EJBI 2018 ISSN 1801 - 5603

European Journal for Biomedical Informatics

Volume 14 (2018), Issue 3

Special Topic

Mastering the Interoperability Challenge

Editors
Bernd Blobel, Philip Scott

18th International HL7 Interoperability Conference (IHIC 2018)

An Official Journal of the European Federation for Medical Informatics

www.ejbi.org

An Official Journal of:

EFMI EUROPEAN FEDERATION for MEDICAL INFORMATICS

Aims and Scope

European Journal for Biomedical Informatics is an online journal publishing submissions in English language, which is devoted towards the progress in biomedical science. The journal welcomes all types of research communications through its rational open access window to maximize the possibilities of global reach for each article. The frequency of issue release is bi-annual, and strictly supported by peer-review process. The primary aim of the journal is to make all the informations freely available irrespective of different socio-economic barriers, which further will help in the progress of science. Moreover, the journal intends to showcase all the timely research information on biomedical informatics and its related studies for the effective contribution in the field of biomedical methods and technologies.

European Journal for Biomedical Informatics has specific instructions and guidelines for submitting articles. Those instructions and guidelines are readily available on the instruction for authors. Please read and review them carefully. Articles that are not submitted in accordance with our instructions and guidelines are more likely to be rejected.

European Journal for Biomedical Informatics accepts manuscript submissions through a mail attachment: editor@ejbi.org

Author Guidelines

This journal follows the guidelines of the International Committee of Medical Journal Editors (www.icmje.org) and the Committee on Publication Ethics (www.publicationethics.org).

Scope of the Journal

European Journal for Biomedical Informatics is dealing with a range of topics which is included under the scopes of biomedical informatics. The focus of the journal is on the major aspects of implementation of information science in Medical sciences, which includes Biomedicine, Ontology, Genomics, Biometrics, Biostatistics, Health informatics, Medical informatics, Clinical informatics, Semantics interoperability, Informatics applications in imaging.


European Journal for Biomedical Informatics
EJBI

Contents

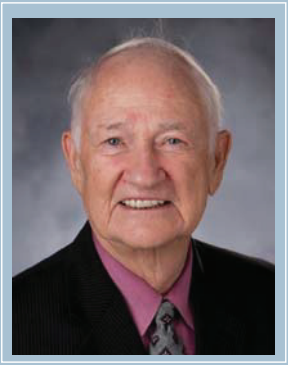

European Journal for Biomedical Informatics
EJBI

18th International HL7 Interoperability Conference (IHIC 2018)

The 18th International HL7 Interoperability Conference (IHIC 2018) will be held from 11-12 July 2018 at the University of Portsmouth, UK. The meeting will be hosted by HL7 UK, with support from HL7 Germany as a permanent IHIC supporter. Other HL7 Affiliates and groups interested in improving healthcare interoperability are invited to support the event.

Keynote Speakers

Prof. Bernd Blobel
Medical Faculty, University of Regensburg, Germany



Prof. William E. Hammond
Director, Duke Center for Health Informatics, Duke University, USA


European Journal for Biomedical Informatics
EJBI

Mastering the Interoperability Challenge

Bernd Blobel[1,2,3]* and Philip Scott[4,5]

[1] Medical Faculty, University of Regensburg, Germany

[2] eHealth Competence Center Bavaria, Deggendorf Institute of Technology, Germany

[3] First Medical Faculty, Charles University in Prague, Czech Republic

[4] Centre for Healthcare Modelling & Informatics, University of Portsmouth, UK

[5] HL7, UK

Correspondence to:

Prof. Dr. habil. Bernd Blobel, FACMI, FACHI, FHL7, FEFMI, MIAHSI
Medical Faculty, University of Regensburg, Germany.
E-mail: bernd.blobel@klinik.uni-regensburg.de

EJBI 2018; 14(3):01-02
Published: July 06, 2018

1 Introduction

This Special Issue of the European Journal for Biomedical Informatics is dedicated to the International HL7 Interoperability Conference 2018 "Mastering the Interoperability Challenge" (IHIC 2018), 11-12 July 2018 in Portsmouth, UK (http://ihic.info/2018). It contains papers selected by an independent peer review process, strictly performed by experts from countries different from the authors' country of residence.

IHIC 2018 is the 18th event of the International HL7 Interoperability Conference series, which was inaugurated in 2000 by the Board of HL7 Germany and its unforgettable Chair and interoperability pioneer Joachim W. Dudeck. The first event in Dresden, Germany, was entitled "Advanced Healthcare Information Standards". While the first conferences were characterized by focusing on CDA, over time, the scope of the conferences has been extended towards all aspects of health information interoperability. The concept of interoperability has dramatically changed from standardized electronic data interchange (EDI) based on data representation at application level, the 7th level of the ISO Open Systems Interconnection stack and basis of the name Health Level 7. Meanwhile, the semantics of shared data as well as service level interoperability, but also domain-specific issues and even social aspects are considered, bringing terminologies and ontologies, but also implementation and conformance challenges on board. The relations to IHE and the FHIR success are especially highlighted at IHIC 2018. So it is just consequent to address also in 2018 both technological and non-technological issues of interoperability.

IHIC 2018 is framed by an Opening Keynote and a Closing Keynote. In the Opening Keynote, titled "Solving the Modeling Dilemma as a Foundation for Interoperability", Bernd Blobel, University of Regensburg (Germany), addresses all levels of interoperability, i.e. technical, structural, syntactic, semantic and organization/service interoperability most of health informatics interoperability standards are limited to, but also non-ICT interoperability such as knowledge-based domain-to-domain interoperability and even skills-based interoperability supporting end-user collaboration. The paper introduces different data model classification systems to analyze widely spread data model based interoperability specifications in comparison with the ISO Interoperability Reference Architecture Model. In his Closing Keynote, Ed Hammond from Duke University (USA) focuses on ICT-specific interoperability specifications and implementations provided by international standards and specifications, thereby especially highlighting HL7 standards and artifacts. In that context, he presents multiple aspects of and perspectives on, interoperability, thereby considering not just technical issues, but also expectations and needs of specific user communities.

IHIC 2018 is structured into four sections: a) Quality Improvement, b) Testing and Implementation, c) Overcoming Local and Global Barriers, and d) Consent and Trust for Care and Research. The papers published in this EJBI Special Issue address different aspects of the interoperability challenge from a theoretical and methodological perspective, usability requirements, professional groups' preferences, process design, semantical ambiguity, and implementation details.

In the first section on Quality Improvement, Peter Seifter and colleagues from HL7 Austria report about clinical decision support systems using data from structured CDA documents collected in the Austrian national EHR solution ELGA (Elektronische Gesundheitsakte - Electronic Health Record). For that purpose, open source platforms such as the Drools business rule management system or the

ArdenSuite software for managing knowledge represented in Arden Syntax Medical Logic Modules have been successfully used to manage the Austrian Patient Summary and the Austrian Microbiology Report. Frank Oemig and Bernd Blobel from HL7 Germany discuss the deployment of FHIR specifications and implementations for standardized quality assurance and control in Germany.

The second section on Testing and Implementation is introduced by a paper from Sebastian Bojanowski and others from HL7 Poland. The authors present the national online platform Tukan – compliant with the IHE Gazelle environment - to publish and to test HL7 CDA Implementation Guides and related HIE profiles for future Polish eHealth services.

Within the fourth section on Consent and Trust for Care and Research, an international Austrian/Canadian team lead by Anna Lackerbauer from Austria developed architecture to implement an interoperable e-consent form for medical treatment using the FHIR methodology.

Ed Conley (UK) and Mathias Pocs (Germany) finally tackled the challenge of personal data protection as foundational to successful eHealth and interoperability implementations. In that context, they highlighted the new European General Data Protection Regulation (GDPR) to be met by all European service providers, but also by all global service providers directly or indirectly serving European citizens. The work aims at enabling GDPR-compliant large scale health information exchange as well as trustworthy research environment within the UK NHS, but also regionally or even globally.

Additionally to the papers presented here, practice reports and implementation experiences will be shared at the conference.

The IHIC 2018 Program is completed by a Panel on Resolving Practical Implementation Issues as well as Tutorials provided on the day prior to the conference.

The Editors whish all interested parties enjoyable reading.

The Guest-Editors are indebted to thank all authors and reviewers for their excellent work. Finally, they thank HL7 International, HL7 UK and HL7 Germany for supporting the event, but also HL7 International and HL7 Germany for financing the Joachim W. Dudeck Award.

Solving the Modeling Dilemma as a Foundation for Interoperability

Bernd Blobel[1,2,3]*and Frank Oemig[4]

[1] Medical Faculty, University of Regensburg, Germany

[2] eHealth Competence Center Bavaria, Deggendorf Institute of Technology, Germany

[3] First Medical Faculty, Charles University Prague, Czech Republic

[4] Deutsche Telekom Healthcare and Security Solutions GmbH, Essen, Germany

Abstract

Introduction: 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.

Methods: The paper discusses classification systems for data models and enterprise business architectures and compares them with the ISO Reference Architecture.

Results and Conclusions: Existing definitions, specifications and standards for data models enabling interoperability are analyzed, and their limitations are evaluated. Amendments to correctly use those models and to better meet the aforementioned challenges are offered.

Keywords

Healthcare transformation; Interoperability; Data models; Knowledge management; Architectures

Correspondence to:

Prof. Dr. habil. Bernd Blobel, FACMI, FACHI, FHL7, FEFMI, MIAHSI
Medical Faculty, University of Regensburg, Germany.
E-mail: bernd.blobel@klinik.uni-regensburg.de

EJBI 2018; 14(3):03-12
Received: May 12, 2018
Accepted: May 23, 2018
Published: July 06, 2018

1 Introduction

In order to support highly distributed, personalized, predictive, preventive, participative, and cognitive care healthcare systems have to provide and to ensure reliable environments. The approach requires the exchange of data in a highly interoperable fashion across different disciplines and domains. The involvement of stakeholders from different specialties and policy domains, offering different levels of knowledge, skills, and experiences to act in different scenarios accommodating different business cases has to be supported by allowing specific methodologies, terminologies, and ontologies to enable analysis, design, implementation, deployment, maintenance, and evaluation of systems within their lifecycle. The management of such highly dynamic, complex, heterogeneous and context-depending business processes, i.e. the execution of ICT (Information and Communication Technology)-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 methods and tools scalable and adaptive to communities, user groups and even individuals, transferring their knowledge, experience, expectations, and intentions into 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]. It covers all levels of ICT-related interoperability from technical interoperability through structural interoperability, syntactic interoperability, semantic interoperability and organization/service interoperability health informatics interoperability standards usually address, but also interoperability beyond ICT-related business cases represented through domain-specific ontologies such as knowledge-based domain-domain interoperability and even skills based interoperability addressing the end-user [4]. Dealing with the data modeling dilemma for enabling interoperability, this paper introduces data model classification systems to analyze widely spread data model based interoperability specifications in comparison with the ISO Interoperability Reference Architecture Model [4], thereby summarizing work published in other context [5, 6].

2 Methods

2.1 General Aspects of Modeling

According to Alter [7], 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. A purpose of developing and deploying models is the creation of knowledge. An outcome of developing mathematical models is that it helps model builders and decision makers understanding the relationships between important sets in a business situation. On the other hand, description and especially the interpretation of real systems are based on knowledge. This aspect is especially highlighted by Langhorst et al. [2], defining 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. A concept shall be uniquely identifiable, independently accepted by experts and users, and has a representation. A concept is a knowledge component that can be specialized and generalized as components can. Knowledge can be represented at different level of abstraction and expressivity, ranging from implicit knowledge up to fully explicit knowledge representation, i.e. from natural language up to universal logic (Figure 1). A key parameter in choosing or creating a proper knowledge representation (KR) is its expressivity. A more expressive knowledge representation language enables an easier and more compact expression of knowledge within the semantics and grammar of that knowlegde representation. However, more expressive languages are likely to require more complex logic and algorithms to construct equivalent inferences. A highly expressive KR is also less likely to be complete and consistent. Less expressive KRs may be both complete and consistent. This is an important advantage of domain-specific terminologies and their underlying ontologies, extensively exploited in good modeling best practices.

2.2 Data Modeling

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 [9]. 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 [10].

3 Modeling Health Systems

3.1 Conceptual Model of Architectural Descriptions

ANSI/IEEE 1471-2000 - IEEE Recommended Practice for Architectural Description for Software-Intensive Systems considers aspects and principles to be considered when modeling information systems [11]. In that context, the importance of the business domain and its mission represented by the domain experts as relevant stakeholders has been highlighted in some detail. The resulting conceptual model of architectural descriptions for such systems is presented in Figure 2.

3.2 Data Modeling Best Practices

Hoberman et al. describe a data model as 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 [12]. 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-

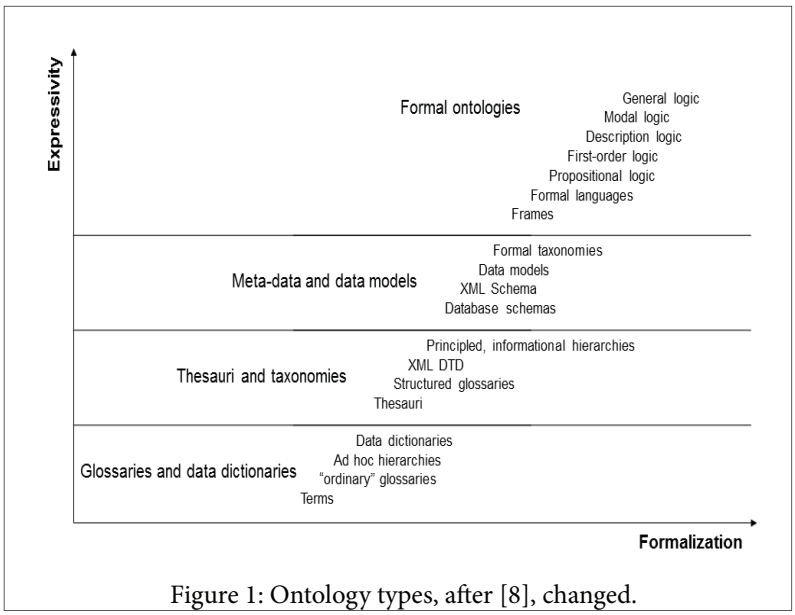


Figure 1: Ontology types, after [8], changed.

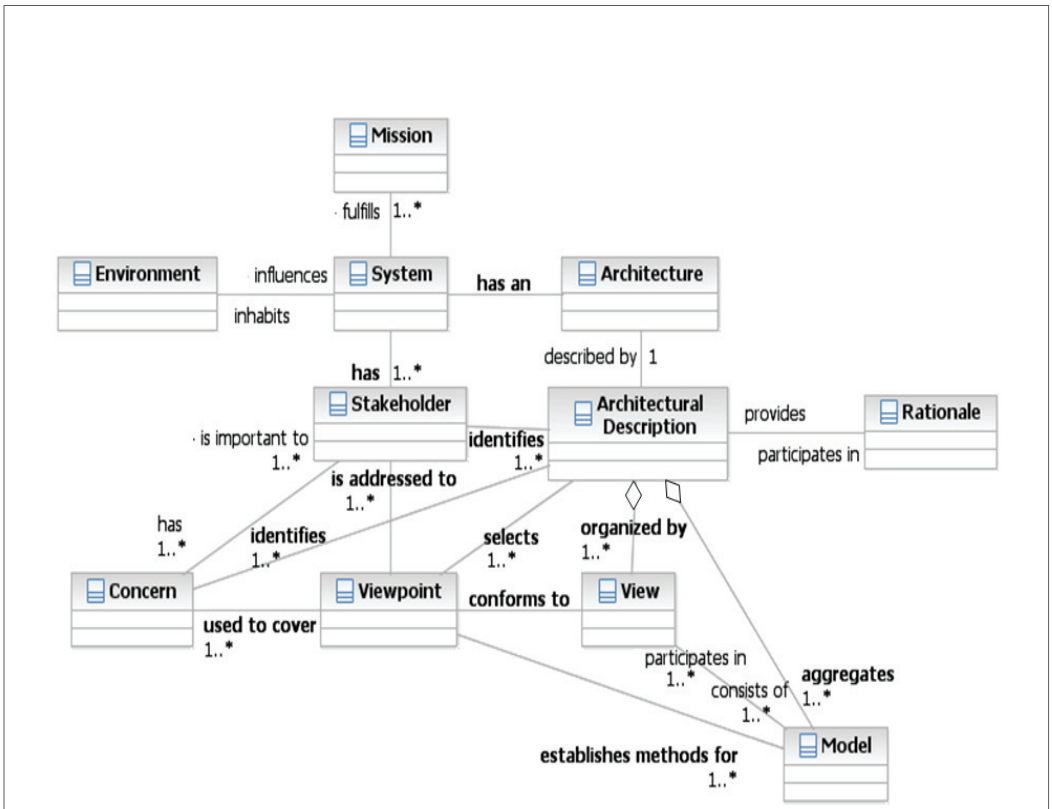


Figure 2: ANSI/IEEE 1471 conceptual model of architectural descriptions [11].

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.

According to Langhorst et al., relevant stakeholders must define the provided view of the business model as well as the way of structuring and naming the concepts of the problem space [2]. Following the ontology-based business integration, thereby 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 [8].

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 [13], 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.

Data modeling enabling advanced interoperability in distributed multi-domain healthcare systems must be guided by domain experts' business models, so representing the main stakeholders perspective, terminology and ontology.

Figure 3 presents the modeling dimensions and the related transformation pathway.

3.3 The ISO Interoperability Reference Architecture

This description of the ISO Interoperability Reference Architercture corresponds to the related text in ISO 13606:2018 Health informatics – EHR communication provided by the first author [14].

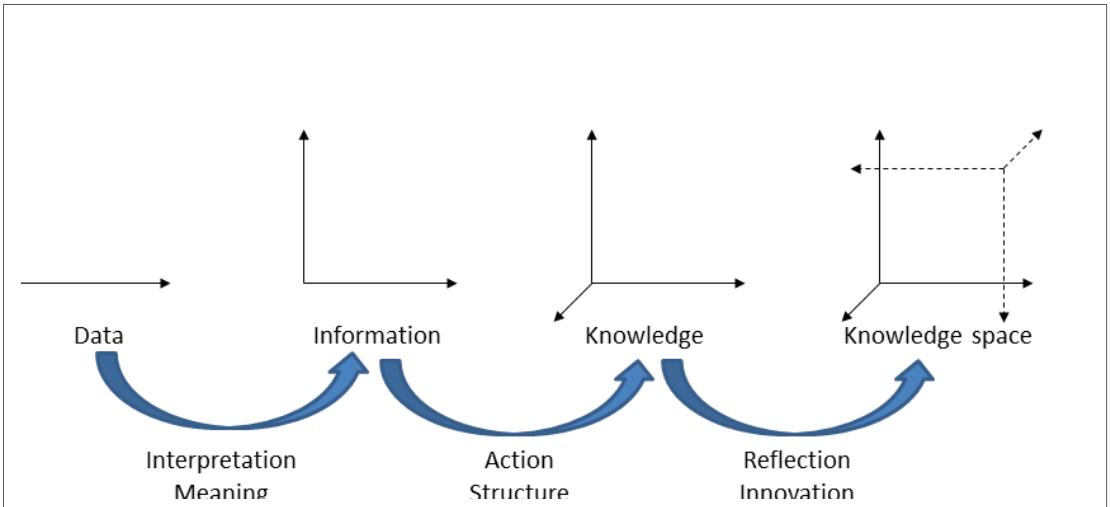


Figure 3: Dimensions of data modeling (after Krogstie [13]).

Meeting the objectives of improving safety, quality and efficiency of care with ICT support requires advancing interoperability between computer systems towards a business process specific co-operation of actors representing the different domains participating in the business case. For that purpose, the agreed domain knowledge, but also individual (language, education, skills, experiences, social and psychological aspects, etc.) and environmental context have to be represented correctly and formally for integration in the ICT system as part of the business system. As the domain experts involved describe specific aspects of that business system in a specific context, using their specific terminologies and ontologies, methodologies and frameworks, the resulting informational representations are quite inconsistent, requiring a peer-to-peer interoperability adaptation process. Adapting existing standardized informational representations of domain-specific use cases as practiced in most current interoperability specifications to changing contexts or including other domains requires another common harmonized informational representation or results in permanent revisions of specifications. A pretty bad example of the latter fact is ISO 13606, which has been revised in more than 15 versions provided over three years.

It is impossible to represent the highly complex, highly dynamic, multi-disciplinary/multi-domain healthcare system by one domain's terminology/ontology or - even worse - by using ICT ontologies. The same holds when using one domain's representational style and models or standards as reference or master all the interrelated components must be adapted to.

The alternative is an abstract domain-independent representation of systems using Universal Type Theory and corresponding logics as philosophers do to describe the universe [15, 16]. The mathematical concept representation in combination with systems engineering methodologies allows representing any system architecturally (i.e. the system's components, their functions and internal as well as external relations) by generically describing its composition/decomposition as well as the aspects (domains) of the system relevant in a specific context (e.g. business case). For correctly and formally representing the concepts and relations of the domain-specific subsystems involved in that business case, those subsystems are represented by their corresponding approved domain ontologies, resulting in a system-theoretical, architecture-centric, top-level ontology driven approach [17, 18]. The reference architecture model can be used recursively, so representing, e.g., the real-world systems' continuum from elementary particles to the universe (Figure 4).

By combining that model with ISO/IEC 10746 [19], the Interoperability Reference Architecture Model (introduced in the nineties as Generic Component Model - GCM) as well as the applicable rules - the Interoperability Reference Architecture Model Framework - (also known as GCM Framework) is completed (Figure 5) [20].

This Interoperability Reference Architecture Model allows consistently transforming and interrelating any domain-specific subsystem's structure and behavior (e.g. domain-specific standards and specifications) by ontologically representing its concepts and relationships at the real world system component's level of granularity. In other words, the domain-specific subsystem (e.g. a domain-specific standard or specification) is re-engineered using the Interoperability Reference Architecture Model, by that way providing a standardized interface to that specification (Figure 6).

Bound to the GCM Framework, inter-domain relationships must happen at the same level of granularity [3]. To get there, intra-domain specializations/ generalizations have to be performed. In summary, the Interoperability Reference Architecture Model supports

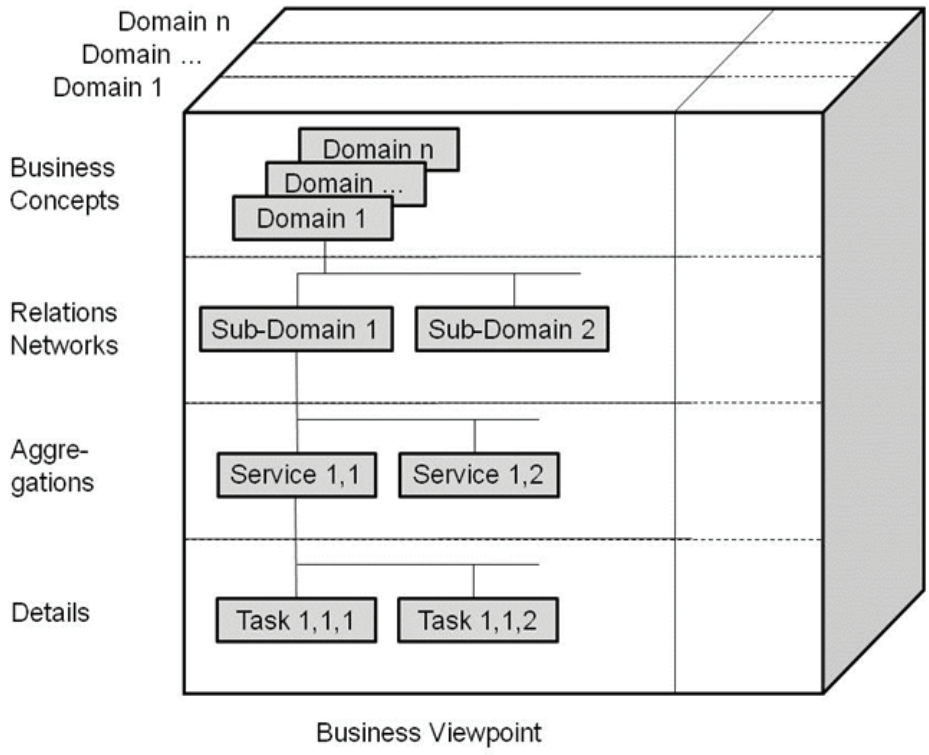


Figure 4: Interoperability Reference Architecture Model granularity levels.

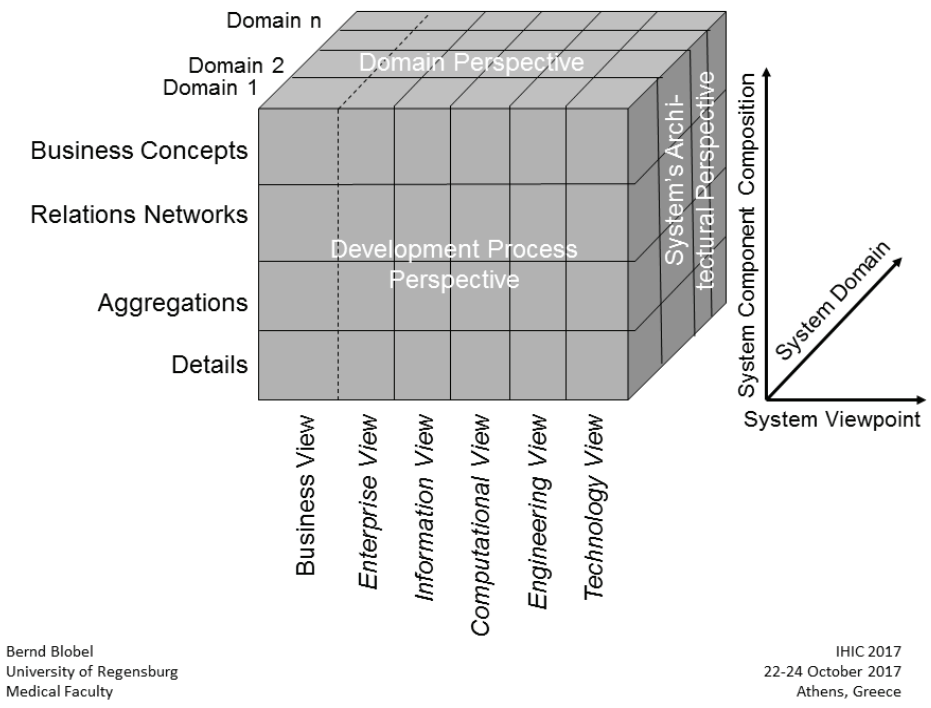


Figure 5: The Interoperability Reference Architecture Model.

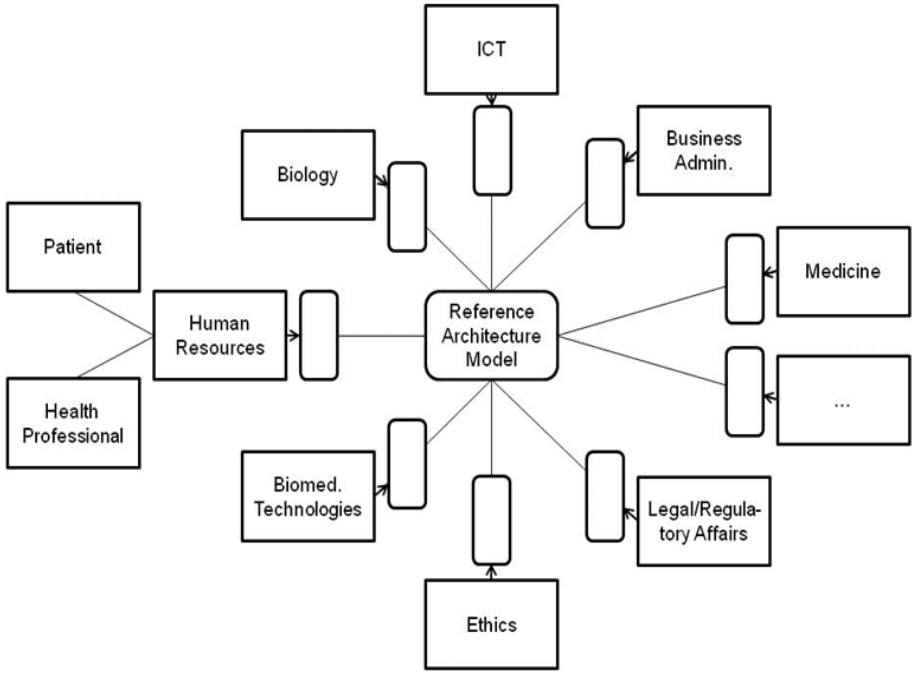


Figure 6: Interoperability mediated by the ISO Interoperability Reference Architecture Model [4].

ontology harmonization or knowledge harmonization to enable interoperability between existing systems, standards and solutions of any level of complexity without the demand for continuously adapting/revising those specifications.

Examples for re-engineering existing standards to provide cross-specification or even inter-disciplinary interoperability can be found in [21, 22] regarding interoperability between HL7v2 and HL7v3 or in [23, 24] enabling use case and domain-crossing interoperability in the context of ISO 13972 Health informatics - Detailed clinical models [25]. The approach has also been adopted for ISO and CEN standards such as ISO 13606-1 Health informatics – EHR communication – Reference Model [14], where the reference model used for all parts has been re-engineered. The feasibility of the Reference Architecture Model and Framework has also been practically demonstrated for automatically designing inter-domain Web services to facilitate multi-disciplinary approaches to Type 2 Diabetes Care management [26]. Several cross-domain ISO specifications, such as ISO 22600 Privilege management and access control [27], ISO 21298 Functional and structural roles [28], or the HL7 Composite Security and Privacy Domain Analysis Model [29] are based on the ISO Interoperability Reference Architecture. A simplification of the model is the basis of the open architectures for national health information systems in developing African countries [30]. The approach also allows a comparative analysis and evaluation of ICT Enterprise Architectures [3].

4 Results

Different interoperability standards, like HL7 Version 3 (including its Clinical Document Architecture – CDA, and its Clinical Information Modeling Initiative - CIMI) [31], openEHR/EN 13606/archetypes [32], OHDSI [33], OMOP [34], ISO 13972 [25], and HL7 FHIR [35], are all claiming to work on enabling and improving interoperability. Unfortunately, their concepts cover diverse aspects in different regards and maturity: communication, system architecture, reference architecture, network access across enterprises, layout/forms structure for data capture, persistency, entity relationship models, and last but not least conformance claims and capabilities. The use of vocabulary like classifications and terminologies, further advanced into knowledge representation in form of ontologies, adds another level of complexity. The dilemma is roughly demonstrated in Figure 7.

In a previous study different interoperability levels from technical through structural, syntactic, semantic, service interoperability knowledge-based to skills-based interoperability are defined [4]. The HL7 V2 EDI protocol, but also HL7 V2/V3 Implementable Technical Specification (ITS) [31] as well as specifications of the observational health data initiatives OHDSI [33] and OMOP [34] define data structure and related data types at the physical data model

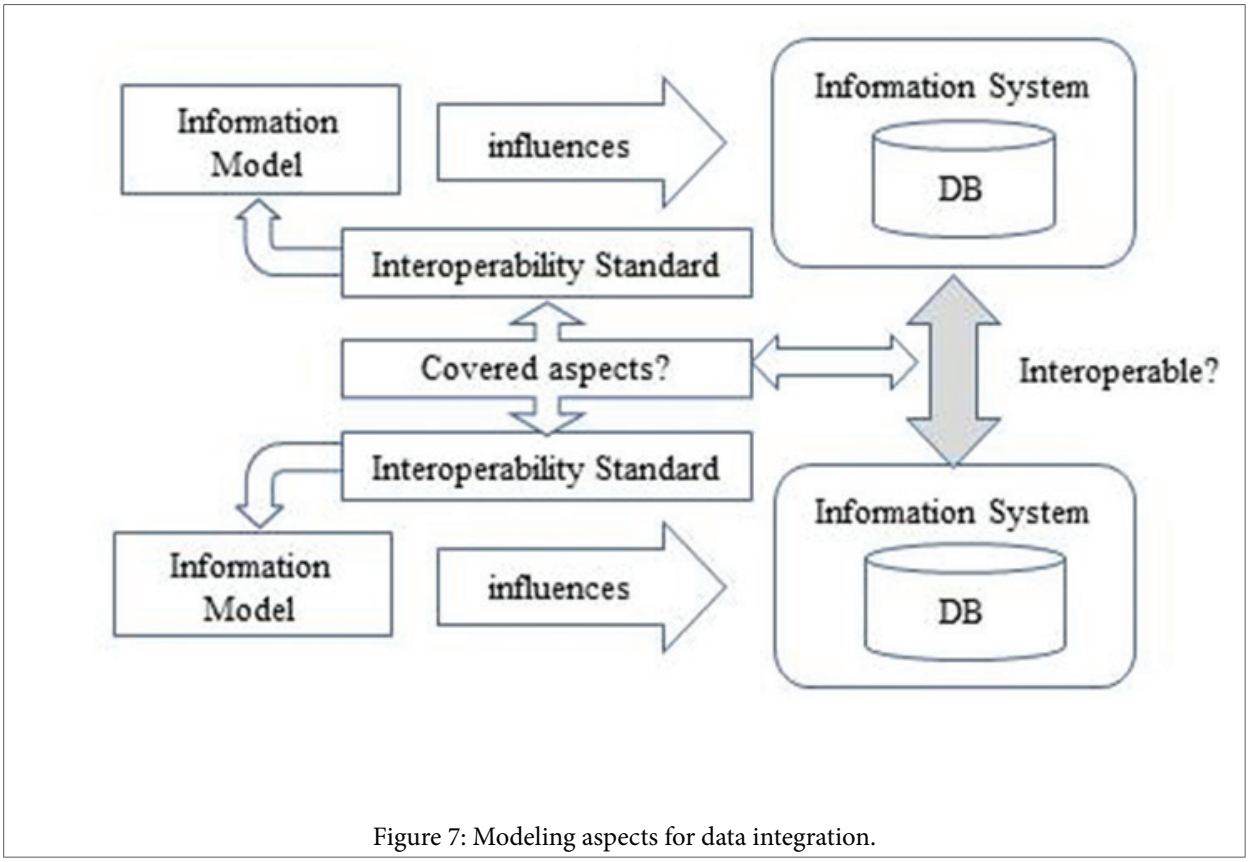


Figure 7: Modeling aspects for data integration.

level, addressing the modeling dimension of the 1-dimensional data approach. With HL7 V3, following the HL7 Development Framework (HDF), the HL7 Reference Information Model (RIM) – also standardized at ISO as ISO/HL7 21731 – has been defined [36]. 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 levels 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 [8] and in [12], 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 [12] and the dimensions of modeling [13] in relation to the system-oriented, architecture-centric, ontology-based, policy-driven ISO Interoperability Reference Architecture Model [4, 6] 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. Also newer developments such as the Federal Health Information Model (FHIM), a project under the Federal Health Interoperability Modeling and Standards (FHIMS) program within the US Federal Health Architecture initiative [37], belong to that 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) [38] or the Communication Standards Ontology (CSO) [39]. 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 [12], Dimensions of Modeling [13], and the ISO Interoperability Reference Architecture Model [4, 6], 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			
High-level data model	Business domains stakeholders	Relevant information and representation & relationships of basic concepts	Knowledge	Enterprise View	DCM, CSO	ISO 10746 ODP-RM	ISO/CEN Interoperability Reference Architecture
Logical data model	Data modelers and analysts	Layout & types of data and object relationships	Information	Information View	HL7 V3 (CMETs), HL7 CIMI, openEHR Archetypes, FHIM		
				Computational View	HL7 FHIR		
Physical data model	Data modelers and developers	Implementation-related and platform-specific aspects	Data	Engineering View			

5 Discussion and Conclusion

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 [38]. For more information, see, e.g., [23, 24]. 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 [9] 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 [9]. 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 [14] and ISO 12967 [40]. Other specifications will follow soon.

In this volume, Ed Hammond presents a very interesting consideration of the interoperability ecosystem [41]. Combining that work with our methodology can help formalizing a multitude of interoperability instances health systems are facing.

6 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, 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 Oct 12-18; Beijing, China: IEEE; 2005. p. 535-540.

[2] Lankhorst M. 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. Eur J Biomed Inform. 2017;13(1): 09-15.

[5] Blobel B, Oemig F, Ruotsalainen P. Data Modeling Challenges of Advanced Interoperability. Stud Health Technol Inform. 2018; 247: 860-864.

[6] Blobel B. Reference Architecture Model Enabling Standards Interoperability. Stud Health Technol Inform. 2017; 235: 401-405.

[7] Alter S. Information Systems-A Management Perspective. Reading: Addison-Wesley; 1992.

[8] Rebstock M, Fengel J, Paulheim H. Ontologies-Based Business Integration. Berlin Heidelberg: Springer- Verlag; 2008.

[9] Seiner RS. Data Modeling is a Form of Data Governance. IDERA. 2017.

[10] Harper J. Lessons learned from big interoperability in government. KMWorld. November/December 2017; 26: 9.

[11] IEEE Standards Association. ANSI/IEEE 1471-2000.

[12] 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.

[13] Krogstie J. Business Information Systems Utilizing the Future Internet. Lecture Notes in Business Information Processing. 2011; 90: 1-18.

[14] International Organization for Standardization. ISO 13606-1:2018 Health informatics. EHR communication Reference Architecture. Geneva: ISO; 2018.

[15] Kamareddine F, Laan T, Nederpelt R. A Modern Perspective on Type Theory. New York: Kluwer Academic Publishers; 2004.

[16] Bloe R, Kamareddine F, Nederpelt R. The Barendregt Cube with Definitions and Generalized Reduction. Inf Comput. 1996; 126(2): 123-143.

[17] Blobel B. Knowledge Representation and Management Enabling Intelligent Interoperability – Principles and Standards. Stud Health Technol Inform. 2013; 186: 3-21.

[18] Blobel B, Brochhausen M. Ontological Foundations of the GCM Reference Architecture Model and Approach. (in preparation).

[19] International Organization for Standardization. ISO/IEC 10746-3:2009 Information technology-Open distributed processing-Reference model: Architecture. Geneva: ISO; 2009.

[20] Blobel B. Architectural approach to eHealth for enabling paradigm changes in health. Methods Inf Med. 2010; 49(2): 123-134.

[21] Oemig F, Blobel B. A Communication Standards Ontology Using Basic Formal Ontologies. Stud Health Technol Inform. 2010; 156: 105-113.

[22] Oemig F, Blobel B. A Formal Analysis of HL7 Version 2.x. Stud Health Technol Inform. 2011; 169: 704-708.

[23] Goossen WTF. Detailed clinical models: Representing knowledge, data and semantics in healthcare information technology. Healthc Inform Res. 2014; 20(3): 163-172.

[24] Goossen W, Gossen-Baremany A, Van Der Zel M. Detailed clinical models: A review. Healthc Inform Res. 2010; 16(4): 201-214.

[25] International Organization for Standardization. ISO 13972:2015 Health informatics -Detailed Clinical Models. Geneva: ISO; 2015.

[26] Uribe Gómez GA. An Architecture-Centric and Ontology-Based Approach to Cross-Domain Interoperability of Health Information Systems for Diabetes Care. Universidad del Cauca, Faculty of Electronics and Telecommunications Engineering, Popayan, Colombia, 2015.

[27] International Organization for Standardization. ISO 22600:2014 Health informatics-Privilege management and access control. Geneva: ISO; 2014.

[28] International Organization for Standardization. ISO 21298:2017 Health informatics-Functional and structural roles. Geneva: ISO; 2017.

[29] Health Level Seven International, Inc. Composite security and privacy DAM.

[30] Moodley D, Pillay AW, Seebregts CJ. Researching and developing open architectures for national health information systems in developing African countries. LNCS. 2012; 7151: 129-139.

[31] www.HL7.org

[32] www.openehr.org

[33] https://www.ohdsi.org

[34] http://omop.fnih.org

[35] Health Level Seven International, Inc. HL7 FHIR.

[36] Health Level Seven International, Inc. HL7 Reference Information Model v. 2.47.

[37] http://fhims.org

[38] Blobel B, Goossen W, Brochhausen M. Clinical modeling- A critical analysis. Int J Med Inform. 2014; 83(1): 57-69.

[39] Oemig F. Development of an ontology-based architecture for ensuring semantic interoperability between communication standards in healthcare (in German). University of Regensburg, Medical Faculty, Regensburg, Germany, 2011.

[40] International Organization for Standardization. ISO 12967:2017. Health informatics. Health informatics service architecture. Geneva: ISO; 2017.

[41] Hammond WE. How do you know when you have interoperability? Eur J Biomed Inform. 2018; 14(3): 13-20.

How Do You Know When You Have Interoperability?

William Ed Hammond[1,2]*

[1] Duke Center for Health Informatics, Duke Medicine, Durham, North Carolina, USA

[2] Duke Clinical and Translational Science Institute, Duke University, Durham, North Carolina, USA

Abstract

This paper first looks at the several definitions of the word interoperability. Next, the paper addresses several of the distinct components that are factors in accomplishing interoperability: semantic interoperability, functional interoperability, stakeholder interoperability, consumer interoperability, business interoperability, privacy and security interoperability, and international interoperability. This paper discusses each of these components in terms of their current state relating to interoperability and suggests what changes might be made to achieve interoperability.

Keywords

Interoperability; Health data standards; Terminologies; EHRs; PHRs; Business models; Consumers

Correspondence to:

W. Ed Hammond PhD, FACMI, FAIMBE, FIMIA, FHL7
Duke Center for Health Informatics, NC 27705, USA.
E-mail: william.hammond@duke.edu

EJBI 2018; 14(3):13-20
Received: May 15, 2018
Accepted: May 22, 2018
Published: July 06, 2018

1 Introduction

Interoperability has been the Holy Grail of Informatics for many decades. Its definitions have been diffuse, and its obtainability has been seemingly impossible. It was an impetus for the creation of Standards Developing Organizations (SDO); the SDOs used the term "plug and play" although we have never reached that point. In the United States, the Office of the National Coordinator (ONC) has made interoperability a focus of its initiatives. "How do you know when you have interoperability?" is almost an unanswerable question; or else it may have innumerable answers. Part of the problem is in understanding the boundaries of the term. If I am able to exchange and understand the exchange of a single data element and that is all I care about, is that an interoperable system? If I can exchange and understand a complete patient summary, is that interoperability? If I can map from a local vocabulary to a common master set of terminologies, does that constitute interoperability?

This paper will not provide an answer to the question in its title. Instead, the paper will provide a fresh look at the many issues of interoperability from both an overall perspective plus a critical analysis of several specific components of interoperability. I will discuss new perspectives of traditional views of interoperability and suggest alternate approaches. The paper proposes that interoperability requires more than the transfer data and even more than the understanding the meaning of that data. Interoperability includes the appropriate use of that data; dealing with privacy and security; dealing with regulations; dealing with quality of the data and with trust; dealing with authentication and authorization; dealing with governance; and dealing with the many stakeholders who have a vested interested in the data and its use.

2 Definitions of Interoperability

The earliest views of interoperability came from IEEE in 1990, defining interoperability as "the ability of two or more systems or components to exchange information and to use the information that has been exchanged" [1]. The ability to exchange information is referred to as functional interoperability, and the ability to use that information is called semantic interoperability. This definition became the driving force for the development of data exchange standards and standard terminologies. My observation is that the word information should be replaced by the word data. That distinction between the word data and information is increasingly important.

Health Level 7 International® uses the IEEE definition of interoperability but adds more detail [2]:

- **"Functional"** interoperability is the capability to reliably exchange information without error
- **"Semantic"** interoperability is the ability to interpret, and, therefore, to make effective use of the information so exchanged.

The recognition of the effective use of the information is an important addition, although no detail is provided about how to do that.

The HIMSS definition [3] builds further on the concepts within the HL7 definition: interoperability means the ability of health information systems to work together within and across organizational boundaries in order to advance the effective delivery of healthcare for individuals and communities. HIMSS defines three levels of health information technology interoperability:

- *Foundational* - allows data exchange from one information technology system to be received by another and does not require the ability for the receiving information technology system to interpret the data.
- *Structural* - defines the structure or format of data exchange (i.e., the message format standards) where there is uniform movement of healthcare data from one system to another such that the clinical or operational purpose and meaning of the data is preserved and unaltered. Structural interoperability defines the syntax of the data exchange. It ensures that data exchanges between information technology systems can be interpreted at the data field level.
- *Semantic* - the ability of two or more systems or elements to exchange information and to use the information that has been exchanged. Semantic interoperability takes advantage of both the structuring of the data exchange and the codification of the data including vocabulary so that the receiving information technology systems can interpret the data. This level of interoperability supports the electronic exchange of patient summary information among caregivers and other authorized parties via potentially disparate electronic health record (EHR) systems and other systems to improve quality, safety, efficiency, and efficacy of healthcare delivery [4].

Dr. Bernd Blobel provides an outstanding discussion of interoperability and its component parts in his article "Standardization for Mastering Healthcare Transformation" [5]. He introduces an Interoperability Reference Architecture Model that provides direct methods for dealing with the different instances of interoperability and includes a discussion of challengres and solutions. In a related paper, Dr. Blobel further discusses challenges, standards, and solutions for EHR systems interoperability [6]. Kevin Heubusch also provides an interesting discussion of Interoperability from the perspective of ONC [7].

From these varied definitions from key, relevant organizations, interoperability is defined mainly from the technical aspects. In truth, interoperability involves a much larger scope of interests. My working definition of interoperability is the ability to share data whose meaning is unambiguously clear, its context understand, and it can be used for whatever purpose – and – the receiver is not previously known to the sender; i.e., an open-loop process. Even so, this is a limited definition that will be expanded below. Figure 1 shows many addition units that have concerns, interests and influence on interoperability.

3 Semantic Interoperability

3.1 ICD

From the moment two persons tried to communicate, semantic interoperability became important. Within the health care industry, semantic communication became important, first between individuals with the same clinical units, then within the same institutions, between clinical departments. With the current interests in data sharing, semantic communications have become increasingly important.

The first international classification was for the list of causes of death, adopted by the International Statistical

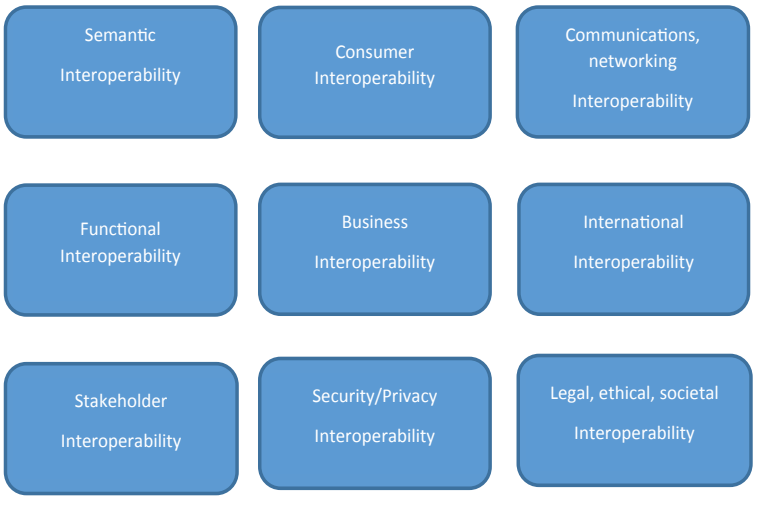


Figure 1: Components contributing to interoperability.

Institute in 1893. WHO was entrusted with the ICD at its creation in 1948 and published the 6th version, ICD-6, that incorporated morbidity for the first time. ICD-10 was endorsed in May 1990 by the Forty-third World Health Assembly and used by more than 100 countries around the world. ICD-10 [8] uses include monitoring of the incidence and prevalence of diseases, observing reimbursements and resource allocation trends, and keeping track of safety and quality guidelines. They also include the counting of deaths as well as diseases, injuries, symptoms, reasons for encounter, factors that influence health status, and external causes of disease.

Although ICD serves an important purpose, it does not address the requirements for semantic interoperability.

3.2 SNOMED-CT

SNOMED CT traces its history back to 1965 with the publication of the Systematized Nomenclature of Pathology (SNOP), published by the College of American Pathologists (CAP) to describe morphology and anatomy. In 1975, under the leadership of Dr. Roger Cote, CAP expanded SNOP to create the Systematized Nomenclature of Medicine (SNOMED). The most widely adopted version of SNOMED was SNOMED II (1979), followed by an expanded revision called SNOMED International (1993). CAP and Kaiser Permanente developed a logic-based version, SNOMED RT, in 2000. The Read Codes developed by Dr. James Reed in the United Kingdom were merged into SNOMED – Clinical Terms in the 1980s. The current version of SNOMED-CT has evolved from that set [9].

Although SNOMED has an increased use in the US and other countries, it can be used only by countries that have paid the licensing fee. Further, the terminology itself does not cover all clinical terminology and is complicated by pre and post coordination. The words of SNOMED are not the words used by professionals to define and describe clinical communication.

3.3 Logical Observation Identifiers Names and Codes (LOINC)

LOINC [10] is widely used internationally for representing clinical results such as laboratory tests, clinical observations, outcomes management and research. LOINC has two main parts: laboratory LOINC and clinical LOINC. Clinical LOINC contains a subdomain of Document Ontology which captures types of clinical reports and documents. Although LOINC comes close to meeting the requirements for semantic interoperability, the multiple names for the same tests creates major issues in exchanging and understanding the data that is represented. A solution for laboratory and other tests would be the assigning a LOINC code when the test was performed by the performing unit based on an agreed standard set throughout the industry.

3.4 Drug Coding

The coding of pharmaceutical products is perhaps one of the most challenging issues in semantic interoperability. In the US, RxNorm [11], developed by the National Library of Medicine, is coming into increased use. RxNorm incorporates data and knowledge from several sources.

A problem with drug coding is that each country has its own formulary and therefore its own drug terminology. Patients typically have drugs prescribed by different providers and at different locations. Bringing these medications together to create a prescription history is critical in the care of a patient. Although medication profiles are one of the most important aspect of a patient's history, it is the most difficult to achieve.

3.5 Addressing Semantic Interoperability

There are a number of reasons why semantic interoperability has evaded us for decades. First, there are many terminology sets currently in use today, with local terminologies being the most common. We seek a solution from common controlled terminologies that are incomplete and do not match the required clinical representation. Most of the terminology coding are influenced by reimbursement and do not represent the terms and the granularity required by direct clinical care. Cancer research is frustrated by the lack of a common vocabulary to support research for example in dealing with tumors. Then, there is confusion about the relationship among vocabularies, terminologies, classifications, nomenclature, ontologies, data models, and data elements. What role does each play in communications and semantic interoperability?

We are trying to solve the problem with what currently exists, rather than understanding what semantic interoperability is about. It is about communication, and we need to communicate at the level of need and understanding. Today's problem with semantic interoperability is that our first approach is mapping between terminologies to a common data model. Mapping inherently loses information. If the mapping does not, then why have two terminologies. Secondly, the cost is high because most terminologies are changing, and the mappings are out of synchronization.

There are a number of efforts to create a common data model as an approach to semantic interoperability. Unfortunately, there are enough common data models in health care so that they become uncommon. Examples include Sentinel, OMOP, i2b2, PCORNet, HL7, NLM VSEC, CDISC, CIMI, CIIC, and many others.

A single, international, data set of data elements offers the best solution in my opinion. The data element can be

knowledge metamodel with a rich set of attributes that can define all knowledge associated with the data element. Figure 2 illustrates typical attributes that might exist to contain the knowledge relating to the data element. Further, a process must be defined and implemented that would permit additional data elements to be added with a vetting process. Such a system would eliminate the need for locally-defined data elements that lose semantic interoperability. With such a common set of data elements defined, every local site should be required to adopt the set within a short period of time. Other authors have addressed this problem of data modeling and its contributions to interoperability [12].

3.6 Data Collection

The intangible component of semantic interoperability is data quality and consequently trusts. Data collection, then, is a necessary consideration for semantic interoperability. First, whenever possible, data collection must be automated. Wearable sensors are a step in that direction and increasingly can measure key parameters related to a person's real time health. If we can determine medical errors after they occur, we need to use the same algorithms in real time to prevent the error from occurring. If we have algorithms that will clean the data, we need to employ those same algorithms as part of the data collection to establish data quality.

3.7 New Data Types

The types of data that have value in clinical decision making has expanded significantly in recent years. In an Institute of Medicine (now the National Academy of Medicine) 2002 publication, J. McGinnis [13] claims that clinical data contributes only 10% information to a person's health index. The other data types are behavioral (40%); genomic (30%); social/economic (15%); and environmental (5%). We now need to create semantic interoperability to include these data types, and we need to define ways to capture and include this data in EHRs.

4 Functional Interoperability

4.1 HL7 International Standards

There are a number of standards developing organizations (SDOs) that have created data exchange standards. HL7, created in 1987, is a leader in the field and has created a continuing progression of standards. The first HL7 standard for the exchange of data is known as v2.n, where the current version is v2.8. Over 95% of hospitals and clinics use some version of HL7 v2 today. A second HL7 standard in wide use today is based on the v3 model-based standard and is known as the Clinical Document Architecture (CDA). CDA defines the structure of certain medical records, such as discharge summaries and progress notes, as a way to better exchange this information between providers and patients. An Implementation Guide based on this standard is in wide use to transfer the Patient Summary. This standard is known as the Continuity of Care Document (CCD).

The most recent HL7 data transfer standard is called the Fast Healthcare Interoperability Resource (FHIR®) standard [14]. It is a web-based standard and uses the REpresentational State Transfer (**REST**). REST is an architectural style that defines a set of constraints and properties based on HTTP.

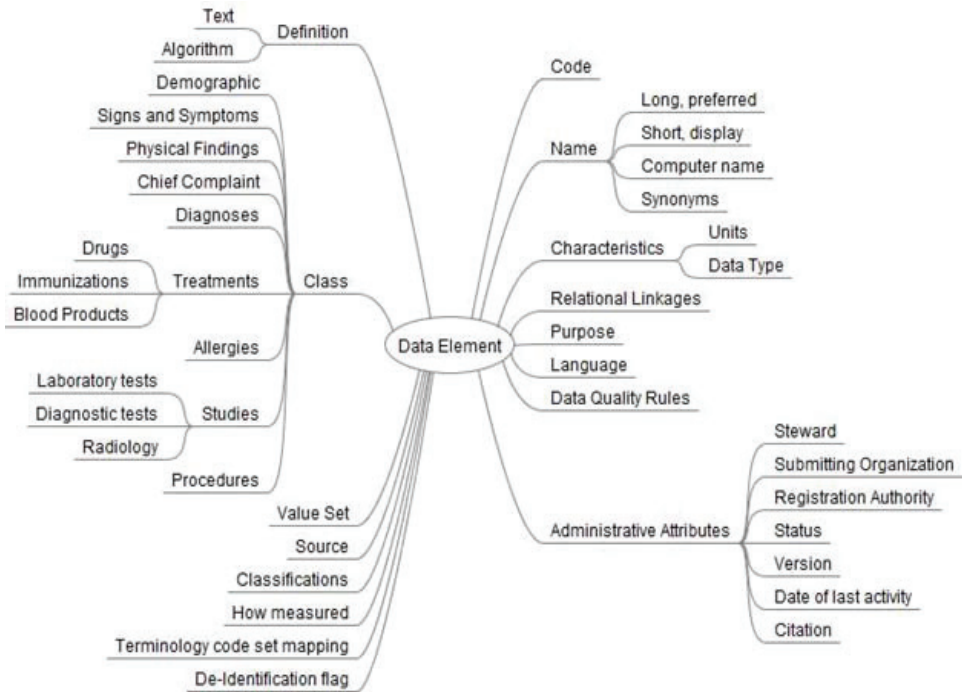


Figure 2: Data element with attributes.

Web Services that conform to the **REST** architectural style, or RESTful web services, provide interoperability between computer systems on the Internet. Facebook, Google, and others use this standard. RESTful systems typically communicate over HTTP verbs (Create/Post, Read/Get, Update, and Delete). FHIR provides interoperability between computer systems over the Internet.

FHIR is built on logical, related compound structures called RESOURCES. Resources consist of small logically discrete units of exchange with defined behavior and meaning. Resources have a known identity and location identified by a Universal Resource Identifier (URI). All exchangeable content is defined as a Resource. There are over 150 different Resources that are intended to cover 80% of healthcare. Examples include Patient, Practitioner, Family History, Care Plan, and Allergy Intolerance. Resources are defined using XML, JSON, or RDF. The core Resources reside in a repository open and free to use for all.

Resources are combined into groups called PROFILES to identify packages of data to address clinical and administrative needs. Parties exchanging data define the specific way they want to use Resources and their relations using Profiles. FHIR is service driven. Profiles define what a particular application needs to communicate based on Resources and Extensions (data elements, self-defined, that are not part of the core set). You only send data that is required for specific purposes. Profiles are used to constrain Resources – that is to define specifically what data is to be sent. Examples of Profiles are for referral of a patient; for populating registries; adverse event reporting; ordering a medication; and providing data to a clinical decision support algorithm such as a risk assessment calculation.

The FHIR standard has the potential to transfer any and all types of data. If FHIR Resources are tightly bound to a global master set of data elements, then functional interoperability becomes more achievable. The remaining barrier to interoperability is extensions. Extensions seem necessary to accommodate the transfer of data beyond the standardized core data elements. Unfortunately, that freedom opens the door to creating innumerable, uncontrolled, and potentially duplicative exchanges. The problem is further complicated by the 80/20 decision. FHIR core resources will address only 80% of the data more commonly required and the remaining 20% will be accommodated by extensions. I propose a better solution would be to define a process in which resources and data elements would be submitted to HL7 to become part of the core set. These additional submissions would be properly vetted and move into the normative standard. As we look into the requirements of the new data types suggested above, and as we consider new requirements such as population health, many of the data elements have not been included in the standard set.

4.2 SMART® on FHIR

SMART is an open standards-based technology [15] that enables developers to create apps that seamlessly and securely run across the healthcare system without requiring specific knowledge about each system. Many clinical apps have been built on this platform and are available through a publically accessible app gallery. SMART on FHIR is a set of open specifications that builds on FHIR API and Resource definitions. FHIR provides the core data models and SMART defines the profiles that carry out the functions of the app. Additionally, SMART uses an authorization model for apps based on the OAuth [16] standard. OAuth permits patients and providers control of their data.

4.3 CDS Hooks

CDS Hooks [17] represents the third type of standard that is important to the use component of functional interoperability. Examination of the causes of medical errors, inconsistencies in care, missed opportunities, and other similar events is the inherent fallacies of humans to perform tasks consistently. CDS Hooks provides a way in which the hooks are inserted into the data flow to trigger external events such as clinical decision support algorithms. Josh Mandel, one of the developers of CDS Hooks, discusses in more detail this functionality and gives a number of excellent examples [18].

4.4 Other Comments on Functional Interoperability

The above section focuses primarily on HL7 standards. Other SDOs, including IHE, ISO, DICOM, IEEE, and others, contribute to the set of standards that have value in enabling functional interoperability. The good news is that most of these other SDOs are working with HL7 and are using FHIR in their standards. The new Gemini Project between HL7 and IHE should more tightly bind those collaborations. For example, IHE is creating profiles based on FHIR. With the new focus on imaging standards, particularly 3D images, close cooperation between SDOs become critically important. Obviously, FHIR resources can be defined to encapsulate these objects.

5 Stakeholder Interoperability

A lesson learned from years of experience is that unless critical stakeholders are engaged and supportive, no new initiatives will change the current system. In its simplest way, I suggest healthcare is a matter of defining the problem, administering the appropriate medication, and monitoring the result. That works only if the payers will pay for it.

Stakeholders are key to interoperability. It is important to know who plans the strategy and who makes the decision about healthcare infrastructure. That may vary among countries – which then may influences differences that must exist in healthcare IT systems. In the US, I think the most influential stakeholders are payers and pharma. Next are the government, specifically FDA, CMS, CDC, and ONC.

Laboratory vendors play a secondary role because that data is critical to defining the problem and measuring the effectiveness of the treatment. Consumers are becoming increasingly important but lack an organization to influence. Beyond that come other government agencies, researchers, health IT venders including EHR vendors, academic medical centers, and providers of care. An important observation is that to maximize the use of technology for better clinical interoperability, the reimbursement process must become secondary to clinical care process.

6 Consumer Interoperability

Ultimately, the purpose of the health care system is the health and well-being of the person. Until recently, the consumer has been the silent and invisible partner of healthcare. We could not have access to our health data, our preferences were never asked, and we were dominated by the healthcare system. That now is a changing world.

We now recognize the value of aggregating a patient's data into a single record - the Patient-Centric EHR. A simple challenge that has been difficult to achieve is a person's medication history. Patients typically will have medications including immunizations administered in more than one site. In the absence of a universal person identifier, there is a significant error rate in identifying a person who has data in several systems. We have created algorithms to identify persons but the error rates are still prohibitive to aggregate patient data with the required accuracy. Data frequently are entered into the wrong person's record. Clinical trials across multiple systems are biased by duplication of records. Most countries do have a universal person identifier; the US does not.

The healthcare process is supported from an IT perspective by EMR or EHR systems, and from a person perspective by a Personal Health Record (PHR). Typically, all functionality related to capturing, analyzing, presenting, and using that data is contained with the EHR. Unfortunately, much data created from the patient is not contained within the EHR. Caregivers and research complain of the difficulty of getting access to the data for any external purpose, such as used in a CDS algorithm or populating a registry. Furthermore, users are limited to the functionalities provided by the vender. Most of the dominant EHR systems are more than 40 years old. Technology has progressed well beyond the technology that exists in these commercial systems. To increase users' interoperability, I propose we create a new approach. We replace the current EHR with a Digital Data System (DDS) whose sole purpose is the intake, storage, and output of data. No functionality other than data in and data out exists. Functionality exists outside the DDS and permits a competitive environment among vendors, incorporates new and changing technology as well as new requirements, and permits specialization among specialties. Furthermore, such an approach would require the use of a common set of data elements for partication in the DDS. To the rest of the world, the DDS could function as a black box, without worrying about its internal workings, only about its performance.

Patients in most countries now have access to their health data either by a full record download or through a browser to view data. Rather than have a PHR, programs would exist to retrieve data from the DDS as required by the patient. It seems that a primary reason for providing access of a patient to their health data is to control that data. I think that misses the point. I want access to that data to better understand my health and how to manage it. First, my institutional EHR system contains data only when I am sick or when I visit for my annual exam. In my case, that data is a set of lab tests, vital signs, a problem list, and some demographic data. I look at it to see what is within normal limits. If a test has results outside number limits, I educate myself about what I can do to bring it back within normal limits. That might be a behavior change, or it might be a visit to the doctor. In any case, it is only a snapshot of my health status. What I want to do is to monitor my status as I live my life with my daily activities. Technology now permits that to happen with technologies including wearable sensors with real time data collection. What I want is a system that analyzes these data streams and makes decisions about my current state. I want a system that puts my data back into my healthcare system and alerts a provider when appropriate.

7 Bussiness Interoperability

Business aspects of the healthcare system dominate all other components of the system. The finances of an institution appropriately drive the Chief Financial Officer. Decisions about with which data may be shared are influenced by the CFO. Policies are driven by the business concerns. Most of the analytics done on health data are performed for business purposes. The balance between financial and healthcare delivery are a critical decision for an institution.

7.1 Governance Interoperability

As organizations share more and more of data and resources, governance becomes an important factor in interoperability. Governance rules are critically important among groups whether health care organizations, SDOs, affiliations, collaborators, or patients. Governance must establish a set of rules precisely defining ownership, flow, what can be done, and other activities.

7.2 Regulatory Interoperability

The purpose of regulations is the safety of the person. Regulations are critical for patient safety. Today we live in a rapidly changing world. Regulations should be reviewed frequently to be sure they are neither too stringent nor too

loose. Interoperability does depend on matching patients across multiple databases. Regulations in the US currently make that impossible to do without an unacceptable error.

8 Private and Security Interoperability

Addressing privacy requirements is one of the challenging problems in interoperability. Both the Health Insurance Portability and Accountability Act (HIPAA) [19], enacted in 1996, and the newly enacted General Data Protection Regulation (GDPR) [20] define rather strict rules that regulate the exchange of data. Both sets of regulations control what data can be exchanged, and particularly identified data. The rules are focused on the privacy of an individual. Patient consent is generally in the use of identified and sometimes de-identified data. I suggest that for interoperability, a backwards linkage to a patient should be possible. If as a result of a clinical trial, new knowledge is generated that would have a positive effect on a patient that link-back would be important for better outcomes. Combining patient data of all types to create big data is critical for clinical research and the discovery of new data.

9 International Interoperability

Interoperability across a country's borders is perhaps the most challenging component to solve. First is the language barrier. Not all concepts translate across languages the same. In some cases, the translation is a physical description of the word. A single language used across all of health care is a possible solution, with demographics such as name and address be expressed in the native language. The obvious choice of such a language would be English, since most countries now teach English throughout all grade and high school levels. Another challenge is the accommodation of culture into the EHR and into treatment. Finally international interoperability accommodation for national drug formularies require continuous mapping from each country's formulary to a master set back to the second country. A global master formulary would save millions of dollars, but that is unlikely to ever happen.

10 Conclusion

It should be clear that I have not answered the question of how do you know when you have interoperability. However, the question does provide an opportunity to increase awareness of all the factors that have some influence on interoperability. From a technical perspective, we have examined more closely semantic functional interoperability. For semantic interoperability, the problem seems to be that too many groups are trying to solve the problems resulting in a non-solution. Controlled terminologies have increasing widespread use and have some intellectual value. Reimbursement, at least in the US, is likely to continue to be the dominating factor in the terminologies that are in use. The gap between what are clinically required vocabularies and the coding for reimbursement will likely be unchanged. We have chosen to use mapping among terminologies to avoid a valued solution. We do workarounds rather than solve difficult problems. Functional interoperability seems to be moving toward a workable solution with HL7's FHIR, SMART, and CDS Hooks. The agreements among the several SDOs will further contribute to functional interoperability.

Stakeholder interoperability is moving towards interoperability in that competitive groups are establishing trust and are defining what they require from the systems. By working together, these groups will define their requirements and share common solutions that can be provided from the SDOs working together.

Consumer interoperability continues to grow in importance and influence. What consumers want will push standardization in mobile devices, in wearable sensors, and in other Internet of Things. There seems to be less pushback on a universal unique personal identifier, and I think it is likely that we will finally adopt a UPID within a couple of years.

New technologies, new healthcare delivery models such as value based care, new policies and other change driving events will demand new business models. Although financial considerations will remain as the strong driving force, business models will better relate to where the other components of interoperability are going. Security will remain a major interest with new steps to contain hacking. Privacy will change to better address what is required for better health for individuals and research to provide new knowledge. As clinical research moves to pragmatic clinical trials using EHR data, the consistency and quality of EHR data will improve. Regulations also will address the use of AI and robotics as part of healthcare delivery. Governance is an essential component of data sharing among institutions.

The prospect for the future looks bright. Quality of life as well as length of life will improve for most of society. Population health will increase the focus on communities and on disparities in those communities. Developing countries will benefit from new technologies and new models of care. Interoperability is good.

11 Acknowledgement

Research reported in this publication was supported by the National Center For Advancing Translational Sciences of the National Institutes of Health under Award Number UL1TR002553. The content is solely the responsibility of the author and does not necessarily represent the official views of the National Institutes of Health.

References

[1] Geraci A, Katki F, McMonegal L, Meyer B, Lane J, Wilson P, et al. IEEE Standard Computer Dictionary: A

Compilation of IEEE Standard Computer Glossaries. New York: IEEE Press; 1990.

[2] www.HL7.org

[3] HIMSS Dictionary of Healthcare Information Technology Terms, Acronyms and Organizations, 2nd Edition, 2010, Appendix B, p190, original source: Wikipedia.

[4] HIMSS Dictionary of Healthcare Information Technology Terms, Acronyms and Organizations, 2nd Edition, 2010, Appendix B, p190, original source: HIMSS Electronic Health Record Association.

[5] Blobel B. Standardization for mastering healthcare transformations-challenges and solutions. Eur J Biomed Inform. 2017; 13(1): 09-15.

[6] Blobel B. Interoperable EHR Systems-Challenges, Standards, and Solutions. Eur J Biomed Inform. 2018; 14(2): 10-19.

[7] Heubusch K. Interoperability: What it Means, Why it Matters. J AHIMA. 2006; 77: 26-30.

[8] www.who.int/classifications/icd/en/

[9] www.snomed.org/snomed-ct

[10] en.wikipedia.org/wiki/LOINC

[11] www.nlm.nih.gov/research/umls/rxnorm

[12] Blobel B and Oemig F. Solving the Modeling Dilemma as a Foundation for Interoperability. Eur J Biomed Inform. 2018; 14(3): 03-12.

[13] McGinnis JM, Williams-Russo P, Knickman JR. The case for more active policy attention to health promotion. Health Aff. (Millwood) 2002; 21: 78-93.

[14] www.hl7.org/fhir/

[15] smarthealthit.org

[16] oauth.net

[17] wiki.hl7.org/index.php?title=201701_CDS_Hooks

[18] www.healthcare-informatics.com/blogs/david-raths/interoperability/what-cds-hooks-interview-josh-mandel-about-fhir-and-clinical

[19] en.wikipedia.org/wiki/Health_Insurance_Portability_and_Accountability_Act

[20] www.theagenci.com/gdpr/

A Comparison of Business Rule Management Systems and Standards for the Implementation of Clinical Decision Support Systems Using Data from Structured CDA Documents

Peter Seifter[1,2]*, Stefan Sabutsch[2,3], Jeroen S. De Bruin[1,4]

[1]Institute of eHealth, Department of Applied Computer Science, FH Joanneum University of Applied Computer Sciences GmbH, Graz, Austria

[2] HL7 Austria, Austria

[3] ELGA GmbH, Austria

[4] Section for Artificial Intelligence and Decision Support, Center for Medical Statistics, Informatics, and Intelligent Systems, Medical University of Vienna, Vienna, Austria

Abstract

Background: Recently, the Austrian Patient Summary (APS) and Austrian Microbiology Report (AMR) document definitions were added to the Austrian national electronic health record repository. How to employ these standardized, structured patient data in clinical decision support (CDS) systems remains an interesting research subject.

Objectives: To create a proof of concept for the emergency use of APS and AMR documents in a variety of CDS platforms and standards in an intensive care setting, in order to generate new insights by linking individual documents to different CDS approaches.

Methods: APS and AMR definitions were stored using MongoDB, a document-oriented database system. For implementation of the clinical use case CDS, we selected the Drools business rule management system, as well as the ArdenSuite software for implementing the HL7 Arden Syntax for Medical Logic Systems standard.

Results: Due to its manifold features, implementation of the CDS use case in Drools could be done quite efficiently without the need for additional coding of technical or algorithmic code, which results in smaller rules. In contrast, the use case implementations in Arden Syntax using the ArdenSuite required additional technical coding, distracting from the medical knowledge implementation. The Arden Syntax as a knowledge definition standard is, however, better understandable due to its resemblance to natural language compared to the Java-like definition language used in Drools.

Conclusion: With the nation-wide availability of structured documents, the foundation for the implementation of clinical decision support systems has been laid. Commonly used open-source platforms offer extensive possibilities for the implementation of CDS systems. This evaluation of modern business rule management systems will advance the implementation of powerful clinical decision support solutions valuable to all stakeholders.

Keywords

Clinical document architecture; Patient summary; Microbiology laboratory report; Decision support systems; Clinical; Decision support techniques

Correspondence to:

Dr. Peter Seifter
Institute of eHealth, Department of Applied Computer Science, FH Joanneum University of Applied Computer Sciences GmbH, Graz, Austria.
E-mail: peter.seifter@fh-joanneum.at

EJBI 2018; 14(3):21-31
Received: March 19, 2018
Accepted: May 19, 2018
Published: July 06, 2018

1 Introduction

As the volumes of medical data generated in healthcare become greater, and the heterogeneity of these data continues to increase as well, the need for support in evaluation and interpretation of these data increases as well. As a result, the demand for clinical decision support (CDS) systems has steadily increased. Over the years, various CDS systems for a large variety of medical specialties and purposes have been created, with varying success, e.g., in infection control [1, 2].

One of the prerequisites for creating successful, interoperable CDS systems is the availability of structured, standardized data sources. Structure in data improves its

usability in digital processing, including CDS system, while semantic standardization through the use of code standards and value sets allow interpretation across healthcare institutes, and even across borders. In our approach, we restricted data sources to documents structured in the Clinical Document Architecture (CDA), an international document markup standard that specifies the structure and semantics of clinical documents for the purpose of exchange between healthcare providers and patients [3]. CDA documents are easy to use because of their in machine-readable format and well-known access methods. Furthermore, due to their XML-based encoding, extraction of individual data elements is quite easy as well, and supported by standardized methods on many development platforms, including Java.

Whereas interoperability standards like CDA express structured aspects of a patient's health, (e.g., a test result, or an overview of a patient's allergies), clinical decision support can be used to create connections between these structured data elements in order to generate new information. In this process, new, higher-level insights in a patient's health or comprehensive views on a patient's health in a specific, expert context, e.g., infection control, are provided through combination and interpretation of data element from individual documents and medical knowledge. To this end, basic elements of a CDA document have to be transformed - depending on the context - for further examination and linkage with other data sources. Furthermore, based on this kind of representation of the source data, a formal knowledge representation system is required for the generation of new knowledge.

In this study, we explore the use of CDA documents for clinical decision support. To this end, we combined two types of CDA-based document standards, the Austrian Patient Summary and the Austrian microbiology lab report. The Austrian Patient Summary contains essential healthcare information intended for unscheduled (e.g. emergency) use, whereas the Austrian microbiology laboratory report provides relevant information for the observation and therapy of bacterial or other microbiological infections. Given an infection control use case in an intensive care setting, we demonstrate how decision support can be applied to combine aforementioned data sources. To enable comparison between different knowledge representation standards in decision support, we implemented the use case in both Arden Syntax [4], a knowledge-based clinical decision support system standard, and in Drools [5], a general-purpose business rule management framework.

2 Methods

2.1 Clinical Document Interoperability Standards

Central to this study is the combination of heterogeneous medical information, which is spread over different clinical documents. CDA is the key to the harmonization of health data structure and CDA Release 2 was stipulated as the relevant document standard. In a nationwide specification process for these CDA documents the main stakeholders of the Austrian health system have developed so called "CDA Implementation Guides" for classes of documents like the "Physician's Discharge Summary", the "Nursing Discharge Summary", the "Laboratory Report" and the "Diagnostic Imaging Report" on a consensual basis.

In this section we provide a short overview of other standard documents, which are used in this study (a comprehensive discussion falls outside of the study scope):

- The Austrian Patient Summary document based on the International Patient Summary document and
- The Austrian microbiology lab report.

International and Austrian Patient Summary: According to HL7 IPS [6], the International Patient Summary (IPS) document is "an electronic health record extract containing essential healthcare information intended for use in the unscheduled or unplanned, cross-border care scenario". In this context, the scope of "essential healthcare information" is defined by the required elements of the IPS dataset, which is a specialty-agnostic, condition-independent minimal and non-exhaustive patient summary dataset, readily usable by clinicians.

The current IPS implementation comprises a set of templates and profiles based on HL7 Clinical Document Architecture (CDA) Release 2 [3] and Fast Healthcare Interoperability Resources (FHIR) [7] resource profiles, with value sets to support standardized coding of data elements. As the use of IPS documents is cross-jurisdictional in nature (both on an international and national level), this implies a need for common templates, and supported value sets based on international (multi-lingual) vocabularies [8].

In 2005 the foundation was laid for the establishment of an Austrian-wide electronic health record system (German acronym: ELGA). Since then, a variety of ELGA-supported structured, standardized document templates [9] were conceived and implemented in the majority of hospitals and at general practitioners, including document templates and resource definitions for electronic prescription, discharge summaries, laboratory report and radiology reports.

The Austrian Patient Summary (APS) [10] document is the latest document structure to have been developed. A patient summary working group was formed which met between 2016 and 2017 and, under the auspices of ELGA management, harmonized the APS content requirements and the terminologies to be used. The resulting APS document definition contains provisions for general demographic patient data (e.g. name, date of birth, gender), a summary of the insights and contents from the medical records of the patient (e.g. current medical problems, allergies and

the degree of intolerance, major surgical interventions, medical implants, inoculations) as well as the current medication. A sample of an APS document is shown in Figure 1.

In accordance with the IPS document definition, APS documents are structured XML documents that comply with the HL7 CDA R2 standard. The CDA Implementation Guide for the APS was created using ART-DÉCOR [11], an open-source tool that supports the creation and maintenance of HL7 templates, value sets, scenarios and data sets. The technical specification is based on the IPS Implementation Guide [6] and was subsequently published via a wiki.

The Austrian Microbiology Lab Report: The CDA implementation guide for the Austrian microbiology laboratory report (AMR) is an extension of the existing CDA laboratory report, which is based on the Integrating the Healthcare Enterprise (IHE) Sharing Laboratory Reports (XD-LAB) Integration Profile [12]. The AMR enables the caregiver to obtain relevant information for the observation and therapy of bacterial or other microbiological infections and to import the structured and coded data into their electronic health record (EHR). At the time of writing, the implementation guide is subject to the normative ballot process of HL7 Austria.

The AMR implementation guide includes sections on general report information, information on the collected test sample, and microbiological laboratory results. The general report information section provides administrative information clinical context on the reported result, e.g., the date of order entry, a patient's suspected diagnosis, requested examinations, and comments on report findings. The section on the collected test sample contains information on the test specimen, e.g., its time of collection, material type, the procedure with which it was extracted, as well as comments on the specimen quality. Finally, the results section contains all relevant result information, e.g., microscopic and macroscopic information, detected pathogens and their antibiotic resistance patterns, infection serology, etc. A document sample of the AMR is shown in Figure 2.

For the standardization of entry values, various coding mechanisms have been used, including SNOMED CT [13], LOINC [14], HL7 value sets (e.g., observation-interpretation, specimen type, …) and ELGA value sets used to complement aforementioned code systems.

Patient Summary

Created 20. Juli 2017 at 21:43 | version: 1

▼ Expand the table of contents ▼ Expand all contents

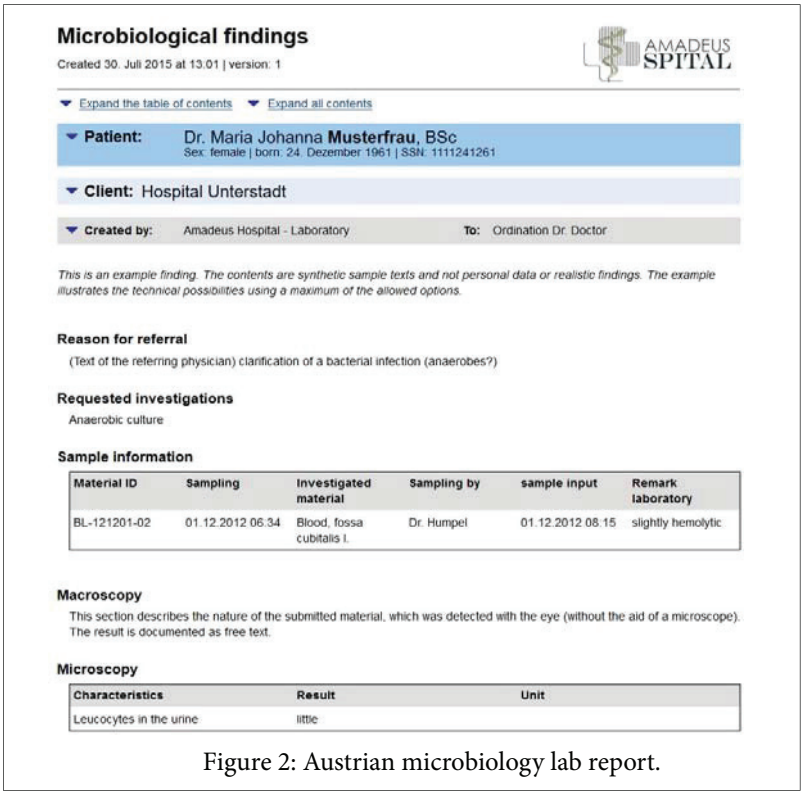
▼ Patient: Dipl.-Ing. Maria Hannelore **Muster**, BSc, MBA
Sex: female | born: 24. Dezember 1961 | SSN: 1111241261 |
Legal representative available

▼ Created by: Amadeus Spital - Surgical Department **Awarded:** (unavailable)

Allergies and intolerances

Triggering substance	Hypersensitivity	
Metamizol	Allergy	• known since 2013-04 • reaction: Larynxödem (severity: medium) • life-threatening • diagnostic confidence: approved • clinical status: consisting • informant: self-declaration of the patient • source: Discharge letter 2015-03-18 Amadeus Spital Salzburg
Peanut (Arachis hypogaea)	Intolerance	• known since 2002-12 • reaction: diarrhea (severity: medium) • reaction: Urtikaria (severity: severe) • non life-threatening • clinical status: consisting • clinical status: consisting • informant: self-declaration of the patient • source: Discharge letter 2015-03-18 Amadeus Spital Salzburg

Figure 1: Austrian patient summary document sample.



Microbiological findings

Created 30. Juli 2015 at 13.01 | version: 1

AMADEUS SPITAL

Expand the table of contents | Expand all contents

Patient: Dr. Maria Johanna **Musterfrau**, BSc
Sex: female | born: 24. Dezember 1961 | SSN: 1111241261

Client: Hospital Unterstadt

Created by: Amadeus Hospital - Laboratory **To:** Ordination Dr. Doctor

This is an example finding. The contents are synthetic sample texts and not personal data or realistic findings. The example illustrates the technical possibilities using a maximum of the allowed options.

Reason for referral

(Text of the referring physician) clarification of a bacterial infection (anaerobes?)

Requested investigations

Anaerobic culture

Sample information

Material ID	Sampling	Investigated material	Sampling by	sample input	Remark laboratory
BL-121201-02	01.12.2012 06:34	Blood, fossa cubitalis l.	Dr. Humpel	01.12.2012 08:15	slightly hemolytic

Macroscopy

This section describes the nature of the submitted material, which was detected with the eye (without the aid of a microscope). The result is documented as free text.

Microscopy

Characteristics	Result	Unit
Leucocytes in the urine	little	

Figure 2: Austrian microbiology lab report.

2.2 Methods and Tools for Knowledge Definition and Inference

For our CDS use case we created a knowledge-based CDS system. Knowledge-based systems are production rule systems, for which a collection of rules and restrictions are defined. These rules are then evaluated with actual patient data (facts) using an inference mechanism.

After evaluation of different tools and standards, we chose to implement our CDS use case with one of the more popular general-purpose, open-source business rule management systems (BRMS) called Drools [5].

Apart from using a general-purpose tool, we also implemented our use case in a standard developed specifically for knowledge-based clinical decision support systems as well: the HL7 Arden Syntax for Medical Logic Systems [4].

The Business Rule Management Platform Drools: Drools is an open-source BRMS platform implemented in Java. Without going too much into the intricacies of Drools, the platform can be described as a collection of tools that permits the decoupling of data and logic, and allows reasoning over data within various business processes, including medical reasoning. As such, Drools requires that a data model is created on which rules are applied. In turn, declarative rules are defined that express constraints on the data model.

For this study, two Drools components are of especial importance. First, the Drools Workbench, which is the web-based user interface that allows authoring and management of business rules and the underlying data model. Second, the Drools Expert module, which is the declarative, rule based, coding and execution environment, which implements both forward chaining (data-driven analysis, based on the Rete / Rete-OO algorithm [15, 16] and backward chaining (goal-driven analysis). Using these two tools and the Drools core, we defined the data model for both the APS and the AMR and implemented the clinical use case.

ArdenSuite CDS platform: Arden Syntax is an HL7 standard for the computerized representation and processing of medical knowledge, e.g., treatment rules, diagnostic decision trees, and risk scores. An Arden Syntax knowledge base commonly consists of multiple modules, called medical logic modules (MLMs) [17, 18]; these MLMs partition the knowledge base in highly-cohesive knowledge artifacts, where each MLM should contain logic pertaining to a single medical decision.

In contrast to Drools, Arden Syntax makes no assumptions about an underlying data model. Instead, in Arden Syntax external data resources can be accessed using read or write statements, or within curly braces, where larger data query and retrieval operations can be defined. Execution of these operations is forwarded to the host system, where it is executed, after which results are returned to the Arden Syntax MLMs.

For this study, we used the ArdenSuite clinical decision support technology platform for the implementation and

execution of our clinical use case [19]. The ArdenSuite comprises an ArdenSuite server which is used for the storage, management, and execution of MLMs. Furthermore, it contains an ArdenSuite integrated development and test environment (IDE), which serves as an authoring and test tool for Arden Syntax MLMs. To access external data sources, the ArdenSuite comes equipped with a standard connector for Java Database Connectivity (JDBC)-compatible databases.

2.3 Clinical Use Case

Consider the following (simplified) clinical use case: An unresponsive patient required emergency heart surgery and was transferred to the intensive care unit (ICU) afterwards. After three days, the patient – still unresponsive – develops clinical infection symptoms, i.e., fever, increased need for vasopressin (noradrenaline). Increased infection parameters (elevated leucocyte counts and C-Reactive protein). To confirm suspicions of sepsis and determine its source, blood samples are taken and sent to the department of clinical microbiology. Simultaneously, the surgeon starts a broad-spectrum antimicrobial therapy (either amoxicillin or meropenem). Two days later, the microbiology test results confirm the presence of Methicillin-Resistant Staphylococcus Aaureus (MRSA), upon which the treatment is changed to either vancomycin or daptomycin.

The aforementioned scenario falls in the scope of intended data use for both the IPS/APS and the AMR. In this case, essential healthcare information is provided for unscheduled (emergency) use. Despite the patient being in an unresponsive state, his/her allergies to various antibiotics, if any, are available in the IPS/APS. This provides invaluable information for both the initial broad-spectrum antibiotics treatment and the later treatment for MRSA: The AMR also provides essential, standardized information on antibiotic susceptibility of the MRSA pathogen which can be compared with available allergy information.

In this study, we implemented alert rules for both decision making scenarios (with respect to antimicrobial therapy) in a knowledge base:

1. The selection of broad-spectrum antimicrobial therapy, motivated and informed by allergies reported in the IPS/APS and the antimicrobial agent selected by the attending physician, and
2. The selection of antimicrobial therapy for MRSA, based on the antimicrobial resistance patterns of the MRSA pathogen reported in the Austrian microbiology report, the allergies reported in the IPS/APS and the antimicrobial agent selected by the attending physician.

For both scenarios, alerts are implemented that warn the physician in case an allergy to an antimicrobial agent is present; the alert also presents the severity of the intolerance (as shown earlier in Figure 1) to allow the physician to make a fully informed decision. Furthermore, in the second use case, a second warning is generated if an antibiotic is selected to which the pathogen has increased resistance (encoded in the data as either "intermediary" or "resistant").

3 Results

For document storage, we used the MongoDB Community Edition [20], a free, open-source, document-oriented alternative for the management of HL7 CDA documents. Although there are various benefits to using MongoDB, our primary reasons were both the availability of a Java driver and a JDBC connector, its ability to store data in flexible document that can change over time ("schema-free"), and its consistent performance in the management of complex CDA documents [21].

Data import from both CDA-based document and subsequent translation to Binary JSON (BSON) format was done using the open-source Model-Driven Health Tools (MDHT) library [22]. Using this library, we created classes for relevant information elements for both APS and AMR, instantiated them for each document in our test data, and serialized them as JSON objects to the MongoDB database.

Separate collections (collections are analogues to tables in relational databases) were created for each document type. Data access to the MongoDB database depended on the CDS implementation method (Drools or Arden Syntax). A graphical depiction of the process described above in Figure 3.

For the implementation of the use case, we used a simplified approach as a proof of concept. In the use case scenarios, medication names for patient allergies and pathogen antimicrobial resistance patterns are matched against a single antibiotic name proposed by the physician. A more advanced approach would take into account transformation of names to and from an ontology or thesaurus of medication families, e.g., as is defined in the Anatomical Therapeutic Chemical (ATC) Classification System, whereby we take into account allergies or resistances to a family of medications or active ingredients in a medication. This approach was deliberately omitted in this paper, and left for publication of the clinical use case implementation at a later stage.

3.1 Drools Implementation

For the implementation of our clinical use case in Drools, we first had to implement persistent objects for the underlying data model based on the APS and AMR document definitions. Based on these object definitions, we automatically generated facts (in the form of messages) that were inserted in the Drools rule engine (Figures 4a and 4b) show the message definitions for the APS and the AMR.

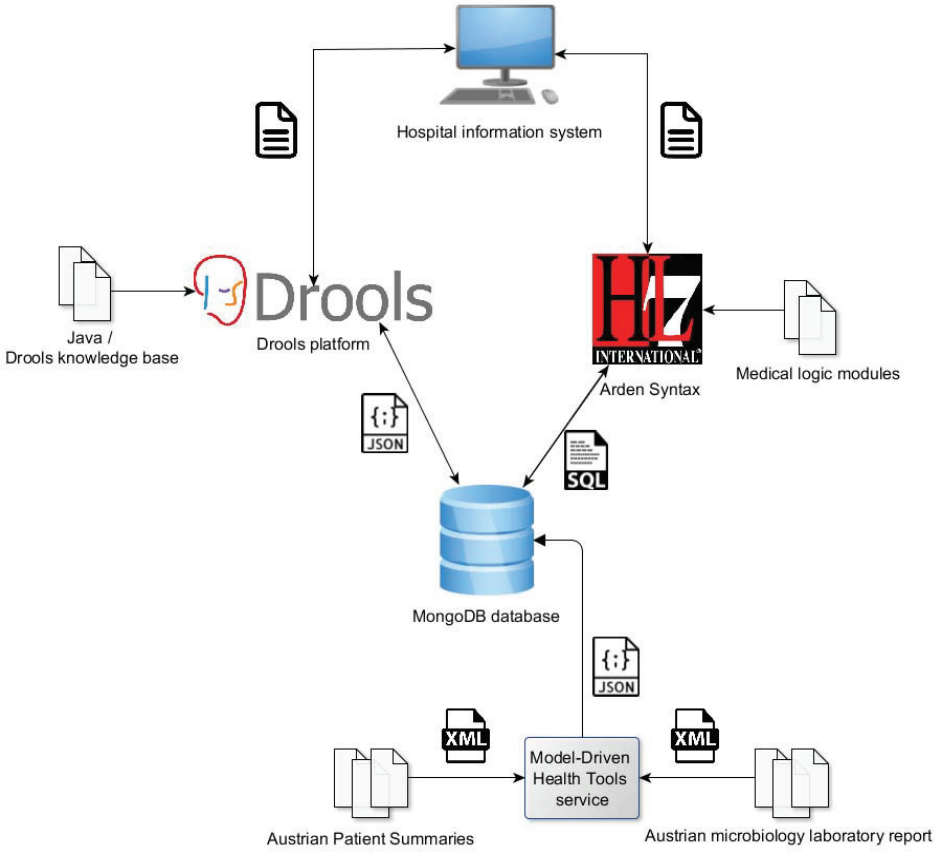


Figure 3: Data and information flow for our study use case implementation.

```
public static class APSMessage {

    private String Substance;
    private String Type;
    private String Reaction;
    private String Intensity;
    private String Severity;
    private AustrianPatientSummary aps;

    // Readers and writers...
    // ...
```

```
public static class AMRMessage {

    private String Pathogen;
    private String Substance;
    private String Susceptib;
    private AustrianMicrobiologyReport amr;

    // Readers and writers...
    // ...
```

Figure 4(a): Austrian Patient Summary fact definition in Java / Drools (4b). Austrian Microbiology Report fact definition in Java / Drools.

To generate facts for Drools, we wrote supporter functions that would unravel list and other containers in Java and thus create a number of facts by Cartesian product. An example: If a new microbiology report was detected, in which antimicrobial resistance patterns for *m* pathogens were tested with n different antimicrobial agents, then m*n facts were inserted; each fact contains a pathogen name, an antimicrobial substance, and a resistance indicator.

Because of the supporter functions, the knowledge base itself could be implemented in a straightforward fashion. Figure 5 shows the implementation of the first use case decision scenario, in a rule called "Find allergy". This rule is only fired if an antibiotic was proposed by the physician and if this antibiotic matches an allergy recorded in the Substance field of the APS message. If so, an *AllergyFound* message is created and passed on.

In this rule, the presence of a substance allergy (the member variable Substance of the message *$apsm*) is checked against *$antib*, the proposed antibiotic by the physician.

Similarly, a rule for the second decision scenario (named "Find resistance pattern") was defined, which is shown in Figure 6. Note that only the alert for increased resistance is implemented here; as Drools checks facts against all rules in the knowledge base, an explicit call to the first rule is not necessary.

In this rule *$antib*, the proposed antibiotic by the physician, is checked against the resistance pattern of a pathogen against the Substance recorded in the AMRMessage *$amrm*.

3.2 Arden Syntax Implementation

The resulting Arden Syntax knowledge base comprises four MLMs. Two MLMs provide supporting functions for both document types; for each document type, we implemented Arden Syntax object definitions, database query definitions and a function that transforms coded database query results into their decoded counterparts and returns those in the defined object. The remaining two MLMs implement the decision logic for both decision making scenarios in the use case using aforementioned support functions. Database access was provided by configuring the ArdenSuite DBConnector for our MongoDB server.

Analog to our Drools implementation, we created objects that model APS and AMR information necessary to our decision making processes. Figures 7a and 7b show (relevant parts of) the Arden Syntax object definitions for respectively the APS and AMR. Using these objects, the two decision making scenarios were implemented. The logic in the MLM for the first scenario is shown in Figure 8. In this MLM, we have to explicitly check for identifications, as we don't have an underlying persistent data model and instead need to use queries. Furthermore, we have to perform various type checks in order to correctly handle the data. As a result, a (list of) allergy objects is returned if allergies were found, otherwise the MLM returns NULL (return code not shown). In this module, the presence of allergies is checked against the parameter antib, again representing the antibiotic agent proposed by the physician.

```
package org.cda.aps.cds.drools

import org.cda.aps.ProposedAntibiotic;
import org.cda.aps.cds.drools.CDS_Drools.APSMessage;
import org.cda.aps.cds.drools.CDS_Drools.AllergyFoundMessage;

rule "Find allergy"
    when
        $antib: ProposedAntibiotic();
        $apsm: APSMessage(Substance == $antib)
    then
        AllergyFoundMessage $afm = new AllergyFoundMessage();
        $afm.setAps($apsm);
        insert($afm);
end
```

Figure 5: Drools rule "Find allergy" for the first use case decision scenario.

```
package org.cda.aps.cds.drools

import org.cda.aps.ProposedAntibiotic;
import org.cda.aps.cds.drools.CDS_Drools.AMRMessage;
import org.cda.aps.cds.drools.CDS_Drools.ResistanceFoundMessage;

rule "Find resistant pathogen"
    when
        $antib: ProposedAntibiotic();
        $amrm : AMRMessage(Substance == $antib &&
                (Susceptib == "R" || Susceptib == "I")
                )
    then
        ResistanceFoundMessage $rfm = new ResistanceFoundMessage();
        $rfm.setAmr($amrm);
        insert($rfm);
end
```

Figure 6: Drools rule "Find resistant pathogen" for the second use case decision scenario.

```
APSobject := OBJECT [
    Substance,  // name of the antibiotic substance
    AType,      // reaction type (allergy, intolerance)
    Reaction,   // (List) of reactions
    Intensity,  // (List) of reaction intensities
    Severity    // seriousness (e.g., life-threatening)
]
```

```
MRRobject := OBJECT [
    Pathogen,   // name of the pathogen found
    Amount,     // amount that was found
    Substance,  // (List) of substance test
    Susceptib   // (List) of susceptibilites (S, I, R)
];
```

Figure (7a): Austrian Patient Summary object definition in Arden Syntax (7b). Austrian Microbiology Report object definition in Arden Syntax.

Finally, in Figure 9, (part of) the MLM for the second scenario is shown. This MLM calls the previous MLM and additionally verifies that any pathogen found in the patient is not (intermediary) resistant to the antimicrobial agent proposed by the physician. Similar to the MLM in Figure 8, we again have to perform some type handling. As a result, a (list of) allergy objects and a list of resistance objects is returned if allergies or resistances were found, otherwise the MLM returns NULL (return code not shown).

In addition to calling the MLM previously discussed, this MLM also verifies that none of the found pathogens in the available microbiology reports are (intermediary) resistant to antib, the antibiotic agent proposed by the physician.

4 Discussion

In this paper we discussed the implementation of a clinical infection-control use case that employs patient data from two new standardized medical documents in Austria: The Austrian Patient Summary (based on the International Patient Summary) and the Austrian Microbiology Report. Data from these documents were extracted, and the clinical use case was implemented using two different knowledge definition standards: the general-purpose platform Drools, and HL7 Arden Syntax for Medical Logic Systems.

Both patient data document types are structured using CDA R2, but some data entries are optional. Furthermore, both document definitions are very recent and still subjective to change. Because of these sources of potential variability in document structure, we chose to store these documents in MongoDB, as there is no rigid schema definition, and because it allows for high loads due to its horizontal scalability (which is useful for data-intensive epidemiological applications). However, these benefits required that some traditional properties of classical data base management systems are changed or omitted (e.g., a lack of table joins). As of yet, little research has been done on the use of MongoDB in EHRs or for the storage of structured medical documents, but initial results are encouraging, showing good and consistent performance in the management of complex CDA documents, even in a non-optimized implementation [21].

```
// apsResults.Allergies can be an object or list of objects.
apsResults := CALL cdaApsMlm WITH lastName, firstName, DoB;

// No Allergies recorded
IF apsResults IS NOT PRESENT THEN
    allergyRetVal := NULL;
ELSEIF apsResults.Allergies IS NOT PRESENT THEN
    allergyRetVal := NULL;
// Only one allergy was found
ELSEIF apsResults.Allergies IS NOT LIST THEN
    IF apsResults.Allergies.Substance IS EQUAL antib THEN
        allergyRetVal := apsResults.Allergies;
    ENDIF;
// A list of allergies was returned
ELSE
    allergyRetVal := ();
    FOR allergy IN apsResults.Allergies DO
        IF allergy.Substance IS EQUAL antib THEN
            allergyRetVal := allergyRetVal , allergy;
        ENDIF;
    ENDDO;
ENDIF;
```

Figure 8: Medical logic module for the first decision scenario.

```
// mrrResults can be an object or list of objects
mrrResults := CALL cdaMrrMlm WITH lastName, firstName, DoB;

// Execute scenario 1
Allergies := CALL APS_Allergies WITH lastName, firstName, DoB;

// No results found
if mrrResults IS NOT PRESENT THEN
    resistanceRetVal := NULL;
// Single result/pathogen recorded
ELSEIF mrrResults IS NOT LIST THEN
    // no resitance test performed
    IF mrrResults.Substance IS NOT PRESENT THEN
        resistanceRetVal := NULL;
    ELSE
        ind := INDEX OF antib FROM mrrResults.Substance;
        IF ind IS NOT NULL THEN
            IF mrrResults.Susceptib[ind] IS EQUAL "I"
            OR mrrResults.Susceptib[ind] IS EQUAL "R" THEN
                resistanceRetVal :=  mrrResults;
            ENDIF;
        ENDIF;
    ENDIF;
ELSE
    resistanceRetVal := ();
    FOR mrr IN mrrResults DO
        IF mrrResults.Substance IS PRESENT THEN
            ind := INDEX OF antib FROM mrrResults.Substance;
            IF ind IS NOT NULL THEN
                IF mrrResults.Susceptib[ind] IS EQUAL "I"
                OR mrrResults.Susceptib[ind] IS EQUAL "R" THEN
                    resistanceRetVal := resistanceRetVal, mrrResults;
                ENDIF;
            ENDIF;
        ENDIF;
    ENDDO;
ENDIF;
```

Figure 9: Medical logic module for the second us case decision scenario.

Our choice for Drools as one of the platforms to implement our clinical use case was motivated by a short online study that we performed on popular open-source decision support tools. In the results of our online search, Drools appeared more often than other tools and platforms that we came across. Moreover, a search on PubMed showed that there are various publications on the successful implementation of CDS systems with Drools [23, 24, 25]. Our choice for Arden Syntax was a straightforward one and follows from the description and purpose of the standard: A clinical and scientific knowledge definition language that is used in a computer-executable format by clinical decision support systems [26].

When comparing our CDS implementations we have to distinguish between the implementation languages and their supporting frameworks. At first sight, the Drools implementation of our clinical use case seems shorter and easier, but this is for a large part because of the underlying Drools platform implementation. ArdenSuite and Drools offer similar functionalities i.e., both offer an IDE, web-based remote deployment, facilities for database access and workflow support in Business Process Modeling and Notation (BPMN). However, compared to the ArdenSuite, the Drools platform is richer in features and offers pre-implemented forward and backward chaining reasoning algorithms, which allow for shorter rule definitions without the need for additional technical or algorithmic implementations in the knowledge base.

A comparison of the knowledge definition languages themselves showed that the Drools rule definition language and its conventions remind strongly of the Perl language, and the Java programming language on which Drools is based. As such, it has a very technical appearance and is therefore harder to understand for those without a technical background. In this respect, the Arden Syntax has a clear advantage; the syntax was created for a broad, potentially non-technical public, thereby supporting operations not only tailored to use in the clinical realm, but also expressed in a syntax resembling narrative, natural language [26]. This makes understanding MLMs easier, which allows the MLMs themselves to serve as a communication device between knowledge engineer and clinician. By avoiding an intermediate representation, a potential source of translation errors is avoided.

There are various caveats and limitations to this study. Although the authors have a thorough understanding of CDA principles and the APS and AMR, and are proficient in the use of MongoDB, Drools and Arden Syntax, the various

data model and knowledge base implementations provided in this study are not guaranteed to be optimal. The comparison provided in this study is a proof of concept that needs further evaluation. Given that only a small rule base has been implemented, the scalability of the used tools has to be verified, both for the data model implementation in MongoDB and certainly for the used CDS platforms and standards. Nonetheless, this study yielded useful experiences and knowledge on the implementation of CDS systems in combination with CDA-based document standards. Furthermore, this study generated hypotheses on combining implementations, possibly using the Arden Syntax a knowledge engineering tool and then translating it to Drools projects to take advantage of the strength of both the Arden Syntax and the powerful underlying Drools implementation. To the authors' knowledge, a proof of concept for such an undertaking exists, but was never followed up on [27].

5 Conclusion

The selection of methods and tools to generate new insights in a clinical environment based on medical data is a key factor for correct and efficient knowledge creation. The need to manage heterogeneous medical information requires the use of interoperability standards for clinical documents. This requirement is met due to the ELGA project and the standardization activities in Austria.

With the nation-wide availability of structured documents, the foundation for the implementation of clinical decision support systems has been laid. Furthermore, linking individual patient health documents enriches the quality and insight of structured patient data.

The implementation of the clinical use case in Drools could be done quite efficiently, and allows for easy entry into the world of medical information management, while the implementation in Arden Syntax produced software solutions which are easy to understand due to its resemblance to natural language.

As future work, we plan to extend and advance this evaluation of modern business rule management systems, in order to advance the implementation of more powerful clinical decision support solutions valuable to all stakeholders.

6 Acknowledgement

The authors are indebted to their colleagues from SDOs such as HL7 for open minded and kind cooperation and valuable support. Furthermore, the authors would like to thank Dr. Martin de Jonge of the Gesundheit Nord GmbH Klinikverbund Bremen for his invaluable assistance in specifying the clinical use case.

References

[1] De Bruin JS, Seeling W, Schuh C. Data use and effectiveness in electronic surveillance of healthcare associated infections in the 21st century: a systematic review. J Am Med Inform Assoc. 2014; 21: 942-951.

[2] Freeman R, Moore LSP, García Álvarez L, Charlett A, Holmes A. Advances in electronic surveillance for healthcare-associated infections in the 21st Century: a systematic review. J Hosp Infect. 2013; 84:106-119.

[3] Health Level Seven International. HL7 Standards Product Brief - Clinical Document Architecture (CDA®) Release 2. http://www.hl7.org/implement/standards/product_brief.cfm?product_id=7

[4] Health Level Seven International. HL7 Standards Product Brief - Arden Syntax v2.10 (Health Level Seven Arden Syntax for Medical Logic Systems, Version 2.10). http://www.hl7.org/implement/standards/product_brief.cfm?product_id=372

[5] Red Hat, Inc. Drools-Business Rules Management System (JavaTM, Open Source). https://www.drools.org/

[6] Health Level Seven International. International Patient Summary - HL7 IPS. http://international-patient-summary.net/mediawiki/index.php?title=IPS_implementationguide_1

[7] Health Level Seven International. Fast Healthcare Interoperability Resources (FHIR), Version 3 (Standard for Trial Use). https://www.hl7.org/fhir/

[8] European Commission. International Patient Summary: Policy, deployment, competencies and standards. European Commission. https://ec.europa.eu/digital-single-market/en/news/international-patient-summary-policy-deployment-competencies-and-standards

[9] ELGA GmbH. Technical Basics. http://www.elga.gv.at/index.php?id=2

[10] HL7 Austria. ILF:Patient Summary Guide. https://wiki.hl7.at/index.php?title=ILF:Patient_Summary_Guide

[11] ART-DECOR Expert Group. ART-DECOR®.https://www.art-decor.org/mediawiki/index.php?title=Main_Page

[12] Integrating the Healthcare Enterprise (IHE). Laboratory Technical Framework. http://www.ihe.net/Technical_Frameworks/#laboratory

[13] Snomed International. SNOMED CT. https://www.snomed.org/snomed-ct

[14] Regenstrief Institute. LOINC -The freely available standard for identifying health measurements, observations, and documents. https://loinc.org/

[15] Forgy CL. Rete: A fast algorithm for the many pattern/many object pattern match problem. Artif Intell. 1982; 19: 17-37.

[16] Sottara D, Mello P, Proctor M. A Configurable Rete-OO Engine for Reasoning with Different Types of Imperfect Information. IEEE Trans Knowl Data Eng. 2010; 22: 1535-1548.

[17] Pryor TA, Hripcsak G. The Arden syntax for medical logic modules. Int J Clin Monit Comput. 1993; 10: 215-224.

[18] Hripcsak G. Writing Arden Syntax medical logic modules. Comput Biol Med. 1994; 24: 331-363.

[19] Medexter Healthcare GmbH. ArdenSuite CDS Platform. https://www.medexter.com/products-and-services/ardensuite

[20] MongoDB Inc. MongoDB 3.6. MongoDB. https://www.mongodb.com/mongodb-3.6

[21] Gomez A, Eijo J, Martínez von Scheidt M, Baum A, Luna D, Quirós F. MongoDB: An open source alternative for HL7-CDA clinical documents management. Proceedings of Open Source International Conference-CISL; 2013; Buenos Aires, Argentina.

[22] Model Driven Health Tools (MDHT). SourceForge. https://sourceforge.net/projects/oht-modeling/

[23] Singh H, Yadav G, Mallaiah R, Joshi P, Joshi V, Kaur R, et al. iNICU-Integrated Neonatal Care Unit: Capturing Neonatal Journey in an Intelligent Data Way. J Med Syst. 2017; 41: 132.

[24] Lardon J, Asfari H, Souvignet J, Trombert-Paviot B, Bousquet C. Improvement of Diagnosis Coding by Analysing EHR and Using Rule Engine: Application to the Chronic Kidney Disease. Stud Health Technol Inform. 2015; 210:120-124.

[25] Ongenae F, De Backere F, Steurbaut K, Colpaert K, Kerckhove W, Decruyenaere J, et al. Towards computerizing intensive care sedation guidelines: design of a rule-based architecture for automated execution of clinical guidelines. BMC Med Inform Decis Mak. 2010; 10: 3.

[26] Samwald M, Fehre K, de Bruin J, Adlassnig KP. The Arden Syntax standard for clinical decision support: Experiences and directions. J Biomed Inform. 2012; 45: 711-718.

[27] Jung CY, Sward KA, Haug PJ. Executing medical logic modules expressed in ArdenML using Drools. J Am Med Inform Assoc. 2012; 19: 533-536.

18th International HL7 Interoperability Conference (IHIC 2018)

The motto of the 2018 conference is:

"Mastering the Interoperability Challenge"

The IHIC is a forum for implementers and scientists to present and discuss concepts, models and implementations for innovative interoperable e-Health solutions. The conference aims to play the role of an interface between science, research and practice with regards to interoperability in the health and social care domain to share experiences and best practices.

It will present scientific papers on the one hand and demonstrations, practice reports, success and failure stories on the other hand.

Topics

- The advancement of interoperability
- Harmonization of interoperability standards and specifications among different SDOs
- Terminology and ontology challenge of interoperability
- Concepts and frameworks for Smart Interoperability Infrastructure Services
- Local, regional or national Electronic Health Record solutions
- Business Intelligence and Clinical Decision Support
- Security, safety and privacy
- Specification and implementation tools
- FHIR and CDA – controversy, coexistence, or synergy?
- "Show me your CDA" – CDA implementations at all levels
- Creating new clinical and integrated care pathways through effective information exchange
- Enabling patient and healthcare providers to interact in the new digital economy (mhealth, Internet of Things, cloud computing and many more)
- Handling patient consent and electronic identity in distributed healthcare settings

Standardizing Medical Quality Assurance and Control in Germany Based on HL7® FHIR®

Frank Oemig[1*] and Bernd Blobel[2,3,4]

[1] Deutsche Telekom Healthcare and Security Solutions GmbH, Bonn, Germany

[2] Medical Faculty, University of Regensburg, Germany

[3] eHealth Competence Center Bavaria, Deggendorf Institute of Technology, Germany

[4] First Medical Faculty, Charles University Prague, Czech Republic

Abstract

In Germany, vendors in cooperation with the responsible organizations have developed a powerful mechanism to supply data for medical quality assurance and control to public health agencies. However, this mechanism is proprietary and limited in some regards. On the other hand, by the current means it comprehensively documents the necessary requirements worth being analyzed in detail so that the shortcomings of the specification may be overcome by facilitating international standards. This paper provides the foundation for this approach and suggests a possible transition by introducing different possible options that can be facilitated independently.

Keywords

Quality assurance; Quality control; Messaging; FHIR; Public health reporting; Rules

Correspondence to:

Dr. Frank Oemig

Deutsche Telekom Healthcare and Security Solutions GmbH,

Friedrich-Ebert-Allee 140, 53113 Bonn, Germany.

E-mail: Frank.Oemig@t-systems.com

EJBI 2018; 14(3):32-36

Received:March 21, 2018

Accepted: May 24, 2018

Published: July 06, 2018

1 Introduction

Introducing communication and interoperability standards in Germany is a difficult task, not only because the market is separated by law into distinct sectors: ambulatory and stationary. Unfortunately, within these sectors the responsibilities are associated with different stakeholders without the necessity to agree on a single solution. In addition, each stakeholder always stresses that his requirements are very specific for the respective sector and are only relevant for the German health system. Therefore, they have the opinion that the migration to an international communication standard like HL7® v2.x, Clinical Document Architecture (CDA®) or Fast Healthcare Interoperability Resources (FHIR®) [1, 2] is inappropriate, not necessary or even impossible.

In Germany, hospitals and - to some extent - also physicians of the ambulatory sector are required by national law to deliver quality assurance (QA) and control data to public health agencies. The overall QA process lies in the responsibility of the national Institute of Quality Control and Transparence in the Health Sector (IQTIG - Institut für Qualitätssicherung und Transparenz im Gesundheitswesen [3]). The IQTIG was authorized by the so-called G-BA (Gemeinsamer Bundesausschuss) [4], which was installed 2004 as a common decision-making organ for the different institutions of the public health sector.

This paper analyses the current features and requirements of quality assurance and control mechanisms in Germany and discusses a possible migration to HL7 FHIR.

2 Methods

The QS-Basisspezifikation (QA base specification), is the result of a joint effort between a group of German software vendors and the institute currently being responsible for QA, IQTIG, and its predecessors, the Bundesinstitut für Qualität und Patientensicherheit (BQS) [5] and the aQua institute [6]. The latter still publishes specifications on a similar basis, e.g. for the clinical cancer registries. Currently, these specifications are distributed (via download) within a ZIP file, that contains - along with the developer documentation and XML style sheets - two Microsoft Access databases providing all the necessary details for assembling and transmitting the QA data.

The MS Access databases for the QA documentation [7] contain a double-digit number of relational tables that

currently represent the information about 26 QA modules (e.g., transplantations, decubitus prophylaxis, etc.), the corresponding forms for data entry (questionnaires) with labels, data elements and their attributes, rules controlling display and processing, coded information (value sets/short lists), format specifications for data export, and corresponding trigger events. In other words, the MS Access DB is used for representing the complete form-based data model in a computable fashion. From a pragmatic perspective, it is a very comprehensive database with a broad coverage of the necessary computational details because all information for data entry and communication is provided and in real use. Consequently, the vendors are able to import the contents of the database into their individual software solutions and to generate with a minimum manual postprocessing executable applications. Figure 1 demonstrates a small snippet of such a form as generated by one of the vendors. It contains structural information (indentation) with text (labels), information with possible data and fields for data entry. The latter is shown as small underscores in Figure 1.

In the first step of the process for generating the export form for QA data, trigger events are defined to indicate and control when a new form must be instantiated. The trigger events are evaluated based on data the specialized QA systems have received from the HIS/CIS like all other subsystems by means of HL7 v2 messages. For instance, they monitor the primary diagnosis and procedure codes as the primary trigger for starting the reporting process. Once a form is triggered (2nd step), the QA system generates the form and instantiates it with data from previously received messages as far as possible. This step requires a proper identification of the corresponding data by manual inspection and analysis of the database contents as a precondition to generate the applications. In the 3rd step, all remaining (missing) data must then be entered manually because it can neither be taken from the already received data nor be retrieved from the originating HIS/CIS because of missing semantic details allowing for corresponding queries. During the data entry process, rules control appearance (visbility) of sections and consistency of entered data. Finally, the assembled data is then converted into the desired export format.

The tables and their relations as contained in the database are examined with regard to their semantic contents and the intended functionality according to the previously mentioned steps.

3 Results

The data contained in the database requires a comprehensive approach allowing for comprising all aspects. The upcoming FHIR standard seems too qualified and appropriate for that purpose.

3.1 Mapping to FHIR Resources

The analysis presented in this section is primarily based on the QA base specification, whose (main) concepts are shown in the left column in Table 1 and in Figure 2. The middle column proposes a mapping to FHIR elements, which is explained in the following with more details.

The QA base specification consists of modules for certain topics, e.g. decubitus prophylaxis. Typically, they consist of a set of hierarchically structured and related forms belonging to that topic. Figure 1 provides a snippet as an example of a top-most "base form". In principle, each form could be represented by an inidividual Questionnaire resource itself. Unfortunately, this approach will complicate consistency and completeness checking across forms so that a representation as individual item groups within an overarching Questionnaire appears more appropriate.

Each form contains several fields, which are usually organized as groups of fields that also allow for nesting. The structure of the fields is maintained as groups with the fields as items within those groups. The content definition of the fields is represented as FHIR DataElement resources to which the Questionnaire items are referencing. This apporach allows for identification and reuse of fields across forms and modules.

Rules are facilitated to control data entry. Plausibility rules support data entry by validating the contents when

Decubitus Prophylaxis			
At least one form must be populated.			
	Decubitus		
	Number of decubitus?	1 - 99	_ _
	Grade and localization		_ _ _ . _ _
	Localization	B = both sides K = no localization mentioned L = left R = right	_
	Decubitus present during admission? (related to the decubitus at this location, independent of grade)	0 = no (no diagnosis present at admission to hospital) 1 = yes (diagnosis present at admission to hospital) 9 = unknown due to missing documentation	_

Figure 1: Part of a data entry form that belongs to the Decubitus Module DEK (translated to English).

Table 1: Mapping QA database to FHIR resources.

QA Database	FHIR	Comment
Module	Questionnaire +Data Element	Alternatively, an extension may aggregate different questionnaires into a group
Form	Questionnaire.item.group	Mapping as embedded element due to module concept.
Fields	Data Elements (Items)	A separate representation allows for re-use that can be used within Questionnaires
Rule	Extension	Captures either FHIR path expressions or original proprietary language
Shortlist	Value Set	
Trigger Export format	EventDefinition (+TaskPlan) Resource bundle or Structure Definition	Ideally, proprietary export formats should be replaced by already existing formats like FHIR resources; alternatively, the original specification may be expressed as Structure Definitions

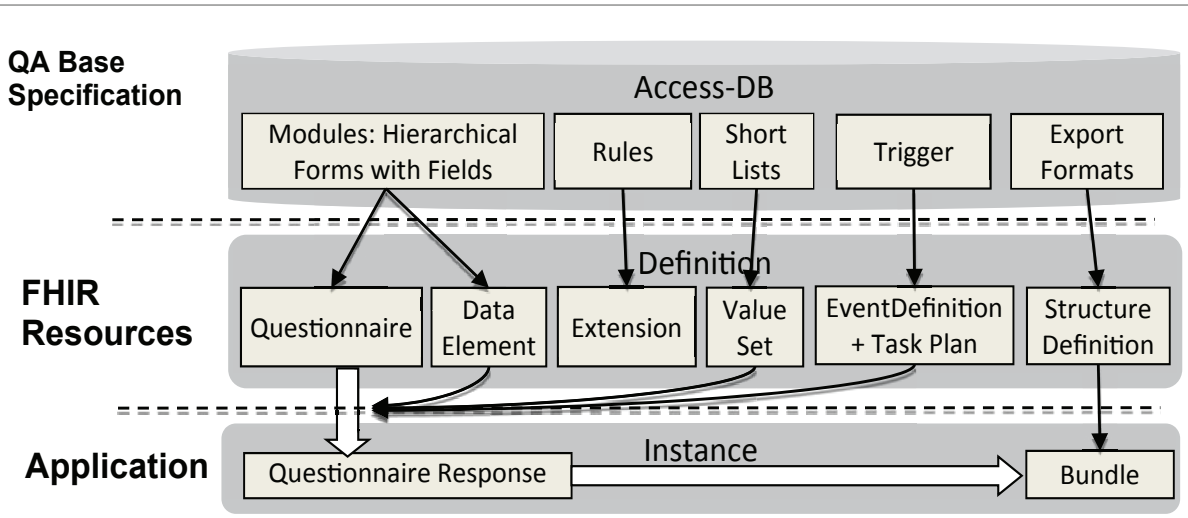


Figure 2: QA Mapping to FHIR Resources and instances.

populating the forms. Rules can apply to either single fields or to previously defined groups of fields. This capability ensures that only valid and complete data can be entered. Other rules trigger the existence of specific parts of the data set depending on the values of other fields of the form. A good example is about pregnancy: This data group is made visible for female patients only. Answering this question with "yes" (checkmarked) may then trigger a field "week of pregnancy" asking for a one- or two-digit decimal number in the range of 0 to 42.

For some fields the allowed values are specified as a shortlist (combobox), maintained as entries in a table of the database. For instance, the field "Localization" can have the values "B" (both sides), "K" (not specified), "L" (left), or "R" (right). Other aspects are the definition and evaluation of trigger events for selecting cases that have to be reported, and the way relevant QA data is generated from the entries in the forms. The trigger mechanism could be represented using the FHIR Event Definition resource.

To ensure data privacy, no personal information of a patient is exposed to the public health agency, i.e., identifying information such as name, address, IDs, demographic information is removed from the data set by not including it in the export specification. Exported QA data can be represented via resource bundles or logical models in form of StructureDefinitions.

3.2 Representing and Mapping Rules

As mentioned earlier, rules are used to validate values entered for a single field or a group of fields. Also, some parts of the form are only shown to the user under certain conditions expressed by rules. This latter kind of rules can be expressed by dependencies represented as attributes of the questionnaire resource ("enableWhen"). But plausibility rules require an expansion to FHIR resources as demonstrated

```
<extension url="http://xx.org/fhir/StructureDefinition/Questionnaire-
validation-expression" >
  <extension url="details" >
    <valueCodeableConcept >
      <coding >
        <code value="G001"/>
      </coding>
      <text value="'Completion date' cannot be a future date"/>
    </valueCodeableConcept>
  </extension>
  <extension url="location" >
    <valueString value="linkId='CompletionDate'"/>
  </extension>
  <valueString value=".where(linkId='Section-G').item.where(linkId =
'CompletionDate').answer.value <= today()"/>
</extension>
```

Figure 3: A rule represented as a FHIR extension.

in Figure 3. The depicted rule expresses that a CompletionDate cannot be later than the value of "today".

Rules can be defined and included into a Questionnaire via extensions - alternative approaches (eg. StructureDefinition) might be viable, too.

However, the set of rules expressed in the MS Access database provide a valuable list of necessary functionalities. For example, a specific field might be required to have at least one value (or all values) from a specific value set, whereas another field might not be allowed to contain certain data. In the current version of the QA base specification, these details are not represented using a formal language with a grammar. Hence, the expansion of FHIRPath language with these functions would benefit from the requirements extracted (reverse engineered) from the MS Access DB whereas FHIRPath would become this base language.

Facilitating FHIR features like external references also overcomes the missing functionality by specifying queries for missing data avoiding reentering the data manually.

4 Discussion

The current QA base specification is a powerful specification that allows for an easy implementation of graphical user interfaces for identifying cases to be reported, filling in the QA forms, and finally submitting the exported pseudonymized data. However, there are several shortcomings of this approach. First, the representation of the specification as a MS Access database is proprietary. The lack of a clearly defined and harmonized semantics makes it difficult to import data from the primary systems, and the reported QA data can only be used for that specific purpose. Acceptance on an international level cannot be expected then.

Although there are first attempts to use the QA base specification for clinical registries like the clinical cancer registries, the limitations of this approach are preserved. Also, the different registries have different requirements, so that a "one-fits-all" approach is not viable. The limitations can be addressed by using HL7 FHIR in combination with a specific

implementation guide describing the specific usage and necessary additions (expansions).

There are different options in progressing towards using international standards. Of course, a full migration appears most useful, but immediately enforces the most workload to convert and transfer the whole specification. Instead, a stepwise migration would also be helpful. It may start with an encapsulation of value sets allowing for re-use within other communication scenarios.

Another useful step would be to replace the proprietary export formats with FHIR resource bundles. A less optimal, but possible solution would be the definition of logical models to represent the individual data elements.

However, a remaining challenge is the replacement of the grammatic-free language designed for the German QA process by FHIRPath expressions. Conversely, an expansion of the FHIRPath/ FluentPath expression language with constructs stemming from real-world use cases might be beneficial to FHIR. This, for instance, includes the possibility of user-defined function, and the inclusion of particular functions from the IQTIG specification.

5 Conclusion

Given the results and possibilities as described above, it appears to be the right point in time to introduce the German specific requirements to the international community, so that an improvement of FHIR resources may benefit from the lessons learned in Germany while improving the QA base specification.

In the case of QA and clinical registries, the ultimate specification to implement is decided by the regulatory bodies and their authorized institutions and not the software vendors. Hence, to take advantage from these options a political debate is necessary involving and convincing the relevant institutions.

6 Acknowledgement

The authors are indebted to thank the international community for investing an unbelievable amount of time working on the details of FHIR.

References

[1] Health Level Seven, International. http://www.hl7.org

[2] Fast Healthcare Interoperability Resources (FHIR), Standard for Trial Use (STU) 3. http://www.hl7.org/fhir

[3] IQTIG, Institute for Quality Assurance and Transparency in Health Care. http://www.iqtig.de

[4] G-BA, Gemeinsamer Bundesausschuss. https://www.g-ba.de

[5] Bundesinstitut für Qualität und Patientensicherheit (BQS). http://www.bqs-institut.de

[6] aQua Institut. https://www.aqua-institut.de/

[7] QA MS Access Database. https://iqtig.org/downloads/spezifikation/Einrichtungsbezogene_QS-Dokumentation/2017/V01/2017_QESUE_EDOK_RB_XML_V01.zip

European Journal for Biomedical Informatics is dealing with a range of topics which is included under the scopes of biomedical informatics. The focus of the journal is on the major aspects of implementation of information science in Medical sciences, which includes Biomedicine, Ontology, Genomics, Biometrics, Biostatistics, Health informatics, Medical informatics, Clinical informatics, Semantics Interoperability, Informatics applications in Imaging.

Original contribution in terms of research, review, short-communication, case reports, are welcomed from the authors, who are engaged in medical or academic research. The articles selected for publication, must have to go through critical process of peer-review as per journal norms. A well-constructed editorial board is the higher most determining authority for the journal, which is formed by highly experienced subject experts in the field of biomedical informatics. Section editors and guest editors are also engaged there to maintain topic specific proper evaluation of research manuscripts.

The selected papers of IHIC 2018 will be published in the Special Issue of the European Journal for Biomedical Informatics.

List of Published Special Issues

Issue	Type	Name
2018/2	event-based	Jana Zvárová Memorial Conference(JZMC 2018)
2017/1	event-based	Re-shaping Healthcare Systems(HL7 IHIC 2017)
2016/1	event-based	Interoperability is more than just technology(HL7 IHIC 2016)
2015/3	topic-based	Apps for Medicine, Health and Home Care - Elements of Safety and Effectiveness
2015/2	event-based	Concept, Models and Implementations for Innovative Interoperable eHealth Solutions (HL7 IHIC 2015)
2014/2	topic-based	Using Information to Improve the Quality of Care in Type II Diabetes in Primary Care
2013/3	topic-based	Sharing Knowledge and Tools for Decision Support in Biomedicine and Healthcare
2013/2	topic-based	Applications of Computational Techniques in Infectious Diseases and Clinical Research
2012/5	topic-based	Semantic Interoperability in Biomedicine and Health Care III
2012/4	event-based	Standards and Solutions for eHealth Interoperability (part 2)
2012/3	event-based	Standards and Solutions for eHealth Interoperability (part 1)
2012/2	topic-based	Supporting eHealth Applications by Legal Systems in Europe
2011/2	topic-based	Education in Biomedical Informatics and Bioinformatics
2011/1	event-based	Semantic Interoperability in Biomedicine and Health Care II
2010/1	event-based	Semantic Interoperability in Biomedicine and Health Care

A Model for Implementing an Interoperable Electronic Consent Form for Medical Treatment Using HL7 FHIR

Anna M Lackerbauer[1, 3]*, Alvin C Lin[2], Oliver Krauss[1], Jason Hearn[3] and Emmanuel Helm[1]

[1] Department of Advanced Information Systems and Technology, University of Applied Sciences Upper Austria, Hagenberg, Austria

[2] Department of Surgery, Faculty of Medicine, University of Toronto, Toronto, Canada

[3] Centre for Global eHealth Innovation, University Health Network, Toronto, Canada

Abstract

Background: For ethical, legal and administrative reasons, patients have to give explicit consent to a medical treatment.

Objectives: This paper identifies the design requirements for electronic treatment consent (eConsent) architecture, and subsequently proposes a model for the eConsent architecture based on the HL7 FHIR® standard.

Methods: Six requirements for the eConsent architecture were identified. A conceptual model for the system was then developed to address the identified requirements using HL7 FHIR.

Results: The proposed concept makes use of the existing consent model of HL7 FHIR, and includes additional resources for presenting the information to the patient. Moreover, it uses the SNOMED CT terminology to enable semantic interoperability with other health information systems.

Conclusions: The proposed eConsent architecture meets the identified requirements. That said, the system is limited by the low maturity of the implemented FHIR resources and the fact that the terminology is currently inexhaustive for the use case. Custom extensions of the used FHIR resources must be considered.

Keywords

Informed consent; Health information systems; HL7; Systematized nomenclature of medicine

Correspondence to:

Anna Lackerbauer
University of Applied Sciences Upper Austria, Hagenberg, Austria.
E-mail: anna.lackerbauer@fh-hagenberg.at

EJBI 2018; 14(3):37-47
Received: March 21, 2018
Accepted: May 21, 2018
Published: July 06, 2018

1 Introduction

This paper identifies the design requirements for electronic treatment consent (eConsent) architecture, and subsequently proposes a model for the eConsent architecture based on the HL7 FHIR® standard [1]. The eConsent architecture comprises template forms, the actual treatment information, the patient consent and the signature of the patient. All four elements are represented using HL7 FHIR resources [2] and can be integrated within a FHIR ecosystem.

1.1 Treatment Consent

For ethical, legal and administrative reasons, patients have to give explicit consent to a medical treatment [3]. The legal requirements for a patient to be capable of giving consent to their treatment are equal across different countries: the patient must *understand* the steps involved in their treatment and he/she must *appreciate* the treatment implications [4, 5, 6]. Some countries additionally define a minimum age for consent [4, 5].

To ensure these requirements are met, the informed consent process involves multiple elements, as depicted in Figure 1. Information is first exchanged during a mandatory discussion between the patient and physician, a session in which the physician may use information materials such as printed content or videos [7]. During the discussion, the patient is educated by the physician performing the procedure about the risks, alternatives and benefits associated with the treatment [8]. Throughout the discussion, the patient is provided with the opportunity to ask questions about the upcoming procedure. Once the patient has been fully informed and their questions have been answered, he/she can decide whether he/she wants to receive the treatment. This choice must be documented and accompanied by a signature from both the patient and the treating physician.

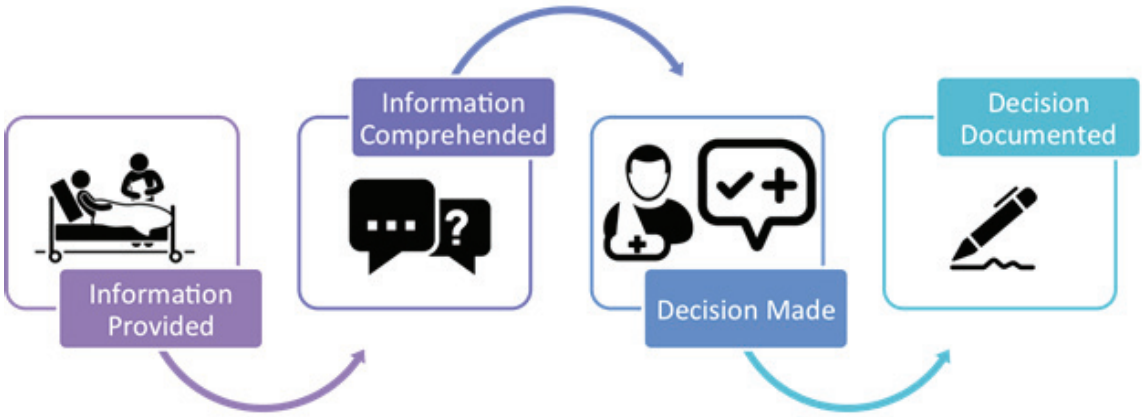


Figure 1: Elements of the informed consent process for medical treatment [5].

There is a large volume of work that has been done to digitize consent for research studies. As part of the data collection tool REDCap (Research Electronic Data Capture), the administration of electronic informed consent is already possible [9]. The literature proposes the usage of electronic informed consent to enhance recruitment in research studies [10].

2 Requirements for Treatment eConsent

Based on legal requirements, expert interviews and the academic literature, the following six requirements for the eConsent architecture were identified: simple creation, easy to understand, multi-language support, signature of various roles, rejection and withdrawal, and interoperability.

2.1 Simple Creation

The information that is given to the patient about the suggested treatment he/she will receive differs for each treatment. It will always contain the same, legally-required sections (i.e. risks, benefits and alternatives [8]), but the content of these sections must be defined individually for each treatment.

Within the eConsent architecture, this content should be entered into the system once for each new treatment. As many different types of treatments exist, the process of inputting new treatment information should require minimal effort. Additionally, the information input should contain all mandatory fields defined by the regulations of the government and/or institution in which the architecture is being used [8, 11]. As an example, the mandatory section *risks* for the treatment information about a laparoscopic appendectomy include information about the risks that apply specifically to that procedure:

- Bleeding,
- Infection,
- Damaging neighbor structures,
- Risk of opening.

A similar requirement concerning the creation of informed consent forms for research studies was addressed by the University of California San Diego Human Research Protection Program with their implementation of an informed consent assistant [12].

2.2 Easy to Understand

According to Schenker et al., patients often have poor understanding of the information they receive as part of the informed consent process [8]. Hall, Prochazka and Fink, as well as Schenker et al., describe multiple ways to increase the patient's comprehension of the provided information, whereas Farrell et al. and Wilson et al. report the effect of multimedia information materials in healthcare [3, 8, 13, 14].

The treatment eConsent prototype should support the following ways to potentially increase the patient's comprehension:

- Use of simple language,
- Manageable amount of information,
- Use of multimedia.

2.3 Multi-language Support

Patients whose native language is not the primary language of the hospital can be disadvantaged due to misinterpreted information [15]. To protect patients from misinterpretation, the eConsent solution should support multilingual information provision.

2.4 Signature of Various Roles

If the patient does not have the capacity to provide consent himself/herself (i.e. if the legal requirements of understanding, appreciation or a possible given minimum age are not met), another person must give consent on the patient's behalf [5, 11, 16]. In Ontario, the person signing on behalf of the patient does not necessarily need to be a relative. For instance, it can also be an attorney for personal

care or another authorized representative [6, 14]. In the case of an emergency, treatment may be administered without consent [5, 11, 17].

2.5 Rejection and Withdrawal

In the event that a patient does not want to receive a suggested treatment, they can opt to explicitly reject the treatment. It is crucial to educate the patient about the consequences of their decision and to document that this information was provided, particularly for urgent or medically-necessary treatment [5]. This is known as *informed refusal*. If the patient has already agreed to the treatment, he/she also has the option to withdraw their consent at any time prior to the surgical time out at the start of the surgery [18].

2.6 Interoperability

According to Palfrey and Gasser, interoperability and flow of information across multiple systems is a powerful tool that is crucial for success, increases innovation and fosters competition [19]. Therefore, the implementation should follow a healthcare standard and should be able to interoperate with other systems. Two use cases that demonstrate the importance of interoperability are as follows:

Patient information (i.e. the name, the date of birth and the patient's hospital record number) should be automatically obtained by an institution's electronic health record (EHR) system.

Structured, coded content describing the treatment information enables the patient and the healthcare provider to easily access and parse the details of a consent agreement.

3 Proposed Architecture

To ensure a standardized data model that is compatible with state of the art technologies such as REST, the chosen standard for the eConsent architecture is HL7 FHIR [1].

FHIR stands for Fast Healthcare Interoperability Resources. It is a new standard that is still under development and is currently released for trial use. The FHIR version that was used in the presented architecture is version STU3. FHIR is meant to be developer-friendly and supports widely-used standards for data-interchangeability and transfer, such as JSON, XML and HTTP. It is also an architecture that is based on the RESTful principles out of the box [1].

The resource-based standard already contains elements to model a patient's privacy consent. The treatment consent use case will be modeled by the responsible FHIR team in the future, and might change the given FHIR Consent resource accordingly [2]. This project tries to use existing resources to address the requirements of eConsent.

Figure 2 depicts the three major components required in the eConsent process:

Treatment information: This is the information that is shown to the patient about the treatment he/she will receive. This information typically includes a description about the procedure, as well as its risks, benefits and alternatives.

Consent: The *Consent* represents the decision a patient makes about whether or not he/she wants to receive the treatment. A later change of this decision (agreement/refusal) is also stored in this component.

Signature: To confirm the decision of the *Consent*, a signature of the patient or the consenting party is obtained.

Before the treatment consent process can start, the treatment-specific *Treatment Information* that is shown to the patient initially has to be created. As mentioned in section 2.1, there are mandatory fields that must be part of each *Treatment Information*. These mandatory fields are defined in the *Template*. The *Template* is modeled using the FHIR Questionnaire resource [2]. Based on one *Template,* there can be multiple *Treatment Information* instances that are modeled by the FHIR QuestionnaireResponse resource [2].

As an example, the *Template* could specify, that every *Treatment Information* of a given country or hospital (depending on the scope of the *Template*) must contain the elements risks, benefits and alternatives. An instance of this *Template* could be the *Treatment Information* for a laparoscopic appendectomy that specifies the risks, benefits and alternatives specifically associated with a laparoscopic appendectomy.

As part of the consent process, during the discussion between the physician and the patient, this *Treatment Information* will be shown to the patient as supporting material. The patient can then decide if he/she agrees or refuses the treatment based on the information he/she receives. This agreement/refusal is modeled by the FHIR Consent resource [2]. The actual signature (i.e. a picture/scan of the signature) that is linked to the Consent resource is represented by a FHIR Provenance resource [2].

3.1 FHIR Resource Mapping

To address the given requirements, the FHIR resources, especially the resources representing the information that is shown to the patient, can be modeled as described in this section.

Simple creation: For each type of treatment (e.g., laparoscopic appendectomy), there should be one specific *Treatment Information* that includes each of the mandatory fields defined in the *Template*. An outline of these mandatory fields is auto-generated from the *Template* Questionnaire and can be displayed as the headings of a form. Similarly, the input elements are auto-generated based on the data type that

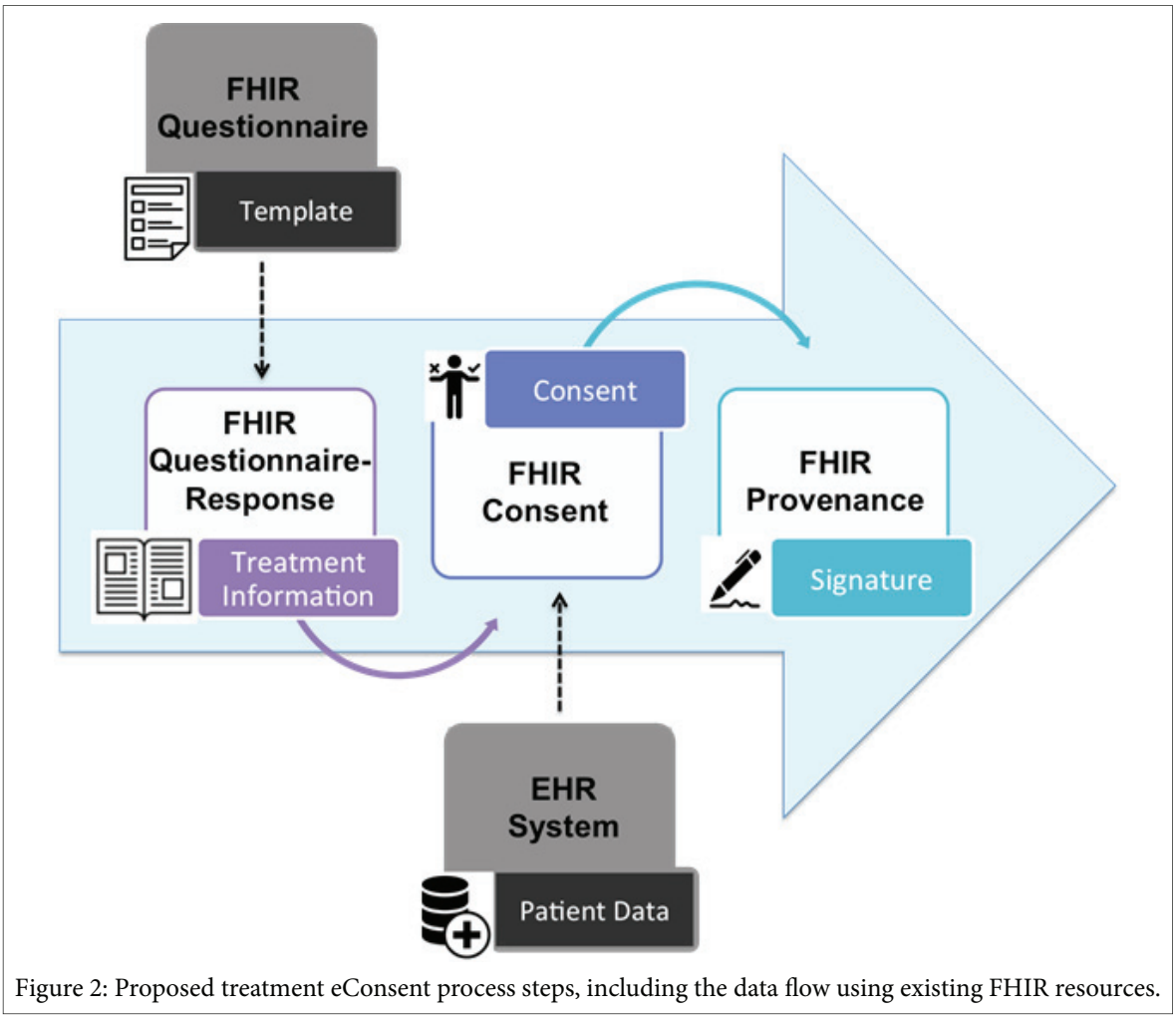


Figure 2: Proposed treatment eConsent process steps, including the data flow using existing FHIR resources.

is defined for that field in the *Template* Questionnaire. Depending on the required data type, the input fields can be displayed in an intuitive way, such as checkboxes or text input.

Figure 3 shows one possibility of how such a *Treatment Information* can be created. In this case, one can see auto-generated input fields of the data type *open-choice* and *text* [20]. The approach depicted in Figure 3 ensures that all required fields of the *Template* Questionnaire will be part of the resulting *Treatment Information* QuestionnaireResponse. As the *Summary* item of *Potential risks* is of the data type *open-choice* with elements predefined by a FHIR ValueSet resource [2], the user is given the option of choosing between common risks and/or creating new risks using the provided text input element.

Manageable amount of information: Too much information on the screen can overwhelm the patient. Accordingly, the content can be split into a summary and a more detailed *Learn more* section. As shown in Figure 4, the patient initially sees a summary of the information. On request (by selecting the *Learn more* button), additional information is provided.

Listing 1 shows how this item of the Questionnaire (*Template*) would appear as FHIR JSON format. A question of the data type *group* defines the overall *Potential risks* question with two sub-items: The *Summary* item for the information that is shown to the patient initially, and the *Learn more* item for additional information. In the auto-generated UI, this pattern can be identified by the following characteristics of the question:

- An item of type *group* as a parent,
- Two sub-items,
- LinkId of the first sub-item ends with "*.summary",
- LinkId of the second sub-item ends with "*.learnMore".

Use of multimedia: Similar to the previous example, the use of multimedia elements mixed with text can be handled by a question of the data type *group* with two sub-items, shown in Listing 2.

As the first part of the question, a textual description is required. In the second part, multiple multimedia items of type *attachment* can be added. The answer-item of the QuestionnaireResponse is an array, thus, on creation of the *Treatment Information,* the upload of multiple elements (e.g., videos and pictures) is possible. Figure 5 shows how this information, composed of text and multimedia-elements, is displayed to the patient.

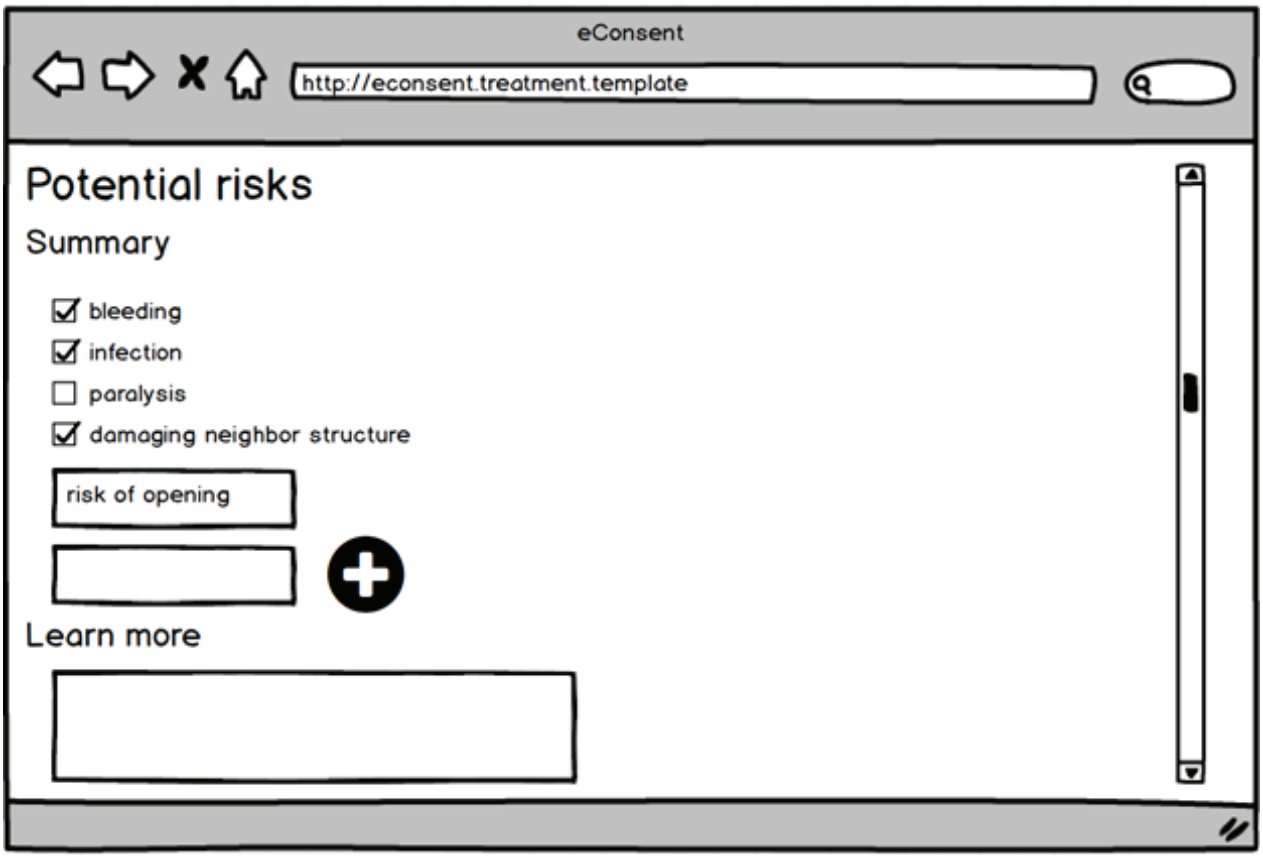


Figure 3: Example UI for the creation of Treatment Information content describing a laparoscopic appendectomy.

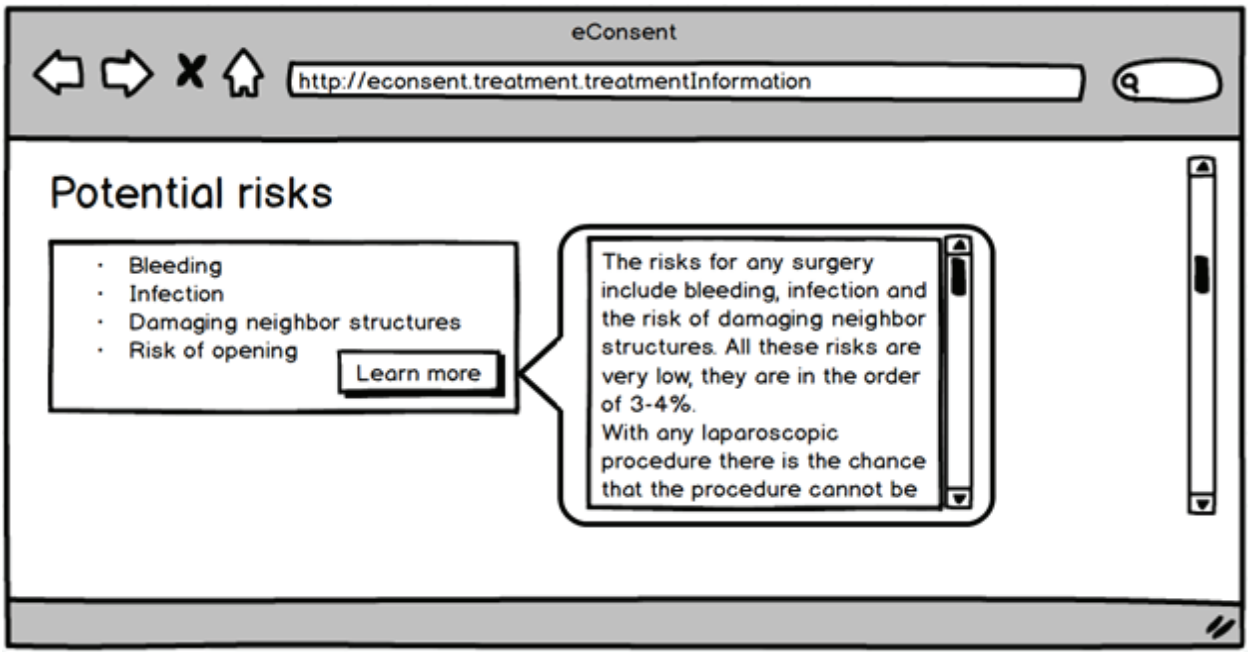


Figure 4: Example UI of a Treatment Information showing the Learn more feature.

```
{
    "linkId": "econsent.treatment.risks",
    "text": "Potential risks",
    "type": "group",
    "item": [{
       "linkId": "econsent.treatment.risks.summary",
     "text": "Summary",
       "type": "open-choice",
     "options": {
          "reference": "ValueSet/risks"
       }
    },
    {
     "linkId": "econsent.treatment.risks.learnMore",
     "text": "Learn more",
     "type": "text"
    }]
}
```

Listing 1: Questionnaire-item representing the *Learn more* scenario.

```
{
   "linkId": "econsent.treatment.description",
   "text": "Description of the procedure",
   "type": "group",
   "item": [{
      "linkId": "econsent.treatment.description.text",
     "text": "Text",
     "type": "text"
   },
   {
     "linkId": "econsent.treatment.description.multimedia",
     "text": "Multimedia",
     "type": "attachment"
   }]
}
```

Listing 2: Questionnaire-item representing the *multimedia* scenario.

Multi-language support: To achieve multi-language support, multiple resources (i.e. Questionnaire and QuestionnaireResponse) for each language can be created.

Alternatively, it may be preferable to store all *Treatment Information* for a given treatment, including the representations of this information in multiple languages, in a single resource. This single-resource option can be achieved by making use of the *translation extension* [21]. The extension can be attached to each item-element of the Questionnaire and the QuestionnaireResponse, so that both the heading and content can be translated into different languages.

Another option would be the use of an integration server, such as NextGen Connect, to deliver a requested resource in the required language. NextGen Connect, formerly known as Mirth Connect, is an open-source engine for HL7 that can be used as an integration server [22]. This option keeps the resources lightweight, similar to using multiple single-language resources, for delivery to a client application. An integration server can either manipulate the existing *translation extensions*, or store translations by item-numbers in the Questionnaire resources [23].

Signature of various roles: Depending on the situation, there can be a third-party individual that signs the consent form on behalf of, or in addition to, the patient. FHIR Consent and Provenance resources already support the case in which a third-party individual signs the form. This is accomplished through the element *Consent.consentingParty*, as well as the *Provenance.agent.who* reference.

Rejection and withdrawal: The Consent.status element is set to *active* for a consent agreement. In case of a refusal, this status is set to *rejected*. If the patient decides to withdraw an already active consent, the Consent resource can be updated and the status will be changed from *active* to *rejected*. Similarly, an already rejected consent can be set to *active* again if the patient agrees to the treatment at a later point in time [24].

Interoperability: To improve the interoperability of the proposed system, ValueSet elements can be represented by codings of a standardized terminology. Ahmadin et al. discuss the representation using narrative text of guidelines for pre-operative assessment with SNOMED CT. They state, that over 70% of the used terms can be represented using that terminology [25].

Right now, SNOMED CT [26] does not offer a specific group of codings for risks as a consequence of a surgical procedure. Nevertheless, there are options to code most of these risks. Given the laparoscopic appendectomy example, the associated risks can be represented by the codes listed in Table 1.

In some cases, such as the *infection* example in Table 1, more than one code can be suitable for a given risk. Other risks, such

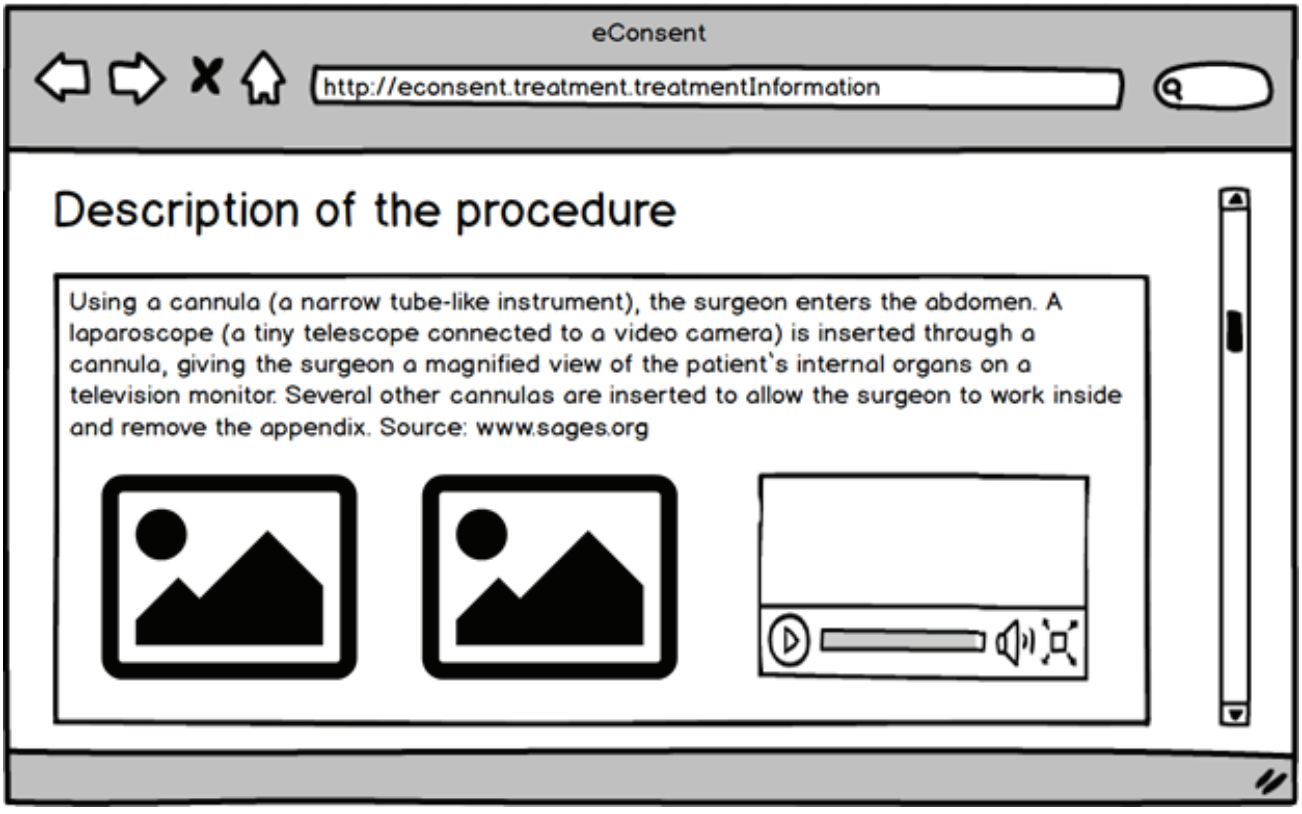


Figure 5: Example UI of Treatment Information content showing the multimedia feature.

as the *risk of opening,* may not have a single representation in the terminology. For any risk without an associated terminology, the default coding shown in the last row of Table 1 can be used. However, this representation is evidently less informative given its lack of specificity.

To have a coding for every risk and to represent them in a more accurate way (i.e. including the qualifiers *at risk at* and *perioperative*), an extension of SNOMED CT might be requested (as discussed in the subsequent section).

4 Discussion

Some underlying conditions discussed below lead to limitations in the proposed architecture. A possible extension of the SNOMED CT terminology is also discussed. Furthermore, the use of the FHIR QuestionnaireResponse as information source is examined.

4.1 Limitations

To allow for compatibility of the proposed eConsent architecture with institutional EHR systems, it was assumed that institutions offer a way to integrate FHIR applications with their systems. This assumption was made on the basis that there are existing solutions that integrate FHIR with an IHE infrastructure (e.g. PIXm [27], PDQm [28], MHD [29]). That said, the number of institutions already using or planning to use an infrastructure that supports these profiles is currently unknown.

The maturity levels of the current FHIR resources must also be taken into consideration. As the FHIR standard is still under development, the maturity level describes the stability of a given resource. More specifically, a resource's maturity level is based on the types and level of review that the resource has received, and can range from 0 (draft) to 6 (normative) [30]. The implemented FHIR version, STU3, is a standard for trial use and none of the defined resources are normative before version R4 (Release 4). Accordingly, the integrated resources may change in later versions and can become incompatible with older versions. Once a resource is part of the normative standard, it is less likely to change and backward-compatibility becomes a requirement [30]. The FHIR Consent resource and the *translation extension* used in the proposed architecture currently possess a maturity level of 1. Accordingly, these resources may change significantly before they become part of the normative standard.

The proposed architecture includes a step for obtaining the patient's signature electronically. This paper discusses technical considerations for the creation of a treatment eConsent architecture using HL7 FHIR. Possible legal limitations, such as cases that mandate a physical consent form or limitations based on data privacy regulations are outside the scope of this paper. Guidelines for obtaining eConsent for research studies exist and permit the use of electronic signatures if they are legally valid [31]. Unlike the informed consent process for research studies, the treatment informed consent process cannot be fully implemented electronically, as the discussion between the physician and the patient must not be replaced [7, 31]. This implies that there is no need to verify the patient's identity in a digital way, as this can be done on site by the physician.

4.2 Terminology

SNOMED CT does not provide a coding for all of the risks as a consequence of a given procedure. Therefore, additional concepts will have to be requested. A potential existing parent element is shown in Table 2.

It might be misleading to classify risks under the parent *medical accidents*. An alternative is requesting additional concepts under a new subset (e.g., suggesting the subset concept name *at risk for perioperative complications*) of the parent *finding of at risk*, as shown in Table 3.

4.3 FHIR QuestionnaireResponse as Information Source

The proposed architecture makes uses of the FHIR QuestionnaireResponse resource to represent the content of the *Treatment Information* that is shown to the patient. Proposals concerning the hierarchy and design of the Questionnaire items (i.e. the *Learn more* and the *multimedia* scenarios that are depicted in Section 3.1) can be used to generate the UI in a more user-friendly way.

Table 1: Possible SNOMED CT representation for treatment risks.

Risk (Narrative)	SNOMED CT Coding	Description
Bleeding	242996005	Accidental hemorrhage during medical care (finding)
Infection	12246311000119109	Infection following procedure (disorder)
Infection	762611002	Infection of organ surgical site following surgical procedure (disorder)
Infection	413590008	At risk of healthcare associated infection (finding)
Damaging neighbor structures	409031004	At risk for perioperative injury (finding)
Risk of opening	-	-
General (Default)	704356008	At risk of healthcare associated complication (finding)

Table 2: Possible SNOMED CT parent for treatment risks: Medical accidents.

SNOMED Code	Definition
269691005	Medical accidents to patients during surgical and medical care (event)

Table 3: Possible SNOMED CT parent for treatment risks: Finding of at risk.

SNOMED Code	Definition
281694009	Finding of at risk (finding)

The aim of the FHIR QuestionnaireResponse resource is to capture a set of answers given to a specific FHIR Questionnaire resource [2]. The proposed usage of the FHIR QuestionnaireResponse resource, to represent compiled information that follows a defined outline and that can be displayed in an auto-generated UI, is not intended by the standard. With the proposed usage of questionnaire items to split up long content and to combine multimedia elements with text, however, this representation is possible. That being said, there are some notable restrictions with the proposed approach:

1. *Multimedia*: Text and multimedia elements can be mixed within a single section (see *multimedia* scenario in Section 3.1.2), however, their order is strictly defined by the item order of the FHIR Questionnaire. This reduces the flexibility with which the text and multimedia elements can be combined. As an example, the order of the items depicted in Listing 2 requires a text input first, followed by at least one multimedia element. Accordingly, the generated UI must show the text prior to any multimedia elements. Allowing for multiple alternating text and multimedia items in the Questionnaire would be an enhancement, but would still fail to provide full control over the order of the content elements. Accepting answer elements of several data types (i.e. by changing the cardinality of the *Questionnaire.item.type* element from 1:1 to 1:*) is one possible solution.

2. *Text emphasis*: The proposed architecture allows for the input of plain text, but it does not allow the user to emphasize content (e.g. bold or italic words). Taking a markup language as the input for an answer of the data type *text* can remedy this limitation. Another option is the expansion of the ValueSet item-type [20] to a *markup* type.

The primary intention of the FHIR QuestionnaireResponse resource differs from the use case proposed in this paper. The introduction of custom extensions or the use of other information sources must be taken into consideration to address the needs of emphasized and flexible content.

5 Conclusion and Future Work

The existing consent model of HL7 FHIR that currently focuses on privacy consent provides the basic elements that are needed to model the electronic treatment consent use case. The FHIR Questionnaire and QuestionnaireResponse resources allow the representation of different structures. These two resources contain the elements necessary to auto-generate a UI containing a form with given headings and dynamic input fields. A notable limitation of this approach is that the order of elements is strict and not flexible when it comes to mixing data types such as text and images within a single section. Furthermore, the proposed system does not support text emphasis, which is needed to further increase content readability.

The usage of codings for some common concepts (e.g., risks) can increase the interoperability of the application. Missing concepts will be requested to be part of the SNOMED CT terminology.

A web-based, open-source prototype that auto-generates the UI for clinicians (to enter new *Treatment Information*) and patients (to sign the consent) are currently under development. The open-source project is available at https://github.com/ehealthinnovation/eConsent. When this prototype is finished, other types of consent, such as informed consent for research studies, will be considered as a continuation of the open-source project. Expert feedback will also be obtained, including feedback from surgeons as the main users of the treatment consent, as well as feedback from researchers concerning the planned research study consent implementation.

Lastly, experiences will be shared with the *HL7 FHIR Consent Directive Project* working group. The working group will also be approached for feedback regarding usage of the draft standard.

6 Acknowledgement

The work is motivated by Dr. Lin's intention of making the consenting process easier for patients, and was realized in cooperation with the University of Applied Sciences Upper Austria. The Centre for Global eHealth Innovation kindly supported the project by providing resources for the development of an open-source prototype. Many thanks to William Chou as the front-end developer of the open-source project, as well as to James Agnew as the founder of the open-source project HAPI FHIR that is used as a basis for the back-end implementation. Also, thanks to Linda Parisien as part of the SNOMED CT International Member Forum for giving advice on the usage and possible extension of SNOMED CT.

Glossary

EHR	Electronic Health Record.
FHIR	Fast Healthcare Interoperability Resources. Standards framework created by HL7.
HL7	Health Level 7. Organization that develops standards concerning electronic health information.
HTTP	Hypertext Transfer Protocol. A protocol that is used for communication in the web.
IHE	Integrating the Healthcare Enterprise. An initiative to improve information sharing in healthcare by promoting the coordinated use of standards.
JSON	JavaScript Object Notation. A lightweight format for data exchange.
REST	Representational State Transfer. A style of architecture to develop web services.
RESTful	A REST-compliant system.
SNOMED CT	Systematized Nomenclature of Medicine - Clinical Terms. A healthcare terminology system including terms, synonyms and codes.
STU3	Standard for Trial Use 3. Current official FHIR version that was published on April 19, 2017.
UI	User Interface.
XML	Extensible Markup Language. A markup language that can be used to store and exchange data.

References

[1] HL7 International. FHIR. FHIR Release 3 (STU). 2017. https://www.hl7.org/fhir

[2] HL7 International. Resource Index. FHIR Release 3 (STU). https://www.hl7.org/fhir/resourcelist.html

[3] Hall D, Prochazka A, Fink A. Informed Consent for Clinical Treatment. Can Med Assoc J. 2012; 184: 533-540.

[4] European Union Agency for Fundamental Rights. Consenting to Medical Treatment without Parental Consent. 2017. http://fra.europa.eu/en/publication/2017/mapping-minimum-age-requirements/consent-medical-treatments

[5] Canadian Medical Protective Association. Consent - A Guide for Canadian Physicians. 2016. https://www.cmpa-acpm.ca/en/advice-publications/handbooks/consent-a-guide-for-canadian-physicians

[6] College of Physicians and Surgeons of Ontario. Determining Capacity to Consent. 2007. https://www.cpso.on.ca/uploadedFiles/policies/policies/policyitems/capacity_consent_july07dialogue.pdf

[7] Canadian Medical Protective Association. Documenting Consent. 2018. https://www.cmpa-acpm.ca/serve/docs/ela/goodpracticesguide/pages/communication/Informed_Consent/documenting-e.html

[8] Schenker Y, Fernandez A, Sudore R, Schillinger D. Interventions to Improve Patient Comprehension in Informed Consent for Medical and Surgical Procedures. Med Decis Making. 2010; 31: 151-173.

[9] Vanderbilt University Medical Center. REDCap About. 2018. https://projectredcap.org/about

[10] Sanderson I, Obeid J, Madathil K, Gerken K, Fryar K, Rugg D, et al. Managing Clinical Research Permissions Electronically: A Novel Approach to Enhancing Recruitment and Managing Consents. Clinical Trials. 2013; 10: 604-611.

[11] National Health Service. Consent to Treatment. 2016. https://www.nhs.uk/conditions/consent-to-treatment

[12] UCSD Human Research Protection Program. Informed Consent Assistant. 2018. https://irb.ucsd.edu/consent.shtml

[13] Farrell E, Whistance R, Phillips K, Morgan B, Savage K, Lewis V, et al. Systematic Review and Meta-Analysis of Audio-Visual Information Aids for Informed Consent for Invasive Healthcare Procedures in Clinical Practice. Patient Educ Couns. 2014; 94: 20-32.

[14] Wilson E, Makoul G, Bojarski E, Bailey S, Waite K, Rapp D, et al. Comparative Analysis of Print and Multimedia Health Materials: A Review of the Literature. Patient Educ Couns. 2012; 89: 7-14.

[15] Schenker Y, Wang F, Selig S, Ng R, Fernandez A. The Impact of Language Barriers on Documentation of Informed Consent at a Hospital with On-Site Interpreter Services. J Gen Intern Med. 2007; 22: 294-299.

[16] University of Innsbruck. Treatment Contract - Medical Liability. 2018. https://www.uibk.ac.at/zivilrecht/buch/kap10_0.xml?section-view=true;section=2

[17] Government of Ontario. Health Care Consent Act. 1996. https://www.ontario.ca/laws/statute/96h02#BK27

[18] World Health Organization. Surgical Safety Checklist. 2008. http://www.who.int/patientsafety/safesurgery/tools_resources/SSSL_Checklist_finalJun08.pdf?ua=1

[19] Palfrey J, Gasser U. Interop: The Promise and Perils of Highly Interconnected Systems. New York: Basic Books; 2012. p. 8-111.

[20] HL7 International. ValueSet Item-Type. FHIR Release 3 (STU). 2017. https://www.hl7.org/fhir/valueset-item-type.html

[21] HL7 International. Extension: Translation. FHIR Release 3 (STU). 2017. http://hl7.org/fhir/extension-translation.html

[22] NextGen Healthcare. NextGen® Connect Integration Engine. 2018. https://www.nextgen.com/products-and-services/integration-engine

[23] Cerlinca M, Turcu C, Cerlinca T, Prodan R. HL7 Messaging Engine with Customizable Translation System. Adv Electr Comp Eng. 2010; 10: 98-101.

[24] HL7 International. Consent State Codes. FHIR Release 3 (STU). 2017. http://hl7.org/fhir/valueset-consent-state-codes.html

[25] Ahmadian L, Cornet R, de Keizer N. Facilitating Pre-Operative Assessment Guidelines Representation Using SNOMED CT. J Biomed Inform. 2010; 43: 883-890.

[26] SNOMED International. The Global Language of Healthcare. SNOEMD CT. 2018. https://www.snomed.org/snomed-ct

[27] IHE. Patient Identifier Cross-Reference for Mobile (PIXm). Integrating the Healthcare Enterprise IHE IT Infrastructure Technical Framework Supplement. 2017. http://ihe.net/uploadedFiles/Documents/ITI/IHE_ITI_Suppl_PIXm.pdf

[28] IHE. Patient Demographics Query for Mobile (PDQm). Integrating the Healthcare Enterprise IHE IT Infrastructure Technical Framework Supplement. 2017. http://ihe.net/uploadedFiles/Documents/ITI/IHE_ITI_Suppl_PDQm.pdf

[29] IHE. Mobile Access to Health Documents (MHD). Integrating the Healthcare Enterprise IHE IT Infrastructure Technical Framework Supplement. 2017. https://www.ihe.net/uploadedFiles/Documents/ITI/IHE_ITI_Suppl_MHD.pdf

[30] HL7 International. Version Management Policy. FHIR Release 3 (STU). 2017. https://www.hl7.org/fhir/versions.html

[31] FDA. Use of Electronic Informed Consent Questions and Answers. Guidance for Institutional Review Boards, Investigators, and Sponsors. 2016. https://www.fda.gov/downloads/drugs/guidances/ucm436811.pdf


European Journal for Biomedical Informatics
EJBI

GDPR Compliance Challenges for Interoperable Health Information Exchanges (HIEs) and Trustworthy Research Environments (TREs)

Ed Conley[1*] and Matthias Pocs[2]

[1] SHiELD Horizon 2020 and Connected Health Cities Projects, AIMES, Liverpool Innovation Park, L7 9NJ, United Kingdom

[2] SHiELD Horizon 2020 Project, Stelar Security Technology Law Research, 21035 Hamburg, Germany

Abstract

Background: We present our current approaches to improving personal data protection in (i) large (regional/national/international) scale health information exchanges (HIEs) and (ii) UK NHS IG toolkit and ISO 27001-compliant trustworthy research environments (TREs) for discovery science communities. In particular we examine impacts of the General Data Protection Regulation (GDPR) on these technology designs and developments and the responses we have made to control complexity.

Methods: The paper discusses multiple requirements to implement the key GDPR principles of "data protection by design" and "data protection by default", each requiring new capabilities to embed multiple security tests and data protection tools in common deployable infrastructures. Methods are presented for consistent implementation of diverse data processing use cases.

Results: We describe how modular compositions of GDPR-compliant data processing software have been used to implement use case(s) and deliver information governance (IG) requirements transparently. Security surveillance analysis is embedded throughout the application lifecycle, namely at design, implementation and operation (run-time) phases. A solution is described to the challenge of integrating coherent research (analytic) environments for authorized researchers to access data and analytic tools without compromising security or privacy.

Conclusion: We recognise the need for wider implementation of rigorous interoperability standards concerning privacy and security management. Standards can be disseminated within low-cost commodity infrastructures that are shared across consortium partners. Comprehensive model-based approaches to information management will be fundamental to guaranteeing security and privacy in challenging areas such as ethical use of artificial intelligence in medicine. The target architecture is still in evolution but needs a number of community-collaborative API developments to couple advanced specifications fulfilling all IG requirements.

Keywords

General Data Protection Regulation (GDPR); Information governance; Secure health information exchange; e-consent; Secure DevOps

Correspondence to:

Dr. Ed Conley Conley, Chief Scientific Officer, AIMES
The Innovation Park, Liverpool L7 9NJ, United Kingdom.
E-mail: ed.conley@aimes.net

EJBI 2018; 14(3):48-61
Received: March 26, 2018
Accepted: May 25, 2018
Published: July 06, 2018

1 Introduction

Our focus within the Horizon 2020 project SHiELD [1] and Connected Health Cities project [2] concerns data protection in health and research information exchange use cases. In particular, we are interested in impacts of the General Data Protection Regulation (EU) 2016/679 [3] also known as GDPR[1] on processing of personal data and on the free movement of such data [4]. This regulation can impose significant penalties for non-compliant data controllers and processors once it comes into force in the spring of 2018. Fundamentally, GDPR aims to provide a set of standardized data protection laws across EU countries. This is intended to make it easier for EU citizens to understand how their data is being used and to raise any complaints. For implementers, it has potential to reduce fragmentation and administrative burdens where business activities flow through local, regional, national and international data exchanges. A full treatment of the data protection principles that drive compliance to the GDPR is beyond the scope of this paper, but an abundance of introductory resources is available on the Internet.

[1] All abbreviations and acronyms used are tabulated at the end of the paper.

The new legal obligation of 'data protection by design' introduced by the GDPR requires data controllers to ensure (and demonstrate) that the traditional data protection principles like data subject rights, lawfulness, fairness and transparency, purpose limitation, data minimization, etc., are supported by technology design as an integral part of the system (GDPR Article 25(1)). GDPR has increased requirements for data controllers[2] to demonstrate compliance: Data controllers must build, implement and be able to demonstrate a comprehensive data privacy compliance programme. They must assess the "likelihood and severity of the risk" of any personal data processing operation relating to any use that "from personal data processing could lead to physical, material or non-material damage". The categories where risks could arise are summarized in Table 1. As a response to these, SHiELD consortium[3] proposes to use an open and extendable architecture with privacy-by-design modelling and embedded risk analysis tools. The aim is to provide systematic protection for storage and interoperable exchange of health data that is scalable across European borders. The exchange use cases are subject to permissions control (electronic consents) by the data subjects, compatible with existing regulatory frameworks. The goal is to ensure privacy, availability and correctness of health data whilst improving trust of patients in the security of their data and its use to address their needs.

This paper is a "current perspective" of challenges when implementing new large-scale infrastructures addressing health care and research domain problems (within the constraints of GDPR and manifold security threats). Data processing architectures are in rapid and continuing evolution, which further challenges for implementers faced with legal constraints. With respect to GDPR compliance, both controllers and processors need to demonstrate status and match their data processing steps to a collaborative IG plan. In this paper we illustrate how controlled reduction of complexity by fitting use cases to a symbolic abstraction set has benefits of increased transparency when applying IG rules across real-world data processing ecosystems. Controlled complexity reduction will become increasingly important as problem-solving data ecosystems scale and federate across the world.

The security and privacy standards landscape relevant to the domains of SHiELD and Connected Health Cities projects is summarised in Table 2. Meeting the challenge of interoperable privacy and security has been described as requiring services and mechanisms that are dynamic, distributed and intelligent [5]. Consistency and cross-compatibility of multiple deployed security and privacy solutions require conformance to international standards that are fit for purpose in complex domains of health and biomedical research. Currently developed international standards that support a path to greater interoperability do exist including standards for privilege management and access control developed by NIST, ISO and HL7 (Table 2). Traditional role-based access control (RBAC) standards are foundational but new specifications e.g. for security and privacy labelling (tagging) of segmented health information can improve interoperability (see also Discussion, Section 4). Comprehensive privilege management and access control (PMAC) principles [6] require explicit, ontology-based, formal (and therefore machine-processable) policies to implement at scale. While considerable theoretical work and a body of standards already exist for PMAC [5, 6, 7, 8], their degree of implementation in real-world solutions is limited (e.g. in large-scale programmes by Kaiser Permanente and

Table 1: Summary of categories where personal data processing could lead to physical, material or non-material damage.

Consequential Risk	Examples
Various losses	Discrimination, identity theft or fraud, financial loss, damage to reputation; loss of confidentiality of personal data protected by professional secrecy
No authorisation	Unauthorised reversal of pseudonymisation
Disadvantage	Any other significant economic or social disadvantage
Deprivation	Where data subjects might be deprived of their rights and freedoms or prevented from exercising control over their personal data
Revelations	Where personal data are processed which reveal racial or ethnic origin, political opinions, religion or philosophical beliefs, trade union membership
Sensitivites	Processing of genetic data; data concerning health, sex life, criminal convictions and offences or related security measures
Personal evaluations	Where personal aspects are evaluated, in particular analysing or predicting aspects concerning performance at work, economic situation, health, personal preferences or interests, reliability or behaviour, location or movements, in order to create or use personal profiles
Vulnerabilities	Where personal data of vulnerable natural persons, in particular of children, are processed
Scaling risks	Where processing involves a large amount of personal data and affects a large number of data subjects

[2] For definition, see later section "Shared responsibilities and roles under the GDPR".

[3] The SHiELD consortium (in alphabetical order) is AIMES, Fondazione Centro San Raffaele (FCSR, Milan), IBM Research (Haifa), IT Innovation (University of Southampton), Metrarc, North West Shared Infrastructure Service (NWSIS, UK NHS), Osakidetza, Stelar Security Technology Law Research, Symphonic Software and Tecnalia. Illustrations of their corresponding interests and expertise are shown in Figure 4.

Table 2: Health informatics standards referenced by SHiELD including security and privacy elements.

Standard (alphabetical)	**Sources: HL7 International, ISO TC 215 Health informatics, CEN TC 251 Health informatics, NIST**
EN ISO 21549-5	Health informatics - Patient healthcard data - Part 5 Identification data
EN ISO FDIS 17523	Health informatics - Requirements for electronic prescriptions
Generic Component Model (GCM)	The Generic Component Model (GCM) as a system-theoretical, architecture-centric, ontology-driven, and policy-controlled approach to privacy and security [6]
HL7 HCS	Context-sensitive segmentation of health information in HL7 International Healthcare Privacy and Security Classification System (HCS) Release 3 [9]
ISO 13606	Health informatics - Electronic Health Record Communication - Part 4 Security
ISO 17090-5	Health informatics - Public-key infrastructure - Part 1 Overview of digital certificate services - Part 2 Certificate profile - Part 3 Policy management of certification authority - Part 4 Digital signatures for healthcare documents - Part 5 Authentication using healthcare PKI credentials
ISO 21298	Health informatics - Functional and structural roles
ISO 22600	Health informatics - Privilege management and access control (PMAC) [5] - Part 1 Overview and policy management - Part 2 Formal models - Part 3 Implementations
ISO 25237	Health informatics - Pseudonymization
ISO 27005	Information Technology -provides guidelines for information security risk management
ISO 27789	Health informatics - Audit trails for EHRs
ISO 27799	Health informatics - Information security management in health using ISO/IEC 27002
ISO TR 18638	Health informatics - Components of education to ensure healthcare information privacy
ISO TS 11633-1	Health informatics - Information security management for remote maintenance of medical devices and MIS - Part 1 Requirements and risk analysis - Part 2 Implementation of an information security management system (ISMS)
ISO/HL7 10781	Health informatics - Electronic Health Record Sytems Functional Model
ISO/IEC TR 80001	Application of risk management for IT-networks incorporating medical devices - Part 1 Roles, responsibilities and activities - Part 2-2 Guidance for the communication of medical device security needs, risks and controls (security capabilities) - Part 2-8 Application guidance - Guidance on standards for establishing the security capabilities identified in IEC 80001-2-2 - Part 2-9 Application guidance - Guidance for use of security assurance cases to demonstrate confidence in IEC/TR 80001-2-2 security capabilities
NIST Security Labels	Security Labels as described in FIPS PUB 188 [36]

the US Veterans Administration). This is a conundrum, given the strong and increasing societal demand for robust privacy and security systems and the large proportions of budgets often apportioned to these aspects. Ability to adapt to rapidly changing environments and wide use case challenges is also essential. In ISO 22600 [7], security and privacy domains are defined by explicit ontologies and policies that can dynamically adapt to changing contextual and environmental conditions and can represent individual preferences at any level of granularity [6, 8]. The HL7 International Healthcare Privacy and Security Classification System (HCS) Release 3 [9] consists of a system-theoretical approach for context-sensitive segmentation of

health information (enabling security and privacy labelling of data segments for machine processing - for potential benefits see Section 4). Of wide utility is also the Generic Component Model (GCM) [6, 8] as a system-theoretical, architecture-centric, ontology-driven, and policy-controlled approach to privacy and security. Constraining the GCM can systematically and formally model any system or subsystem of actors (persons, organizations, but also devices, applications, or components) in reusable segments bound to context-specific rules.

2 Methods

2.1 Approach to Secure Cross-Border Exchange (SHiELD)

The SHiELD use cases are based on cross-border health information exchange (HIE) via a national contact point to relay source system messages in respective countries. The approach to GDPR-compliance is direct implementation of the key GDPR principles of "data protection by design" and "data protection by default". Data protection tools are embedded in a common HIE infrastructure that is deployable by Secure DevOps technology [10]. The information exchange infrastructure is based on an extended OpenNCP architecture [11] itself implementing components of EpSOS [12]. The secure exchange of health data across borders is driven from a set of use cases (see below). Secure DevOps offers unique advantages for software deployments, semi-automating APIs that work over large geographies for source and receiving system connections. There are also advantages for reduction of infrastructure costs, efficiency of upgrades and security tool co-provision as part of the distributive model.

Detailed technical descriptions of SHiELD's privacy-by-design innovations such as security risk modelling, enhanced digital permissions (consent), and enforcement mechanisms shall appear elsewhere. The consortium is working together to specify procedures for privacy-by-design in eHealth interoperability solutions, refining and deploying infrastructure, preparing legal recommendations (for policymakers, regulators and standards bodies), engaging in threat modelling and designing risk mitigation tools. Other approaches identify security requirements and provide automated analysis of data structures to identify sensitive elements vulnerable to specific threats. The overall objective is to enable systematic protection of health data against threats and cyber-attacks.

2.2 Approach for Analytic/Research Data Processing (Connected Health Cities)

In the Connected Health Cities (CHC) project [2] we have considered scalability of multi-EHR system (i.e. regional-to-national scale) information exchange implementations coupled to analytic or research data processing. A key challenge is standardising information governance (IG) at scale (source systems collectively serving c.5 to 7 million patient population sizes). In these designs, patient permissions concerning data use and access can be created electronically (within any source system connected to the exchange). As part of well-established policy-based access control mechanisms [13, 14, 15] these are consistently enforced, independent of the information requesting system. Without system-wide consistency, complexities and ambiguities of interpretation (e.g. diverse consent models applying slightly different sharing rules) can compromise personal data protection. The CHC project is focused on development of scalable Learning Health Systems. These often require significant data processing to enact analytical and research processes (see example below). A challenge for creating consistent data processing infrastructure partly comes from variability in the technical specifications needed to meet complex information governance requirements. There are also significant differences in the way multiple (independently implemented) electronic consent solutions work. Implementers of consent apps often "begin again" and solve only the immediate (local) problems for information sharing or use (making no reference to existing interoperable standards for setting and enforcing policy-driven electronic consents). As a result, local consent applications frequently do not interoperate and compliance to the GDPR becomes more difficult to achieve in practice.

The core concept of Information Governance (IG) requires some definition. Health information governance is complex and has often been contextualised to a geographic region or legal system. The UK NHS legal framework governing the use of personal confidential data in health care includes the NHS Act 2006, the Health and Social Care Act 2012, the Data Protection Act, the Human Rights Act and the UK Data Protection Bill (equivalent to GDPR). The law allows personal data to be shared between those offering care directly to patients and protects patients' confidentiality when data about them are used for other purposes. GDPR impacts the entire spectrum of NHS uses with a "legal basis for data processing" needing to be established for all data flows. While analyses concerning quality of care, what treatments work best, commissioning of clinical services, public health service planning can use non-identifiable datasets, analysis that need to use personal identifiable data may require consent of the patient with some well-defined exceptions.

2.3 Legal and Standards Compliance as Basics of Security and Privacy

The cross-European scope of the SHiELD project has referenced an important body of standards work alongside the GDPR (Table 2) and reviewed in the wider context of interoperability [16]. In addition, organisational measures need to be taken in response to the legislative requirements from the GDPR. Whereas many elements and principles

already existed according to previous EU legislation (e.g., data minimisation, lawfulness, supervision by data protection authorities, purpose limitation, etc.), some have been introduced by the GDPR such as accountability, data protection impact assessments, data protection by design, data portability, one-stop shop, etc. For example, the GDPR requires certain elements and principles to be included in organisational measures such as binding corporate rules (for controllers [17] and processors [18] - see also Section 4.3, Figure 3). Key topics such as privacy policies, privilege management and access control have been specifically addressed by IMIA and EFMI Security Work Groups, but also at HL7, ISO, CEN standards development organisations. Joint IMIA (Security in Health Information Systems) and EFMI (Security, Safety and Ethics) work groups have recognised challenges in trustworthiness in the security and safety of solutions and infrastructure deployed. A joint workshop "Personal Health Data –Privacy Policy Harmonization and Global Enforcement" highlighted privacy concerns by presenting different cases and approaches to develop a mechanism for a global healthcare information certification framework.

The CHC project requires a vendor-neutral framework based on interoperability standards as a solution for consent. In a foundational project (miConsent, [19]) the implementation standards in the HL7 Consent Directive, IHE BPPC [14] and APPC [20] have been evaluated. Currently, we have not yet fully evaluated for suitability in SHiELD or CHC health information exchanges the HL7 FHIR Consent Directive [15] or Consent2Share [16] frameworks. Further work is underway within SHiELD and CHC (i) critically evaluating whether or not blockchain is a security technology compatible with the GDPR's "right to erasure" (see also Section 4.3, Figure 3c). SHIELD partners are developing OpenNCP-associated API's that will support cross-border interoperable consent statements. Tools also are also being developed in order to simplify use of XACML [21] (a general-purpose access control policy language) in health information exchange and these will be described in future publications. A number of methodological improvements for security and privacy interoperability are discussed in Section 4.

3 Results

3.1 Consistent Matching of Information Governance Requirements to Data Processing

A graphical method was used to map required information flows within a limited number of privacy zones [22]. We nominated privacy zones as Care Zone, Non-care Zone and Research Zone [23] but additionally incorporated a use-case driven Trustworthy Research Environment [24] for data analytics. The Trustworthy Research Environment (commonly abbreviated to TRE) is a fully-implemented system supporting secure, regulated (authenticated researcher) access to datasets and tools. We emphasise its name as "trustworthy" not "trusted" as it is designed according to a set of principles that are deserving of trust or confidence and as such are more dependable or reliable. We recognize a continual process for design improvements. The term "trusted" is an absolute which may not be defensible, for example in the event of a breach. If real-world breaches occurred, "trustworthiness" would mean that an immediate and effective mitigation measure would be put into place (by virtue of the security risk modelling tools, see below). TRE's work with underlying health information exchanges and standalone sources of data that require specific (bespoke) processing. Current generation TRE designs use virtual machine (VM) and secure network technology to implement interoperable interfaces, databases, data processing routines and transformation operators driven by IG requirements. The next-generation approach (also embedded in SHiELD) will also use Secure DevOps technology to deploy modular data processing infrastructure builds within a coherent technical architecture. A key innovation is the coupling of use case data flows (Figure 1a) to reduce complexity privacy zones describing the information governance requirements (Figure 1b) to actual data processing infrastructure needed to deploy the entire end-to-end system with use-case to use-case consistency (Figure 1c).

3.2 Creation of Trustworthy Research (Analytic) Environments

The role of the Trustworthy Research Environment (TRE) implementing compute applications [25] governs legitimate researcher access to data collections and the invocation of permitted analytic services. In the current-generation TRE, this was achieved through:

i. **Data provisioning** where data is stored in accordance with the NHS IG toolkit and ISO 27001:2013 standards; these compliance processes are not a 'one-off' but a matter of continuous improvement, vigilance and organizational awareness. Data provisioning also provides workflow infrastructure enabling production of data pipelines, which automate extract transformation and dataset preparation.

ii. **Analytics provisioning** is secured using two-factor authentication over VPN, providing (for example) university-based analysts with access to an eight core / 32GB RAM data science virtual desktop and access to software packages including common statistical packages and geospatial software. TRE's can be used in many areas of the research enterprise including collaborative drug target prioritisation, medication repurposing and stratification of populations into cohorts for personalized medicine, exposome/ adverse event registration, comparative treatment effectiveness and pharmacovigilance.

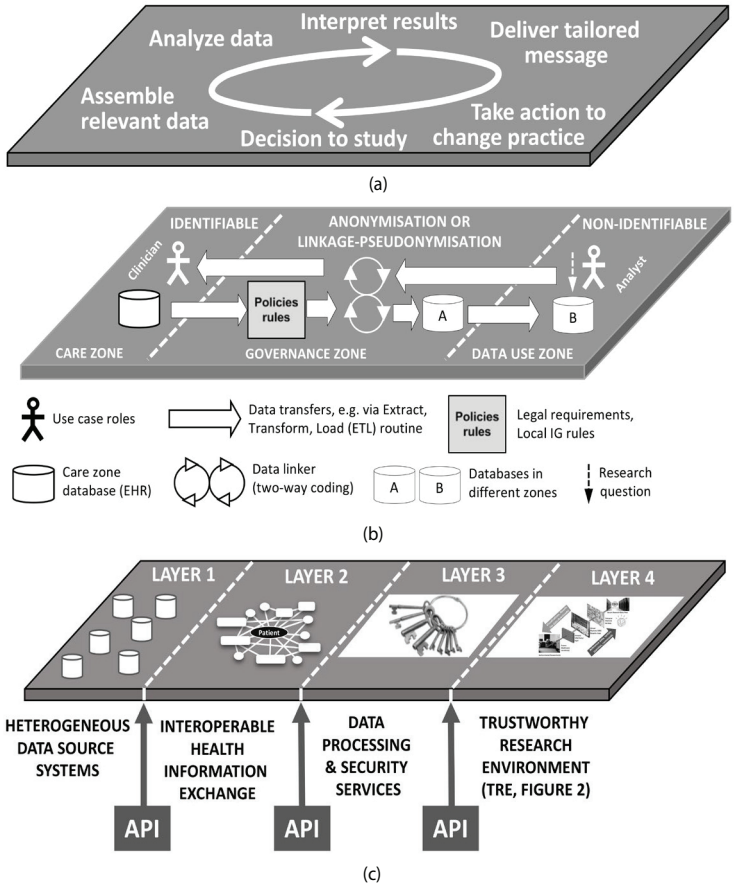


Figure 1: Consistent matching of use case information governance requirements to data processing; (a) An example use case (for learning health) of cross-system information flows; (b) Defined privacy zones for the total (end-to-end) cross-system data paths (IG model); (c) Four-layer architecture enabling end-to-end flows to enact the use case. Modular data processing infrastructure (Layer 3) and research environment (Layer 4) tooling is mapped to the IG model. APIs can permit bidirectional flow if the IG model permits. See text for management of contiguity of security within and between the four layers.

The design of TREs assumes re-use (large-scale hosting) of existing cohort data (for retrospective studies) and admission of electronic health record system data for prospective studies complying with the GDPR. The "Researcher View" of the TRE is illustrated in Figure 2. Security standards (above and Table 2) are maintained across different levels of a four-layer architecture with approved flows and privacy zones managed by an end-to-end IG plan. Currently TRE data processing components are selected and installed manually according to UK NHS IG toolkit and ISO 27001-compliant trustworthy research environment guidelines. Cross-border data exchange components in SHiELD will be selected and hosted using DevOps technology (Section 3.5). Where explicit consent is required for storage, sharing or use of personally identifiable data it is managed by electronic consent documents implemented using the IHE BPPC (Basic Patient Privacy Consents) profile. Permissions for sharing or use specify the access control within the health information exchange.

Common Interfaces (APIs) Coupling Analytic Information Flows: With reference to Figure 1, information flows from source systems exploiting a common API (all source systems need to agree standards to develop an interoperable "web-of-care" [26]). Source systems will likely include patient-identifiable data, within a classic health information exchange (HIE) that is designated as Layer 2 in the 4-Layer architecture (see Section 3.1, Figure. 1c, right-hand side). To facilitate cross-system exchange common APIs – i.e. using agreed standards to which all connected participants conform - are critical to scaling interoperability. In healthcare and patient-facing systems, interfaces use HL7 (V2, V3, FHIR) sometimes employing IHE profiles where they fit (e.g. IHE MHD for mobile clients) or web service interfaces (WS, SOAP). If further data processing is required (data transfers, storage, linkage/coding and security analytics) this constitutes Layer 3 (data processing) services. The specification of these is critical to ensure information processing compliant with legislation and the IG plan (Figure 1b). For example, there are obligations in GDPR for providing "opt-in, informed, free choice" consents, a mechanism to revoke such permissions, enable personal data portability and support statements on data holdings. In order to scale, all such features need to be facilitated as part of modular data processing services. Secure DevOps technology (Section 3.5) will assist information

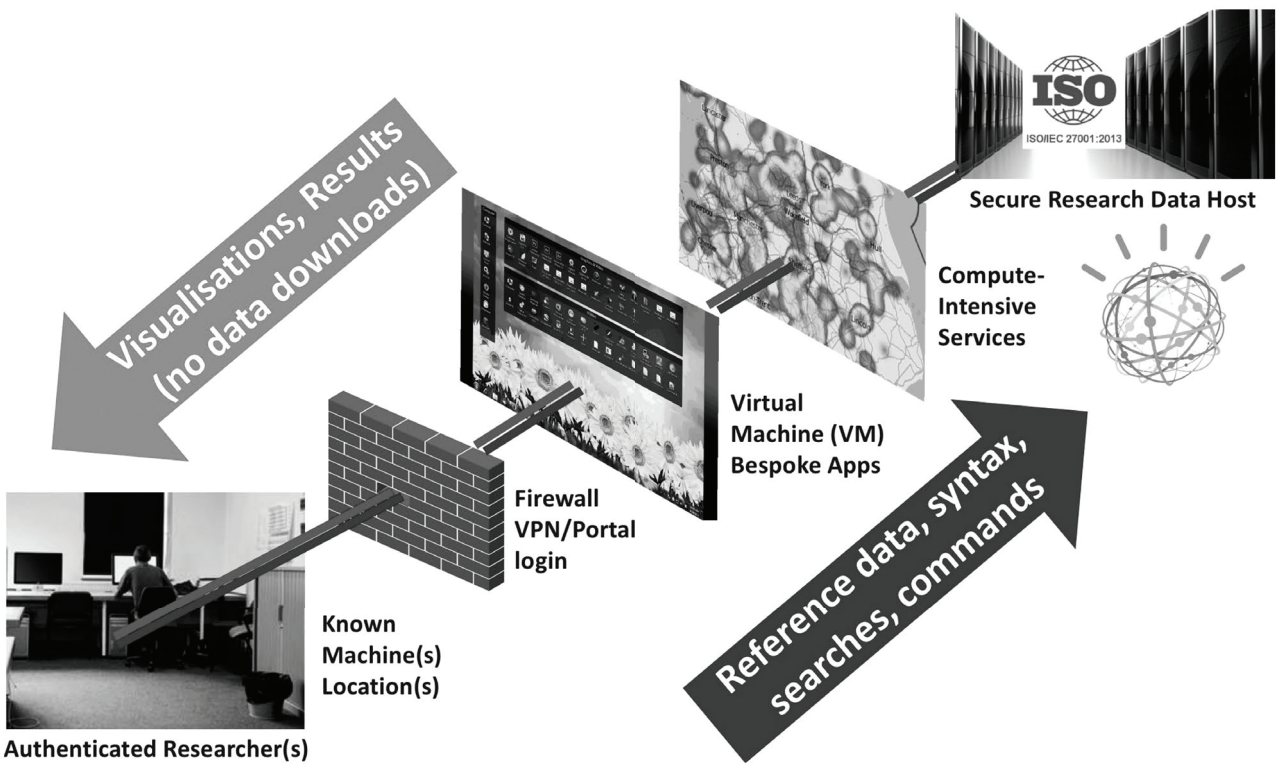


Figure 2: Researcher view of a current generation data-protective Trustworthy Research Environment (TRE).

system designers optimising selection of services and security validation tools in both design-time and run-time environments.

3.3 Comprehensive Security Threats Modelling and Mitigation for Use Cases

The SHiELD consortium implements a wide variety of risk mitigation tools in the context of its cross-border information exchange use case scope. It has introduced comprehensive security threat modelling and testing directly into the development process. The understanding comes from comprehensive intelligence of known descriptions of risks e.g. as described by ISACA [27] plus those in the collective experience of the SHiELD partners.

- Security risk mitigation approaches currently within the project include:
- **Asset inventory** - comprehensive records kept of assets and applications.
- **Configuration management** - Vulnerability modelling activities act as a comprehensive reference. Configuration tools are evaluated for capabilities in log management and additional threat analysis, intrusion detection and network vulnerabilities (for example: Puppet, Salt, Ansible, Chef and API-driven tools).
- **Counteraction measures** - Threat-associated rules that trigger threat counteraction mechanisms; these prevent unauthorized access, loss of data and cyber-attacks.
- **Documentation of policies/procedures** - Policies need to cover all steps of the production release process and need to be available to auditors.
- **Cross-border regulatory management** - Maintaining compatibility with regulations in different countries with data is being exchanged.
- **Logging of access and activity during development** - Timestamped code modifications against each developer, e.g. provided by Cucumber and Jira.
- **Introduction of novel security technologies** - Data hiding/masking and sensitive data analysis; anonymisation/pseudonymisation; provision of data and privacy protection to detect and prevent emerging threats such as inference attacks including cryptographic methods to prevent conventional attacks.
- **Peer review processes** - All code is peer reviewed with explicit rules regulating the independence of code approvers.
- **Performance** - Metrics are created with paths to solve problems.
- **Releases/deployments verification** - Automated releases require consistent deployment architecture for serving repeatable scalable processes described by

a use case (see Figure 1). Deployments use a rule base for consistency, but any design-time security mitigations need to be verified in the operational phase

- **Security experts** - Included as part of the stable development and deployment team.
- **Security training for developers** - Training for correct application of tests and external validation procedures.
- **Software module dependency tracking** - For reuse of fully defined blocks of code (modular computational workflow) to minimize opportunities for insecure code injection.
- **Streamlining processes** - Minimising errors through increased automation and raised quality; i.e. fewer code approvals but more trustworthy, continuous improvement.
- **Test types** - Static, dynamic, interactive and runtime application of security tests (evaluating tools such as Veracode, Waratek, Contrast Security, Fortify).
- **Traceability of lessons learned** - Tracking past software errors and mitigations.
- **Vulnerability points analysis** - Access control-related, device-related, consent-related; security tool assessments will adopt a continuous approach to analysing gaps.

3.4 Privacy-Protecting Legal Compliance Actions

The SHiELD health information exchange implements a number of legal and privacy-enhancement and security actions. These include Article 25 of GDPR for health data exchange, using documents of the Article 29 Working Party on Data Protection (e.g. on EHR) and European technical standardisation of "privacy by design" and obligation to "data protection by design and by default". These actions cover the GDPR data protection principles such as data minimisation, technical privacy measures such as pseudonymisation in response to the potential privacy impacts from automatic health data exchange.

Relevant international and European standardisation (ISO, CEN) is identified and addressed, for example, ISO/AWI 22697 ‚Health informatics - Application of privacy management to personal health information'. The collaboration agreements with standardisation bodies are approved by CEN-CENELEC/JTC 13 Cybersecurity and data protection. HL7 standards (e.g. CDA, V2, V3 and FHIR) are used in the technical implementations for documents and interfaces. Where appropriate, IHE profiles of standards such as Basic Patient Privacy Consents (BPPC) and IHE MHD (Mobile access to Health Documents) are employed. The project is currently developing architectural enhancements to the ePSOS/OpenNCP data exchange architecture (including extensions that address process models to handle incremental privacy threats and inference attacks, see Section 3.3).

Use cases for SHiELD include (i) chronic disease involving European travel with continuous monitoring and linkage of personal health data with secure exchange (ii) an emergency use case (e.g. stroke and loss of consciousness, with a "break glass" scenario to access records). In these use cases, patients are given the opportunity to consult their health data without having to reveal their identity to cloud operators that may be linked to previous consultations. SHiELD implements the only example of a technical and organisational measure that the GDPR [3, 4] offering pseudonymisation designed to achieve data minimisation (a prime example of ‚data protection by design' as cited in GDPR Article 25(1)). Like other legal obligations such as ‚accountability' (GDPR Articles 5(2), 24(1), which suggests to implement a data protection management system) the obligation of ‚data protection by design' is also subject to feasibility and risk-based conditions. Any data controller needs to take into account of:

The state-of-the-art - this may begin with standards such as ISO 25237 (health informatics pseudonymisation) or any future application of ISO/IEC 20889 (privacy-friendly de-identification techniques) in addition to guidance by the data protection authorities (for example [28, 29] are considered for relevance).

The nature, scope, context and purposes of processing - the use case descriptions need to be detailed to inform controllers accordingly.

Risks of varying likelihood and severity for rights and freedoms of natural persons posed by the processing - therefore, the European Charter of Fundamental Rights [30] has been analysed in the project.

Overall, a way of implementing ‚data protection by design' including pseudonymisation is the IG model. Both concepts are being used in the SHiELD project with a view to making subsequent proposals into technical standards bodies in the domain of health informatics [31], cybersecurity and data protection [32]. One currently undefined role is that of a scalable trusted third party (TTP) actor for generation and management of pseudonymisation keys. SHiELD needs to conduct external discussions in order to come up with meaningful recommendations for this functionality.

3.5 Impacts of Secure DevOps Technologies for End-to-End System Deployments

The SHiELD project [1] infrastructure development plan cites "Secure DevOps" methodology i.e. semiautomatic compilation of code, including deployment and testing with embedded security surveillance tools. This is a principal approach for raising security standards in health information exchange. Multiple security interventions can be embedded into the design and development phases. Currently the OpenNCP information exchange source code is being

analysed for performance with a range of security tools. Code and infrastructure elements will also be comprehensively tested at run-time with monitoring tools that can detect potential vulnerabilities (reports being generated for the developer/end user). Releasing software that has security vulnerabilities is a retrograde step. The Secure DevOps approach creates fundamental value for enabling reusable deployments meeting security and legal compliance requirements. It would impact every aspect of development, testing, integration, deployment and operations team work. It also represents a move to increasing automation of the agile application development process and deployment on to highly-scalable platforms. Such semi-automated approaches improve ability to deploy modular infrastructure builds specified to fulfil an IG plan (see Figure 1b, c).

4 Discussion

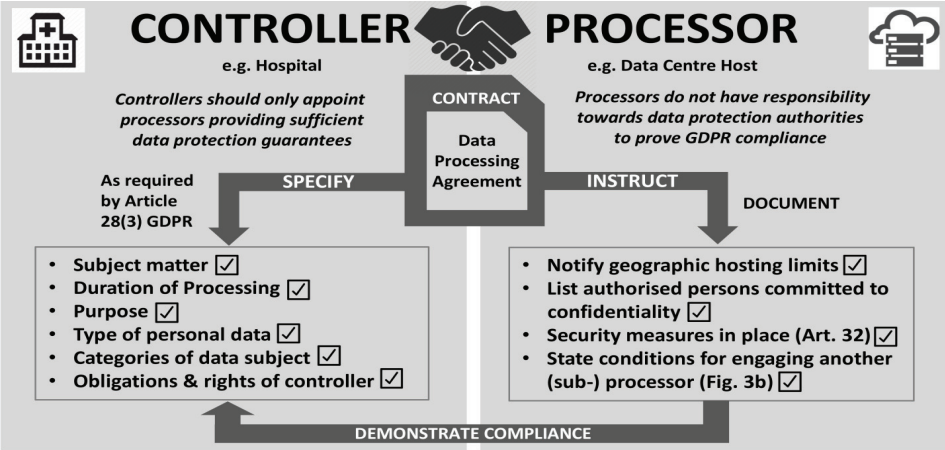
4.1 Raising "Trustworthiness"

Overall, our approaches follow (i) an open and extendable architecture supported by (ii) security mechanisms, (iii) privacy-by-design modelling (iii) risk analysis tools and (iv) Trustworthy Research Environments for research or analytic applications. The aim is to provide systematic protection for the storage, exchange and use of health care data across European borders and in distributed research projects. Within a SHiELD point to point information exchange, data use is controlled by the data subject, compatible with regulatory frameworks and compliance to the GDPR. The consortium members have a common focus on privacy, with improved availability and accuracy of data. This aims to raise the level of trust patients will have in the security of their data and its use to address their needs. This aim directs our focus on solving data security and privacy threats in different phases of the application lifecycle, namely, design, implementation and operation (run-time) using the methods and technologies.

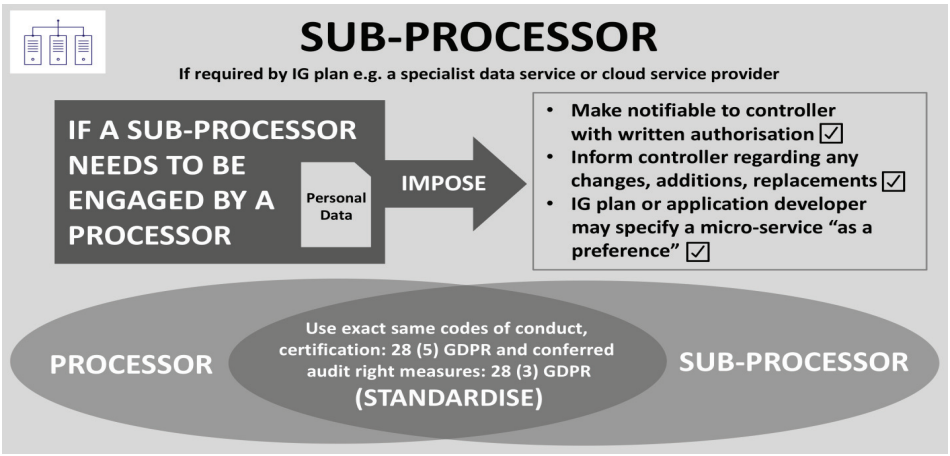
The wide range of test/validation checks exemplifies a key shift in the importance of security and data protection regulation concerns. There is a commitment for "attention to detail" as it is well known that simple mistakes and "weakest links" can easily create security vulnerabilities. Establishing security checklists is only a start point – a rigorous solution requires continuous evaluation and collaboration (for the system to meet fitness-of-purpose). The challenges in this paper can only be met by critical shifts in culture where security and data protection become the responsibility of all members of collaborating organisations using SHiELD outputs. In the future, these approaches are foundational for rational, secure and ethical approaches using artificial intelligence (AI) and personalized medicine [33]. For example, it is widely recognized that AI has untapped potential to improve reliability of diagnoses, higher quality prognostic indicators with applications in medicine [34]. Initiatives such as the 100,000 Genomes Project [35] already show the power of data combination governed by common data models from across multiple settings.

At current status, the projects *per se* are providing the Trust Framework - establishment of trust between the sender and the receiver systems. Trust is static i.e. established prior to any exchange through mutual participation, but in future a dynamic trustworthiness is needed, meaning that the conditions of the exchange and governing policies are negotiated at runtime. In this case, the expression and conveyance of policy includes the security labels applied to shared information and the application of privacy protections, markings and handling instructions bound to the exchange policies. In order to be effective, the Trust Framework must be legally binding and can apply retrospectively to the exchange pattern of publish and subscribe. The reliability of labelling solutions (next section) depends on the trustworthiness of the labelling entity and involved authorities including related accreditation and certification processes. The "cross-border" record sharing of SHiELD Health Information Exchange and "cross-domain" use of Trustworthy Research Environments makes the Trust Framework around core infrastructure critical (see role of the core OpenNCP infrastructure in Section 4.4, Figure 4).

Security/Privacy Labels to Model Use Cases, IG Zones, Data Processing Infrastructure: The "use case to IG zoning to infrastructure build" relationships shown in Figures 1a, 1b and 1c requires a substantially-researched interoperability framework in order to scale. One beneficial approach is security labels. These are markers bound to a resource, which connect an information object to a set of security and privacy attributes. The HL7 HCS specification defines Confidentiality labels, Sensitivity labels, Integrity labels, Compartment and Handling Caveats labels. The four labels (tags) can enable security and privacy rules about specific health information objects. Handling caveat labels convey dissemination controls and information handling caveats such as obligations and refrain policies to which an IT resource custodian or receiver must comply. Overall, Security Policy Information Files define which security labels are valid and how they can be checked against the Clearances – through these innovations, privilege and access control management in health information systems can be automated. The HL7 HCS Security Labels are described in NIST FIPS PUB 188 [36]. Operationalising the HCS is assisted at runtime by a Security Labelling Service [37] and the Privacy and Protective Services. The latter enforces obligations by applying various transforms to the response package including masking, redaction, annotations, anonymization or pseudonymisation based upon rules. If this standard was applied to the scheme illustrated by Figures 1a, 1b and 1c, objects can be reused by an access control system to support access decisions (e.g. matching classification labels



(a)



(b)

(c)

Figure 3: Shared responsibilities and roles under the GDPR; (a) Role of the data processing agreement and other expectations between data controllers and processors under the GDPR; (b) Conditions applying if the data processor needs to invoke a separate data processing service to fulfil the use case and IG requirements; (c) Data subject rights that the GDPR seeks to uphold.

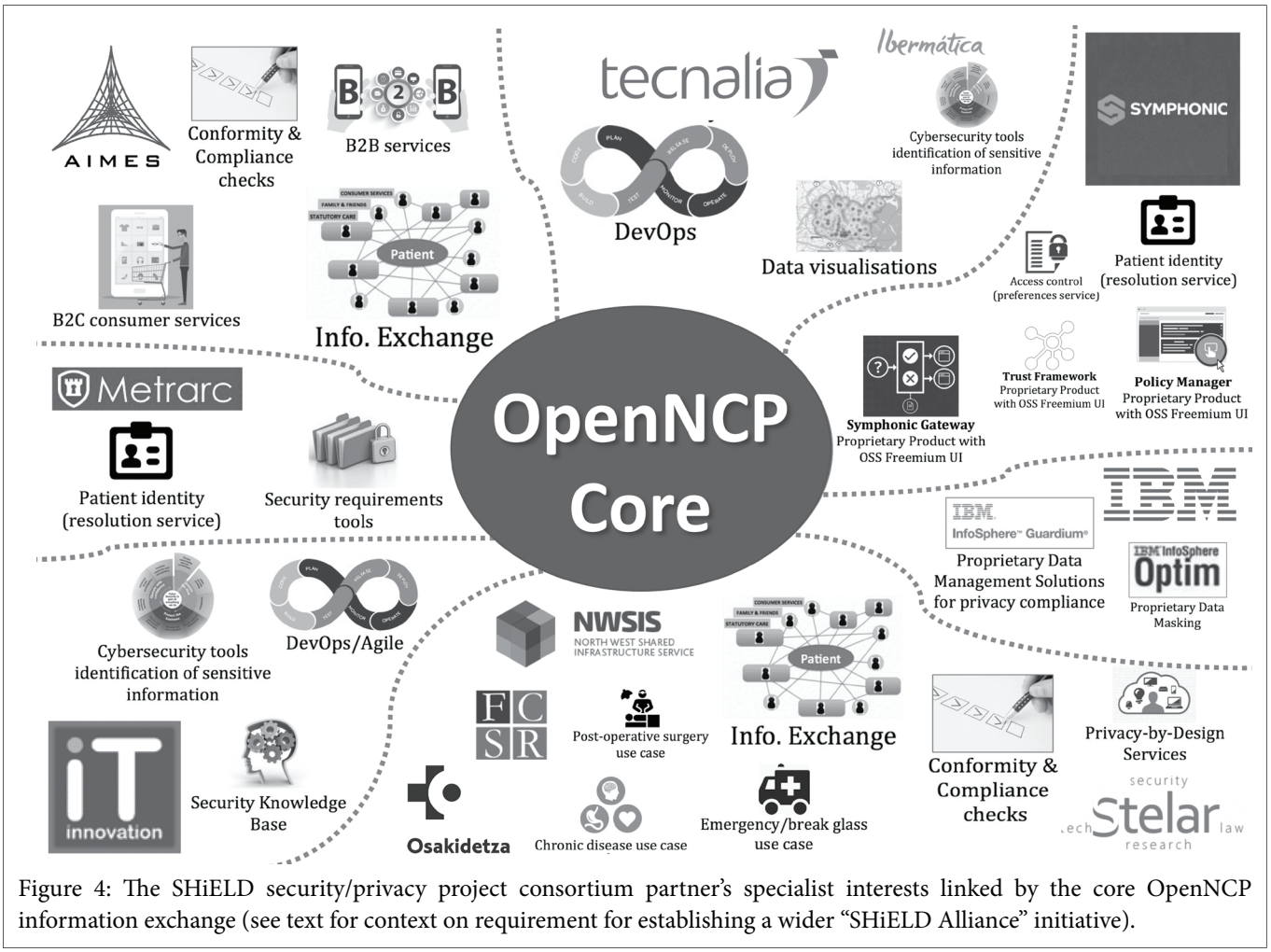


Figure 4: The SHiELD security/privacy project consortium partner's specialist interests linked by the core OpenNCP information exchange (see text for context on requirement for establishing a wider "SHiELD Alliance" initiative).

to clearances or other attributes specified by a security policy). These policies can be dynamic (e.g. in patient preferences) so HCS labels are applied at runtime (rather than being permanently stored with information objects). The runtime approach ensures the most current policy and trust framework (controlling the information exchange between sender and receiver) are enacted. Currently information exchange is bound by conventional IHE BPPC transactions, but their shortcomings are recognised. For example bespoke policy formulation is highly complex and the transactional nature of access control can become fragile as numbers of systems joined to the exchange increases.

4.2 Monitoring of System Privacy/Security Compliance

To date, privacy compliance checklists have been developed for organisations that are considered to be data controllers. Checklists are not as rigorous as the new GDPR obligation of data protection by design. GDPR brings accountability (not just responsibility) which means new requirements to demonstrate compliance. The data protection by design legal obligations address data controllers who may need to ensure the obligations are transferred to the suppliers. In the context of future SHiELD-based service use, data controllers could be hospitals, while data processors could be IT companies. Health organisations may act alone or as a joint buyer consortium creating supply tenders. They would specify data protection requirements and the SHiELD approaches could help meet the specifications. This would need to cover both the data protection side (coinciding with GDPR) as well as cybersecurity (digital security, information security, IT security, ISO27000-series) aspects and the legal basis for models of consent where these influence geographic scalability. The latter frequently depends on whether the GDPR permits national deviations; sometimes there are no extra permissions for the national legislators (e.g. a SHiELD "break glass" use case which concerns the vital interest of the data subject; in this case the national legislators cannot deviate from the GDPR rules according to GPDR Article 6 (1)(d), (2)). The concept of scalability is also tied to "legal interoperability".

4.3 Impacts of Shared (Contractual) Responsibilities under the GDPR

Irrespective of the approaches in this project some generic impacts also need addressing within a discussion of impacts. The applicable scopes of the GDPR is large - the

sum of national populations across the EU itself. Within the UK[4] for example, there will be more than 60 million data subjects (persons who have data stored about them) and approximately 500,000 data controllers (companies or organisations which store data about data subjects). The GDPR was intended to harmonise Europe's data protection laws. However, its flexibility and scope will likely create differences on how it is applied. Whereas a data controller is someone who "determines the purposes and means of the processing of personal data" (GDPR Article 4(7)), a processor is "any person who processes personal data on behalf of the controller (GDPR Article 4(8); other than a person who is an employee of the controller)". One of the major changes in the GDPR is that data processors have specific obligations. For example, if a processor fails to report a data loss to their controller, then the processor can be subject to regulatory action from the data protection authority (e.g. the Information Commissioner), and this is not possible under the Data Protection Act in such a strict way. To clarify these overriding issues, Figure 3a outlines some of the shared responsibilities between data controllers and processors as overriding considerations. Figure 3b summarises the relationship between processors and sub-processors. Finally the objective of guaranteeing data subject rights is annotated in Figure 3c. An organisation is likely to hold a data processor role if it does not decide the goals and means of data processing of the health data itself. It may host and maintain an infomation platform, but unless it is processing data for its own purposes, it is unlikely to be a data controller. A processor has much less responsibility towards data procesing authorities to prove compliance with data processing law. Data processors are not the first line of contact for Data Subject rights (GDPR Articles 12-22). It does, however have responsibilities to keep minimal records of processing it carries out for data controllers.

4.4 Scaling Security/Privacy Standards in Real-World Implementations

We set out to write a "current perspective" of challenges when implementing new large-scale infrastructures addressing health care and research domain problems (within the constraints of GDPR and manifold security threats). We acknowledge it is not a completed set of work but early communication (dissemination) is vital as there is a community-building aspect to the project. For example, a comprehensive list of security risk mitigation approaches (as described in Section 3.3) will require a community-based interoperability approach to sustain and refine. A key difficulty is how multiple interested parties (many of them competitors in the market and from mixed sectors of expertise) can move forward coherently and in control producing high quality pre-competitive guidelines that are actually implemented into interoperable products. One proposal for sustaining international coherence is the formation of a health data security and privacy "alliance" that would act to consistently implement standards in a shared non-proprietary infrastructure (e.g. the core OpenNCP ecosystem). Common components, interfaces and methodologies would be agreed, and incremental technical and policy developments could take place within implementation projects. The SHiELD consortium's journey shows that this proposition is challenging but given a collaborative ethos it is not impossible. Similar implementation initiatives have already taken place in other health market sectors (e.g. the Continua Health Alliance for personal health devices) resulting in coherent use case management, certification and test, policy alignment, technical working groups and ultimately shared interoperability guidelines (across hundreds of competing companies). In our projects, a wide set of expertise has also been essential to generate and critique cross-community solutions. Figure 4 illustrates how diverse expertise and interests of the current consortium partners have formed around the common OpenNCP infrastructure. This problem is certainly challenging. Of the list of practicable security/privacy standards (Table 2) some of which have reached HL7 International realm-specific Implementation Guide (IG) status for the US realm [38] and practically demonstrated e.g. in the Consent2Share project [16] which is under evaluation in our Connected Health Cities project. Cultural and legal specificities can act as a barrier for direct reuse of standards across national realms and adaptations are necessary to accommodate these and individual's needs.

[4] Despite Brexit, the UK will be implementing essentially all of GDPR into UK national law via the UK Data Protection Bill published on 14 September 2017.

5 Acknowledgement

The authors gratefully acknowledge project funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 727301. We also thank Peter Gryffoy for analysis and insights underlying Figure 3 and to Marie Messenger and Michael Walker for proofreading; we also gratefully acknowledge the Connected Health Cities (CHC) programme commissioned by the Northern Health Science Alliance (NHSA) and funded by the UK Department of Health

Abbreviations and Acronyms Used

AI-Artificial Intelligence, **API**-Application Programming Interface, **APPC**-Advanced Patient Privacy Consents, **BPPC**-Basic Patient Privacy Consents, **CEN**-Comité Européen de Normalisation (European Committee for Standardization), **CHC**-Connected Health Cities, **DevOps**-Development Operations, **EFMI**-European Federation for Medical informatics, **GCM**-Generic Component Model, **GDPR**-General Data Protection Regulation, **EHR**-Electronic Health Record, **HCS**-International Healthcare Privacy and Security Classification System, **HIE**-Health Information Exchange, **HL7**-Health Level 7, **FHIR**-Health Level 7 Fast Healthcare Interoperability Resources, **HL7 V2**-Health Level 7 Version 2, **HL7 V3**-Health Level 7 Version 3, **IG**-Information

Governance, **IHE**-Integrating the Healthcare Enterprise, **IMIA**-International Medical Informatics Association, **ISACA**-Information Systems Audit and Control Association, **ISO**-International Standards Organisation, **MHD**-Mobile Health Documents, **NHS**-National Health Service, **NIST**-National Institute of Standards and Technology, **OpenNCP**-Open National Contact Point, **PMAC**-privilege management and access control, **RBAC**-Role-Based Access Control, **R&D**-Research and Development, **SOAP**-Simple Object Access Protocol, **TRE**-Trustworthy Research Environment, **TTP**-Trusted Third Party, **VM**-Virtual Machine, **WS**-Web Services.

References

[1] SHiELD - European Security in Health Data Exchange. http://projectshield.eu/

[2] Connected Health Cities Project. https://www.connectedhealthcities.org

[3] GDPR, Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation), Official Journal of the EU. 2016 L 119, page 1.

[4] Pocs M. Will the European Commission be able to standardize legal technology design without a legal method? Comput Law Secur Rev. 2012; 28: 641-650.

[5] Blobel B. Davis M, Ruotsalainen P. Policy Management Standards Enabling Trustworthy pHealth. Stud Health Technol Inform. 2014; 200: 8-21.

[6] Blobel B, Nordberg R, Davis JM, Pharow P. Modelling privilege management and access control. Int J Med Inform. 2006; 75: 597-623.

[7] International Organization for Standardization. ISO 22600 Health informatics - Privilege management and access control. Geneva: ISO; 2006.

[8] Blobel B. Ontology driven health information systems architectures enable pHealth for empowered patients. Int J Med Inform. 2011; 80: e17-e25.

[9] HL7 International Inc. HL7 Healthcare Privacy and Security Classification System (HCS) - Release 3. Ann Arbor: HL7 International; May 2013.

[10] Yasar H, Kontostathis K. Secure DevOps Process and Implementation. In: Cybersecurity Development (SecDev); 2016 Nov 3-4; Boston, USA: IEEE; 2016. p. 166.

[11] Fonseca M, Karkaletsis K, Cruz IA, Berler A, Oliveira IC. OpenNCP: a novel framework to foster cross-border e-Health services. Stud Health Technol Inform. 2015; 210: 617-621.

[12] Moharra M, Almazán C, Decool M, Nilsson AL, Allegretti N, Seven M. Implementation of a cross-border health service: physician and pharmacists opinions from the epSOS project. Fam Pract. 2015; 32: 564-567.

[13] Integrating the Healthcare Enterprise (IHE). Basic Patient Privacy Consents (BPPC). https://wiki.ihe.net/index.php/Basic_Patient_Privacy_Consents

[14] HL7 International Inc. HL7 FHIR Consent Directive - current version of consent resource. http://build.fhir.org/consent.html

[15] Consent2Share open source software supporting an online consent process. https://bhits.github.io/consent2share/

[16] Blobel B. Standardization for Mastering Healthcare Transformation - Challenges and Solutions. Eur J Biomed Inform. 2017; 13: 09-15.

[17] Article 29 Working Party on Data Protection (WP29). Working Document setting up a table with the elements and principles to be found in Binding Corporate Rules (updated) (WP 256). Brussels 2017.

[18] Article 29 Working Party on Data Protection (WP29). Working Document setting up a table with the elements and principles to be found in Processor Binding Corporate Rules (updated) (WP 257). Brussels 2017.

[19] miConsent: Patient-controlled information sharing (legacy specification project combining HL7 and IHE standards). http://www.hl7.org.uk/marketing/archive/2012/120518_miconsent.asp

[20] Integrating the Healthcare Enterprise. IHE Advanced Patient Privacy Consents (APPC). https://wiki.ihe.net/index.php/Advanced_Patient_Privacy_Consents

[21] OASIS. A Brief Introduction to XACML.https://www.oasis-open.org/committees/download.php/2713/Brief_Introduction_to_XACML.html

[22] Delaney BC, Curcin V, Andreasson A, Arvanitis T, Bastiaens H, Corrigan D, et al. Translational Medicine and Patient Safety in Europe: TRANSFoRm - Architecture for the Learning Health System in Europe. Biomed Res Int. 2015; 2015: 961526.

[23] Kuchinke W, Ohmann C, Verheij R, van Veen E, Arvanitis T, Taweel A, et al. A standardised graphic method for describing data privacy frameworks in primary care research using a flexible zone model. Int J Med Inform. 2014; 83: 941-957.

[24] Lea N, Nicholls J, Dobbs C, Sethi N, Cunningham J, Ainsworth J, et al. Data Safe Havens and Trust: Toward a Common Understanding of Trusted Research Platforms for Governing Secure and Ethical Health Research. JMIR Med Inform. 2016; 4: e22.

[25] Sintero Server - Simplifying interoperability for distributed collaborative health care. http://docplayer.

net/6258059-Sintero-server-simplifying-interoperability-fordistributed-collaborative-health-care.html

[26] Conley E. Realising the Interoperable ‚Web-of-Care'. Eur J ePractice. 2013; 19: 48-64.

[27] ISACA -Information Systems Audit and Control Association. http://www.isaca.org/

[28] Article 29 Working Party on Data Protection (WP29). Working Document 01/2012 on epSOS (WP 189). Brussels 2012.

[29] Article 29 Working Party on Data Protection (WP29). Working Document on the processing of personal data relating to health in electronic health records (EHR) (WP 131). Brussels 2007.

[30] Charter of Fundamental Rights of The European Union (2000/C 364/01) http://www.europarl.europa.eu/charter/pdf/text_en.pdf

[31] CEN/TC 251 Health informatics. https://standards.cen.eu/dyn/www/f?p=204:7:0::::FSP_G_D:6232&cs=18CA078392807EDD402B798AAEF1644E1

[32] CEN-CENELEC/JTC 13 Cybersecurity and data protection. https://standards.cen.eu/dyn/www/f?p=204:7:0::::FSP_G_307986&cs=1E7D8757573B5975ED287A29293A34D6B

[33] Blobel B, Lopez DM, Gonzalez C. Patient privacy and security concerns on big data for personalized medicine. Health Technol. 2016; 6: 75-81.

[34] Hamet P, Tremblay J. Artificial intelligence in medicine. Metabolism. 2017; 69: S36-S40.

[35] The 100,000 Genomes Project. https://www.genomicsengland.co.uk/the-100000-genomes-project/

[36] National Institute of Standards and Technology. NIST FIPS 188 - Standard Security Label for Information Transfer. Gaithersburg, Maryland, USA; 1994.

[37] HL7 International Inc. HL7 Version 3 Standard: Privacy, Access and Security Services; Security Labelling Service, Release 1. 2014

[38] HL7 International Inc. HL7 Implementation Guide: Data Segmentation for Privacy (DS4P), Release 1. Ann Arbor: HL7 International. 2014.

Interoperability Specifications and Conformance Testing Services Made Available on the Tukan Platform

Sebastian Bojanowski[1,2]*, Roman Radomski[1,2], Marcin Grudzień[1,3], Daniel Matras[1,4], Paweł Masiarz[1,5] and Maciej Łańko[1,6]

[1] 1HL7 Poland

[2] HIT Inn, Warsaw, Poland

[3] Lux Med, Warsaw, Poland

[4] Comarch Healthcare, Kraków, Poland

[5] National Centre for Healthcare Information Systems, Warsaw, Poland

[6] Marshal Office of Małopolska Region, Kraków, Poland

Abstract

In 2018 Polish health IT community faces critical challenges related to the national eHealth agenda. The mission of HL7 Poland is to establish a community around interoperability standards and integration profiles with health IT vendors, medical providers and public authorities. One of the activities that would support this goal is providing tooling for specification publication and implementation validation. HL7 Poland has started a project to create a central hub for specifications and tools. The Tukan is an online platform dedicated to Polish healthcare IT community, where national specifications for interoperability are published together with a set of testing tools supporting their implementation. The platform is based on software components originating from various sources: open source release of IHE Gazelle components, development tooling of Polish National Implementation Guide of HL7 CDA, ART-DECOR platform software components, HAPI FHIR reference implementation for FHIR STU3 standard and Central Authentication Server (CAS) software components. Tukan platform is ready to be used as an environment supporting peer-to-peer testing in connectathon-like events. The pilot phase of Tukan platform has shown that there is a significant interest in testing services, especially when there are official specifications of interoperability standards published. In 2018 the first Polish connectathon will be held by HL7 Poland on Tukan platform in cooperation with national and regional projects.

Keywords

Tukan platform; Connectathon; IHE Gazelle; Polish HL7 CDA IG; Interoperability testing; Validation; HL7 CDA; HL7 FHIR; IHE profiles

Correspondence to:

Sebastian Bojanowski,
HIT Inn, Warsaw, Poland.
E-mail: bojanowski@iehr.eu

EJBI 2018; 14(3):62-64
Received: March 18, 2018
Accepted: May 16, 2018
Published: July 06, 2018

1 Introduction

In 2018 Polish health IT community faces critical challenges related to the national eHealth agenda. The pilot of central ePrescription system, the official recommendation of IHE profiles use and continuous development of the Polish National Implementation Guide for HL7 CDA [1], all of them shape the national perspective of eHealth for the next years. For the first time, Polish vendors have shown quite substantial interest in HL7 FHIR® [2] standard. HL7 Poland, in cooperation with national and regional authorities as well as key organizations of health IT vendors and medical providers, put an effort to supply of tools supporting implementation of interoperability standards.

2 Objectives

The mission of HL7 Poland is to establish a community around interoperability standards and integration profiles with health IT vendors, medical providers and public authorities for the purpose of boosting standards adoption and implementation. One of the activities that would support this goal is providing tooling for specification publication

and implementation validation. This tooling should be open and easily accessible for HL7 Poland members, and for the rest of interested parties in some extent. This platform of support should leverage HL7 standards and IHE profiles in a great extent and use globally approved tooling, configured to local requirements, to help especially those vendors, who are not able to participate in official events like connectathons organized by IHE or HL7. It should be collaborative effort to build a trusted environment for regional, nation-wide and international interoperability testing.

3 Methods

To secure a consistent approach to standards and their national specifications, HL7 Poland has started a project to create a central hub for specifications and tools. The Tukan is an online platform dedicated to Polish healthcare IT community, where national specifications for interoperability are published together with a set of testing tools supporting their implementation, to serve as focus point of all the efforts of vendors and regulators. The platform is ready to be used as an environment supporting peer-to-peer testing in connectathon-like events.

The Tukan platform is based on software components originating from various sources:

- Open source release of IHE Gazelle [3] components,
- Development tooling of Polish National Implementation Guide of HL7 CDA,
- ART-DECOR [4] platform software components,
- HAPI FHIR [5] reference implementation for FHIR STU3 standard,
- Central Authentication Server (CAS) [6] software components.

IHE Gazelle base components are used to support basic communication platform to allow secure, SSL-based, IHE ATNA [7] conformant service endpoint publication and reliable proxy-based peer-to-peer testing. As far as IHE profiles are concerned, we have focused on IHE XDS.b [8], as a key architectural concept in clinical document cross-enterprise exchange. For that reason, we have used XdsTools [9] component, which has been developed by NIST and as well is a part of IHE Gazelle platform, to expose the simulator of XDS.b conformant API.

Regarding validation of HL7 CDA [10] conformant documents, we have used the tool created during development process of Polish National Implementation Guide of HL7 CDA. The component is based on validation artifacts generated from ART-DECOR environment and is deployed on the same eXist XML database engine as DECOR services. The reason why we have used our own tool instead of IHE Gazelle EVS Client is the better integration with the different versions of the Polish specification. The next step in development of the Tukan platform will be customization of the IHE Gazelle ObjectsChecker for Polish realm related clinical document validation.

As a form of promotion of the HL7 FHIR standard, we have built the STU3-conformant reference server instance using HAPI FHIR implementation. We have imported and published all conformance resources related to definition of the base structures, v2 code tables, v3 vocabulary domains and FHIR value sets. We have also imported additional terminology resources in the form of value sets derived from DECOR specification of Polish National Implementation of HL7 CDA. It was done for the purpose of initial launching of the FHIR-based terminology service.

Despite the validation and integration testing functionality, the Tukan is also designed to be specification publication platform. ART-DECOR environment and FHIR server are the key components in that field.

From the technical perspective, having extensibility and scalability in mind, all Tukan platform services are deployed as isolated, Docker-based containers [11] in the Linux environment. The platform itself is the main repository of the container images. Any service can be easily replicated to many computing nodes if needed, and effortlessly deployed to other infrastructure, including various cloud service providers.

4 Results

The project started in June 2017 and the first Tukan services have been made available online in September 2017. 28 organizations, mostly software vendors, but also some medical providers, universities and local authorities, for 6 weeks were able to participate in the pilot phase. HL7 CDA validator attracted most teams; several hundreds of test CDA documents were validated. Till March 2018 the total number of organizations, that use Tukan, reached 40 and Tukan have been chosen as a platform for nation-wide connectathon-like event, that will take place in Warsaw in autumn of 2018. Some regional projects have also shown their interest in using Tukan as independent platform of reference.

5 Conclusion

The pilot phase of Tukan platform has shown that there is a significant interest in testing services, especially when there are official specifications of interoperability standards published. It is the best way to improve the quality of implementations and to increase maturity of the specifications. We have established Polish community interested in interoperability standards and integration profiles. The Tukan platform will expand to provide increasing number of validation services, with collaborative effort of HL7 Poland members.

References

[1] Polish National Implementation Guide for HL7 CDA. Release 1.2. http://art-decor.org/art-decor/decor-project--plcda-

[2] Fast Healthcare Interoperability Resources (FHIR)/ Release 3.0.1 (STU3). http://hl7.org/fhir

[3] IHE Gazelle Framework documentation. https://gazelle.ihe.net/gazelle-documentation/

[4] The ART-DECOR Expert Group. ART-DECOR Main Site. http://art-decor.org/mediawiki/index.php/Main_Page

[5] HAPI FHIR reference implementation for Java documentation. Release 3.2.0. http://hapifhir.io/docindex.html

[6] Aporeo Central Authentication Service documentation. Release 5.2. https://apereo.github.io/cas/5.2.x/index.html

[7] IHE IT Infrastructure Technical Framework Volume 1 (ITI-TF1) Integration Profiles, Revision 14.0. July 21, 2017: 73-86.

[8] IHE IT Infrastructure Technical Framework Volume 1 (ITI-TF1) Integration Profiles, Revision 14.0. July 21, 2017: 86-125.

[9] NIST Document Sharing Test Facility. http://ihexds.nist.gov/

[10] HL7 Clinical Document Architecture, Release 2.0. http://www.hl7.org/v3ballot/html/infrastructure/cda/cda.htm

[11] Docker Documentation. https://docs.docker.com/