Medical Document Modelling and Validation for the Latvian National Electronic Health Record System

Andrej Dubrovsky

Datorzinibu Centrs JSC, Lacplesha 41, Riga, Latvia

Abstract. In this article the author considers the case of mass code generation approach used for the analysis and implementation of Latvian Electronic Health Record System. The problem that appeared and was solved successfully is to validate the documents that are sent to the central storage system by a medical institution against the technology standard and the legal requirements.

Keywords: CDA, HL7, Latvian Electronic Health Record.

1 Project and Subject Description

This paper describes the experience of implementing the Latvian National EHR (Electronic Health Record) System in the time period from 2011 to 2013. The author's role in this project was a Project Manager from the supplier's side. The project itself was implemented in broad cooperation with other Latvian software service companies working as subcontractors. Another similar project was also run in parallel with this one by the Latvian National Health Service as part of the National e-Health Programme.

The e-Health programme in Latvia, similarly to other European countries [1], states as its goal the broad use of technology as a way to make healthcare more accessible for a patient and more efficient for the entire society. The way to achieve this target is to minimize work duplication and to reduce information delays and errors to allow doctors to get the most relevant information about a patient in real time or close to it.

The stated goals require a change in doctor-patient interaction business processes that are related to information storage and retrieval. The data, which are now recorded and stored on paper, are to be entered into the central storage system and should be retrievable by other practitioners or by a patient himself/herself. It is evident that this shift will require retraining and changing work habits of many general practitioners, as well as other doctors who do not use the Information and Communication Technologies (ICT) in their daily work. Fortunately, the trend in computer skills of doctors is in direction of greater use and understanding of technology.

Even when a doctor is ready to process medical information electronically, the question of interoperability remains. The medical landscape in Latvia consists of many independent private and public organizations, each with its own understanding

of information that should be stored and transmitted. They all have their own distinct business processes and no single format of data to be exchanged.

Some large hospitals and their IT providers create solutions to store and exchange medical data; however, those projects and undertakings remain regional and concentrate on a particular hospital and not the country as a whole.

2 Problems and Ideas

The main way of achieving the stated e-health medicine digitalization goals is the idea of centralized medical information storage, accessible both for all Latvian practitioners and for Latvian people domestically and cross-border.

Latvian e-health systems are organized into hierarchical architecture stack consisting of: a common medical secure messaging platform allowing for common authentication, authorization, auditing and secure messaging; business systems, namely EHR (Electronic Health Record) and e-Recipes; user-interface – health portal system or Health Service Provider IT systems.

In contrast with larger countries, all Latvian e-Health systems are centralized and not regional [2]. The centralized EHR allows for greater information coverage for all Latvians and doctors regardless of the location of a patient at the moment. By comparing a centralized solution with interconnected regional systems, it is easy to see the benefits in the area of maintainability and data consistency – because data fragmentation is impossible in this case. This solution has also its drawbacks as any centralized data storage is prone to the disaster risk. Therefore, care was taken in ensuring redundancy in system software and hardware to minimize this risk.

Technologically speaking, Latvian e-Health in general and EHR in particular are based on HL7 [3] messaging and document storage standard. This decision was taken after analysing the alternative approach – ES 13606 [4] standard. It was found that both standards meet the Latvian national EHR data storage and processing needs, but HL7 protocol stack is more mature and its technology support is better, allowing for faster and more reliable implementation of e-Health programme.

As in many other EHR systems, the main storable item in the Latvian EHR is a medical document created and signed by a practitioner. Documents that are stored in EHR are the electronic representation of traditional paper-based documents created during medical encounters and kept in practitioner's archives. HL7 standard defines the form of such an electronic document – it is CDA (Clinical Data Architecture) [5] standard. This is a format that enables the easy expandability and flexibility for implementers of EHR. However, technological steps must be taken to assure that medical document content is meaningful. Any EHR system should have tools to validate received documents to be correct according to business rules and legal requirements (for example, a discharge letter should have final clinical diagnosis in it, and it should conform to ICD-10 "International Statistical Classification of Diseases and Related Health Problems 10th Revision" [6] standard).

From the legal side, big part of Latvian medical documents is described with their information content in the single piece of legislation –Regulations of the Cabinet of Ministers No. 265, adopted on 4th April 2006 [7].

Different medical documents described in regulations reflect different areas of health business processes. Therefore, they have their own, unique internal structure tailored to a certain subject area. Historically, requirements for the documents were created one by one, and there was no single body governing the standardization process.

The result of this kind of organic growth – documents use many different classification schemes, and sometimes in different documents there are different classifications for the same data (even standardized as a diagnosis – not always ICD codes are used to record them).

One possible solution to this problem is to analyse all the documents, identify similar parts and then harmonize them by working out the unified data content. That would require adding or removing some of the data content to one or many documents in question, as well as unifying classification schemes used. In case when documents represent different areas of healthcare, content unification can be difficult. The difficulties are changing classification schemes on the established IT system as a prerequisite to filling in newly harmonized documents, as well as reflecting this change in legal framework, by altering the body of law. Both of these processes can be long and costly for all parties involved.

From the other hand, having done document reengineering like stated, there is a possibility of obtaining a unified document validation algorithm because individual document elements are validated using once-written software components. That allows reusing each created piece of validation code throughout many documents by using it wherever appropriate data are encountered in documents. For example, let us assume that there is a component able to validate the diagnosis structure. If all the documents that need to include a diagnosis use this exact structure, it will allow reusing the validation component for all of them. Thus, to assure the validation of all possible documents the library of validation components is needed. The library and the algorithm, which can select an appropriate validation component for the document fragment, can solve the problem.

3 Selected Solution and History of Its Implementation

At the initial stage of Latvian EHR project, the original approach to document modelling was to analyse the information structure of each document and divide it into pieces from which then reconstruct the harmonized documents for the use in e-Health project.

It quickly became obvious that there were certain problems with the approach stated above. Firstly, to assure that all documents are truly built of standardized components it would be necessary to have deep and broad cooperation with experts in the field of medicine, which was not readily available to the project to the extent required for that. Secondly, this approach would require changes in the regulations, which were better to be held minimal to minimize risks for the project. Then, there were problems with the Health Service Providers using in their IT systems the existing classifiers and document structures, all of which would need to be changed as a result of document harmonization efforts.

Considering all this, a decision was made to base the system on current documents used in Latvia. This approach removed the need for broad legal changes related to reengineered document solution. Changes in the existing business processes or classifiers used will still be needed, but in much smaller number. The need for industry experts and cooperation between interested parties was minimized, too.

The drawback of having a legacy document structure is a need to be able to validate each part of each document that multiplies the amount of modeller's work because there is much less overlap in document structure in this case. A more complex document model will require a more validating code, and, naturally, more development and testing effort.

It should be mentioned here that the increase in model complexity is 'technical' – in the sense that data needed to create the model are already available in the existing documentation and from subject area experts. There is little new development needed in the business analysis. On the other hand, there are more efforts of 'technical' system analysis, but those two effects are at least compensated one another.

There are other areas of software development where the decision to capture situation 'as is' causes problems to be overcome. The validation code that is to be written and tested is at least 2–3 times bigger than in the case when similarities in the structure of documents are found and exploited. Taking into account that software development is nonlinear due to limited interpersonal communication throughput, this amount of code increase bears a risk of project being over its budget and schedule.

Due to challenges mentioned above, it was decided that the only way to create all the required medical document models and achieve the necessary quality level for the validation components was to focus on the code generation. There were two possible open source technologies considered for this purpose – namely Mohawk college HL7 framework [8] and MDHT eclipse IDE medical documentation modelling extension [9].

The MDHT toolset was chosen due to better overall readiness for the task size we had, and broader project support from its community. It is written in Java language, while the Latvian EHR project uses .NET stack. However, the integration required was simple considering the amount of work saved by using mass code generation both for the analysis and development/debugging.

As a result of decisions made, the modelling process of medical documents is as follows:

1. The document in question is analysed using the existing legal base. Necessary improvements are discussed with the National Health Service and documented as additional requirements.
2. The model of document is made, presented to a customer and agreed on. For this step we used regular UML class diagrams, having for each HL7 document element its own requirement model with all data types and constraints.
3. When the model documentation (which contains all the structure elements of future document and all the constraints that will be needed for it) is accepted by a customer, the physical model of document is developed. Physical model reflects the business model into MDHT environment.
4. When the physical document model is developed in MDHT, its validation code is generated into common *.jar file that contains all the other document validators. This *.jar executable is then deployed into the EHR background document

processor, and is also a redistributable software component, available for medical software vendors who will be developing interfaces with national EHR.

5. The last step in the process is the feedback loop of testing the resulting validator component with respect to the business model and also initial legal and business requirements. The bugs found are processed at a stage where they are encountered – modifying a business or physical model, then regenerating the validation code and redeploying into the solution.

4 Lessons Learned, Ways to Improve

In general, the code generation approach selected is success in the sense that it has allowed implementing the validation code for documents required by a customer within time and budgetary bounds.

It is possible to estimate the gain of using the code generation approach for document validators. From the industry experience, we can assume that waterfall project coding percentage is 40% of the overall effort. By replacing coding with generating, it is possible to reduce this percentage to ~10% of the overall effort (so 30% economy). The other effect of this approach is the reduction of debugging because our experience has shown that the main error source is the model itself but not the implementation process. It is difficult to numerically estimate the reduction of debugging, but comparing to the past experience and other parts of the project where code generation was not used, this was up to ~50% of debugging effort, which is ~10% of overall effort. Thus, the total estimated effect on the budget is at least 40% reduction in overall effort needed to implement validator part.

Another less tangible effect is the reduction of business analysis. From the project management point of view, this allows reducing uncertainty and improving planning. From the other hand, the use of this technology discouraged potential streamlining of medical record keeping in Latvia, as there are many documents in need of restructuring or being dropped altogether.

A better solution would be to choose a hybrid approach; namely, first it would be necessary to select only the most necessary documents, then to standardize their content and to generate the validation code. This would not only ensure that all relevant medical data are recorded in the central database, but also the digitalization of health industry is used as an opportunity to improve efficiency by reengineering some healthcare business processes.

Another possible area of improvement lies in the overall e-health programme architecture. All work was contracted in large chunks, which required long and complex business and system analyses in the initial phases of projects with few opportunities to correct the possible errors or pursue opportunities uncovered.

A better approach here would be to go the agile way – contracting smaller amounts of work as appropriate. This would bring benefits such as speeding up systems to the market; allowing for test-drive solutions and easier change of the course in the middle of a project. Agile approach is, of course, compatible with the document modelling approach we selected – as individual documents are not strongly connected with one

another; they are practically independent and their modelling can be made in stages without a tangible impact on the project as a whole.

5 Conclusion

In the course of the project it has become apparent that document validation code generation is a viable software development practice, if used within its limitations. In case, when the subject area permits the use of code generation, the resulting economy can justify additional expenses for both learning and adopting the technology and for coping with its limitations. However, by allowing capturing and retaining the status quo, the technology removes one of the drivers to change and improve business processes. It does not prevent a change by itself, but leaves an option for business-as-usual as an alternative to those parties involved in the decision-making process.

References

1. eHealth Action Plan 2012-2020 - Innovative healthcare for the 21st century, http://ec.europa.eu/health/ehealth/docs/com_2012_736_en.pdf
2. European countries on their journey towards national eHealth infrastructures (Final European progress report), http://ehealth-strategies.eu/index.htm
3. Health Level Seven International, http://www.hl7.org/
4. The CEN/ISO EN13606 standard, http://www.en13606.org/the-ceniso-en13606-standard
5. CDA® Release 2, http://www.hl7.org/implement/standards/product_brief.cfm?product_id=7
6. International Statistical Classification of Diseases and Related Health Problems 10th Revision, http://www.who.int/classifications/icd/en/
7. Medicīnisko dokumentu lietvedības kārtība, http://likumi.lv/doc.php?id=132359
8. Model-Driven Health Tools, https://www.projects.openhealthtools.org/sf/projects/mdht/
9. The Everest Framework, http://www.mohawkcollege.ca/Assets/ideaworks/case-studies/eHealth+-+MARC+Everest+Framework.pdf