Formal Design of Electronic Public Health Records

Diego M. LOPEZ and Bernd BLOBEL

eHealth Competence Center Regensburg, University of Regensburg Medical Center, Germany

Abstract. EHR systems have evolved from management of medical and patient records to the management of comprehensive health records including information about any observed health states e.g., social, economic and environmental conditions; and process such as public health surveillance, health promotion, prevention, education, etc. The paper discusses the analysis and design of Electronic Public Health Records (EPHR) according to the advanced state of knowledge in methodologies, models, techniques and tools for the specification of EPHR systems. A formal component-based architectural approach, based on internationally agreed terminologies, healthcare standards and software engineering de facto standards is presented.

Keywords: Computerized Medical Records Systems, Public Health, Software Design, HL7, Rational Unified Process, UML.

Introduction

For establishing efficient and high quality care of citizens, comprehensive and accurate information about states and processes directly and indirectly related to individual's health must be provided and managed. The Electronic Health Record (EHR) goes beyond patient's healthcare, recognizing the concept of “health” as a general term implying, social, economic and environmental conditions which influence individual and collective well-being.

From the health and medical informatics discipline point of view, the above multidisciplinary nature of EHR has not yet been completely accepted. In this context, the emerging discipline of Public Health Informatics (PHI) comes into scene. PHI recognizes the need for information architecture for Public Health Information Systems (PHIS), and it is driving some efforts in this direction. Currently, both disciplines are evolving separately, but the knowledge generated in the other, should be mutually acknowledged. The specification of EHR differing the Electronic Healthcare Record (EHCR) -oriented to healthcare systems-, and the Electronic Public Health Record (EPHR) -oriented to public health information systems- is a key for the integration/cooperation between the healthcare (clinical healthcare) and the public health domains.

1 Corresponding Author: Diego M. Lopez., ETSI Telecommunication. Universidade de Vigo. Vigo. Spain, Email: dmlopez@det.uvigo.es. The corresponding author is Ph.D. student affiliated to the University of Vigo and supported by the EU Programme Alban, identification number E03D16197CO and the University of Cauca under contract number 136 October 2003.
In the paper, the analysis and design of EPHR according to the advanced state of knowledge in methodologies, models, techniques and tools for the specification of EHCR systems is discussed. The approach is a formal component-based architectural methodology, based on internationally agreed terminologies, healthcare standards and also software engineering de facto standards.

1. **Formal Design of EHCR Systems**

Developing EHCR systems is not a simple task. From the software development team perspective, it requires the acknowledgement of advanced state of knowledge about methodologies, models, techniques and tools from systems analysis, software engineering, information sciences, etc. The software development process involves the participation of not only software developers, but also end-users, analysts, system integrators, project managers, etc., which have different views on the systems and the problem domain. It also requires a wide experience and knowledge about the healthcare “business” and the particularities of the healthcare domain. From the EHCR requirements perspective, advanced EHCR systems have to cope with the challenge of openness, scalability, flexibility, semantic interoperability, portability, distribution at Internet level and conformity with international standards. Furthermore, they have to be trustworthy and driven by the business process.

A component-based architecture is the most appropriate means for approaching the specification of EHCR systems with the aforementioned requirements. Using component-based architectures provides a number of solutions to the root causes of software development problems [1][2]:

1. Components facilitate resilient architectures.
2. Modularity enables a clear separation of concerns among elements of a system that are subject to change.
3. Controlling the iterative and incremental development of a system throughout its lifecycle and the different perspectives on the system from the involved stakeholders supports acceptability and usability.
4. Reuse is facilitated by leveraging standardized frameworks (such as COM+, CORBA, and EJB) and commercially available components.
5. Components provide a natural basis for configuration management.
6. Visual modeling tools provide automation for component-based development.

The Generic Component Model (GCM) [2] provides an architectural framework for designing advanced EHR systems, able to integrate EHCR-oriented systems to clinical care systems on one hand and EPHR-oriented systems to public health systems on the other. This abstraction model considers the systems in three dimensions, as shown in figure 1. The first dimension reduces the complexity of really inter-relating domains by separating them. In the figure, the model shows the separation between the healthcare domain and the public health domain, necessary for mutual interoperability and collaboration. The second dimension reduces the structural complexity of systems by decomposing them. As a result, the granularity of the system can be increased from business concepts over relationship networks and basic services/functions up to details such as basic concepts. The third dimension reflects the viewpoints of ISO 10746
2. Approaches on Designing EPHR

Only recently, with the emerging of the Public Health Informatics discipline, the need for information architecture for Public Health Information Systems was formally acknowledged. The IMIA National Agenda for Public Health Informatics [4] has recommended the definition of an information architecture that included longitudinal, person-based, integrated data repositories; similar to those defined by current EHR systems but somewhat broader including the requirements for public health activities.

In this direction, some related efforts are currently underway: the definition of the public health reporting domain in the HL7 v3 standard, the HL7 Special Interest Group on Public Health and Emergency Response (PHER SIG), and the respective Public Health Data Standards Consortium (PHDSC) Ad Hoc Task Force on Electronic Health Record - Public Health (EHR-PH).

2.1. HL7v3 Public Health Reporting Domain

The public health reporting domain information model (PORR_DM100001) is part of the HL7 v3 standard defined in the HL7 version 3 normative edition 2005 [5]. The PORR_DM100001 includes messages and documents that are specifically designed to support reporting and investigation in the public health context. The public health reporting domain is organized in three topics:

– Case investigation topic dealt with through case reporting: HL7/ANSI approved standard.
– Structured product labeling topic managed by structured product labeling, release 1: HL7/ANSI approved standard.
– Structured product labeling, release 1 implementation guide: HL7 approved informative document.

The use of this domain information models (DIM) in the specification of E-PHR systems is illustrated in section 4.

2.2. HL7 Interest Group: Public Health and Emergency Response (PHER SIG)

The PHER is an interest group within the HL7 Consortium and pursues the development of standards for the exchange of population-based health information, concretely messages for public health and emergency preparedness warrants. Specific areas of interest include event detection, outbreak investigation, human and animal population health monitoring, disease/condition case reporting, environmental observations related to health issues, emergency coordination, and legal issues including chain of custody and isolation/quarantine [6].

The PHER is currently in the process of being balloted towards a standard for the public health reporting domain: The investigation request topic (PORR_INVROO), under the umbrella of the HL7 Patient Care Technical Committee. The investigation request topic covers messaging requests for investigation of exposures and other public health events. It also covers communication of investigation subject demographics, exposure information, and data regarding the specifics of the public health event.

2.3. PHDSC Ad Hoc Task Force on Electronic Health Record-Public Health (EHR-PH)

The EHR-PH committee is part of the Public Health Data Standards Consortium [7], a non-for-profit organization founded by a confederation of public health agencies, professional associations, public and private sector organizations and individuals primary in the US, with the overall goal of promote data standards in public health. The purpose of the corresponding task force was to describe and evaluate the public health perspectives on the EHR initiatives, concretely on the HL7 EHR functional model [8]. The task force achieved an across-mapping between the HL7 EHR functions and the well-known core public health functions (assessment, policy and assurance). The cross-mapping demonstrated that, at high level of abstraction, the core public health functions are well represented in the HL7 EHR functional model. However, a more granular level of cross-mapping is needed to assure the ability of the model to support public health work and data flows. Some use cases and future projects are suggested in the initiative. Methodological approaches or real E-PHR systems have not been developed, however.
3. Formalizing the Design of E-PHR

A methodological approach for the analysis, design and implementation of E-PHR based on the GCM presented in section 3, is described in figure 2. The E-PHR system is specified according to the RM-ODP viewpoints. For each ODP perspective, a set of models is defined in order to describe, with different levels of granularity, the system’s components. The process of composition and decomposition of components is guided by the Rational Unified Process (RUP) [1]. The models to describe components are defined using models from the Unified Modeling Language UML 2.0 standard [9] and the HL7 Development Framework (HDF) specifications [10].

![Figure 2: The proposed modeling process](image)

The Rational Unified Process is a software development process that provides guidance about the “who”, “what”, “how” and “when” in the context of system’s components. The RUP covers the entire software development lifecycles by the definition of workflows. A workflow is configured by a set of roles (who), activities (how), and artifacts (what) necessary to transform the user requirements into a software system. The unified process identifies six core workflows (technical workflows): Business modeling, requirements, analysis and design, implementation; test, and deployment. To define how the process rolls out over time, the RUP identifies four iterative phases (inception, elaboration, construction, and transition) [1]. Each phase is implemented as a set of workflows where, according to the requirements of the EPHR system to be developed and the software development team involved; the number of iterations, roles, activities and artifacts are configured and extended. In the approach, all technical workflows are used in the different viewpoints but changing the emphasis depending on the viewpoint, the iteration and the phase of the RUP process. The emphasis in the enterprise viewpoint is on the business modeling workflow, both the Information viewpoint and the Computational viewpoint are stressed on the Requirements, Analysis and Design workflows, and the emphasis in the Engineering and Technology viewpoints is on the Implementation, Test and Deployment workflows.

System architecture can be represented as a set of artifacts describing the different levels of granularity of the system’s components. The modeling approach in Figure 2 describes the main artifacts to touch each of the RM-ODP viewpoints. UML is used
because it is the facto standard for the specification, visualization, and documentation of software systems models. In the approach, the UML Class Diagrams, Object Diagrams, Component Diagrams, Package Diagrams, Deployment Diagrams, Use Case Diagrams; Activity Diagrams and Sequence Diagrams are deployed.

Furthermore, the composition and decomposition of components is also supported by the models defined in the HDF specification. HDF is an emerging standard proposed by the HL7 Modeling and Methodology Technical Committee. It documents the processes, tools, actors, rules, and artifacts relevant to development of all HL7 standard specifications, not just messaging as its predecessor the Message Development Framework MDF. Our approach re-uses some of the HDF artifacts and HL7 version 3 specifications (RIM, D-MIM, R-MIM, CMET). Some activities and artifacts defined in HDF Chapter 2 (Requirements Gathering and Analysis) and HDF Chapter 3 (Modeling Analysis and Harmonization) will be used as reference for the design of the Static Domain Model, the Dynamic Domain Model and the Glossary in the Enterprise viewpoint. Furthermore, those activities and artifacts are used as reference for the design of the Models of static and views Models of dynamic views in the Information and Computational viewpoints as shown in Figure 2.

4. A Use Case in Public Health Surveillance

In this section, the specification of an EPHR system is illustrated by the development of a scenario in public health: the reporting of measles diseases. Measles accounts as one of most infectious diseases and one of the principal causes of Infant Mortality in developing countries. The requirements for this scenario are elicited from the WHO recommended standard for measles surveillance [11] and the Surveillance Program for communicable diseases (SIVIGILA) in Colombia [12]. The main models and artifacts in the Enterprise Viewpoint are detailed below.

4.1. Models in the Enterprise Viewpoint

The RM-ODP Enterprise Viewpoint offers a perspective on the EPHR system and its environment, describing the system’s purpose, scope and policies. In terms of the RUP workflows, it constitutes a Business Modeling process where the processes of the public health organization in charge of the reporting of measles diseases are modeled.

4.1.1. Scope Definition

The reporting of measles events flows through three main levels in a National health system: The local, regional (states or provinces) and national level (Federal level). Considering that the information flows are quite similar, the scenario is centered on the first one:

- The reporting from the Local Public Health Authorities to the Regional Health Authorities
4.1.2. Business Uses Cases

Figure 3 shows the high-level use cases for the selected scenario. The use cases Receive Case and Deliver Report are described below.

Use Case: Receive Case
Actor: Healthcare Provider
Description: This use case is initiated by an actor Healthcare Provider. The healthcare provider (hospital, clinic, health centre, health post, etc) in a district (city, county, municipality or town) sends to the Local Public Health Authority an aggregated report of clinical cases of measles in its institution. The reporting is achieved on monthly bases and the deadlines and reporting forms are defined in advance by the Local/Regional/National Public Health Authorities.

Use Case: Deliver Report
Involved actors: Regional Public Health Authority
Description: This use case is initiated by a Report Receiver. The Local Public Health Authority sends, monthly and after to a fixed deadline, a consolidated report of measles cases in its locality to the Regional Public Health Authority. The report contains aggregated data of number of cases during the last month.

4.1.3. Dynamic Domain Model

The Dynamic Domain Model visualizes the dynamic aspects of the business process that the E-PHR system supports. UML 2.0 activity diagrams are typically used for business process modeling. Figure 4 describes the activity diagram for the use case Deliver Report. In the diagram, the flow of activities and objects involved in the process of request, generate and send a consolidated measles report is detailed. An analogous Activity diagram is designed for the use case Receive Case.
4.1.4. Static Domain Model

The purpose of the Static Domain Model is to identify the key concepts in the domain. According to the Business Modeling workflow, it corresponds to the Business Object Model and is normally represented as a UML Class Diagram. The UML Class Diagram could be designed in a simpler way, based on the objects identified on the Activity diagram and the entities from the Use Cases descriptions. However, in the example, the enterprise concepts are harmonized with the classes (concepts) defined in the HL7 information models (D-MIM, R-MIM and CMET). This process is adapted from the HL7 normalization process proposed in HDF Chapter 3 (Modeling Analysis and Harmonization). The advantage of this earlier HL7 harmonization process is the facilitation of subsequent semantic interoperability. The process is developed as described in the following.

4.1.4.1. Initial identification of the domain concepts

Domain concepts are initially identified from the use cases, actors’ descriptions and the activity diagrams. Afterwards, the domain concepts are classified according to the RIM core classes (act, entity and role) in order to facilitate the further mapping to HL7 classes. The classification of the main concepts identified in the measles scenario is shown in table 1.
Table 1. Main concepts identified in the Measles scenario

<table>
<thead>
<tr>
<th>Identified Domain Concepts</th>
<th>Act</th>
<th>Entity</th>
<th>Role</th>
</tr>
</thead>
<tbody>
<tr>
<td>Health Care Provider</td>
<td></td>
<td>x</td>
<td></td>
</tr>
<tr>
<td>Local Public Health Authority</td>
<td></td>
<td>x</td>
<td></td>
</tr>
<tr>
<td>Regional Public Health Authority</td>
<td></td>
<td>x</td>
<td></td>
</tr>
<tr>
<td>Locality</td>
<td></td>
<td></td>
<td>x</td>
</tr>
<tr>
<td>Receive Report</td>
<td></td>
<td>x</td>
<td></td>
</tr>
<tr>
<td>Clinician</td>
<td></td>
<td></td>
<td>x</td>
</tr>
<tr>
<td>District</td>
<td></td>
<td></td>
<td>x</td>
</tr>
<tr>
<td>Reporting to the Local Public Health Authority</td>
<td></td>
<td>x</td>
<td></td>
</tr>
<tr>
<td>Generate Report</td>
<td></td>
<td></td>
<td>x</td>
</tr>
</tbody>
</table>

4.1.4.2. Identification of Reusable HL7 Specifications

The second step in the process of harmonizing the Domain Information Models (DIM) is to identify the HL7 information models which better describe the domain under specification. The HL7 information models are identified by comparing the system’s scope with those ones defined in the HL7 Standard specifications. The reusable HL7 information models to re-use are D-MIM, R-MIM and CMET.

The HL7 domain that better matches the scope of the Measles Reporting system is the Public Health Reporting Domain Information Model (PORR_DM100001), defined in the HL7 Version 3 normative edition 2005 [5]. The topic within the domain that best fits the E-PHR systems scope is the Case Investigation Topic (Case Reporting – HL7/ANSI approved standard).

4.1.4.3. Specialization of HL7 models

The selected case notification topic defines R-MIM: the Notifiable Condition Report RMIM (PORR_RM100001UV01). That R-MIM captures the information needed to support case reporting between different jurisdictional levels within the public health system. The last step in the harmonization process is the specialization (cloning) of classes from this R-MIM. For the sake of specialization, a cross-reference is created between the domain concepts in the scenario and the R-MIM. Each domain concept (based on the initial classification in Table 1) is matched against the R-MIM model elements. Candidate elements from the R-MIM model may include classes, attributes, relationships, datatype, datatype property, vocabulary domain, value set, or coded terms. Cross-reference MS Excel spreadsheets as proposed in HDF Chapter 3 are used.

The Domain Information Model is finally designed and refined using HL7 tooling set (R-MIM Designer). The R-MIM Designer allows to specialize (clone) the Notifiable Condition Report RMIM and to generate a Visio Model for the scenario. The resulting Domain Information Model for the reporting of measles diseases scenario is shown in figure 5.
4.1.5. Glossary

The glossary is obtained from the HL7 vocabulary defined in the RMIM (PORR_RMI000001UV01). To facilitate this task, the HL7 v3 standard documentation provides a navigable HTML documentation hyper-linked to the graphic representation of the HL7 models. The HL7 compliant vocabulary should facilitate the use of other HL7 models and further semantic interoperability.

5. Discussion and Conclusions

Designing EHR systems is a complex process that requires an extensive knowledge about the health domain but also about Information and Communication Technologies, essential to address the eHealth systems requirements of openness, scalability, flexibility, semantic interoperability, portability, distribution at Internet level and compliance of international standards.

A methodological approach for the analysis, design and implementation of EPHR has been shown. The design model uses the Generic Component Model as architectural framework, by differentiating the domain of the EPHR systems from the health Care domain; and abstracting their components according the RM-ODP viewpoints. For each ODP perspective, a set of models are defined in order to describe, with different levels of granularity, the system’s components. The process of composition and decomposition of components is guided by the Rational Unified Process and the models to describe those components are defined using models from the Unified Modeling Language UML 2.0 standard and the HL7 Development Framework specifications.

Information Architectures for public health Information Systems are also discussed. A critical need for the definition of information architectures and methodological approaches – similar to those defined by current EHR systems but somewhat broader including the requirements for public health activities – is identified.
Figure 5. Domain Information Model for the reporting of measles diseases scenario
6. Acknowledgments

The authors are in debt to their colleagues from HL7, ISO TC 215, CEN TC 251, and OMG/CORBA for kind co-operation.

References