support users in maintaining and evolving knowledge models to meet their needs. Finally, tools are needed to facilitate the medical investigation process and help with the annotation of images and 3D models.

Meaningful semantic annotations/associations can make the knowledge contained in medical sources (such as MRI, X-rays, CT etc.) available in a structured way, allowing both accurate and focused retrieval and knowledge sharing. Moreover, this knowledge can be used to provide valuable services; for example, it could help the diagnostic procedure, the therapy planning and all kinds of different assessments by the medical team, tasks which usually consume doctors' valuable time.

Knowledge-based methods have an enormous potential to manage, access and share the increasing amount of visual information produced. Merging ontologies (which provide a generic knowledge/ information framework) with computer-aided diagnosis systems would result in a solution targeting patient specific information. For example, ontology-driven knowledge could be used to improve the segmentation and analysis process as well as the follow-up and treatment of a patient. In addition, significant benefits can also be foreseen regarding the visualization of the patient specific data in a multi-scale, multi-modal and multi-perspective environment.

In this context, we propose a platform composed of loosely coupled components, either web-based or standalone, that could support the medical investigation process and could provide different views on the data in a multi-scale collaborative working environment. Our work intends to define a framework for capturing invaluable expert knowledge that is mostly undocumented or implicitly contained in medical data, and therefore hard to be reused or automated. Our goal is to foster semantically augmented systems and services for clinical decision-making, research and learning.

ACKNOWLEDGEMENTS

This work is supported by the FP7 Marie Curie Initial Training Network "MultiScaleHuman: Multi-scale Biological Modalities for Physiological Human Articulation", contract MRTN-CT-2011-289897. The authors would like to thank all the MSH partners and especially Softeco Sismat S.r.l. and CNR IMATI for their valuable help and support.

REFERENCES

AIM@SHAPE, FP6-IST Network of Excellence: Advanced and innovative models and tools for the development of semantic-based systems for handling, acquiring, and processing knowledge embedded in multidimensional digital objects. (Online), Available: http://www.aimatshape.net (20 Oct. 2013)

Attene, M., Giannini, F., Pitikakis, M. and Spagnuolo, M., 2013. The VISIONAIR Infrastructure capabilities to support research. *Computer-Aided Design and Applications*, Vol. 10, No.5, 2013, pp.851-862.

BioDigital Human (Online), Available: https://www.biodigitalhuman.com (20 Oct. 2013)

Catalano, C. E., Robbiano, F., Parascandolo, P., Cesario, L., Vosilla, L., Barbieri, F., Spagnuolo, M., Viano, G., and Cimmino, M. A., 2012. Exploiting 3D part-based analysis, description and indexing to support medical applications. In *Workshop on Medical Content-based Retrieval for Clinical Decision Support, in conjunction with MICCAI*, Nice, France.

Chen, L., Martone, M.E., Gupta, A., Fong, L. and Wong-barnum, M., 2006. OntoQuest: Exploring Ontological Data Made Easy. *Very Large Data Bases - VLDB* , pp. 1183-1186.

Doi, K., 2007. Computer-aided diagnosis in medical imaging: historical review, current status and future potential. *Computerized Medical Imaging and Graphics*, 31: 198–211.

International Classification of Diseases, Ninth Revision, Clinical Modification (ICD-9-CM). [Online], Available: http://www.cdc.gov/nchs/icd/icd9cm.htm (20 Oct. 2013)

Jena ontolgy API. (Online), Available: http://jena.apache.org/index.html (20 Oct. 2013)

Koo, S., Gold, G.E., Andriacchi, T.P., 2005. Considerations in measuring cartilage thickness using MRI: factors influencing reproducibility and accuracy. *Osteoarthritis and Cartilage,* 13: 782–789.

Magnenat-Thalmann, N., Gilles, B., Delingette, H., Giachetti, A. and Agus, M., 2007. Towards the Virtual Physiological Human. In *Eurographics 2007*, Prague, Czech Republic.

MyCF browser. (Online), Available: http://www.mycorporisfabrica.org/ (20 Oct. 2013)

Palombi, O., Bousquet, G., Jospin, D., Hassan, S., Revéret, L. and Faure, F., 2009. My Corporis Fabrica: a Unified Ontological, Geometrical and Mechanical View of Human Anatomy. In *2nd workshop on 3D Physiological Human, 3DPH2009*, Zermatt, Switzerland. LNCS, Spinger, Vol. 5903, pp. 207–219.

Pitikakis, M., Vavalis, M. and Houstis, C., 2012. Semantically interoperable three-dimensional scientific objects. *The Knowledge Engineering Review (KER)*, Vol. 27, No 01, 2012, pp. 33-55.

Rodríguez-González, A., Hernandez-Chan, G., Colomo-Palacios, R, 2012. Towards an ontology to support semantics enabled diagnostic decision support systems. *Current Bioinformatics*. 7(3):234–245.

Rosse, C. and Mejino, JLV, 2003. A reference ontology for biomedical informatics: the foundational model of anatomy. *J Biomed Inform*, Vol. 36, pp. 478–500.

Spackman, K., 2000. Snomed rt and snomed ct. promise of an international clinical terminology. *MD Comput.,* Vol. 17(6):29.

Vasilakis, G., Garcia-Rojas, A., Papaleo, L., Catalano, C.

E., Robbiano, F., Spagnuolo, M., Vavalis, M. and Pitikakis, M., 2010. Knowledge-based representation of 3D media. *International Journal of Software Engineering and Knowledge Engineering (IJSEKE)*, Vol. 20, No. 05, pp. 739-760.

Virtual Soldier Project. (Online), Available: http://www.virtualsoldier.net/summary.htm (20 Oct. 2013)

Voxel-Man. (Online), Available: http://www.voxel-man.de/3d-navigator/inner_organs/ (20 Oct. 2013)

Zygote Body browser. (Online), Available: http://www.zygotebody.com/ (20 Oct. 2013)