Design and Implementation of Clinical Databases with openEHR

Pablo Pazos Gutiérrez\textsuperscript{a,b,c}, Koray Atalag\textsuperscript{d}, Luis Marco-Ruiz\textsuperscript{e}, Erik Sundvall\textsuperscript{f,g}, Sérgio Miranda Freire\textsuperscript{h}

\textsuperscript{a} openEHR en Español, \textsuperscript{b} Asociación Chilena de Informática en Salud, \textsuperscript{c} CaboLabs
\textsuperscript{d} The University of Auckland and openEHR New Zealand
\textsuperscript{e} Norwegian Centre for Integrated Care and Telemedicine, University Hospital of North Norway
\textsuperscript{f} Linköping University, Sweden
\textsuperscript{g} Region Östergötland, Sweden
\textsuperscript{h} State University of Rio de Janeiro, Brazil

Abstract

Databases for Clinical Information Systems are difficult to design and implement, especially when the design should be compliant with a formal specification or standard. The openEHR specifications offer a very expressive and generic model for clinical data structures, allowing semantic interoperability and compatibility with other standards like HL7 CDA, FHIR, and ASTM CCR. But openEHR is not only for data modeling, it specifies an EHR Computational Platform designed to create highly modifiable future-proof EHR systems, and to support long term economically viable projects, with a knowledge-oriented approach that is independent from specific technologies. Software Developers find a great complexity in designing openEHR compliant databases since the specifications do not include any guidelines in that area. The authors of this tutorial are developers that had to overcome these challenges. This tutorial will expose different requirements, design principles, technologies, techniques and main challenges of implementing an openEHR-based Clinical Database, with examples and lessons learned to help designers and developers to overcome the challenges more easily.

Keywords:
Electronic Health Records; Clinical Data Repositories; Clinical Database Design; openEHR; Clinical Information Systems.

Tutorial Description

The main motivation of this tutorial is to help reducing the gap between the openEHR formal specifications and concrete database implementation. The rest of the tutorial will be focused on presenting proven solutions and lessons learned to create openEHR-based Clinical Databases.

We'll not only focus on clinical data storage, but how that that is queried and used for health care (documentation, order management, Clinical Decision Support, healthcare plan follow-up for chronic diseases and other conditions, etc.), and secondary uses (clinical research, education, public health, quality assurance, and complying with medico-legal requirements). This tutorial will limit its treatment of secondary uses of clinical data to a brief overview. The openEHR community has proposed another tutorial that focuses on this topic. The current proposal can be considered the first of a two-part tutorial, here the focus is on practical way of implementing clinical databases for primary uses of clinical data, and the other focused on secondary uses, database exploitation and data reuse.

The main content of the tutorial will include an overview of different database technologies (relational, document-oriented, key/value) and techniques to actually design and implement openEHR-based databases, including lessons learned from implementation experiences and good practices that may help to tackle implementation challenges. We'll overview the key concepts of the openEHR specifications that affect database design (EHR, Folder, Composition, Entry, Item Structure, Data Value) and the openEHR dual-model (information + archetypes). This will help attendees to understand how to design Clinical Databases to create systems that provide services for storing health records and query those records in the “openEHR way”, by creating queries that use clinical concept definitions (Archetypes) instead of relying on a specific database technology (SQL, XQuery, JSON-based queries, OO queries, etc.), making queries a generic, reusable and shareable artifact between implementations.

We’ll explain how to implement some functional requirements related to health record versioning (to support modifications, amendments and corrections at the database level), clinical data validation and indexing to improve query performance.

Will end with the conclusions about the different approaches to Clinical openEHR Database design and implementation.

During the tutorial, audience participation and discussion will be encouraged.
Tutorial Structure

This tutorial will be organized by the following structure, that is based on the openEHR Clinical Database Design and Implementation course content [4].

Introduction

- Requirements for data access
  - EHR purposes from ISO 18308 [5], guiding different uses of clinical data in healthcare (primary and secondary uses).
  - Information System classification by data-base uses (write-oriented, read-oriented, read/write-oriented).
  - Organization and access to clinical information (chronologic, by health problem, by medical specialty, by logical or physical location, by episode).
- Requirements for data structures
  - Paper-based health record structure analysis.
  - Minimal EHR content from ISO 18308.
  - Clinical records organization and main entities from the openEHR Information Model (Free text vs. semi-structure vs. fully structured clinical records, support for different data types and temporal series, times in the EHR).
- Requirements for EHR audit.
- Design principles for openEHR databases
  - Introduction to Archetypes and Operational Templates (big archetypes).
  - Data and metadata in the database.
  - Generic data storage and data query design using openEHR Archetypes and Information Model.
  - OpenEHR Information Model details (compositions, sections, entries, structures, items, data values).

Technologies and Techniques

- Different kinds of database management systems (relational, document-oriented, object-oriented, key/value, graphs, multidimensional, etc.)
- The Object-Relational Mapping technique
  - Mapping classes, relationships and inheritance for an Object Oriented model into a relational database.
  - Database schema generation techniques and normal forms.
  - Examples of implementation.
- openEHR data in XML and JSON
  - Communication and storage of XML and JSON documents.
  - Tools to work with XML and JSON.
  - Examples of implementation.
  - Working with openEHR Archetypes in software
    - Introduction to the Archetype Definition Language (ADL).
    - Parsing and using archetypes for data validation and terminology extraction.

Advanced concepts

- EHR change control and versioning, implementation at the database level.
- Data indexing and performance.
- Synchronization / Replication for high availability and backup.
- System integration and the Extraction-Transformation-Loading (ETL) process using openEHR.

Specific educational goals

The main goal of this tutorial is to shorten the gap between the openEHR specifications and real world Clinical Database design and implementation for software developers.

Secondary goals include: sharing experiences and lessons learned, enabling discussion about Clinical Database implementation (a topic that is not discussed or taught often enough), and spreading the use of eHealth standards at the database level on Clinical Information System development.

Expected outcomes

After the tutorial, we expect attendees to have a better understanding of the openEHR specifications and scope, to know the main requirements of clinical data structures and database technologies that meet those requirements. Also to learn techniques and good practices that can be used to create different kinds of openEHR-compliant clinical data repositories.

Expected Attendees

Software Developers, Database Managers, Software Architects, Technical Leaders, Decision Support Specialists, and Business Intelligence Experts, participating in healthcare informatics projects or planning to.

Tutorial Speakers

Pablo Pazos Gutiérrez, Ingeniero en Computación (UdelaR); Clinical Informatics Consultant at CaboLabs Medical Informatics and Standards, Lecturer at Asociación Chilena de Informática en Salud (ACHISA), Coordinator of the openEHR-ES community, Qualified member of the openEHR Foundation - Localization Programme Member for Latin America; Montevideo, Uruguay.
Koray Atalag, MD, PhD, FACHI: Senior Research Fellow in Biomedical Informatics at Auckland Bioengineering Institute and National Institute for Health Innovation at the University of Auckland, New Zealand; openEHR Localisation Programme Leader; openEHR New Zealand representative.

Luis Marco-Ruiz, M.Sc.; Norwegian Center for Integrated Care and Telemedicine, University Hospital of North Norway; Department of Clinical Medicine, Faculty of Health Sciences, University of Tromsø; Qualified member of the openEHR Foundation; Tromsø, Norway.

Erik Sundvall, MSc, PhD - Researcher, Department of Biomedical Engineering, Linköping University and information architect in Region Östergötland, Sweden. Erik coordinates the openEHR software program: http://openehr.org/programs/software/

Sergio Miranda Freire, DSc, Associate Professor, Department of Health Information Technology Education, State University of Rio de Janeiro, Brazil.

Sergio and Erik have research interests in scalability and storage solutions for archetype-based EHR systems and have been testing and evaluating different noSQL storage and query approaches for openEHR data.

References


