A Survey on Multiple Patient Data Semantic Conflicts and the Methods of Electronically Data Exchange Advantages and Disadvantages

Rageed Hussein Hussan Al-Hashemy

Almustansiryah University-College of Education-Computer Science

Abstract: *In last few years heterogeneous healthcare information such as electronically patient data has gains a great attention especially from clinicians, researchers, health care organizations. The government of united stat of America had urged for widespread adoption of Electronic Health Records or Medical Records (EHRs) within the next decade shows the need for electronic patient data exchange. The more number of health data systems the more of heterogeneity nature of electronic patient data conflict increase which led to the important of integration issues. All healthcare centers have their own aspects in their Independent healthcare information systems like type of medical reports, terminology, methods of connection between hospital sectors, data base architecture and semantics which is different from one place to another. Consequently the conflict arises as a result from heterogeneous data source. This survey focuses on multiples patient data semantic conflicts and the methods of electronically data Exchange advantages and disadvantages.*
Keywords*: Semantic conflicts, Healthcare system, EHR*

I. Introduction

The incredible development in health societies nowadays led to more thriving in various kinds of medical information systems and patient's data Systems in recent years. These systems are required to support data comes from several resource such as , hospitals , physicians, healthcare centers and laboratories each of which collect and keep medical data for healthcare delivery [Claus B. et al., 2004]. Despite the fact that, a variety of concepts of Independent healthcare information systems in healthcare centers like type of medical reports, terminology, methods of connection between hospital sectors, architecture of data base and semantics which is different from one system to another that may create a data semantic conflict after all . A collection of documents that include a wide range of papers and medical information related to one patient, which supporting communication and decision making in on practice with a daily basis and having different users and purposes is called a medical history report [Wyatt JC, 1994]. Health and medical information recording requires experts to access patients data system that may be distributed across internet, contained a various kinds of documents, papers and electronic formats, such as description, structured, coded, and multimedia entries [Kalra D., 2006].

To smooth the process of health care data exchange between hospitals and laboratories and other organizations such as pharmacies and payers, medical data needs to be organized, integrated, and analyzed since these data are held in various existing heterogeneous data sources that are spread across the network. Data source in medical data systems comes with its own features, structures, semantics, data formats, concepts and method of access in spite the fact that the broad spectrum of information is accessible over the web as the primary entry which make a challenging problem in supporting ad-hoc queries through multiple data sources that are retrieved from experiments .Semantic conflict, which is known by conflict, occurs between two semantically different concepts, or in other hand using different terms to describe the same concept in most databases. It is one of the most challenging conflicts that are usually aroused in integration of heterogeneous data sources. Data across constituent databases are different but may be related. Recognition of semantically related data in various databases and finding the resolution of the differences among the semantically related data are considered a heavy burden that falls on scientists to exchange between the data formats, resolve conflicts resolving, data integration, and interpret results manually in order to have visible use of patient data record.

Thus, resolving semantic conflicts in electronic patient data provides a standardized method for query translation and heterogeneity resolution to reduce the burden on scientists to be familiar with all contents and structures of the medical information sources.

II. Literature Review

Reconciling semantic conflicts in integrating heterogeneous data sources has been an active research area, because it helps to improve the consistency and robustness of integrated data sources especially during ad-hoc queries over the web. The heterogeneous information resource over the Internet has been repeatedly considered as the biggest obstacles before such an ambitious aim.

Standardized data plays a critical role in exchanging information accurately over different sites and users. [Qamar R. and Rector A., 2007a], mentioned the first step towards achieving standardization of medical data is data matching to codes in controlled terminologies. Proposed archetypes by the open EHR organization are used for modeling to define some clinical concepts which conform to the open EHR Reference Model (RM). [Beale T. and Hear S., 2005] mentioned that the expression is inherited from the RM in the form of structured constraint statements. The proposed purpose of archetypes is to empower clinicians or healthcare employee to define the content, semantics, as well as data-entry interfaces of systems separately from the information system [Beale T. and Hear S., 2005]. Three issues would be encountered when looking up matches in archetype terms Systematize Nomenclature of Medicine which called (SNOMED, which is a comprehensive terminology that provides clinical and medical content and expressivity for medical documentation and reporting): a) determining the semantics of use, b) determining the source of semantics, and c) misspelling leading to incorrect or no matches.

A methodology proposed by [Qamar R. and Rector A., 2007b] used to perform data mapping function in data standardization and developed a Model Standardization Using Terminology (MoST) system to test the methodology using open RHR Archetype Models and SNOMED. Methods of context and non-context with lexical and semantic techniques were employed to find matches and the most appropriate matches resulting from filtering process were presented to the modeler. There were several issues encountered when resolving conflicting semantics of archetype terms and SNOMED concepts, for example, the "autopsy examination" belonged to an observation archetype [Qamar R. and Rector A., 2007b]. The clinical modeler categorized it as a proposed meaning of 'observation' or 'finding' [Qamar R. and Rector A., 2007b]. However, SNOMED had categorized it as procedure, which upon manual examination was considered to be an appropriate categorization given the context due to the basic differences in the modeling strategies of the two models.

Health Level 7 is the most widely used which is a non-profit volunteer organization that develops specifications for being a messaging standard that enables disparate applications of healthcare to exchange keys sets of clinical, medical and administrative data. Its improve care delivery, optimize workflow, reduce ambiguity, and improve transfer of knowledge among stakeholders, like healthcare staff, agencies of government, the vendor community, whose fellow Standard Developing Organizations (SDOs), and patients. HL7 version 2 is popular used nowadays. [Veli B. et al., 2005], motioned that HL7 Version 2 compliant does not involve direct interoperability between healthcare systems .To remedy this problem, HL7 has developed Version 3, called Reference Information Model (RIM). HL7 Version 3 produces Hierarchical Message Definition (HDM), which defines the schema of the messages based on the RIM classes. However, an HL7 Version 3 message does not interoperate with HL7 Version 2 message where semantic mediation is needed to address the problem of interoperability.

Recently ontology mediation has been an active research area. [Veli B. et al., 2005] proposed a semantic mediation of exchange messages to improve interoperability among healthcare information systems. Exchanged messages are transformed into Web Ontology Language (OWL) ontology instances and then mediated through an ontology mapping tool which is called OWLmt that uses engine of Web Ontology Language -QL to reason over the instances source ontology through a GUI. Although this paper claimed that the proposed framework is comprehensive, but its only demonstrates the way to mediate between HL7 Version 2 and HL7 Version 3. There is no further work on mediating the other incompatible healthcare standards.

 Semantic conflict occurs either between two semantically different terms or to describe the same concept in most data base, which is the main conflict that cause semantic interoperability failure in electronic data exchange systems [Lee et al., 2010]. Semantic conflicts have been classified into two levels of conflicts, data level and schema level conflicts, six common types each; schema level conflicts types are; naming conflicts, entity identifier conflicts, schema isomorphism conflicts, generalization conflicts, aggregation conflicts and schematic discrepancies. The data level conflicts six types are data precision conflicts, known data value reliability conflicts, data representation conflicts, data unit conflicts, and spatial domain conflicts [Lee et al., 2010]. Matching algorithm helps in tackling the limitation of the previous work.

[Ricardo J. C. et al., 2007] reviewed on papers published between years of 1995-2005 by examining the approaches used to integrate patient data from heterogeneous data sources. From the result the preview showed that HL7 was the most common used messaging standard, while direct database access and web services were the most widely used communication method. Message passing emphasizes both of the remoteness of the object and the missing values of the code body which will be executed [Ricardo J. C. et al., 2007]. The difference between message passing and conventional concept of procedure calls or operation invocation is the reliance on open Internet standards. One key omission from those reviewed papers is the lack of error detection.

Unambiguous data exchange can be defined as computable semantic interoperability (CSI). [Mead C. N., 2006] showed an overview of the prime constructs of HL7, the Reference Information Model (RIM). There are four pillars that motivated the creation of HL7 version 3. However, it's still not sufficient to achieve the goal for CSI. For example, the four pillars say little or nothing about critical issue, such as enterprise-wide person

identity management; security, auditing or consent services; or terminology cross-mapped semantic relationship [Mead C. N., 2006]. These critical tasks are left to the vendor organizations or other IT resources.

In order to develop a referral letter system for electronic exchange of medical information [Yong H., 2008] designed a prototype model by utilizing HL7 clinical document architecture (CDA) in Japan with the local standard. [Yong H., 2008] expressed the corresponding concepts as standard data items in HL7 CDA as well as found that most referral module content's defined in (MML) Medical Markup Language could be represented in the HL7 CDA model to provide worldwide standard to meet the needs of the local clinical. However, the research on CDA with local standard is just a beginning; there are still a lot of interoperability problems that need to be tackled including extension of the document type scope more than the referral letter, and combination with local standard other than MML. Figure 1.0 shows the diagram which summarized the papers under the survey with the advantages and disadvantages of their work.

Another and Year	Title	Advantages	Disadvantages
Beale T. and Hear S., 2005	Archetype definition and principles	1-Archetypes are used directly for computational purpose. 2-Archetypes enable domain experts to be modeled in a formal way by experts.	1- In order to be useful, Archetypes should define coherent, whole informational concepts from the domain. 2- Each archetype has a limited size.
Mead C. N., 2006	Data Interchange Standards in Healthcare IT- Computable Semantic Interoperability: Now Possible but Still Difficult, Do We Really Need a Better Mousetrap?	1-HL7 V3 provides a number of tools to assist developers in building RIM-compliant. 2- Provides machines with unambiguous semantics for each data element transferred.	1-software is limited in its ability to disambiguate a symbol that can point to more than one concept. 2- it does not allow use of the methodology of a top-down for defining each data interchange structure
Veli B. et al., 2005	Artemis Message Exchange Framework: Semantic Interoperability of Exchange Messages in the Healthcare Domain	Using of Web Ontology Language -QL engine which enables the mapping tool to reason over the source ontology instances while generating the target ontology instances according to the definition of mapping patterns graphically.	This paper only demonstrates the way to mediate between HL7 Version 2 and HL7 Version 3. There is no further work on mediating the other incompatible healthcare standards.
Qamar R. and Rector A., 2007a	Semantic Issues in Integrating Data from Different Models to Achieve Data Interoperability	The more archetypes bound to SNOMED the more standardized data entry process.	Care must take into consideration to address the problems arising from the disparities as there is wide use of objectives and different models.
Qamar R. and Rector A., 2007b	Semantic Mapping of Clinical Model to Biomedical Terminologies to Facilitate Data Interoperability	The mapping methodology discussed in the paper was successfully tested and evaluated against the MoST system.	Archetype terms are not very suitable for generating post-coordinated codes, as they are mostly post-coordinated at the time of modeling. Manual mapping processes need to be replaced with intelligent semi or automated systems to take over most of the tedious search tasks.
Ricardo J.C. et al., 2007	Reviewing the integration of patient data: how system are evolving in practice to meet patient needs	This paper appraises studies examining the different approaches to integrating patient data from heterogeneous IS. And how systems are evolving in practice to meet the needs of patient, professional and organization.	1- Lack of detail reported in most of the articles. 2- Some of the papers it was difficult to determine if they were describing the same project or not.
Yong, 2008	A prototype Model using Clinical Document Architecture (CDA) with a Japanese Local Standard: Designing & Implementing a Referral Letter System	This paper attempted to promote a better harmonization of HL7 CDA and MML.	The research on using CDA with local standards is just a beginning and much work remains to be done, including extension of the document type scope more than the referral letter, and combination with local standards other than MML.
Lee.Yi 2010	Resolving semantic interoperability challenges in XML schema matching	The rules proposed in this paper not only maintain the semantic interoperability of information exchange, but also prove the correctness and reliability of data exchange over the internet.	This paper focuses on schema level conflicts and leaving data level conflicts for future work.

Figure 1.0 Comparison Table for the Methods of Electronically Data Exchange Advantages and Disadvantages

III. Conclusion

The rapid growing of Healthcare Information System has led to the issues of electronic data exchange in medical and healthcare systems across heterogeneous data source as each of the data source presents some semantic conflicts in several levels. Medical and patient systems are developed by different groups or organizations that work separately from each other. Each set of patient records is inherently different in the semantic of representation. In order to improve the communication and medical and healthcare related data, integration of healthcare information system plays a vital role. Therefore, standardization is the most popular method in electronic data exchange [Lee et al., 2010]. The majority of developed countries apply this technique in Health level 7 and archetype while less developed countries still used manual mechanizes or outdated schema.

Integrate data from heterogeneous sources is taken a great effort because the individual feeder medical and healthcare systems usually differ in several aspects. There are many solutions being proposed to these problems in last few years. Hence, a number of standardization efforts are progressing to address the semantic conflicts in each electronic patient data However; it's not realistic to force all of the healthcare organizations to conform to a single standard, Therefore, extensive works for facilitate electronic patient data exchange at semantic level remains to be developed.

References

[1]	Beale T. and Hear S., 2005	Beale T, Hear S.: Archetype definition and principles.(Revision 0.6), March 2005
[2]	Claus B. et.al, 2004	Claus Boyens, Ramayya Krishnan and Rema Padman: On Privacy-Preserving Access to Distributed Heterogeneous Healthcare Information, 2004
[3]	Kalra D., 2006	Kalra D.: Electronic Health Record Standards. *Methods In Med 2006, 45(1):136-144*
[4]	Lee.Yi . et al., 2010	Chiw Yi Lee • Hamidah Ibrahim • Mohamed Othman • Razali Yaakob: Resolving semantic interoperability challenges in XML schema matching, 2010.
[5]	Mead C. N., 2006	Charles N.Mead: Data Interchange Standards in Healthcare IT-Computable Semantic Interoperability: Now Possible but Still Difficult, Do We Really Need a Better mousetrap?, *Journal of Healthcare Information Management — Vol. 20, No. 1,* 2006.
[6]	Qamar R. and Rector A., 2007a	Rahil Qamar, Alan Rector. Semantic Issues in Integrating Data from Different Models to Achieve Data Interoperability. Medinfo Conference 2007, Brisbane, Australia.
[7]	Qamar R. and Rector A., 2007b	Rahil Qamar, Alan Rector. Semantic Mapping of Clinical Model Data to Biomedical Terminologies to Facilitate Data Interoperability. Healthcare Computing 2007 Conference, Harrogate, U.K.
[8]	Ricardo J. C. et al., 2007	Ricardo J Cruz-Correia, PedroM Vieira-Marques, Ana M Ferreira, Filipa C Almeida, Jeremy C Wyatt and Altamiro M Costa-Pereira: Reviewing the integration of patient data: how system are evolving in practice to meet patient needs, *BMC Med Inform Decis Mak. 2007 Jun 12;7:*
[9]	Wyatt JC, 1994	Wyatt JC: Clinical data systems, Part 1:Data and medical records. *Lancet 1994, 344:1543-1547*
[10]	Veli B. et al., 2005	Veli Bicer, Gokce B.Laleci, Asuman Dogac, Yildiray Kabak: Artemis Message Exchange Framework: Semantic Interoperability of Exchange Messages in the Healthcare Domain, SIGMOD Record, Vol. 34, No. 3, Sept. 2005
[11]	Yong H., 2008	Huang Yong, Guo Jinqiu and Yohsio Ohta: A ptototype Model using Clinical Document Architecture (CDA) with a Japanese Local Standard: Designing and Implementing a Referral Letter System, 2008