An Architecture to Support Clinical Knowledge Sharing Among Medical Community via Internet

D.F. Pires\textsuperscript{1} and R.A.I. Halah\textsuperscript{2} and R. Tinós\textsuperscript{3} and E.E.S. Ruiz\textsuperscript{3}

\textsuperscript{1}Departamento de Ciência da Computação, Faculdades COC, Ribeirão Preto, Brasil.
\textsuperscript{2}SH Clinicas, Ribeirão Preto, Brasil.
\textsuperscript{3}Faculdade de Filosofia, Ciências e Letras de Ribeirão Preto, Universidade de São Paulo, Av. Bandeirantes, 3900, 14040-901 Ribeirão Preto, SP, Brasil.
dfpires@imagcom.org

Abstract. An architecture to support the development of Electronic Health-care Systems (EHS) integrated to a Knowledge Manager System (KMS) is proposed in this paper. The proposed architecture enables the exchange of clinical data and knowledge among the medical community via the Internet. Web Services, XML and SOAP are the web technologies used to implement the knowledge exchange between an EHS and a KMS. A decision tree is used in the Computer-aided Diagnostic System (CaDS) as a knowledge inductor to support and enable integration. Results are also presented from a designed minimum requirement EHS and KMS which were implemented to store data from patients suffering from acute abdominal pain.

1. Introduction

Computer-aided Diagnosis System (CaDS) has becoming, depict some barriers [Short et al., 2004], an important resource to improve quality of services provided to patients and also to reduce administrative costs for clinics and hospitals ambulatory services [Krause et al., 2004]. Techniques proposing the use knowledge representation and knowledge acquisition in CaDS have been largely studied by the scientific community and have also been an area of great interest for commercial applications [Jiang et al., 2003]. These techniques, known in general as Machine Learning techniques, can help to define possible diagnosis based on a given paradigm of learning.

Parallel to the development of CaDS, Electronic Health-care Systems (EHS) have becoming a fundamental tool to help clinical information managing [Ammenwerth et al., 2003]. This can be justified by the possibility of a physician, using an EHS, to digitally store and retrieve clinical information. The use of EHS in medical practice may also facilitate the use of computer as providers of more general medical services, for example: services to retrieve completed clinical histories, services to search in large data volumes, searching for epidemiological data, services to provide quantitative analyses and others.

Usually EHS stores its data in a database, or system file, of unique access. This situation could create a limited scenario to exchange clinical experience among physicians, scenarios that might be considered crucial sometimes [Horvath et al., 2004]. The possibility of information exchange among physicians encourages a collaborative environment
research [Pratt et al., 2004] to support the sharing of EHS clinical cases and knowledge within medical community.

A Knowledge Manager System (KMS) involves the application of organization resources and information technology systems to manage knowledge strategically [Quaddus and Chu, 2005]. Specifically concerning the knowledge management, a KMS could be applied to support the management of EHS clinical cases and knowledge sharing activity via Internet. At this context, some new KMS management activities could be possible, and will be demonstrated by this work, as follows:

1. A KMS could be used and controlled by a physician discussion board;
2. A KMS could be used to specify a set of signals, symptoms, and diagnosis particularly related to a pathology which would be specified based on the physician experience or on the frequency these signals, symptoms and diagnosis appear. This set would be used and shared by physicians of same speciality. An EHS graphical interface could also be automatically configured with question modules referring to the same signals, symptoms and diagnosis suggested;
3. A KMS could also be used to specify a set of real experience clinical cases. Based on this set of clinical cases, a knowledge base can be built;
4. A KMS would be responsible to manage and to moderate situations as follows: (a) a physician proposes a new signal, a new symptom or a new diagnosis to the community; and (b) a physician intend to insert a new clinical case that he/she considers interesting. In both cases, the medical community is responsible to decide weather to enter or not these data in the KMS.

A Web scenario with EHS and KMS integrated, enhanced by a collaborative approach, both supported by a CaDs, some new and interesting situations would be possible, such as:

1. A shared case database could be formed from a physician discussion board about diagnostic knowledge;
2. A junior physician could have a second opinion from a shared case database before to decide for an specific diagnostic case;
3. A physician could have a second opinion from two different knowledge bases, a first one from the discussion board and a second opinion could be provided by his/hers own knowledge database;
4. A physician who participates in this collaborative knowledge community could also share his/hers knowledge with others and/or retrieve knowledge from some other specific participating physician, creating new opportunities for a second opinion. It is also possible for a physician to insert a new case in a shared case database. The decision of whether to insert or not is controlled by the community via a moderator software agent. This situation is particularly important because of its potential to help the community to discover new diagnostic rules and inferences. The proposed architecture has also a potential to enhance discoveries about new pathologies, and can also be used to facilitate accurate statistical findings on health-care.

Eich [Eich and Ohmann, 2000] presents an implementation of a central server to support clinical decision. His work is particularly concerned with abdominal pain and it does not investigate clinical knowledge sharing. Nardon [Nardon, 2003] discusses how RDF (Resource Description Framework) and DAML +OIL (DARPA Agent Markup Language, Ontology Inference Layer), a semantic markup language for Web resources, and Web technologies can be combined with deductive database systems query language to share clinical data and knowledge bases to solve complex queries. Web services technologies have also been used to share DICOM images among the medical community.
with the help of the emerging P2P (Peer-to-peer) technology, as done by Pisa and coworkers [Pisa, 2003].

In this context, this article presents an architecture to support a collaborative environment to share EHS knowledge and clinical cases summaries with a KMS via the Internet, integrating both systems and promoting a collaborative environment. To implement a test application adjacent to the proposed architecture, the first step was to define a set of minimal requirements, in other words, a minimum set of signals and symptoms, for the EHS implementation, which would enable a physician to diagnose acute abdominal pain. The same signals and symptoms will be used by the KMS. Three web services were also defined and implemented to support both, the collaborative work and, the knowledge exchange between the EHS and KMS. With the 40 acute abdominal pain signals and symptoms chosen, 14 different diagnosis are possible to be inferred by the CaDS. A symbolic learning paradigm was elected as the inference machine behind the CaDS which in turn uses the decision tree approach as the learning technique. The use of a decision tree is justified by its easy logic case interpretation. Web Services technology is used to implement the EHS–KMS integration. Both systems support a CaDS, enabling, as far as the reviewed literature, an original architecture integrating EHS, KMS and CaDS, crucial systems in the forthcoming health-care management.

This paper is organized as follows: next session presents the proposed architecture, which supports two types of databases: one is used to store clinical cases information (summaries) and the other is used to store clinical knowledge. Here the term ‘clinical knowledge’ refers to the decision trees generated from clinical training cases by the inference machine. In Section 3 the proposed architecture implementation is presented, including a case study. Results obtained from this implementation are presented in Section 4. Finally, in Section 5, discussions, conclusions and future works are presented.

2. Proposed Architecture

The proposed architecture represents a integration of two health-care systems, an Electronic Health-care System (EHS), and a Knowledge Manager System (KMS). The first health-care systems is to be used generally by a physician community while the second one should be used by a physician discussion group to control the activity of sharing knowledge and clinical cases in the community. The EHS has access to four different databases: 1) a physician’s local database of clinical cases; 2) a local database of inferred knowledge; 3) an optional database of imported clinical cases from another clinical case database, and finally; 4) an optional imported knowledge database.

On its turn, the KMS has access to other three distinguished databases to control the activity of sharing the knowledge, they are: 1) the received clinical cases database; 2) a database of shared clinical cases; and finally; 3) a shared knowledge database.

The CaDS is one of the tree web services used and available by the components of the proposed system architecture.

Figure 1 illustrates the proposed architecture, which shows:

**EHS** : the EHS is a computational application which stores, retrieves and is used to update the physician’s patient information in a local database. At a local level, the EHS is also integrated to the KMS and supported by the CaDS.

**CaDS** the CaDS is the actual computer-aided diagnosis system used by both knowledge databases, the local one and the imported knowledge database. The CaDS is also used by both clinical case databases, again, the local one and the imported knowl-
Figure 1: An architecture to support the exchange of clinical knowledge

edge database. The latter can be chosen by the physician. CaDS could be used by the EHS and the KMS too.

**The EHS local cases database** is a set of clinical cases codified into signals, symptoms and diagnoses information acquired during consultations;

**The EHS local knowledge database** is the decision tree used to infer clinical decision. The knowledge representation is built based on local cases information.

**The EHS imported case database** is formed either with the help of the physicians community through their case database (within the same medical field) or by the Knowledge Manager System case database. It is important to notice that an imported case database should be kept separated from a local case database, so that a physician can have a second opinion from two or more different trustily sources.

**EHS’s imported knowledge database** can be created from other physician’s knowledge databases or from the integrated Knowledge Manager System’s knowledge database. As the in the latter, the imported knowledge database should be kept apart from the local knowledge database;

**The Knowledge and Clinical Cases Manager System** It can be classified as a repository manager and moderator of the following: the shared cases database, the received cases database, and the knowledge database. This system is responsible for the creation of a set of standard cases and to generate new knowledge cases from them. It is also responsible for receiving new cases from physicians and to consider (moderate) if these cases should or should not be included in the shared cases database.

**Manager System’s Received Cases Database** It is created either by clinical cases sent by physicians who cooperate with the Manager System or by new cases not present in the Manager System’s shared cases database.

**The Manager System’s Shared Cases Database** It is a database formed by a set of signals, symptoms, and diagnosis defined by the physician’s discussion board. The Manager’s tasks includes receiving new entries and to moderate the decision of whether or not this case should be part of the shared cases database. A more complete description of the moderator’s tasks will be seen later in this paper.

**The Manager System’s Knowledge Database** It is built using the shared cases databases. For each new case which enters the database, a new knowledge database for that particular needs to be created.

Next, the use of the proposed architecture will be discussed. Its usage may lead to four different scenarios, here presented, and pictured in Figure 2.

**Situation 1** : Suppose a physician is attending a patient and is he/she is using a particular EHS without KMS integration. After fulfilling the signals and symptoms about
its patient, having an early diagnostic hypothesis about a diagnosis, it is possible the physician would like to ask for a second opinion. For this second opinion he/she may connect the particular EHS to the KMS, send the clinical case data, and be replied with a possible second opinion diagnosis. See Figure 2–A. It is also possible that a physician, initiating his/her medical activity, can use the knowledge produced by physician discussion boards to support his/her decision.

**Situation 2**: A physician may routinely store his/hers patient’s clinical cases using the provided EHS. This continuous procedure will enable the creation of a local case database and also, a local knowledge database. See Figure 2–B1. When the physician has finished fulfilling patient’s signals and symptoms and is about to make a first diagnosis hypothesis, he/she may ask for a second opinion using the connection to a KMS. This second opinion can be accomplished by his/her local knowledge database or by the knowledge database manager system, as seen in Figure 2–B2. This way, before deciding for a final diagnosis, a physician could compare a diagnosis produced by his/her local knowledge with a diagnosis suggested by a physician discussion board. It is important to notice that if the local database and the manager’s system database are not separated, the local medical experiences from the physician would certainly injure knowledge manager system credibility.

**Situation 3**: A physician would like to share one particular case, or even all of his/hers clinical cases (knowledge database), with others. In this situation a physician may send his/hers cases to the manager system. View Figure 2–C1. All these cases are included in the received database, as seen in Figure 2–C2. Next, a moderator, which in this architecture is a software agent, verifies the possibility of the new different case to be included in the shared database (Figure 2–C3. This situation might be implemented using a moderator software agent which sends this particular case to every participating physician. When receiving the new case, physicians can vote if the new case can or cannot be included. The moderator then analyzes if the new case has sufficient positive grade to be considered. If the new case is included, a new decision tree has to be created. This situation can be very interesting to define new diagnostic rules and may also help to uncover new pathologies. Another interesting case in this situation is the one which the manager system can obtain statistical data about particular diseases.

**Situation 4**: A physician’s EHS could connect to the KMS and obtain another physician case database or knowledge database. View Figure 2–D1. An imported case database and/or an imported knowledge database may then be created. In this context, a physician can obtain a second opinion diagnosis and compare it with 3 differ-
Aiming to implement a computer application adjacent to the proposed architecture, Web Services were defined and implemented. It was developed a simple EHS integrated to a KMS to share acute abdominal pain clinical knowledge. Both systems have been implemented the web services created. This study case is discussed in the following section and it is used to test the system viability.

3. Initial Considering Implementation

Forty signals and symptoms related to acute abdominal pain have been applied to develop the study case. They are: sex, age, abdominal pain, location, character initial, character changes, character final, severity, severity changes, radiation, radiation location, exacerbating factors, alleviating factors, bowel movements, dyspareunia, vaginal discharge, alcoholism, anorexia, vomiting, vomiting character, jaundice, tachycardia, abdominal auscultation, hydration, distention, temperature, murphy, bloomberg, migration, initial location, last menstruation, rebound pain, WBC (White Blood Cells), Shift to Left, hematuria, amylase, alkaline, SGOT (Serum Glutamic-Oxaloacetic Transaminase) and SGPT (Serum Glutamic-Pyruvic Transaminase). The signals can assume continuous values, as temperature and age, or discrete values, like vomiting yes, no, severity discomfort, moderate, severe, and abdominal auscultation normal, metallic sounds, hypoactive, hyperactive, no matter. The 14 possible diagnosis were: appendicitis, cholecystitis, peptic ulceration, gastroenteritis, diverticulitis, pancreatitis, hepatitis, intestinal obstruction, pelvic inflammatory disease, mesenteric thrombosis, dysmenorrhea, ectopic pregnancy, ovarian pain, and unknown. This set of signals, symptoms, and diagnosis was initially chosen because they are frequent in acute abdominal pain pathology [Porto, 1991].

Java objects oriented language was elected to implement EHS, KMS, and CaDS systems. For the CaDS the WEKA API [Witten and Frank, 2000] was also used to implement the decision tree machine learning technique. The algorithm used was C4.5.

The Apache AXIS Java Web Services\(^1\) tool was used to create the necessary Web Services software structure. It was also used the MySQL\(^2\) database manager system to store the clinical cases and knowledge.

3.1. Defining and Implementing the Web Services

Web Services\(^3\) specification is recommended by W3C (World Wide Web Consortium). It aims to promote and to make easy the communication between two or more computer applications developed using different programming language and/or different computer architecture. To improve interoperability between web application, this specification uses some other technologies, as: WSDL\(^4\) (Web Service Description Language) to describe services interfaces, SOAP\(^5\) (Simple Object Access Protocol) to format XML messages and to call remote object using HTTP, and UDDI\(^6\) (Universal Description, Discovery and Integration) to describe services metadata, to discover business, and to integrate business services using the Internet.

\(^{1}\)http://ws.apache.org/axis/
\(^{2}\)http://www.mysql.org
\(^{3}\)http://www.w3.org/TR/ws-arch/
\(^{4}\)http://www.w3.org/TR/2004/WD-wsdl20-20040803/
\(^{5}\)http://www.w3.org/2000/xp/Group/
\(^{6}\)http://www.uddi.org
In order to define a Web Service, the developer needs to decide what kind of service application should be offered. In this work, three services were initially specified during the analysis and the project software phase: a service to create a knowledge database using the clinical cases database; a service to classify (obtain a diagnosis) using a set of signals and symptoms based on the manager system’s shared knowledge database; and finally a service to classify using a set of signals and symptoms based on a local or imported knowledge database. The last two services basically implement the CaDS. The three web services implemented are:

- `public void createDBFile();`
- `public String classifyCenter(signals-and-symptoms-value)`
- `public String classifyLocal(signals-and-symptoms-value, local-knowledge-database)`

The KMS and the EHS can make use of the first and second web services. Only EHS uses the last web service. In the way the Web Services are used by two different kind of client application, a business chain is defined. Better discussion about it will can be find in sections 3.3 and 3.4.

### 3.2. WSDL (Web Service Description Language)

The WSDL was used to make a description interface of those 3 services presented earlier. This way, any computer program can read and make use of them.

The WSDL file are available at [http://143.107.220.180:8080/arq/ArqServiceImpl.jws?wsdl](http://143.107.220.180:8080/arq/ArqServiceImpl.jws?wsdl). Some important consideration should be done about WSDL file: each service is defined as an object message; each message has a response and a request part; each part has their parameters and, each parameter is strongly typed; each type is limited to be an object type, like a `String`, an `Integer`, a `Float`, or a `Double`. All these types are serializable by default. Java Web Services implementation claim for that. These services will be available beyond some UDDI registry, not yet defined, to a software’s search service.

### 3.3. Implementing KMS

A module to support the activity of creating, updating, and removing signals and diagnosis of acute abdominal pain was also developed. Using this module, a physician discussion board on acute abdominal pain has defined those 40 signals and 14 diagnoses already presented. Figure 3 presents an example of a set of acute abdominal pain signals and symptoms.

Another module, where physician discussion boards can inserts, removes, deletes, and updates acute abdominal pain clinical cases, was designed. Using this module, it was inserted 211 clinical cases to train and test the CaDS. Figure 4 shows this frame window that explore a structured mode to the input and output signals, and diagnosis information. The frame window was divided in 6 parts, which are: Patient Identification\(^7\), Complaint/Period – Part 1, Complaint/Period – Part 2, Physical Exam, Laboratory Exams, and finally, Diagnostic (diagnosis).

These modules (Figure 3 and 4), presented in the KMS tool, can help a physician discussion boards to create new sets of pathology information or to update an existing set while the system is working.

Having defined signals and symptoms and inserted all clinical cases, a physician discussion board can generate the main knowledge database system. This is accomplished as the first web service is called `createDBFile()`, which was implemented to connect to a

---

\(^7\)Patient identification is not sent to Knowledge Manager System.
MySQL database, selecting all signals and diagnoses values, and creating a knowledge database, represented by an XML file. Another scenario can be visualized when a physician discussion board decides to create a new pathological clinical case. Before deciding for a diagnosis for this disease, they would want to see the classification results based on the knowledge database. In this scenario, the KMS needs to call `classifyLocal()` service.

Following, any physician discussion board can create their Knowledge Manager System software and make use of `createDBFile()` and `classifyLocal()` web services to generate and share their knowledge.

### 3.4. Implementing EHS

To support a physician located at his/her office inserting, deleting, removing, and updating his/her acute abdominal pain clinical cases’ patient, a frame window illustrated by Figure 5 was created. This window frame is similar to the physician discussion board.

There is another resource presented in the EHS implementation, which is illustrated in Figure 6. After filling all fields and have chosen, as an example, the appendicitis diagnosis, it is possible that he/she wants to be supported by the KMS knowledge database for a second opinion. As can be observed in Figure 6, the diagnosis is the same. The physician could have chosen the local knowledge database to compare its decision based on his/her decision, or one imported knowledge database’s physician. After CaDS has supported the physician, clinical data are stored into a local case database with the physician final diagnosis. Finally, if necessary, physician could view its new statistics and decision tree local knowledge.

At this context, physician EHS software can make use of all web services available. EHS could call `classifyCenter()` service sending signals and symptoms parameters chosen. At this first situation, EHS clinical case evaluated would be classified using KMS knowledge database. Then, diagnosis results are returned. Physician EHS can make use
of a third web service too. The system make a call to \texttt{classifyLocal()} service, sending signals and symptoms parameters, and its knowledge database too. So, this service will classify its signals and symptoms based on the physician knowledge database. Then, diagnosis results are returned. Another case, where \texttt{classifyLocal()} service could be used, is to classify using an imported knowledge database. Finally, an physician´ EHS can make use of this \texttt{createDBFile()} message. This situation could be employed when its necessary to generate a new knowledge database based on their clinical case database.

Based on study case experience, any physician could create their EHS software and make use of the \texttt{createDBFile()}, \texttt{classifyLocal()} and \texttt{classifyCenter()} web services to generate their knowledge and ask for a second opinion. Or any physician discussion board could create their KMS software and use correspondent web services to create signals/symptoms and diagnosis, to create a knowledge database, and to share it with medical community.

4. Results

Results obtained considering the implementation presented in previous sections were considered positive. During case study development with a coloproctologist specialist physician, 211 real clinical cases were inserted. These 211 cases were used to construct a decision tree. The 10-fold cross-validation technique was used to validate the diagnosis system, with 98,75% of tests cases been correctly classified. A second test was done with a set of 26 abdominal pains clinical cases obtained in SH Clinics´ database. On this test, the result match with 97,55% of the cases were correctly classified.

Technologies used to implement the proposed architecture like Web Services, SOAP, WSDL and Java were considered suitable. Their tools are available at Web and are free. Performance aspects were considered satisfied, although compiled and interpreted java language aspect.

5. Conclusions and Future Work

The results obtained with acute abdominal pain pathology were considered positives. In the future, other experiments with other pathologies will be generated with whole new sets of training cases. An aspect already demonstrated in the specialized literature became evident in this investigation: decision tree creation activity is influenced by training cases quality. When just representative clinical cases are used to create decision tree, errors percentage number are lower.

The architecture presented in this paper and, partially implemented throw a acute abdominal pain case study, can originate four different situations that promote the activity of sharing clinical knowledge and cases by physician via the Web. While physician change experiences with each other, new diagnostic rules and new pathologies can be discovery with physicians´ suggestions. This way, it´s necessary to investigate how cases that doesn´t have these new signals or new diagnostics and was used to create a decision tree not updated, would affect a new and updated decision tree construction.

Based on the EHS and KMS clients´ software implementation, and experiences obtained with web services definition and implementation too, this work next steps are:

- The authors will be in discussion soon with a health institute, in order to to offer to the resident medicine students use the EHS and KMS to create new pathology clinical cases and to test them;
- To create web services to implement shared clinical cases situations
To promote the use of Web Semantic and ontology to enter clinical knowledge using RDF concepts and existing RDF extensions.

Acknowledgement

Tinós’ research work has been supported by FAPESP (JP 04/04289-6).

References


---

8http://www.semanticweb.org
9http://www.w3.org/TR/rdf-schema