DESIGN OF A HEALTH CARE ARCHITECTURE FOR MEDICAL DATA INTEROPERABILITY AND APPLICATION INTEGRATION

Christina Catley[1] and Dr. Monique Frize[1, 2]

[1] Department of Systems and Computer Engineering, Carleton University, Ottawa, ON, Canada

[2] School of Information Tech & Eng., 161 Louis Pasteur, University of Ottawa, Ottawa, ON, Canada

Abstract- **The Internet is changing the face of medical research. The current world of isolated research and proprietary data encodings is evolving into a future of standardized medical databases and integrated medical applications, such as clinical decision support systems. This paper explores the use of XML, and its associated Schema Language, to enhance sharing of medical data. XML enables data portability and will reach its full potential when the medical community develops a standardized basis for medical schema content and shares these schemas in recognized repositories. Our research group is currently harnessing XML's standardization potential by designing a standards-compliant, medical information infrastructure that will allow for seamless integration of all our clinical decision support tools.**

Keywords- **eXtended Markup Language, Clinical Decision Support Systems, integration architecture, metadata**

I. INTRODUCTION

Medical system development relies heavily on the use of medical databases. Most of these databases are institution specific and as such have heterogeneous data models. The majority of medical software handles the problem of non-standardized databases by manipulating the data into a form that can be interpreted by the program in question. This practice severely limits the Internet's potential for remote sharing of medical data within a global context. The Internet provides connectivity – not the context within which semantics can be shared.

Medical informatics experts agree that an 'integrated vision' is needed in order to facilitate the shared usage of electronic patient data [1]. McDonald expresses this idea eloquently: "*The sources of patient data that do exist… reside on many isolated islands"*[2]. Intensive standardization efforts are required before medical databases can become the shared commodity envisioned by researchers and health care professionals. An ideal medical database would then have broad applicability and would follow standardized, global processes that are universally acceptable

II. METHODOLOGY

Our team's primary objective is to develop artificial intelligence-based software, such as the Case-Based Reasoner (CBR) and Artificial Neural Network (ANN), to be used in clinical decision support systems (CDSSs). Our long-term goal is to link our software tools to the Children's Hospital of Eastern Ontario's (CHEO's) Neonatal Intensive Care Unit (NICU) online data repository. The real-time data received from the online data repository must be processed in a variety of different contexts, since ideally the data will be accessible through ANN, CBR or other clinical decision support systems dependent upon the physician's preference. This greatly increases the need to extend the applicability of our medical data through strict conformance to accepted medical standards. As expressed by Dr. Robin Walker, head of CHEO's NICU: "The ultimate objective would be to transmit infant data from the hospital to diverse, external ANN or CBR systems, and to receive a reply in real-time". Our system will serve as a foundation to bridge this and other communication impediments caused by a lack of conformance with medical standards.

A. XML's Role in Medical Standardization

Our group is exploring the use of XML, employed in conjunction with the Unified Medical Language System's (UMLS) Knowledge Sources, as a means to solve the system challenges posed by weak or non-existent medical database standardization. The UMLS Metathesaurus represents semantic information about biomedical concepts.

In order to understand XML's benefits, it is important to identify which system challenges it best solves. XML's main feature is that it enables data portability by separating the presentation of the data from the content. XML is a good candidate technology when the goal is to interchange self-describing data from one application to another over communication channels such as the Internet.

As the use of Internet technology in medicine becomes more prevalent [3], such an interchange capability has the potential to become a very important health care enabler. Widom stated that: *"The exciting promise of XML is that it turns the web into a global database"* [4]. Our aim is to capitalize on this promise by sharing medical data and associated access to enterprise applications, such as clinical decision support systems, both internally and with other research groups and health care institutions.

B. XML Schema Language

As described, XML provides a self-describing syntax for heterogeneous data structures [5]. XML also provides a means of attaching metadata to describe data representations and semantics, through the use of the XML Schema Language, proposed as a candidate recommendation by the World Wide Web Consortium in October 2000 [6]. XML Schema Language supports the rich data types associated with object-oriented programming languages and provides support for data validation to ensure that the XML data conforms to its attached schema. This rigorous data checking is imperative in medical applications where data integrity is essential. In addition, XML Schema Language supports schema inheritance. This feature is of particular

use to our group's ANN application since it allows automatic generation of new schemas from existing schemas after our ANNs eliminate variables.

Domain-specific groups, such as those within the medical community, can define original content by creating their own schemas to meet the data exchange needs of their particular fields. The schemas are transmitted with the data to be interpreted by XML parsers and subsequently directed to enterprise applications, thereby creating self-describing data. The important point is that the platform-neutral XML tags do not use a proprietary format, but rather an open and self-describing one, with the goal that: "*XML will be to data portability what Java is to application portability"* [7]. XML schemas can then be stored in repository structures for publication and discovery purposes within the medical informatics community.

It is essential to realize that without support, conformance, and acceptance from medical groups, XML-enabled documents and widely published schemas will not overcome data barriers caused by schema diversities. This implies considerable content standardization of XML medical tags to prevent the proliferation of several overlapping tags attempting to represent the same information. Publishing schemas in recognized registry/repository structures can foster such metadata standardization.

III. Results

A prototype system is under development based on 'XML-enabling' our medical databases and publishing the relevant XML Schema Language metadata in a publicly accessible repository. This will provide our researchers with secure web-enabled access to all our databases and allow external groups and affiliated organizations to run their XML-enabled data through our CDSSs. The system design advocates an open source technological approach. The relational database management system used to store the data was implemented using MySQL and accessed remotely using Java Database Connectivity (JDBC). All code was written using Sun's Java 2 SDK and the Java APIs for XML processing. The system extracts data from the MySQL database into a new XML document, and conversely parses existing XML documents into the database for future processing and storage using Apache's IBM supported Xerces 4J parser. The XML data is to be integrated with our MATLAB-based artificial neural network tool and our case-based reasoning tool.

IV. Discussion

Our goal is to enable application integration by creating an XML-enabled interface between all of our group's clinical decision support systems. Currently, the ANN and the CBR are separate entities with complex and different pre-processing requirements. Our aim is to design a transparent interface that will allow seamless processing of XML data by a wide variety of clinical decision-support tools.

Much has been written in the literature regarding the application of XML to integrate distributed information sources and to promote information exchange within hospital information systems. However, the majority of this work is concerned with promoting shared and standardised access to information stored within the Computer Patient Record (CPR). One notable extension was reported, where a CBR system interacts with data stored within heterogeneous CPR repositories [8]. The CPR data is transformed into operable CBR-oriented cases that can be processed through the CBR system to provide web-enabled diagnostic services. To the authors' knowledge an XML-based health care architecture designed to support integration between CDSS software is a new application of this technology.

V. Conclusion

Through our team's affiliation with CHEO's neonatal intensive care unit, we have the opportunity of testing and validating our system in a real hospital environment. This will involve an iterative design approach to accommodate feedback from Neonatal Intensive Care Unit physicians and medical researchers. This system could be deployed in the NICU, allowing physicians to feed real-time patient data into diverse clinical decision support systems in order to gain insight into difficult cases and to make more knowledgeable decisions regarding the potential survival of very premature infants. The system also has great potential for the education and training of physicians.

References

[1] S. Flesca, S. Greco, E. Zumpano, "Modeling and querying XML data," *International Database Engineering & Applications Symposium*, 2000, pp. 275–283.

[2] C.McDonald, "The barriers to electronic communication and how to overcome them," *J. Am. Med. Inform. Assoc.,* vol. 4, 1997, pp.213-221.

[3] N. de la Cruz, C. Kahn, "Extensible markup language in health care: integration of structured reporting and decision support," *J. Am. Med. Inform. Assoc.*, Supp., 1998, pp. 725-730.

[4] J. Widom, "Data management for XML: research directions," *IEEE Data Engineering Bulletin*, 22(3), Sept. 1999, pp.44-52.

[5] L. Seligman, A. Rosenthal, "XML's impact on databases and data sharing," *Proc. IEEE Comp.,* June 2001, pp.59-67.

[6] XML Schema Part 0: Primer, accessed: March 2002, http://www.w3.org/TR/xmlschema-0/

[7] A. Lear, "XML seen as integral to application integration," *IT Pro*, Sept/Oct 1999, pp.12-16.

[8] S. Manickam, A. Abidi, "Extracting clinical cases from XML-cased electronic patient records for use in web-based medical case based reasoning systems. *Proc. Medinfo*, 2001, pp.643-647.