Building Continents of Knowledge in Oceans of Data: The Future of Co-Created eHealth
A. Ugon et al. (Eds.)

doi:10.3233/978-1-61499-852-5-860

Data Modeling Challenges of Advanced Interoperability

Bernd BLOBEL[a,b,c,1], Frank OEMIG[d] and Pekka RUOTSALAINEN[e]
[a] *Medical Faculty, University of Regensburg, Germany*
[b] *eHealth Competence Center Bavaria, Deggendorf Institute of Technology, Germany*
[c] *First Medical Faculty, Charles University Prague, Czech Republic*
[d] *Deutsche Telekom Healthcare and Security Solutions GmbH, Essen, Germany*
[e] *School for Information Science, University of Tampere, Finland*

Abstract. Progressive health paradigms, involving many different disciplines and combining multiple policy domains, requires advanced interoperability solutions. This results in special challenges for modeling health systems. The paper discusses classification systems for data models and enterprise business architectures and compares them with the ISO Reference Architecture. On that basis, existing definitions, specifications and standards of data models for interoperability are evaluated and their limitations are discussed. Amendments to correctly use those models and to better meet the aforementioned challenges are offered.

Keywords. Healthcare transformation, interoperability, data models, knowledge management, architectures

Introduction

Healthcare systems around the world are on the move towards highly distributed, personalized, predictive, preventive, participative, and cognitive care. Such approach requires the involvement of many sovereign stakeholders from different policy domains, representing different disciplines, using different methodologies, terminologies, and ontologies, offering different levels of knowledge, skills, and experiences to act in different scenarios accommodating different business cases in multiple businesses. Such business systems set big challenges on analysis, design, implementation, maintenance, and evaluation within the systems' lifecycle. The management of such highly dynamic, complex, heterogeneous and context-depending business processes, i.e. the execution of IT-supported business operations from a business process expert's view, must be formalized [1,2] to enable automation of the business process management. A system-oriented, architecture-centric, ontology-based modeling approach based on ontology languages, repositories, reasoners, and query languages provides scalable and adaptive methods and tools for machine-accessible representation and manipulation of business knowledge [1]. Such approach has been developed by the authors and standardized at ISO and CEN [3,4]. Dealing with the data modeling challenge for interoperability, this paper introduces data model classification systems to analyze widely spread data model

[1] Corresponding Author. Prof. Dr. habil. Bernd Blobel, FACMI, FACHI, FHL7, FEFMI, MIAHSI; Medical Faculty, University of Regensburg, Germany; Email: bernd.blobel@klinik.uni-regensburg.de

based interoperability specifications in comparison with the ISO Interoperability Reference Architecture Model [4].

1. Methods

According to Alter [5], a model is a partial representation of reality. It is restricted to attributes the modeler is interested in. Defining the pragmatic aspect of a model, the interest is depending on the addressed audience, the reason and the purpose of modelling the reality and using the resulting model for a certain purpose and for a certain time instead of the original. Langhorst et al. [2] defined a model as an unambiguous, abstract conception of some parts or aspects of the real world corresponding to the modeling goals. Hereby, the domain of discourse, the business objectives, and the stakeholders involved have to be defined. The relevant stakeholders define the provided view of the model as well as the way of structuring and naming the concepts of the problem space. First capturing key concepts and key relations at a high level of abstraction, different abstraction levels should be used iteratively, where the first iteration is performed in a top-down manner to guarantee the conceptual integrity of the model. This requires meeting design principles such as orthogonality, generality, parsimony, and propriety [6].

Data modeling is frequently described as a series of processes to define data requirements for supporting business processes by enabling all related process decisions, so defining the system behavior to meet the business objectives. Depending on the level of abstraction, we distinguish conceptual, logical and physical data definition representing the informational components of the considered ecosystem [7]. Especially for managing complex multi-domain ecosystems, the definition of business cases and involved assets including a comprehensive metadata repository and accurate quantifiers as well as data governance management is impossible without deploying the business domains’ ontologies [8].

2. Modeling Health Systems

According to Hoberman et al., a data model is a visual representation of people, places and things of interest to a business, and is composed of a set of symbols that communicate concepts and their business rules [9]. For data modeling enabling advanced interoperability in distributed multi-domain healthcare systems, we follow Hoberman et al. [9] in their four levels approach. Starting point is the definition of the business, thereby aligning its scope and the common interest of the different stakeholders from different domains involved. The resulting very-high-level data model represents scope, requirements and related basic concepts of the business case. The high-level data model defines the relevant information and the representation and relationships of the basic concepts. The logical level data model describes in more detail the layout and types of the data as well as the object relationships. At this level, data modelers and analysts enter the stage, while the former levels are accommodated by domain experts. However, for properly managing data governance as discussed later on, business domain experts should be involved throughout the project lifecycle. The physical level data model considers ICT paradigms and related platforms, addressing implementation-related aspects relevant for storing, processing and communicating information such as architectures and principles of relational versus non-relational databases, communication protocols, Web services, representation styles, etc.

Another approach for interrelating the different model levels uses the dimension of modeling from the 1-dimensional data modeling through information modeling, knowledge modeling up to the four-dimensional knowledge space representation [10], allowing for transformation between the different representation levels. The knowledge dimension covers the knowledge of one domain. The knowledge space dimension represents multiple domains' concepts and their relations, so enabling their mapping. The higher the dimension the more the modeling process is dominated by business domain experts. Figure 1 presents the modeling dimensions and the related transformation pathway.

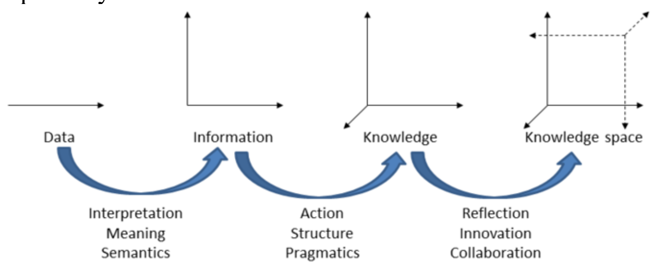

Figure 1. Dimensions of data modeling (after Krogstie [10])

3. Results

In [4], different interoperability levels from technical through structural, syntactic, semantic, service interoperability knowledge-based to skills-based interoperability are defined. The HL7 V2 EDI protocol, but also HL7 V2/V3 Implementable Technical Specification (ITS) as well as specifications of the observational health data initiatives OHDSI and OMOP define data structure and related data types at the physical data model level, addressing the modeling dimension of the 1-dimensional data approach. With HL7 V3, following HL7 Development Framework (HDF), the HL7 Reference Information Model (RIM) – also standardized at ISO as ISO/HL7 21731– has been defined [11]. That way, business case related data exchange via messaging, documents or services was defined, using ICT ontologies and therefore ICT concepts to reflect the business case. The related data model level is the logical one, considering the modeling dimension perspective of the 2-dimensional information approach. When representing the business concepts deploying the knowledge and methodologies of the involved domain experts expressed using their terminologies and ontologies, the high-level data model (or in the three level metrics the conceptual data model) must be exploited. Regarding the modeling dimension, the 3-dimensional knowledge model applies here. The challenge of advanced interoperability for personalized, preventive, predictive, participative and cognitive care and precision medicine can only be managed by very-high-level data models, or the 4-dimensional knowledge space modeling approach, respectively. The four stages modeling dimensions roughly correspond to the modeling levels and their relations to specs as presented in Table 1.

As stated both in [6] and in [9], the described top down approach is inevitable when developing new, complex and interoperable health systems solutions. When adopting solutions within a well-defined business framework, a combination of top down and bottom up modeling processes is possible. The importance of ontologies has been declared in many papers. However, some just refer to the IT part of the interoperability, so addressing the ontology stuff just with IT ontologies such as the Web Services Modeling Ontology (WSMO) [1]. Table 1 summarizes the described data model levels [9] and the dimensions of modeling [10] in relation to the system-oriented, architecture-centric, ontology-based, policy-driven ISO Interoperability Reference Architecture Model [4] with its different model viewpoints. In the rightmost column, some sample standards and their association with the corresponding level or view is presented. Starting with platform specific specifications at the physical data model level, most of the so-called "higher level" standards must be placed on the 2nd level. Only a few reflect the conceptual level of business and domain knowledge to reach the 3rd data model level such as Detailed Clinical Models (DCM) [12] or the Communication Standards Ontology (CSO) [3]. Currently, just the ISO/CEN Interoperability Reference Architecture Model and standards including it fulfill the 4th level requirements, covering all modeling levels and dimensions.

Table 1. Comparing Data Model Levels [9], Dimensions of Modeling [10], and the ISO Interoperability Reference Architecture Model [4], applied to specification examples

Data Model Level	Modeling Actors	Model Scope	Dimension of Modeling	Interop. Reference Architecture	Examples	
Very-high-level data model	Business domains stakeholders	Scope, requirements and related basic concepts of business case	Knowledge space	Business View		ISO/CEN Interoperability Reference Architecture
High-level data model	Business domains stakeholders	Relevant information and representation & relationships of basic concepts	Knowledge	Enterprise View	DCM, CSO	
Logical data model	Data modelers and analysts	Layout & types of data and object relationships	Information	Information View	HL7 V3 (CMETs), HL7 CIMI, openEHR Archetypes	
				Comp. View	HL7 FHIR	
Physical data model	Data modelers and developers	Implementation-related and platform-specific aspects	Data	Engineering View		
				Technology View	HL7 V2/V3 ITS, SQL, OHDSI, OMOP	

4. Discussion and Conclusions

Despite the definition and standardization of architecture models for enabling advanced interoperability [4], many standards and specifications still rely on data models for managing that challenge, however ignoring or even incorrectly claiming to overcome the related limitations demonstrated in this paper. This does not just apply to the aforementioned specifications such as the RIM-based solutions, but is also a concern in managing clinical models such as the HL7 CIMI approach [12]. For more information see, e.g., [13, 14]. Not just the presented classification systems, but also standard modeling conventions and data modeling best practices advise in overcoming the problems in data modeling and data governance management. The data modeling best practices [7] require getting the right people timely and properly involved in defining requirements. Furthermore, appropriate metadata must be recorded including core definitional qualities from physical attributes in the database or communication protocol context through any type of policies up to business terminology and business process management. Third, also the business understanding must be harmonized. That way, data modeling is a form of data governance from the definition through the production and the usage of data [7]. The data use includes risk management by protecting sensitive information and managing compliance. Details around data governance will be managed in another paper in preparation. All those data modeling best practices address more or less business domain experts and only partially information scientists, who currently wrongly dominate the process. To enable business process management and related decision support, the crucial level of data modeling is the very-high-level data model, equivalent to the 4-dimensional modeling process. Thus, the performed analysis justifies the interoperability approach of a system theoretical, architecture centric, domains ontology based and policy driven model [4] as approved by ISO TC 215 and CEN TC 251 and realized or in process in ISO 13606 and ISO 12967 [15, 16]. Other specs will follow soon.

Acknowledgement

The authors are indebted to thank their colleagues from HL7, ISO TC 215 and CEN TC 251 for their kind and constructive support and cooperation.

References

[1] Hepp M, Leymann F, Bussler C, Domingue J, Wahler A, and Fensel D. Semantic business process management: a vision towards using semantic web services for business process management. In: IEEE 10th International Conference on e Business Engineering (2005), pp 535-540.
[2] Lankhorst M, et al., Enterprise Architecture at Work. The Enterprise Engineering Series. Berlin Heidelberg: Springer- Verlag; 2009.
[3] Blobel B, Oemig F. The Importance of Architectures for Interoperability. Stud Health Technol Inform 2015;211:18-56.
[4] Blobel B. Standardization for Mastering Healthcare Transformation – Challenges and Solutions. European Journal for Biomedical Informatics, 2017; 13,1:9-15.
[5] Alter S. Information Systems – A Management Perspective. Reading: Addison-Wesley Publishing Company; 1992.
[6] Rebstock M, Fengel J, Paulheim H. Ontologies-Based Business Integration, Springer-Verlag, Berlin Heidelberg, 2008.
[7] Seiner RS. Data Modeling is a Form of Data Governance. IDERA 2017.
[8] Harper J. Lessons learned from big interoperability in government. KMWorld November/December 2017;26,9.
[9] Hoberman S, Burbank D and Bradley C. Data Modeling for the Business: A Handbook for Aligning the Business with IT using High-Level Data Models. Bradley Beach, NJ: Technics Publications, LLC.; 2009.
[10] Krogstie J. Business Information Systems Utilizing the Future Internet. LNBIP 2011;90:1-18.
[11] HL7 International Inc. HL7 Reference Information Model v. 2.47. http://www.hl7.org/implement/standards/rim.cfm
[12] Blobel B, Goossen W, Brochhausen M. Clinical modeling – A critical analysis. International Journal of Medical Informatics 2014;83,1:57-69.
[13] Goossen WTF. Detailed clinical models: Representing knowledge, data and semantics in healthcare information technology. Healthcare Informatics Research 2014; 20,3: 163-172.
[14] Goossen W, Gossen-Baremany A, Van Der Zel M. Detailed clinical models: A review. Healthcare Informatics Research 2010; 16,4: 201-2014.
[15] International Organization for Standardization. ISO 13606-1:2017 Health informatics. EHR communication Reference Architecture. Geneva: ISO; 2017.
[16] International Organization for Standardization. ISO 12967:2018. Health informatics. Health informatics service architecture. Geneva: ISO; 2018.