Digital Health Innovation for Consumers, Clinicians, Connectivity and Community
A. Georgiou et al. (Eds.)

doi:10.3233/978-1-61499-666-8-100

An Implementation of Clinical Data Repository with openEHR Approach: From Data Modeling to Architecture

Lingtong MIN[a], Juanjian LIU[a], Xudong LU[a,1], Huilong DUAN[a] and Qingzhi QIAO[b]
[a] *College of Biomedical Engineering & Instrument Science, Zhejiang University, China*
[b] *Department of Information Technology, Shanxi Dayi Hospital, Taiyuan, Shanxi, China*

Abstract. Since electronic healthcare records are widely implemented in the hospitals, the explosion of the clinical data brings great opportunities to secondary use. Clinical data repository (CDR) plays an important role in clinical data using, including healthcare service, research and management. With evolving of clinical knowledge, it is a challenge for CDR to meet more and more requirements with adaptability and flexibility. In order to overcome this challenge, this paper proposed a solution that implemented a CDR with openEHR approach from data model to implementation. This study modelled archetypes according actual CDR requirements and implemented a CDR system in a tertiary hospital. Furthermore, this study developed data applications to facilitate healthcare service and research based on the CDR. The results of CDR system and data application demonstrate that openEHR approach could meet the continually evolving need of systems.

Keywords. Clinical data repository, openEHR, data model, implementation

Introduction

Since electronic healthcare records are widely implemented in the hospitals, the explosion of the clinical data brings great opportunities to secondary use. Clinical data repository is a centralised database that collect and store administrative and clinical data from heterogeneous data sources and open access to users [1], such as organisations, researchers, managers, physicians etc. It is well established that CDR has played an important role in biomedical data using.

With development of biomedical information technology, evolving of clinical knowledge, more and more requirements of data using are emerging. In order to meet these requirements, many initiatives or organisations have built CDRs for specific purposes [2-7]. As clinical knowledge is evolving continuously, requirements of systems are changing subsequently. As a result, CDR has to face the challenges of adaptability and flexibility. However, the conventional methods use a single-level data model to build CDR, which hard-code information and knowledge in data models and systems. In other words, the conventional methods build CDR in a specialised way but not a dynamic way.

[1] Corresponding Author.

In addition to this, healthcare data is generally too complicated, flexible, and changeable to capture a universal, comprehensive and stable schema of information, which is the foundation of the entire data using architecture. Highly-specialised systems cannot acclimate to the evolution of healthcare data requirements, which require systems to embrace a dynamic, state-of-the-art, rapidly evolving information infrastructure to facilitate data using.

OpenEHR focuses on systems and tools necessary to the computation of complex and constantly evolving health information at a semantic level based on the two-level modelling method. The openEHR approach is a potential solution to the problems that CDR systems faced. In two-level data modelling, openEHR divides data model into reference model (RM) and archetype model (AM) [8]. RM is comprehensible and stable information model, which is account for components of information systems. AM accounts for defining the clinical content, which is representation of clinical knowledge and defined by domain experts. AM consist of archetypes and templates, the benefits of AM includes: knowledge-enabled systems, knowledge-level interoperability, domain empowerment and intelligent querying [9]. Feasibility of representing clinical content models by openEHR RM and AM has been proven and legacy content models can be transformed into archetypes automatically [10].

This study implemented a CDR in a tertiary hospital with openEHR approach, including data modeling, implementation architecture and case implementation.

1. Method

1.1. Data Modeling

Clinical knowledge manager (CKM) is an online repository for managing openEHR archetypes and templates, which consists of 426 archetypes, 17 templates and one term set. Archetype defines the semantics of information system, which facilitates interoperability of semantics and knowledge. Template specifies particular groups of archetypes to use for a particular purpose, corresponding to a screen form. Among the existing archetypes in CKM, there are 47 archetypes on the published status that means the content and structure of archetypes are relatively stable at the time when they were set to published status. Whereas there are many existing archetypes into CKM, they can't meet all the requirements of clinical practices. As a result, new archetypes may inevitable when CDR system was built with openEHR approach.

This study built archetypes referencing the method proposed in authors' early research according the clinical requirements [11]. Base on the early study [11-12], we refine the data modelling method to facilitate reasonability and feasibility. Besides that, this study extends the scope of data modelling. Firstly, analysed the existing legacy database schema or database specification, including number of tables, fields and relationship. Then combined fields according the semantics of them, combination would happen when the fields have the same semantics. Based on the results, clinical experts abstracted clinical concepts and mapped into archetypes within CKM. Among the procedure of mapping, if the clinical concepts can be covered by existing archetypes, the archetypes would be used directly without modification. Modification and new operation happened when the clinical concepts and existing archetypes were mismatched.

1.2. Implementation Architecture

Although there are several projects to implement openEHR standard or map openEHR data model to data storage, such as LinkEHR [13], Think!EHR [14], etc. They are relatively closed for end users. In other words, they are developer-oriented rather than user-involved advocated by openEHR. The end users can control neither the data nor the display layout of systems, they only account for defining archetypes rather than get involved in design and development of systems. User-involved and adaptability are crucial for openEHR, two-level data modeling is a prerequisite for that. Domain experts can involve in and control systems by defining archetypes and templates. In addition to this, an architecture for bridging the gap between knowledge (archetypes) and systems is necessary.

In order to achieve the goal of user-involved and adaptability, this study described an open architecture for openEHR implementation that is illustrated in **Figure 1**. This architecture mainly consists of archetype relational mapping (ARM) [15], archetype access service and application designer. ARM is a practical data persistence solution for archetype-based systems, which bridges the gap between archetype models and relational database. Taking into consideration limitations of ARM, this study proposed the template relational mapping (TRM) that mapped templates into relational database rather than mapped archetypes into relational database. It has been demonstrated that relational database has an excellent performance for data persistence in many electronic healthcare record (EHR) projects and it is still the mainstream choice for EHR systems. It is necessary to map archetypes into relation database schema for implementation of systems with openEHR approach.

After data modelling, we got the general healthcare concepts and encapsulated them into archetypes. In contrast to archetype corresponds to general clinical concepts, template corresponds to local use by assigning at least one archetype to an archetype group for specific application purpose. According the rules for mapping, TRM mapped the templates that defined for meeting specific application requirements into relational database schema, including table names from the ontology in the archetype, field names from the full path of the items in the template, key, foreign keys and so forth.

Besides the TRM, the architecture includes access service and application designer, they are both based on the archetype model. Because the data persistence is based on the archetype-driven method and openEHR provides a query language at the semantic level, named archetype query language (AQL) [16], the data access service can be executed through archetypes based on the archetype-driven method. Although users could manipulate clinical data by AQL, in order to facilitate development of systems developers always adopt application programming interface (API) rather than AQL. Then it is inevitable to generate API based on archetype model, including archetype model parsing relationship mapping based on rules and automatic generation. By this access service, users can manipulate the clinical data to follow their thoughts without taking the technology of database into consideration, just control the clinical data by controlling archetypes.

It is well known that the screen form of application can be generated by templates. However, this generated screen form is relatively static that the elements in the form are arranged in a fix order. In other words, users could not control the display style. Nevertheless, the display style means represent something meaningful to users and has great influence on the healthcare service. This study supports an application designer that help users to design the display style of clinical data by drag-and-drop operation on the templates.

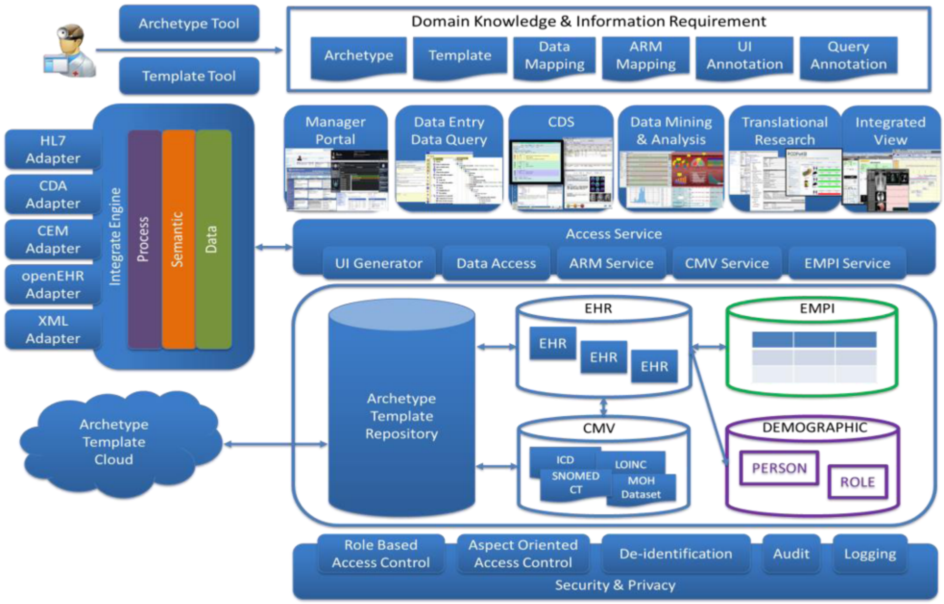

Figure 1. Architecture of implementation for the openEHR approach

1.3. Case Implementation

In order to verify the feasibility and usability of the openEHR implementation mentioned above, we carried out a Casestudy by building a CDR in a tertiary hospital with the data modelling method and implementation architecture.

The hospital chosen in the case is a tertiary hospital that opened since 2011 and has 2000 beds and 5000 encounters per day. By using relatively new technologies, the development of biomedical information in this hospital is very fast, which provides the technical basis for this case implementation. Moreover, the kinds of electronic clinical data are relatively complete and complex, which provides the data basis for this case study.

Firstly, we constructed a domain experts group for defining archetypes and templates to cover the scope of the CDR requirements, which consists of clinical experts, physicians, nurses, managers and developers. By using the data modelling method mentioned above, this study built 73 archetypes and 65 templates to cover the requirements of CDR in this tertiary hospital, among these 13 archetypes were used directly, 45 new archetypes were built and 15 archetypes are modified. By using the TRM method, this study generated the schema of database with 85 data tables in a SqlServer database.

Then we integrated clinical data into the CDR from existing information systems in this tertiary hospital, including Hospital management information system (HMIS), Laboratory information system (LIS), Radiology information system (RIS), Picture archiving and communication system (PACS), Pharmacy management system (PMS), Operation management system (OMS) and Computerised Physician Order Entry (CPOE). Finally, based on the CDR, we developed an application to help physicians view clinical data of a patient with a patient-center pattern by accessing the CDR rather than clinical data scattered in heterogeneous data sources, the clinical data includes demographic, imaging examination, order, operation, diagnosis and laboratory test data.

2. Result

According to the feedback from the hospital, it has been proved that the CDR built with the implementation proposed in this study is feasible and usable. In addition to this, based on the CDR this study also developed a web application to manage the clinical data for research with the physicians. Due to the requirements of research, the application for research need to has the characteristic of adaptability. This study expanded the application with the proposed architecture to meet the new requirements, including screen form and clinical data content. It has been represented that the architecture can adapt to the knowledge evolving.

3. Discussion

Although the proposed implementation has been proved that it can response to the challenges caused by knowledge evolving, there are some limitation existing both on data model and architecture.

Firstly, archetypes and templates build a bridge for communication between clinical experts and software developers, which facilitates development of biomedical informatics. Although this study can build archetypes and templates to meet the requirements of the clinical practices, the standardisation, theoretical basis of this data modelling method remains uncertain. In addition to this, due to the limitation of modelling tools, technical threshold and modelling specification, it is not easy for users to set out to define data models appropriately.

By employing an archetype-driven approach, TRM generates data schema for data persistence with archetypes and templates according to the mapping rules, which put clinical experts in the position of controllers who decide how to storage the clinical data. Furthermore, TRM can update the database any time without disruption of database. Although TRM can provide better performance than Node+Path database, in order to acquire automation, it has some problems about performance and structure in contrast to the conventional relational database.

4. Conclusion

This paper introduced an implementation of openEHR approach, including data modelling and implementation architecture. By a Casestudy for building a CDR with the implementation solution, the results indicate that this implementation is feasibility against clinical practices and supports the characteristics of adaptability and user-involved. Although there are some limitations in this implementation, it will be an important step toward implementation and localisation of openEHR approach.

References

[1] HIMSS, *HIMSS Dictionary of Healthcare Information Technology Terms, Acronyms and Organizations*, Healthcare Information and Management Systems Society, 33 W. Monroe, Suite 1700 Chicago, 2013.

[2] R. S. Evans, J. F. Lloyd, and L. A. Pierce, "Clinical use of an enterprise data warehouse," in *AMIA Annual Symposium Proceedings*, 2012, p. 189.

[3] T. Botsis, G. Hartvigsen, F. Chen, and C. Weng, "Secondary use of EHR: data quality issues and informatics opportunities," *AMIA summits on translational science proceedings,* vol. 2010, p. 1, 2010.

[4] C. G. Chute, S. A. Beck, T. B. Fisk, and D. N. Mohr, "The Enterprise Data Trust at Mayo Clinic: a semantically integrated warehouse of biomedical data," *Journal of the American Medical Informatics Association,* vol. 17, pp. 131-135, 2010.

[5] H. J. Lowe, T. A. Ferris, P. M. Hernandez, and S. C. Weber, "STRIDE–An integrated standards-based translational research informatics platform," in *AMIA Annual Symposium Proceedings*, 2009, p. 391.

[6] S. Rea, J. Pathak, G. Savova, T. A. Oniki, L. Westberg, C. E. Beebe*, et al.*, "Building a robust, scalable and standards-driven infrastructure for secondary use of EHR data: the SHARPn project," *Journal of biomedical informatics,* vol. 45, pp. 763-771, 2012.

[7] L. R. Waitman, J. J. Warren, E. L. Manos, and D. W. Connolly, "Expressing observations from electronic medical record flowsheets in an i2b2 based clinical data repository to support research and quality improvement," in *AMIA Annual Symposium Proceedings*, 2011, p. 1454.

[8] T. Beale, S. Heard, *Architecture Overview*, http://www.openehr.org/releases/1.0.2/architecture /overview.pdf, Mar. 2016.

[9] T. Beale and S. Heard, *Archetype Definitions and Principles*, http://www.openehr.org/releases /1.0.2/architecture/am/archetype_principles.pdf, March. 2016.

[10] R. Chen, G. O. Klein, E. Sundvall, D. Karlsson, and H. Åhlfeldt, "Archetype-based conversion of EHR content models: pilot experience with a regional EHR system," *BMC medical informatics and decision making,* vol. 9, p. 33, 2009.

[11] L. Min, L. Wang, X. Lu, and H. Duan, "Case Study: Applying OpenEHR Archetypes to a Clinical Data Repository in a Chinese Hospital," in *MEDINFO 2015: EHealth-enabled Health: Proceedings of the 15th World Congress on Health and Biomedical Informatics*, 2015, p. 207.

[12] M. d. Santos, M. Bax, and D. Kalra, "Dealing with the archetypes development process for a regional EHR system," *Applied clinical informatics,* vol. 3, pp. 258-275, 2012.

[13] *LinkEHR Platform*, http://www.linkehr.com/, Mar. 2016.

[14] *Think!EHR Platform*, http://www.marand-think.com/, Mar. 2016

[15] L. Wang, L. Min, R. Wang, X. Lu, and H. Duan, "Archetype relational mapping-a practical openEHR persistence solution," *BMC medical informatics and decision making,* vol. 15, p. 1, 2015.

[16] T. Beale, H. Frankel, and C. Ma, *Archetype Query Language (AQL)*, http://www.openehr.org /releases/QUERY/latest/docs/AQL/AQL.html#latest_issue_date, Mar. 2016.