A Proposal of Clinical Decision Support system Architecture for Distributed Electronic Health Records

Shaker H. El-Sappagh[1,2], Samir El-Masri[3]
[1]College of Science, King Saud University, Saudi Arabia
[2]Faculty of Computers and Information, Minia University, Egypt
[3]Department of Information Systems, College of Computer and Information Sciences, King Saud University, Saudi Arabia

Abstract - *Improving the quality of healthcare, reducing medical errors, guarantying the safety of patients are the most serious duty of the hospital. Electronic Health Record (EHR) was introduced to achieve these goals. HER has a very large data source which can guide and improve the clinical decision making process. In this paper, we will propose a distributed Clinical Decision Support System (CDSS) architecture which satisfies the compatibility, interoperability, and scalability objectives of EHR. The proposed framework will take advantages of EHR, data mining techniques, clinical databases, domain experts' knowledge bases, and available technologies and standards to provide decision making support for the healthcare personnel.*

Keywords: Data Mining; Knowledge Management; Clinical Decision Support Systems (CDSS); Electronic health record; Health informatics.

1 Introduction

Healthcare faces multiple problems, including high and rising expenditures, inconsistent quality, and gaps in care and access. Because of this, health care services represent a major portion of the government spending in most countries [1].

Healthcare information technology, especially EHRs, have been thought to be possible solution to healthcare problems. EHRs help administrators, physicians, nurses, researchers and healthcare personnel. EHR provides a complete, integrated and consistent view about patient conditions. However, the volume of data is very large and is increasing continuously. Healthcare personnel need to take all of the patient medical history in to consideration; they also need to connect this information together and take advices from domain experts. This huge amount of data cannot benefit physician without a helping automated system. This system can analyze these data, connect it together, integrate it with knowledge from domain expert, and search for needed knowledge - if it is possible - in other connected systems. This system is CDSS. In this paper, we tried to build a complete architecture for this system. The proposed model will take an order and initial diagnose from healthcare personnel and provide a decision support in understandable form based on existing knowledge. The system will integrate off-line standardized knowledge bases from domain experts with online knowledge extracted continuously from EHR and clinical databases and provide applicable decisions support. The paper is organized as follow. Section 2 discusses related work. Section 3 explains the research problem. In section 4 we define CDSS. The proposed framework for CDSS is discussed in section 5. The conclusion is shown in section 6.

2 Related Works

2.1 EHR standards

Many organizations provide EHR standards that standardize the structuring, implementing, sharing, integration and interoperability in EHR environment. Some of them are ISO, CEN, CFR, ASTM, HL7, NEMA, ONCHIT, etc. Also, coding systems are critical to build a shared EHR because the new environment connects heterogeneous systems each with different terminologies. Some organizations that provide these standards are AMA, IHTSDO, CMS, WHO, etc.

2.2 Data mining and Artificial Intelligence (AI)

Applying data mining and AI techniques on EHR data has many opportunities to improve the delivery, efficiency, and effectiveness of health care [12][13], such as operations management, preventive healthcare, chronic disease treatment and prevention, association analysis, evidence-based treatment, population tracking, etc. If CDSS depends only on the Knowledge Base (KB) derived from knowledge expert, then it will be inactive and not applicable. EHR contains a very large and historical dataset which change continuously and contains useful hidden knowledge. As a result, data mining and AI services should be embedded in the active CDSS system to continuously update the CSDD's knowledge base by the most recent patterns from EHR and clinical databases.

2.3 Knowledge representations in medical domain

Because there are many sources and uses for medical knowledge, many international methodologies and standards for representing medical and healthcare body of knowledge are integrated. Clinical workflows (clinical guidelines) are used to represent human-based medical knowledge through rule-based or flow-based guideline techniques. Furthermore, mined knowledge can be automatically extracted from clinical databases and/or EHR through data mining and AI techniques to be incorporated into human-generated knowledge in order to enhance their decision-making processes.

Both types of knowledge can be represented as logical conditions, rules, graphs/networks, or structural representations [5]. Predictive Model Markup Language (PMML) and GLIF (Guide Line Interchange Format) are examples of knowledge representation languages which are used to acquire and integrate knowledge. Also there are many tools for knowledge acquisition and representation as Unified Medical Language System (UMLS) [6], Protégé [7], GLARE [8], PROforma [9], Asbru [10].

2.4 Service Oriented Architecture (SOA)

SOA has been widely adopted to solve the interoperability of the involved heterogeneous distributed EHR systems [2][3]. It plays a key role in the integration of heterogeneous systems by the means of services that represent different system functionality, independent from the underlying platforms or programming languages, and interact via messages exchange. *Web services* also play critical role in systems interoperability.

Web services technology is defined as a systematic and extensible framework for application-to-application interaction built on top of existing web protocols. These protocols are based on XML [11] and include: Web Services Description Language (WSDL) to describe the service interfaces, Simple Object Access Protocol (SOAP) for communication between web services and client applications, and Universal Description, Discovery, and Integration (UDDI) to facilitate locating and using web services on a network [4].

3 The Research Problem

Building CDSS will improve the quality and efficiency of healthcare [17]. These systems will be more practical when coupled with Computerized Physician Order Entry (CPOE). It contains a set of knowledge bases (one in each hospital) extracted off-line from domain experts. If CDSS only depends on these knowledge bases it will be inactive and will become not applicable. The solution is to continually update these knowledge bases to make CDSS more active. At each site, new knowledge will be discovered and added to knowledge base from (1) new expert knowledge discovered by research, (2) data mining engine connected to local EHR and clinical databases. This action will make CDSS more active by including the most recent knowledge from active databases. Because knowledge base must be in specific domain such as heart diseases, the proposed framework will be distributed with co-operative and integrated knowledge bases. Each knowledge base in each hospital will be in specific domain. At each hospital, CDSS will build patient profile from patient's medical history and current diagnose, and it will use its local knowledge base to make decision. If CDSS cannot take decision by using its local knowledge base, it can send some data to other sites to consult its specialized knowledge bases. Other sites will response by some knowledge that helps CDSS to make more accurate decision. The goal of this paper is to propose a distributed CDSS framework that achieves (1) build co-operative knowledge bases from different domain experts' knowledge and most recent academic researches, (2) Standardize knowledge into XML format before storage, (3) connect data mining engine to EHR and clinical databases to continuously mine the most recent and applicable knowledge and adds it to local knowledge base, (4) CDSS can consult specialized knowledge bases in other institutions for other relevant knowledge, (5) before starting to take decision, CDSS collects all patient EHRs from all sites, integrates it with current diagnose, standardizes it and enters it to the inference engine, (6) assure interoperability by converting all patient data and knowledge in to standard XML format. This way, we build a complete, interoperable, active, distributed and continuously learning CDSS system.

4 Clinical Decision Support System

CDSS are interactive computer programs which are designed to assist physicians and other health professionals [14]. It helps in drug prescription, diagnosis and disease management to improve services and reduce risks and errors. It can check for patient drug allergies, compare drug and laboratory values, evaluate the potential for drug-drug interactions, suggest drug alternatives, block duplicate orders, suggest drug doses, routes, and frequencies and provide recommendations. Also, CDSS can provide clinical knowledge and best practice standards and guidelines for inexpert physicians. CDSS must be integrated with EHR and CPOE system which is connected to other HISs (laboratory, radiology, billing, etc). The basic components of a CDSS include medical knowledge base and an inference mechanism (usually a set of rules derived from the experts and evidence-based medicine) and implemented through medical logic modules based on a language such as Arden syntax or using artificial neural network as in figure 1 [15].

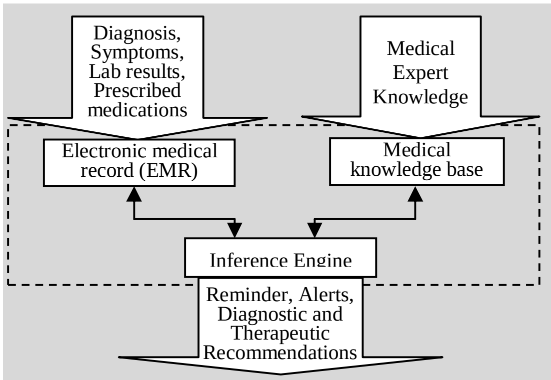

Figure 1: General model of CDSS

CDSS provides recommendations based on the available patient specific data (EHR) and medical facts (knowledge base). It has ten levels of automation ranging from L1 where all decisions made by humans to L10 where computer makes all decisions.

The EMR is continuously updated, so the knowledge bases must be continuously updated be discovered knowledge from domain expert and discovered knowledge from EMR.

5 Proposed CDSS Framework

We assume that EHR architecture and connectivity exists, and we will integrate the distributed CDSS architecture with it. The proposed architecture of CDSS is independent. It does not depend on and does not affect by the architecture of EHR or HIS. Moreover, the architecture is scalable. We can add any number of knowledge bases, EHRs, or clinical databases to the architecture using available standards and technologies.
Previous CDSS was separate system from the healthcare systems. This way, it will require physician to manually activate it, log in to it, and reenter redundant patient data. This process will make CDSS not applicable and waste time. Also CDSS will depend on the entered data which may be inadequate or contain errors. The needed CDSS will be directly integrated with the healthcare system's CPOE component, it will be activated automatically, collect the needed data from patient order, ask for unknown parameters, and make recommendation on time. Figure 2 [16] show the three phases in the decision making process.
Phase 1 (knowledge preparation) uses data mining techniques to extract knowledge from electronic healthcare data and store it in knowledge base. *Phase 2* (knowledge interoperation) takes the patient data that need decision making and translate it in to standard XML form (CDA) and make PMML encoding of the knowledge from knowledge base (KB). *The last phase* takes the previous standardized data and knowledge and makes decision. Figure 3 shows our proposed CDSS framework. It will operate as follows:

5.1 Knowledge Bases Building

The first step is to build the initial KBs. Constructing KBs of the CDSS is a crucial task that determines the success of the CDSS in general [19]. The goal is to collect the medical knowledge from the relevant sources (domain expert, EHR and/or clinical databases, and research), systemize it and represent it in a formal human understandable and computer-interpretable manner. In this framework the three services or components responsible for generating and standardizing knowledge to populate the standard XML KBs are:

1- *Knowledge Extraction Module (KEM)*, it is responsible for extract knowledge from domain expert. There are many ways to represent this knowledge.
2- *Data Mining Engine (DME)*, it is responsible for mine both EHR and clinical databases.
3- *PMML Encoding Module (PEM)*, it employs PMML to encode the generated knowledge in to standard XML based document to achieve interoperability goal between knowledge discovered from different HISs and knowledge from domain experts. The XML schema for each document describes the input data items, data mining algorithm specific parameters, and the final mining results.

The challenges in constructing and maintaining the knowledge bases are numerous. Firstly, for specific domain in one hospital, the KB is built from domain expert's knowledge off-line. KEM can take variety of methods and techniques to build knowledge base [5]. KEM Then passes the knowledge to PMML encoding module to translate it to XML form. In each hospital, CDSS will have specialized KB according to the field of the hospital, and the distributed framework will make these KBs co-operative.

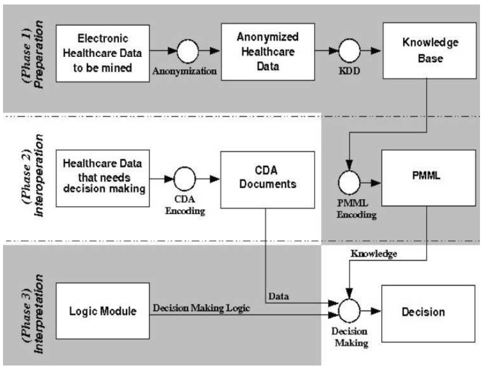

Figure 2: Health care Knowledge management framework. The shaded areas designate off-line parts.

Assuring that KBs are up to date is critical to make CDSS active and continuously learning and therefore applicable, and this can be achieved by:
(1) Continuously update KBs by new domain expert or research knowledge,
(2) Applying data mining techniques and algorithms on local clinical databases and EHR to discover hidden and non-trivial patterns and update the KB by these results.

This way, the CDSS will provide the most up to date and the most applicable knowledge. DME has two processes:
(1) Data Preparation Engine which identifies the task relevant data from clinical databases and HER after triggers from there sources to apply data mining process, removes the healthcare data attributes that can identify a patient or reveal their private data (Anonymization) and performs data selection, cleaning and transformation.
(2) KDD which performs the actual data mining operations. Finally, the results are assessed in terms of usefulness, validity, and understandability.

EHR is important source for medical knowledge. It contains a longitudinal and history of patient clinical and diagnostic data. This makes EHR a good place for applying data mining and AI techniques. Also EHR attributes are selected carefully which add another advantage. This process is continuous because EHR is updated continuously. Any update to EHR will trigger the DME to discover new knowledge then pass it to PEM to standardize and store it in KB.
Another source of knowledge is the clinical databases because it contains detailed data about patient. The DME is triggered to search for new knowledge in the updated databases as with HER. This way we assure that KB contains the complete, most recent, accurate and applicable knowledge.
The domain expert knowledge and the data mining discovered knowledge are passed to PEM module to be standardized in XML format and stored in knowledge base.

5.2 CDSS Supporting CPOE

Healthcare personnel use the CPOE for prescription. Previously, the Health Information System (HIS) was depending on the paper-base prescribing or poor or unstructured notes in a separate system (order entry system). Also, the order entry system was collecting only administrative data not medical, clinical or diagnostic data. The needed system will use electronic prescribing system which allow the writing of e-prescribe. Additionally, human errors and mistakes are expected when writing the prescription. With the existence of CDSS integrated with CPOE system, CDSS will not only provide recommendations for treatment, but it also can check for errors or shortage of data and notify physician before proceeding with decision support. There are many methodologies for building user interface for CPOE. It may be a series of questions and answers [18]. Another methodology uses the standard paper-base forms to build data entry templates and adds features relevant to decision support. Web-based order entry forms also can be used.

5.3 Framework Execution Steps

After building knowledge bases the CDSS is now ready to guide and help healthcare personnel. The execution of this framework will work as follow:

1- In an on-line operation, Healthcare personnel enters patient Universal ID (UID) which identify the patient nation-widely, and enters subject data or current diagnose (i.e., healthcare data that needs decision making).

2- UID passes to HRS (History Retrieval Service) in the local hospital, and travels via a secure network channels to all hospitals.

3- Each HRS in each hospital checks whether this patient has an EHR in its hospital or not.

4- If the patient has no record then the service returns message indicating that, else there are many methodologies for implementing the service to retrieve patient record. It may be implemented to retrieve the last N visits, visits within specific period, specific disease's related data, etc.

5- The returned records will be collected and filtered by Accumulator and Filter service which produce the patient profile.

6- Patient profile is integrated with the current diagnose and entered to CDA Encoding service which standardize the patient medical and diagnoses data into standardized XML-based CDA.

7- The encoded PMML knowledge from local KB and CDA document from CPOE provides the interoperability of knowledge and data in our framework in the sense that CDSS will be independent of the proprietary data format of the involved healthcare providers.
Now we have a complete view about patient's current and previous conditions.

8- The encoded patient profile enters as input to local Knowledge Engine (KE) which make inference of diagnose, determine the correct medicines, etc, as discussed in section 4.
9- KE can be programmed by any AI methodology as artificial neural network. It can access, query, and interpret the data and knowledge that flow from CPOE and KB respectively. Decision making is carried out in 3 main steps, retrieving the right data fields from the data source (CDA); applying the knowledge's models to the data; and eventually taking an action or a set of actions based on the results of this application. For example, if the module was invoked at a decision step in a guideline, it may branch to a specific path; or it may simply display the results in the form of a reminder or an alert.

10- According to the complexity of the problem and according to the specialization of KE, KE may need to consult the other site's KE of its problem if it has shortage in available knowledge or if it is not specialized in this problem. This is done by sending the CDA or specific fields from its site to all or set of other sites that use the same technologies, interfaces, standards, services, and terminologies. All of the helping KEs determine the relevant knowledge and send it to the requesting KE.

11- This way KE will take decision based on the initial physician diagnose, EHRs, and knowledge from its local KB other KBs. Also, it will use KB which contains the most recent knowledge. This way we ease the process of developing KBs because each KB will be specialized in specific domain and KEs will co-operate or consult each other according to patient profile to make the most accurate decision.

12- The final results of the KE will be displayed to healthcare personnel by the Knowledge Representation (KR) module. It will be used to communicate the final results to physician. According to the level of automation in CDSS, the KR may:

1- Display recommendation in the form of images, texts, sounds, videos, etc.
2- Require physician's decision about the final diagnose and actions. The physician has the choice to refuse, alter or accept the given support. If the physician accepts the support, CDSS will send an order to: the pharmacy to prepare the medicine and give it the treatment policy, laboratory system to prepare for specific tests, radiology system to be ready for some rays tests, etc.
3- Request additional data to be entered again into CPOE.

13- The CDSS may make many diagnoses with different probabilities and physician can choose the best. Also, data mining and machine learning can predict the likelihood for any future problem in health of community.
We expect that this framework will provide the most accurate and applicable decision support, and will achieve great integration between HIS and decision support processes. Also, the proposed model is fully automated. The physician only enters the patient UID and initial diagnoses, and CDSS returns decision supports. Moreover, the architecture is component-based. Each component of architecture is pluggable and reusable.

6 Conclusion

In this paper, we proposed a novel knowledge management framework for distributed health care systems that incorporate the knowledge extracted by data mining techniques with knowledge from domain experts with EHR data into health care information systems for decision making support. The model successfully integrates CDSS into the workflow of the HIS. This process fast the physician operations and reduce the level of error. We use many standards as CDA and PMML to achieve the system interoperability and integration to enhance healthcare decision making environment. Our model is fully automated. It only needs the patient universal ID and the physician's initial diagnose, then the system collects all patient EHRs from all hospitals, standardize it, and introduce it to the Knowledge Engine which make intelligent decision support.
This model depends on a set of knowledge bases located in different hospitals. Each is specialized in specific domain and the distributed CDSS architecture facilitates the integration and cooperation of KEs in case of patients who have complex medical or diagnostic problems. KE may send the patient profile or specific data to other sites for consultation.
This model also assures that knowledge base is continuously up to date to allow the CDSS to produce

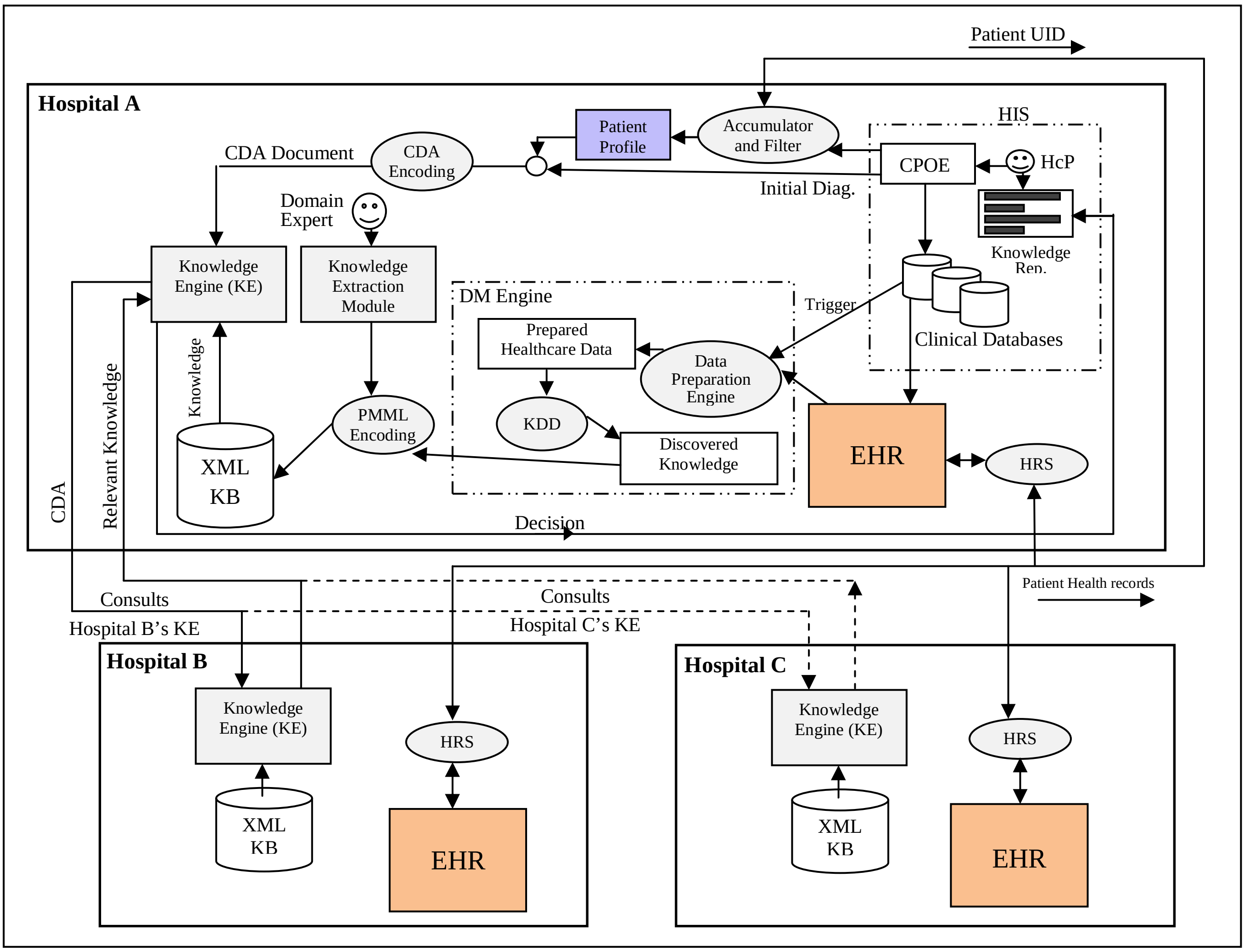

Figure 3: Proposed Distributed CDSS architecture

an applicable recommendations and actions. If the result recommendations are not represented to physician with a correct way, then it will have less benefit. As a result of that, the model has a module to represent results from KE in a meaningful way which allows physicians to make fast and accurate decisions. The next step is to implement this framework.

7 References

[1] Canadian institute for health informatics (cihi) project. Available: http://www.cihi.ca/.

[2] Hahn, C., Jacobi, S., Raber, D.; "Enhancing the Interoperability between Multiagent Systems and Service-Oriented Architectures through a Model-Driven Approach"; Web Intelligence and Intelligent Agent Technology (WI-IAT), 2010 IEEE/WIC/ACM International Conf.; Volume: 2; Page(s): 415; 2010.

[3] Maciel, R.S.P.; David, J.M.N.; "WGWSOA: A Service-Oriented Middleware Architecture to support Groupware Interoperability"; IEEE 11th International Conference in Computer Supported Cooperative Work in Design (CSCWD); Page(s): 556; 2007.

[4] Jean Bacon, Ken Moody; "Toward open, secure, widely distributed services"; Magazine Communications of the ACM; Vol. 45; Issue 6; 2002.

[5] Guilan Kong, Dong-Ling Xu, Jian-Bo Yang; "Clinical Decision Support Systems: A Review on Knowledge Representation and Inference under Uncertainties", International Journal of Computational Intelligence Systems, Vol.1, No.2 (May 2008), 159-167.

[6] Bodenreider, Olivier (2004), "The Unified Medical Language System (UMLS): integrating biomedical terminology", Nucleic Acids Research, 32, D267-D270.

[7] Protégé official web site, Feb 2010. Available: http://protege.stanford.edu/
[8] Terenziani P, Montani S, Bottrighi A et al. "The GLARE approach to clinical guidelines: main features". Stud Health Technol Inform. 2004;101:162-6.
[9] Open Clinical: PROforma, Feb 2010. Available: http://www.openclinical.org/gmm_proforma.html
[10] Open Clinical: ASBRU, Feb 2010. Available: http:// www.openclinical.org/gmm_asb ru.html.
[11] w3c consortium. Available: http://www.w3.org/XML/
[12] Ramakrishnan N., Hanauer D., Keller B.; "Mining Electronic Health Records"; IEEE Computer Society; Vol. 43; Issue: 10; PP. 77 – 81; 2010.
[13] Eugenia G. Giannopoulou, "Data Mining in Medical and Biological Research," ISBN 978-953-7619-30-5.
[14] Gamberger D., Prcela M., Jovic A., Smuc T., Parati G., Valentini M., Kawecka-Jaszcz K., Styczkiewicz K., Kononowicz A., Candelieri A., Conforti D., Guido R.; "Medical knowledge representation within Heartfaid platform." In Proc. of Biostec 2008 Int. Joint Conference on Biomedical Engineering Systems and Technologies, pp.205-217.
[15] Aleksovska-Stojkovska L. Loskovska S.; "Clinical Decision Support Systems: Medical knowledge acquisition and representation methods"; Electro/Information Technology (EIT) IEEE International Conference; on pages: 1; 2010.
[16] Reza Sherafat Kazemzadeh, Kamran Sartipi; "Interoperability of Data and Knowledge in Distributed Health Care Systems"; Proceedings of the 13th IEEE International Workshop on Software Technology and Engineering Practice (STEP'05); Page(s): 230; 2005.
[17] June Eichner, Maya Das, NORC at the University of Chicago; "Challenges and Barriers to Clinical Decision Support (CDS) Design and Implementation Experienced in the Agency for Healthcare Research and Quality CDS Demonstrations"; AHRQ National Resource Center for Health Information Technology; 2010
[18] Liang Xiao, Gráinne Cousins, Lucy Hederman, Tom Fahey, Borislav Dimitrov; "The Design of an EHR for Clinical Decision Support"; 3rd International Conference on Biomedical Engineering and Informatics(BMEI 2010).
[19] Jeong Ah Kim, InSook Cho, Yoon Kim; "CDSS (Clinical Decision Support System) Architecture in Korea"; International Conference on Convergence and Hybrid Information Technology (ICHIT); Page(s): 700 – 703. 2008.