



Conceptual Model for Clinical Data Format Analogous to Clinical Data Warehouse

Kanchan Badwe

Student, MBA IT, SICSR, Affiliated To Symbiosis International University (SIU), Pune-16

Syed Khizer

Assistant Professor, Department of Computer Science, College of Computer, Qassim University, Al-Qassim, Kingdom of Saudi Arabia (KSA)

Pravin Metkewar

Associate Professor, MBA IT, SICSR, Affiliated To Symbiosis International University, Pune-16

ABSTRACT

The Clinical Data Warehouse is a result of collecting data spread throughout its many system and utilising data warehouse technology in medical field. Collection of data into a consolidated, organised and accessible database for analysis, reporting and research purpose makes it a better usage for clinical data in order to support decision making.

The challenges include clinical data format, business analysis, data integration, data quality and ETL process. In this we will focus on actual problem of data format and propose solution for the same. In the paper we will discuss the issues and challenges to address successful data warehouse for medical organisation which provide knowledge for effective decision making and problem solving.

KEYWORDS : Data Warehouse (DWH), Clinical Data Warehousing (CDWH), Data Quality, ETL process.

1. INTRODUCTION:

The first data warehouses technology was developed in 1980 due to lack of information provided by online application systems that were implemented but were rarely integrated. Data warehouse integrate data from two or more operational system present in one or more organisations. The integration process include three steps: (i) implementing a unified model which can store information from single databases, (ii) Transformation of data into developed model before piling them into data warehouse, (iii) extraction of data from source database and integrating and accessing it to obtain required knowledge which one particular source cannot provide. However, integration also includes set of hardware and software components which make it easy for analysis of vast data which helps for decision making and research. Integration process also includes architecture and tools to collect, analyse, clean and present information. Nowadays clinical fields is becoming a famous research domain for DWH and Computer Science. CDWH is a process of integrating medical data from administrative and operational medical system. CDWH platform includes research, reporting and study planning, and improve the value of decision making and intervention on timely basis. CDWH also facilitates storage, increases standard of real time decision process and enhance analysis.

2. PROBLEM DEFINITION:

The CDWH technology aims to identify the relationship in clinical data, discovering disease trends, improving patient's outcome, evaluation of performance of different treatment protocols used and to provide information to users in areas ranging from research to management. The medical data collected on regular basis is stored in various systems which includes statistical information, medical information system, and laboratory information system. During patients visit clinical data is stored in various medical systems.

The types of data consist of:

- i) Clinical Information-collecting data about patients life habits, which enhances the capabilities of data analysis.
- ii) Demographic Information-collection once to provide analysis of rich data.
- iii) Diagnosis Information-Brief about diagnosis procedure.
- iv) Treatment Information-Brief about treatment which includes type, procedure and the risk involved.
- v) Laboratory Information-Testing the following results.

In this context we are required to highlight the issue of clinical data format. In this format medical data produces new issues and challenges to DWH technologies.

The medical data consist of following requirements:

1. Requirement of ethical and legal constraints.
2. Issues persisting to medical information system regarding sensitive data, different storage formats and privacy issues.
3. Substantial amount of data about patients along with treatment details and clinical conditions is required.
4. Clinical data concept contains various types of data: numeric and quantitative, Image (MRI and radiology), Ultrasound, text and qualitative format, Signal data (ECG) and generic and so on.

Mining of data mentioned above requires transformation phenomenon that is developed specially to deal with specific characteristics of medical data. Aggregation of data is necessary where clinical data is complex and poorly characterized.

3. EXPERIMENT BASED RESEARCH METHODOLOGY:

Clinical data is a major resource for research on health and medical. Collection of clinical data is either through course of ongoing patient care or by formal clinical trial program.

The XML based on clinical and experimental data exchange also known as XCEDE based schema provides ample metadata pyramid for storing, describing and documenting the data produced by scientific studies. Current version 2.0, the XCEDE schema consist of stipulations for the interchange of accurate data between databases, analysis tools, and web services. It serves as a framed metadata pyramid, stores information relevant to different aspects of an experiment (project, subject, protocol, etc.). Each level also stores data source information that allows for a traceable record of processing and/or changes to the respective data. The schema supports the needs of different data modalities and expresses types of data not originally visualised by the developers. The XML files consist of a unique technique for transport and interchange of experimental data, enabling import/export between heterogeneous databases, discovering specialized web services, local stash of exploring information within data collections, and formation of human and machine distinct depictions of the ac-

tual data.

XCEDE is constructed to reserve scientific data and Meta data from various sources and format.XML schemas have XCEDE origins in it and was discovered to facilitate the shift and storage for different types of data that includes but is not limited to clinical,behavioural,demographic and image data.

4. HOW XCEDE SDCHEMA CAN HELP IN ANALYSING CLINICAL DATA FORMAT USING XML:

In the above figure we can select a single subject and have multiple operations perform through that subject. For our proposed solution to perform successfully we need to select a particular clinical data and start analysing by performing operation on it through xml.When we come to certain predictive analysis only then we can make decisions. If more combinations of operations need to be performed to analyse the clinical data and come to a certain predictive decision this is one the method we can use. The multiple operations can later be grouped and named as group x or any other name you wish to give. If the decision is to be made on different analysis of clinical data they can be made into a different project.

5. CONCLUSION:

As the saying goes "Try and Fail, but don't Fail to try".Through the experiment mentioned above we can conclude that the requirement and integration issues may be solved. Clinical data format is of complex nature and a difficult data type that requires complex transformation processes. So we can conclude that till we reach a point where we can say that clinical data issues are solved experimentation is the best that we can do.

6. FUTURE SCOPE:

The field of clinical analytics is hastily growing and vast variety of technologies are implemented in the hands of users related to this field. Some trends such as growing settings of care in the data warehouses. Other trends such as patient care will require adequate management and processes related to clinical implementation.

CLINICAL DATA AND THE GROWING SCOPE:

Setting of care

Structured form of clinical data elements

Flow measures of patients

Analysis based on Xml based schema XCEDE

7. REFERENCES:

1. <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3836560/>
2. <http://software.incf.org/software/xml-based-clinical-experiment-data-exchange-schema>
3. https://scholar.google.co.in/scholar?q=research+papers+on+clinical+data+formats&hl=en&as_sdt=0&as_vis=1&oi=scholar&sa=X&ved=0ahUKEWjuiNOqta_OAhX-JpY8KHcjKB5wQgQMUTAA
4. http://www.sersc.org/journals/IJUNESST/vol7_no5/22.pdf.