pHealth 2020
B. Blobel et al. (Eds.)

doi:10.3233/SHTI200617

The Value of Domain Information Models for Achieving Interoperability

Frank OEMIG[a1] and Bernd BLOBEL[bcd]
[a] *Deutsche Telekom Healthcare and Security Solutions GmbH, Bonn, Germany*
[b] *Medical Faculty, University of Regensburg, Germany*
[c] *eHealth Competence Center Bavaria, Deggendorf Institute of Technology, Germany*
[d] *First Medical Faculty, Charles University Prague, Czech Republic*

Abstract. The value of data models in general and information models in specific has been evaluated by many scientific papers. UML as one modelling notation has documented its value as a foundation for precise specifications. Analyzing implementation guides for data exchange, they rarely include or are based on information models but simple data sets, if at all, as simple technical representation thereof. This paper wants to argue in favor of information models as a basis for creating interoperability specifications using a quite simple example and to include – or at least reference – them when providing implementation guides. The reader is invited to transfer this example to even more complex scenarios.

Keywords. Information models, data sets, interoperability, data exchange, FHIR

Introduction

The Keynote published in this volume highlights the importance of cross-domain knowledge sharing for advanced interoperability [1]. Understanding the business system and the concerns to be addressed is the pre-requisite of any end user communication and cooperation for meeting intended business objectives and therefore starting point in the system development process. Performing the necessary communication and cooperation requires a practical implementation sharing and processing related data. So, this paper focuses on the data sharing interoperability paradigm based on data sets, which should be founded on information models and lead to implementable technical representations.

To determine, whether interoperability has been established and achieved, exchange of data must be established first. For that purpose, implementable technology specifications as technical models must be provided. This paper is asking for the type of base specifications allowing for interoperability. Frequently, the opinion dominates that information models, i.e. "a representation of concepts and the relationships, constraints rules, and operations to specify data semantics." [2], as logical models are too complicated and therefore unnecessary for data exchange, and requirements of healthcare providers are adequately covered by data sets. Therefore, most implementation guides for data exchange only provides hierarchic data element lists – named data sets (Figure 1) – leaving out proper information models.

[1] Corresponding Author, Frank Oemig, DTHS, Friedrich-Ebert-Allee 140, 53113 Bonn, Germany; Email: Frank.Oemig@t-systems.com.

This paper does not evaluate the value of information models or the way they are represented in general, but just discusses them as basis for providing data exchange specifications. From the authors' experience, quite a lot of projects start their definitions with data sets that are collected on the fly. In some cases, the data collection is performed directly in form of XML schemas. This paper explains the need of information models and their importance for interoperability, if data is extracted from transmissions (like messages or documents), facilitating the loss of information as a simple example.

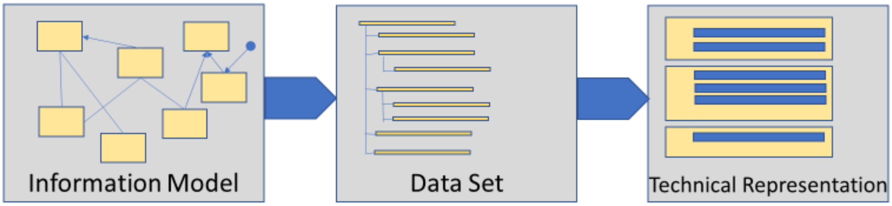

Figure 1. Derivation Process

1. Goal

In 1987, Health Level Seven (HL7) [3] started with data exchange specifications at application level [4] independent of implementation details [5], implicitly incorporating information models. Even in such a scenario, interoperability requires more than just identifying and accessing information such as "birth date" in a record, especially when considering different levels of interoperability [1, 6, 7]. Information contexts within and across systems must be included. Interoperability also happens between specifications when reusing parts. In [8], the interoperability question was answered according to Figure 2. The graph exemplifies that data exchange involves more than just two information and communication (IT) systems. Instead, interoperability of specifications is an essential and integral part. Spanning data exchange across multiple systems increases the complexity of the problem space. Most people will remember the children's game "Chinese whispers" that is an appropriate analogy for this problem.

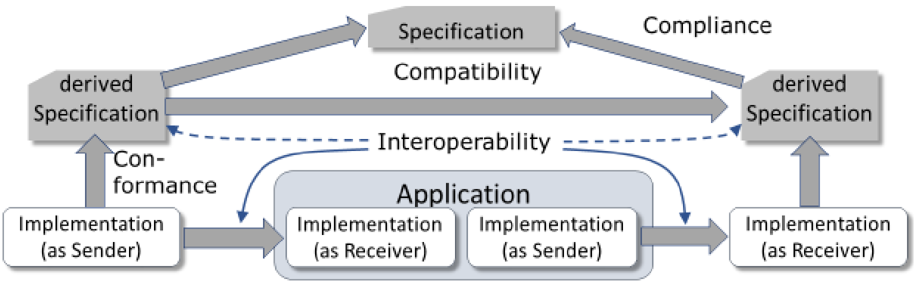

Figure 2. Definition of Terms

Usually a lot of freedom is provided in the form of optional elements that are not always implemented in the same way. These optionalities can be eliminated, but diminish the acceptance, and the specifications can no longer be used universally in different use cases and contexts, which contradicts the aforementioned goal of re-use. From this point of view, optionalities are an essential part of a good specification. In Figures 2 and 3, the dependencies of the specifications are symbolized by thick vertical arrows, which represent corresponding refinements and additional constraints. This vertical hierarchy can be used to slowly introduce further requirements allowing for a broader set of use cases. Horizontally oriented arrows express the constraints for conveying the information. Compatibility must apply in order to create interoperability (white arrows in Figure 3). The reuse of components of a specification simplifies reaching the goal [9].

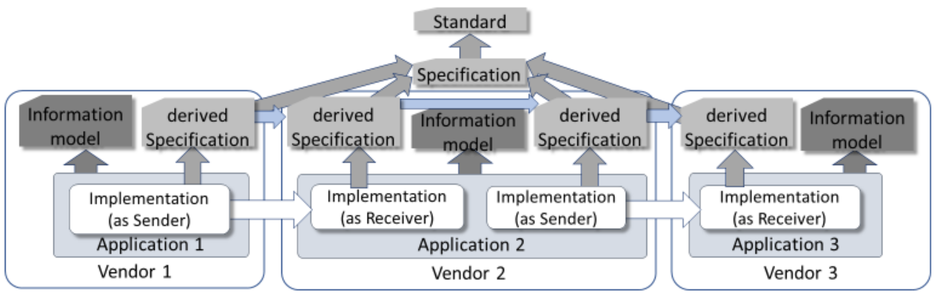

Figure 3. Using Models and Specifications

Despite the aforementioned insights, applications are usually developed for a specific purpose based on use-case- and domain-specific information models (dark grey), which are unfortunately not always explicitly presented. Usually, those information systems are developed independently of any data exchange specification. This is especially true for form-oriented applications that focus on user-oriented frontends supported by an appropriate toolbox. But even then, an information model can be extracted from a form that documents the relationships between the individual data elements. For example, a procedure could be justified by a reason as a reference to one or more coded diagnoses. Whether this relationship refers to one or more diagnoses in the explanatory statement has implications for the interoperability of applications. In Figure 4, these facts are expressed by two different information models which we will use for our examination. Each manufacturer could present such a model for its application, even if it is not explicitly documented. In summary, an application satisfies its own information model, while the data exchange must obey a foreign specification provided by somebody else. So, a conflict is pre-programmed.

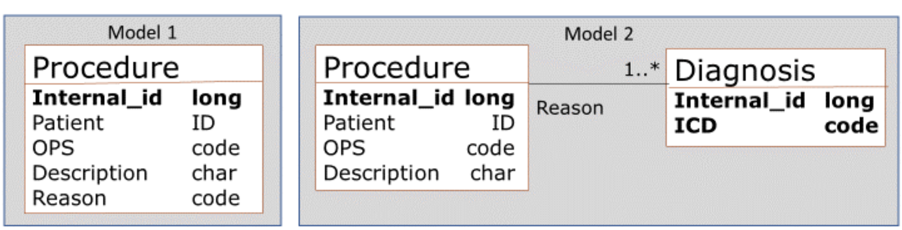

Figure 4. Example Models (1 + 2)

2. Data Exchange

Let us take the two models from Figure 4 as an example. Both models allow for recording a (coded) reason, but different in their approach. Technically, both can be equally well motivated depending on external requirements. Therefore, there is no argument to prefer one above the other.

The analysis regarding loss of information in three communicating systems performed by possible combinations of the example models presented in Figure 4 (left column of Tab.1) in terms of missing general interoperability is shown in the middle column of Table 1, given that backwards communication of the received information does not count.

Table 1. Interoperability among Models

Model Combination	Interoperability	Reason
1-1-1	Yes	All systems are identical.
1-1-2	Yes	System 3 as recipient allows more than both senders. Changes in system 3 become problematic when communicating back.
1-2-1	Depends	System 2 as recipient allows for more than system 1. No problems if no changes in system 2.
1-2-2	Yes	Systems 2 and 3 allow for more than system 1, so that a change in system 2 is unproblematic. Changes become problematic when returning data.
2-1-1	No	System 1 allows for more reasons, which may be eliminated in system 2.
2-1-2	No	See 2-1-1. System 3 may no equalize any more.
2-2-1	No	See 2-1-1.
2-2-2	**Yes**	All systems are identical.

If data exchange is considered between these systems in terms of interoperability, thereby providing an information model as well, the result slightly changes. In Tab.2 the data exchange for different model combinations (left column) is enforced according to exchange specifications obeying to specific models (right-most columns). A possible second assessment (after slash) considers modifying the data in the second system before forwarding it to the third.

Table 2. Interoperability among Models based on a data exchange format

Model Combination \ Data Exchange	1-1	1-2	2-1	2-2
1-1-1	Yes (1)	Yes (2)	Yes (2)	Yes (2)
1-1-2	Yes (2)	Yes (2)	Yes (4)	Yes
1-2-1	Yes (2, 3)	Yes (2, 3) / No (5)	Yes (3, 7)	**No** (5)
1-2-2	Yes (2, 3)	Yes (2, 3) / Yes	Yes (3, 7)	Yes
2-1-1	No (6)	No (6)	No (8)	No (8)
2-1-2	No (6)	No (6)	No (8)	No (8)
2-2-1	No (6)	No (6)	No (5)	No (5)
2-2-2	**No** (6)	**No** (6)	**No** (9)	Yes

Notes (in brackets):
1. All models are the same.
2. No loss because no system can do more than the data exchange specification requires.
3. The systems do not make any internal changes to the data.
4. The first system cannot do more than finally needed.
5. Data loss in the third system.
6. Data loss when transferred from the first system.
7. The exchange specification allows more than is used by System 1. The second specification restricts the exchange of data to what the first system has already supplied.
8. Data loss in the second system.
9. Data loss during transfer to the third system.

Despite the existing general interoperability of the individual systems (bolded in Tab.1), interoperable data exchange is not possible (bolded in Tab.2), if the exchange specification does not fit. If further systems are incorporated or multiple conformance constructs as well as semantic misinterpretations are considered, interoperability declines further.

3. Use of Information Models

How can information models be used to establish interoperability between applications in combination with a data exchange format which requires a hierarchic, acyclic representation allowing for sequentialization? Figure 5 provides as an example a class diagram that represents data elements to be considered for oncological data exchange [10].

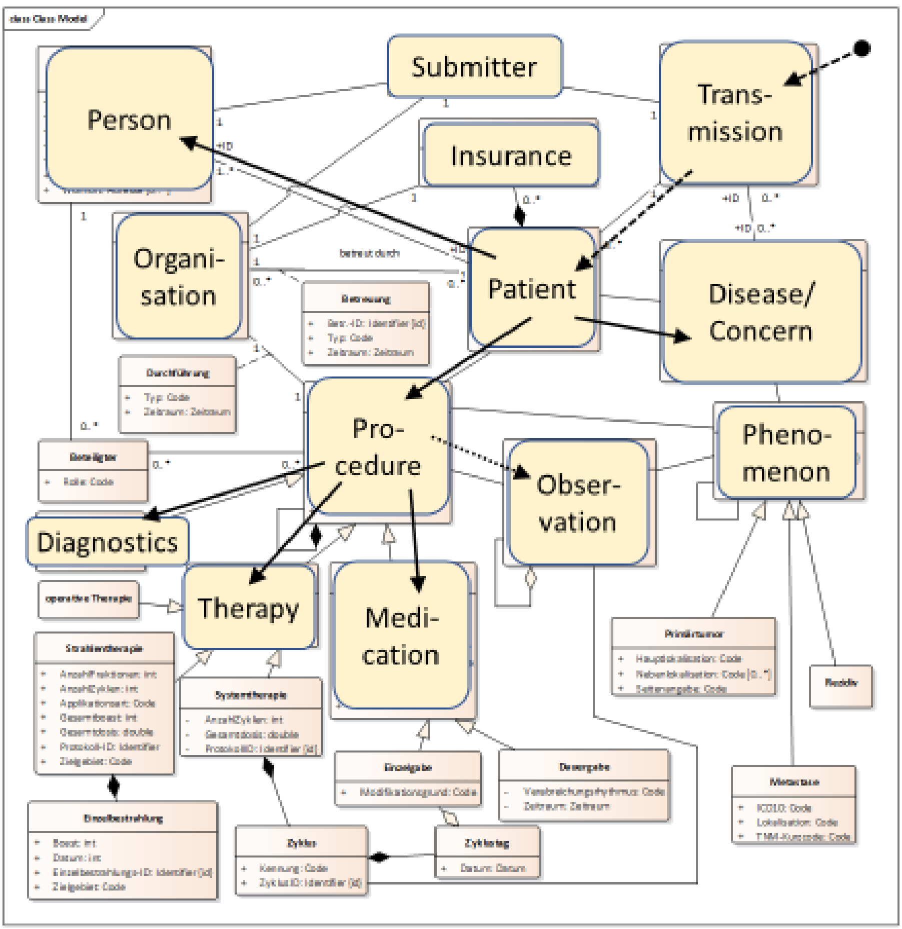

Figure 5. Example Information Model (from oncological data exchange)

In our example, the model is intended to establish the links between administrative information and diseases and their treatment in form of observations. (According to [1] this model should be split into different domains but that is not done for the sake of simplicity. Also, the specific details within that example are of minor importance and thus unreadability is intentional.) This example shows that a result ("observation") can only exist – or accessed – through a procedure (dotted arrow)—or vice versa, because a result without linking it to a procedure does not make sense. Every data exchange has to consider this issue.

At the same time, it determines the type of relationship of those two elements. Deriving a data exchange specification from this model is relatively easy because only the existing paths need to be traversed like it is done with the black arrows. In our example, there is only one entry point for a message (dashed arrow), enforcing the assignment to a patient as the most important subsequent relationship.

Traversing the structure passing the patient a decision must be made whether this data exchange is for an administrative notification or transmission of special clinical information, such as reporting a disease or even carrying out a treatment, which then splits into an examination or therapy (solid arrows). This differentiation must be reflected in specific message bodies which is mirrored in accompanying message's metadata. This allows for introducing integrity checks that are not possible with a simple hierarchical specification. This way a specific walk-thru results in a hierarchic structure that can be used for transmission. In other words, other entry points with different walk-thrus will result in different hierarchic structures obeying the same requirements. Hierarchic structures without an underlying information model will not allow for deriving (reconstructing) the original relationships.

4. Data Sets

Different paths result in different hierarchies, which also include different levels of detail depending on the objective. Thus, for a reference a simple identification is sufficient. An update of details requires complete information, which does not have to be specified repeatedly. From this point of view, the so-called data sets are an extraction from an information model (left-hand side of Fig.1), which are compiled for a specific exchange purpose. Thus, data sets are the link between an overarching specification that enables interoperability and a precise implementable specification. Given our simple example for loss of information above, application architects have to implement the information model, whereas message designers have to extract the information in focus.

HL7 has already demonstrated these dependencies through the Service-Aware Interoperability Framework [11] and the Enterprise Compliance and Conformance Framework [12], which are based on the ISO Reference Model for Open Distributed Processing (RM-ODP) [13].

5. Representing Information Models

Information models can be used in the examined form preserving the structures (e.g. Figure 4). This is especially preferable for larger or more complex models. For smaller models like blood pressure, it is possible to pre-coordinate the associated information and combine them into a special concept collection (Figure 6).

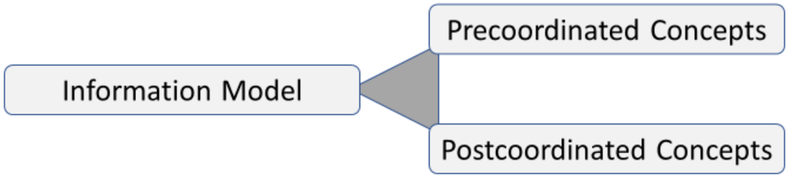

Figure 6. Equivalence of Information Models

In this way, these can be labelled meaningfully, so that they become directly and easily usable as "value sets". This is an approach that applies to LOINC [14]. Alternatively, the various aspects can be post-coordinated by appropriate coding systems and grammar. The latter can be found with Snomed CT [15]. The preferred way depends on the use case and the data exchange standard. On the positive side, HL7 version 2.x, Version 3/CDA® [16] and FHIR® [17] stand out supporting all variants.

6. A word about FHIR

FHIR is a logical and consistent step forward from an XML or JSON perspective. The integration of RESTful Services, the FHIR Mapping Language with FHIRpath, FHIR Cast, CDS Hooks [18], SMART Health IT on FHIR [19] and referencing information units is advantageous, so that a total harmonization takes place and new system architectures are made possible. Nevertheless. it is also a step backwards in terms of content and standardization: the shift of basic requirements from the standard to implementation guides is a double-edged sword since it increases acceptance while simultaneously softening conformance constructs. Reverse engineering of various implementation guides tries to reapply uniformity, if this is (still) possible. At least, the hidden complexity is rediscovered and made visible in HL7 Version 3 similar nomenclatures [16]. The so-called Structure Definitions in FHIR are an excellent basis for progressing from a data set based on a uniform technology and common tools, but it is not a substitute for information models for the aforementioned reasons. The FHIR Graph Definitions will go into that direction, but it will be very challenging to introduce its value afterwards.

7. Summary

The basics outlined above demonstrate that domain information models are necessary to achieve interoperability and thus must be consensually developed in order to get sustainable implementation guides. Mind maps or data sets as well as FHIR profiles are lacking the necessary understanding of the underlying relationships. Only explicit and proper information models – ideally based on UML – have the necessary formal precision.

For clarity and to avoid misunderstandings, we are not talking about a large information model, but rather about domain-specific small information models that are easy to implement (DCM, [20]). This is the only way to create real added value. Past attempts by the community show that large information models are doomed to failure [16]. The necessary information models must be developed and specified by competent professional societies providing the knowledge on an abstract level as a genuine (UML) information model. Interoperable technical implementations can then be derived, both in applications and for data exchange. All presented facts and statements hold for the growing move towards higher-level interoperability as presented in [1].

References

[1] Blobel B, Routsalainen P, Oemig, F. Why Interoperability at Data Level is not Sufficient for Enabling pHealth? Stud Health Technol Inform. 2020; 273.

[2] Tina Lee Y. Information modeling from design to implementation. National Institute of Standards and Technology. 1999.

[3] HL7 International Inc. http://www.hl7.org.

[4] International Organization for Standardization. ISO/IEC 7498:2000 – Information technology – Open Systems Interconnection. Geneva: ISO; 2000.

[5] HL7 v2.xml Specification. 2002. www.hl7.org.

[6] Blobel B, Oemig F. The Importance of Architectures for Interoperability. Stud Health Technol Inform. 2015; 211: 18-56.

[7] European Interoperability Framework (EIF), https://joinup.ec.europa.eu/collection/nifo-national-interoperability-framework-observatory/eif-european-interoperability-framework-0, last accessed July 2020.

[8] Oemig F, Blobel B, Helmer A, Birkle M, Comparison of different solutions for interoperability in medicine, HL7 Mitteilungen, Issue 41/2018.

[9] Oemig, Snelick: Healthcare Interoperability Standards Compliance Handbook, ISBN 978-3-319-44839-8, 2016.

[10] Übermittlung onkologischer Daten auf Basis von CDA R2. http://wiki.hl7.de/index.php?title=IG:Übermittlung%20onkologischer%20Daten , 2013.

[11] HL7 International Inc. HL7 SAIF. Service-Aware Interoperability Framework, www.hl7.org

[12] HL7 International Inc. HL7 ECCF. Enterprise Compliance and Conformance Framework, www.hl7.org

[13] International Organization for Standardization. ISO/IEC 10746 Open distributed processing - Reference model. Geneva: ISO; 2009.

[14] LOINC. Logical Observations, Identifiers, Names and Codes, www.loing.org

[15] SNOMED International. SNOMED CT. Systemized Nomenclature of Medicine – Clinical Terms, www.snomed.org

[16] HL7 International Inc. HL7 Version 3, www.hl7.org

[17] HL7 International Inc. HL7 Fast Healthcare Interoperability Resources (FHIR), Release 4, http://www.hl7.org/fhir

[18] CDS Hooks. https://cds-hooks.org

[19] SMART Health IT on FHIR – Connecting health system data to innovators' apps, http://smarthealthit.org

[20] International Organization for Standardization. ISO/TS 13972:2015 Health Informatics - Detailed Clinical Models (DCM). Geneva: ISO; 2015.