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

doi:10.3233/SHTI200611

Why Interoperability at Data Level Is Not Sufficient for Enabling pHealth?

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

Abstract. Multidisciplinary and highly dynamic pHealth ecosystems according to the 5P Medicine paradigm require careful consideration of systems integration and interoperability within the domains knowledge space. The paper addresses the different aspects or levels of knowledge representation (KR) and management (KM) from cognitive theories (theories of knowledge) and modeling processes through notation up to processing, tooling and implementation. Thereby, it discusses language and grammar challenges and constraints, but also development process aspects and solutions, so demonstrating the limitation of data level considerations. Finally, it presents the ISO 23903 Interoperability and Integration Reference Architecture to solve the addressed problems and to correctly deploy existing standards and work products at any representational level including data models as well as data model integration and interoperability.

Keywords. pHealth ecosystem, 5P Medicine, architecture, knowledge representation, knowledge management, modeling, integration, interoperability

Introduction

For improving safety and quality of care as well as efficiency and efficacy of care processes under the current demographic, social, and economical constraints, healthcare systems undergo organizational and methodological paradigm changes. The first paradigm change describes the transition from organization-centric through disease-specific process controlled to person-centric care, and the latter that from the phenomenological perspective of general care through dedicated care for patient groups with specific clinically relevant conditions towards personalized, participative, preventive, predictive precision medicine (5P Medicine). Systems and services following the 5P Medicine paradigm will be in the course of the paper called pHealth systems and pHealth services. Appropriate pHealth services provided to the subject of care are defined by its individual health status, conditions and expectations, as well as genetic and genomic dispositions in their personal social, behavioral, environmental, and

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

occupational context. Table 1 summarizes objectives and characteristics of pHealth and methodologies and technologies for meeting them [1].

Table 1. pHealth objectives, characteristics and methodologies/technologies to meet objectives, after [2]

Objective	Characteristics	Methodologies/Technologies
Provision of health services everywhere anytime	• Openness • Distribution • Mobility • Pervasiveness • Ubiquity	• Wearable and implantable sensors and actuators • Pervasive sensor, actuator and network connectivity • Embedded intelligence • Context awareness
Individualization of the system according to status, context, needs, expectations, wishes, environments, etc., of the subject of care	• Flexibility • Scalability • Cognition • Affect and Behavior • Autonomy • Adaptability • Self-organization • Subject of care involvement • Subject of care centration	• Personal and environmental data integration and analytics • Service integration • Context awareness • Knowledge integration • Process and decision intelligence • Presentation layer for all actors
Integration of different actors from different disciplines/domains (incl. the participation/ empowerment of the subject of care), using their own languages, methodologies, terminologies, ontologies, thereby meeting any behavioral aspects, rules and regulations	• Architectural framework • End-user interoperability • Management and harmonization of multiple domains including policy domains	• Terminology and ontology management and harmonization • Knowledge harmonization • Language transformation/ translation
Usability and acceptability of pHealth solutions	• Preparedness of the individual subject of care Security, privacy and trust framework • Consumerization • Subject of care empowerment • Subject of care as manager • Information based assessment and selection of services, service quality and safety as well as trustworthiness • Lifestyle improvement and Ambient Assisted Living (AAL) services	• Tool-based ontology management • Individual terminologies • Individual ontologies • Tool-based enhancement of individual knowledge and skills • Human Centered Design of solutions • User Experience Evaluation • Trust calculation services

More details can be found in the papers published in the pHealth conferences series [2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12], in other proceedings [13, 14, 15, 16, 17] and in specific journal publications or books such as [1, 18, 19, 20]. The pHealth service delivery is inseparably bound to the ongoing technological and methodological paradigm changes such as mobile, micro-, nano-, bio- and molecular technologies, big data and business analytics, artificial intelligence and autonomous systems including robotics, etc., but also new computing technologies such as cloud, cognitive and edge computing. Table 2 presents a more complete list of technologies, methodologies and principles for transforming healthcare ecosystems to pHealth. The described organizational, methodological and technological paradigm changes affect granularity, complexity and maturity of health systems as well as the maturity levels of the underlying processes, resulting in the transition from empiric through evidence-based to translational medicine, or in other words as transition from an observational (subjective) through an analytical (objective) to an integrated pHealth approach.

Table 2. Technologies, Methodologies and Principles for Transforming Healthcare Ecosystems

• Mobile technologies, biotechnologies, nano- and molecular technologies • Big data and business analytics • Integration of analytics and apps • Assisting technologies → Robotics, autonomous systems • Natural Language Processing → Text analytics → Intelligent media analytics • Knowledge management (KM) and knowledge representation (KR) → Artificial intelligence (AI) → Artificial common (general) intelligence → Intelligent autonomous systems • Security and privacy, governance, ethical challenges, Education → Asilomar AI Principles • Cloud computing, cognitive computing, social business	• Edge computing as a "family of technologies that distributes data and services where they best optimize outcomes in a growing set of connected assets“ (Forrester Research), • Virtual reality and augmented reality, thereby blurring “the boundaries between the physical and digital worlds“ (Gartner). • Creation of IoT-Platforms and app-ecosystems • Patient-generated health data ecosystem → multiple, dynamic policies • Web content management → Digital experience management • Data bases → NoSQL technologies → Data warehouses → Graph DBs → Data lakes • EHR includes genomic data • Specifications → Implementation → Tooling → Testing → Certification

1. Methods

The 5P Medicine approach as introduced above requires advanced communication and cooperation between different disciplines and their actors (persons, organizations, devices, applications, components), thereby bridging between their different perspectives and contexts, the related different methodologies, terminologies and ontologies, but also their different levels of maturity, knowledge, experiences, skills, etc. The paper addresses requirements and solutions for knowledge-based interoperability necessary for multidisciplinary complex and dynamic pHealth ecosystems and demonstrates the limitations of data sharing and health information exchange (HIE) paradigms. For that purpose, it considers knowledge representation (KR) and knowledge management (KM) from different perspectives such as theory of knowledge, cognitive sciences and philosophy, abstraction and expressivity of languages and scope-specific grammars, enterprise architectures, modeling good practices, data modeling, system development processes, etc., thereby discussing opportunities and limitations of approaches. It provides solutions for overcoming the gap between the current practice of static models and the need for dynamic, flexible, non-deterministic and adaptive models representing the pHealth paradigms. An advanced and comprehensive solution, the interoperability and integration reference architecture model and framework, developed by the first author and standardized at ISO/TC215, is introduced and discussed in some detail.

2. Foundations of Knowledge Representation

There are multiple definitions of knowledge. Alter defines knowledge as "a combination of instincts, ideas, rules, and procedures that guide actions and decisions" [21]. The Merriam-Webster Online Dictionary defines knowledge as "the sum of what is known: the body of truth, information, and principles acquired by mankind" [22]. According to Davenport, knowledge is "information combined with experience, context, interpretation, and reflection. It is a high-value form of information that is ready to apply to decisions and actions" [23]. Other definitions focus more on the human mind's aspect. The different knowledge classes such as classification-based knowledge, decision-oriented knowledge, descriptive knowledge, procedural knowledge, reasoning knowledge, or assimilative knowledge should not be detailed here.

The representation of reality following the theory of knowledge or cognitive theory requires the move from cognition/sense-perception to conceptualization [24]. According to Doerner, domain knowledge consists of reproducible and reliable models of a domain, i.e. models which can be justified and repeatable formulated in the domain of discourse (discipline) [25]. Knowledge of a discourse domain, representing that domain's perspective on reality to facilitate reasoning, inferring, or drawing conclusions, must be created, represented, and maintained by domain experts using their methodologies, terminologies and ontologies.

Since the late 1990ies, ontologies have been a key strategy in encoding, using and sharing knowledge in computer sciences [26]. Following Guarino, an ontology, can be seen as the study of the organization and the nature of the world independent of the form of our knowledge about it [27]. From the modeling perspective, three levels of knowledge representation are distinguished and must be consecutively processed: a) epistemological level (domain-specific modeling), b) notation level (formalization, concept representation), c) processing level (computational, implementations) [25]. A model is thereby defined as a representation of objects, properties, relations and interactions of a domain, enabling rational and active business in the represented domain. The generalization of domain-specific epistemological models requires their transformation into a universal KR notation. The outcome must be validated on the real world system and thereafter adopted if needed (Figure 1) [25].

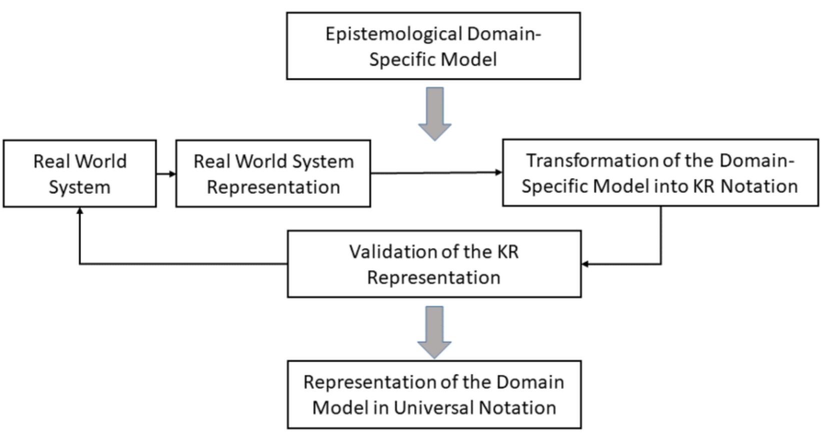

Figure 1. Knowledge representation development process (after [25]).

Deploying KR techniques such as frames, rules, tagging, and semantic networks, a good KR has to manage both declarative and procedural knowledge.

Another concept in the cognitive theory is intelligence, which is established by four fundamental principles: data, information, knowledge, and wisdom (frequently summarized as intelligence). The basis of intelligence is data, i.e. measures and symbols of the world gathered through observation and investigations. Collected by various sensors and organs, data are represented as external signals, facts, quantities, etc., forming the structural part of intelligence. At the next level, data are interpreted producing information as the second principle of intelligence. Attaching meaning to data represents the semantic level of intelligence, enabling decision-making. At the third level, knowledge allows the subjective, purpose- and context-specific interpretation of information to derive actions. Wisdom provides awareness, insight, moral judgments, and principles to validate/improve existing as well as to create new knowledge. Knowledge modeling combines data or information into a reusable format for the purpose of preserving, improving, sharing, aggregating and processing knowledge to simulate intelligence [28].

3. Knowledge Representation Principles and Practices

KR is first of all a substitute or surrogate for a real thing to allow reasoning about the world rather than taking action in it. For a deeper understanding of knowledge representation, it is useful to look at how knowledge is composed and what the relevant properties of knowledge and its components are. Floridi provides a definition that is closely linked to his theory of both information and data. Hence, it is of particular interest to the knowledge representation and knowledge management communities. In Floridi's theory, knowledge is the meaningful associate of factual semantic information. The latter is defined as follows: "*p* qualifies as factual semantic information if and only if *p* is (constituted by) *well-formed, meaningful,* and *veridical data*." It is important to note that Floridi's notion of *well-formed data* is not restricted to refer to syntactically well-formed data in the sense of grammar, which would point to a more digital interpretation of what he means by data. He specifically states that he refers to anything that determines the form, construction, composition, or structuring of something, be it a machine, a movie, a painting, or a garden. He uses the term syntax, which is what determines whether something is well-formed in an extremely broad sense. In addition, his definition refers to the *meaningfulness* of the data. Meaningful data is also called semantic content. Semantic content can be either instructional (describing actions to be taken, typically to achieve a certain aim) or factual (describe the state of affairs in the world). Floridi's definition entails a number of aspects relevant to our investigation: a) Factual semantic information (and thus knowledge) is not restricted to statements or sentences. Pictures, diagrams, etc. may also be part of knowledge. Hence, the use of "veridical" over "true". b) False information is, indeed not information, and hence not part of knowledge. [29]

Similarly, Schulz distinguished between ontological knowledge (axioms that are universally true), symbolic knowledge (statements about properties and meaning of signs of language), factual knowledge (statements about concrete entities and their relationships) as well as contingent knowledge (probabilistic and default statements, uncertainty) [30].

To enable reasoning about the world, KR provides a set of signs to represent the objects, phenomena and processes of interest. The assigning of the relationship between sign and entity that is pointed to by the sign is charging the signs with meaning. This is typically done using the methodologies of model theory and in the process ontological commitments are formed and expressed. KR provides a theory of reasoning, composed of (a) the representation's fundamental conception of intelligent reasoning, (b) the set of inferences the representation sanctions (the proof theory), and (c) the set of inferences it recommends. Furthermore, KR supports pragmatically efficient computation by properly organizing information to facilitate making the recommended inferences. Finally, KR is a medium for human expression by defining the language to represent the world. For realizing those functions and objectives, KR deploys knowledge representation technologies such as rules, logic, frames, semantic nets, etc. [31].

Nonaka and Takeuchi have analyzed the dynamics of knowledge creation, particularly the importance of tacit knowledge and its conversion into explicit knowledge. Externalization is a process of converting tacit knowledge into explicit concepts through the use of abstractions, metaphors, analogies, or models [32].

4. The Language Challenge of Knowledge Representation

This section summarizes insights provided in different papers such as [16, 33, 34]. According to [35], an ontology is a content specification of a particular domain, where the domain can be of any kind. Thus, an ontology can be described as "an explicit specification of a conceptualization" [36] or as "a shared explicit formal specification of domains of knowledge" [37]. According to Gruber [38], an ontology provides a formal explicit specification of a shared conceptualization of a domain of interest. It describes an ordering system of entities of a domain with their concepts, functions, and relations. A concept is a knowledge component the expert community has agreed on. A concept must be uniquely identifiable, and independently accepted by experts and users. It has a representation and can be specialized and generalized. Thereby, knowledge can be represented at different levels of abstraction and expressivity, ranging from implicit knowledge (tacit knowledge) up to fully explicit knowledge representation, i.e. from natural language up to universal logic. Expressivity is the key factor when selecting an appropriate KR. A more expressive knowledge representation language enables an easier and more compact expression of knowledge within the KR semantics and grammar. However, there is a trade-off between expressivity and practicality [39]. Expressive languages probably require more complex logic and algorithms to construct equivalent inferences. This leads to a complexity problem of formal language and reasoning systems with the lack of computability, at the same time losing the consistency of the language system. In summary, highly expressive KRs are less likely to be complete or decidable, while less expressive KRs may be complete or decidable. Therefore, natural languages are not only efficient in representing meaning, shared knowledge, skills, and experiences assumed. They also provide an optimum between restriction to special structure and generative power enabling the rich and nevertheless decidable representation of real-world concepts, supported of course by common sense knowledge. This is one of the reasons for representing facts and knowledge about a system and its domain-specific subsystems, their architecture and behavior by deploying natural-language-based domain-specific terminologies and concepts using domain-specific ontologies, extensively exploited in good modeling best practices. Context-free grammars, originally used to define the structure of syntax for programming languages, cannot express static semantics. One way to overcome these limitations is the deployment of context-sensitive grammars, by Chomsky proposed for describing NL, thereby embedding constraints and contexts in their rules. Alternatively, other grammatical models have been introduced to enhance the expressive power of programming languages. Syntax and static semantics of a small typed programming language has been defined by a grammar with contexts, however facing the complexity and consistency problem [40]. In summary: The difference between formal and natural languages – or between logic and natural language – is semantics, but even more pragmatics. There is a sound relation between logic and NL if logic is understood in a minimalist sense: providing a semantics for logical words, which needs to be enriched by pragmatic processes [41]. For more information on KR and KM see, e.g., [5, 16, 33].

When implicit knowledge contained in ontologies is made available and new knowledge is derived, the use of first-order-logic-based languages leads to the problem of non-determinability or semi-determinability. Languages are determinable if their characteristic function can be fully or at least partly calculated. Therefore, for knowledge management, determinable, sufficiently expressive first-order-logic subsets are used: the description logics. They provide well-defined formal semantics without variables in the formalism of predicates. Concepts, roles and individuals, as existent in the real world, can be described independently of each other [37].

Humans' communication is further enriched by means such as semantics enhanced by gestures, mimics, social or psychological signals. To reduce the gap between human and technical communication especially in the demanding healthcare domain, multimodal interfaces have been tested [42, 43]. Landgrebe and Smith have also highlighted the limitations of machine communication vs. human dialog and human language [44].

The Chomsky Hierarchy as important contribution to the formal language theory (FLT) describes a set of classes or types of grammars with different levels of constraints for production rules resulting in different levels of complexity of the related languages [45]. The level of constraints applied to grammars, and so the level of complexity of the language built, decreases from Type 0 (unrestricted grammar / computably enumerable languages) through Type 1 (context-sensitive grammar / context-sensitive language) and Type 2 (context-sensitive grammar / context-free language) up to Type 3 (regular grammar / regular language), so forming a subset of the prevailing lower-typed grammar and lower-typed language. Meanwhile, the Chomsky Hierarchy has been refined into additional sub-classes through extending the class of context-free languages and regular language by mildly context-sensitive language and sub-regular language subclasses [46].

Because of their simpler syntactical and semantical properties, programming languages follow a higher-typed grammar than natural languages. However, the terms "context-sensitive" and "context-free" used here should not be mixed up with the non-FLT terms used in the description of systems, use cases and scenarios.

5. Good Modeling Practices

The representation of the system covering the subject of care and the processes of analyzing and managing his or her health must consider all levels of its structural granularity and related behavioral aspects [1]. Structurally, we have to consider the continuum from elementary particles to population, functionally deploying the methodologies of the multiple disciplines providing perspectives on pHealth ecosystems.

A model is a simplified reflection of reality. It conceptually represents empirical objects, phenomena and processes in a logical and objective way. Typically, a model deals just with some aspects of reality. Therefore, Alter defined a model as partial representation of reality restricted to attributes the modeler is interested in according to the purpose of modeling, the addressed audience, etc. [21]. Similarly, Langhorst et al. [47] introduced a model as an unambiguous, abstract conception of some parts or aspects of the real world corresponding to the modeling goals. Consequently, two models of the same phenomenon may be essentially different due to differing requirements of the model's end users as well as behavioral, conceptual or contextual differences (e.g. knowledge, experiences, skills, etc.) among the modelers and to contingent decisions made during the modeling process [48]. For overcoming related problems when integrating different KR models, good modeling design principles such as orthogonality, generality, parsimony, and propriety [34] must be met. This requires that the relevant stakeholders shall define the provided view of the model, including the way of structuring and naming the concepts of the problem space. After capturing key concepts and key relations at a high level of abstraction, further abstraction levels will be used iteratively. However, the first iteration must be performed in a top-down manner to guarantee the conceptual integrity of the model [47]. The demonstrated good modeling practices hold for every new use case or new context to be represented including new aggregation of models.

6. Data Modeling

A data model is a visual representation of the people, places and things of interest to a business. It is used to facilitate communication between business people and technical people. A data model is composed of symbols that represent the concepts that must be communicated and agreed upon [49].

There are different levels of data models [49, 50]. On top, called the Very High Level Data Model, is the External Information Level or viewpoint. It describes the reality from a particular perspective for a particular purpose. It addresses business experts / users. At the next level, called the High Level Data Model, the Conceptual Information Level determines the relevant information and defines the basic concepts and their relationships, i.e. the business logic. That way, the business requirements can be collected and presented, and the basic concepts can be understood. The High Level Data Model addresses business users as well as problem analysts. The Logical Data Model (Logical Information Level) represents the platform-independent models, addressing business analysts and architects/designers. The Physical Data Model (Physical Information Level) represents the platform-specific models, addressing architects, database administrators and developers. 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 [51], 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. Sirpur proposes a three level data modeling architecture, consisting of a conceptual model, a logical data model, and a physical database model [52]. A similar approach is proposed by CIMI Corporation [53] with the three data modeling phases conceptual model, logical model, and physical model, however integrating in the conceptual model things like organizations, people, facilities, products and application services, that way presenting somehow a Very High Level Data Model. However, this is done inconsistently by mixing domain ontologies with ICT ontology at one model level.

7. Business Modeling and Enterprise Architecture

Most of the early business modeling and enterprise architecture specifications have been developed by companies and industry consortia. Also, formal Standards Developing Organizations developed data modeling architectures. Many of them follow the conceptual model of architectural description defined in IEEE 1471:2000 [54], which has further evolved to the conceptual model of an architectural description according to ISO/IEC/IEEE 42010:2011 [55] (Figure 2).

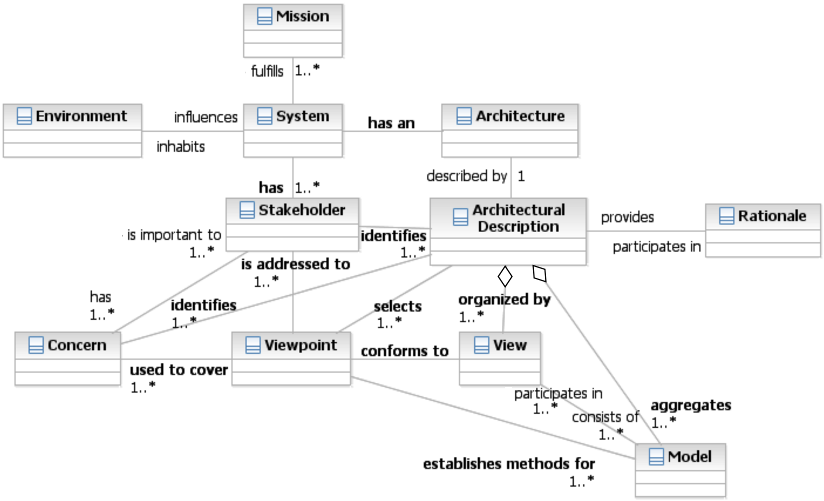

Figure 2. Conceptual model of architectural descriptions [13], changed after [55])

IEEE 1471-2000 [54] and its generalization to ISO/IEC 42010:2011 [55] define an architecture as "the fundamental organization of a system embodied in its components, their relationships to each other, and to the environment, and the principle guiding its design and evolution." According to those standards, a stakeholder is an individual, team, or organization (or classes thereof) with interests in, or concerns relative to, a system. Following this definition, an advanced approach to intelligent, knowledge-based ecosystems has been developed and currently standardized at ISO/TC215 as ISO 23903 Health Informatics – Interoperability and Integration Reference Architecture [56].

Another architectural approach to represent development processes of systems and solutions for open distributed processing is ISO/IEC 10746 [57]. It defines the representation of different views from the enterprise perspective describing the (IT) system in question (enterprise view), its platform-independent informational representation (information view) and computational aggregation (computational view), followed by the platform-specific engineering view (implementation) and technology view (deployment). The different views are represented through different languages with different grammars depending on the intended level of complexity and semantics to properly consider context, implicit knowledge, etc. Table 3 provides examples of representation tools (languages and grammars) to intendedly represent the different views.

Table 3. Representation tools for the different ISO 10746 viewpoints

Viewpoint	Language/Grammar
Business View	Business Process Modeling Language (BPML)
Information View	Unified Modeling Language (UML), Object Constraint Language (OCL)
Computational View	
Engineering View	Programming languages with different levels of grammar
Technology View	

An overview of alternative architectural models and frameworks such as the Zachman Framework for Enterprise Architecture [58], Reference Architectures for Service-Oriented Architecture (SOA) specified by The Open Group [59] or IBM [60], The Open Group Architecture Framework (TOGAF) [61], and many others is provided and in details discussed in [7].

8. A Reference Architecture Model and Framework for Representing pHealth Ecosystems

Translational medicine requires the advancement of communication and cooperation from data level (data sharing) to concept/knowledge level (knowledge sharing). This chapter shortly introduces the ISO 23903 Interoperability and Integration Reference Architecture model and framework for advanced interoperability in pHealth ecosystems. ISO 23903 is a system-oriented, architecture-centric, ontology-based, policy-driven approach including development processes following ISO 42010 and ISO 10746. Thereby, it extends the latter and details OMG's Model Driven Architecture (MDA) Computation Independent Modeling [62]. The approach enables cross-domain concept/ knowledge sharing, mapping as well as integration of, and interoperability between, independently developed specifications and related products by transferring proprietary into standardized concept representations without requiring any revisions of those artifacts (Figure 3). The approach shows similarities with communication standards such as HL7, EDIFACT, ebXML, etc., enabling data sharing between independently developed applications with proprietary data format by transferring them into a standardized EDI format, however on the highest interoperability level (knowledge-based and skills based) presented in Table 4.

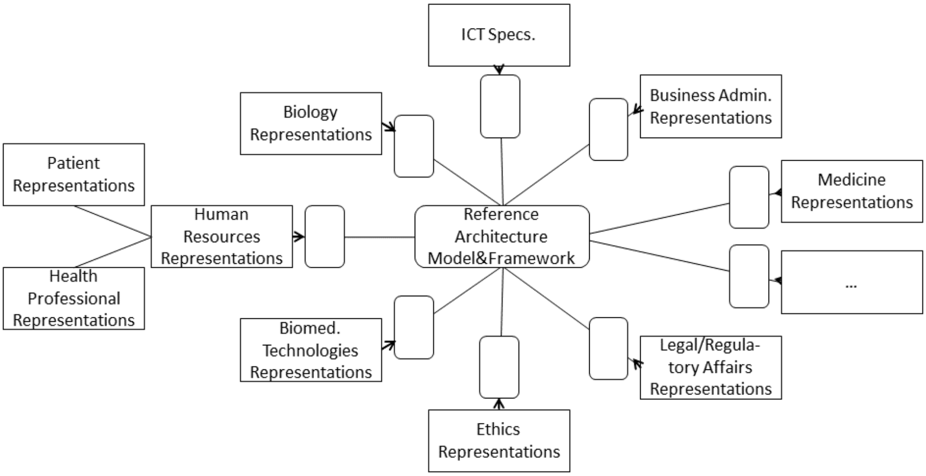

Figure 3. Interoperability and Integration Mediated by the ISO Interoperability and Integration Reference Architecture Model [56]

Table 4. Interoperability Levels [56]

Information Perspective		**Organization Perspective**
Interoperability Level	**Instances**	**Interoperability Level**
Technical	Technical plug&play, signal & protocol compatibility	Light-weight interactions
Structural	Simple EDI, envelopes	Data sharing
Syntactic	Messages and clinical documents with agreed vocabulary	Information sharing
Semantic	Advanced messaging with common information models and terminologies	Knowledge sharing at IT concept level in computer-parsable form → Coordination
Organization/ Service	Common business process	Knowledge sharing at business concept level → Agreed service function level cooperation
Knowledge based	Multi-domain processes	Knowledge sharing at domain level → Cross-domain cooperation
Skills based	Individual engagement in multiple domains	Knowledge sharing in individual context → Moderated end-user collaboration

All systems, work products or artifacts, developed independently under different perspectives and contexts (e.g. standards, components, FHIR resources, etc.), have to be formally represented according to ISO 23903 to interrelate them with each other for integration or interoperability. For that purpose, business systems must be modelled in a granularity or complexity appropriate for the considered use case. Thereby, each domain contributing to the business system in the context of the specific use case is represented at four generic levels of specialization/generalization. The components and intra-domain relations are instantiated using the corresponding domain ontologies. For facilitating the semantic integration of domain ontologies by harmonizing the representations and enabling the necessary concept mapping, top-level ontologies such as ISO/IEC 21838 Information Technology – Top-level Ontologies [63] are deployed. Top level ontologies provide at a very abstract level general concepts that are common to all domains, that way supporting the re-engineering of existing or the development of new ontologies. For properly representing semantics and pragmatics, appropriate grammars or logics have to be used. The outcome is the conceptualization of the real world business system and use case in question as the ISO 23903 Business View, which is thereafter transformed into the corresponding ISO 10746 RM-ODP views, deploying the representational tools (view ontologies) and thereby following the good modeling practices (Figure 4). According to the ISO 23903 framework, concepts can only be mapped and transformed at the same level of specialization/generalization. For getting there, the components have to be specialized or generalized before. The process of re-engineering existing systems for integration and/or interoperability is shown in Figure 5.

ISO 23903 also enables the harmonization and integration of existing architecture models and frameworks as well as reference architectures (RAs) such as IoT RA, Industry 4.0 RA, etc. Ongoing projects frequently ignore the presented principles, challenges and limitations by trying just to aggregate “related” information or data models or to perform term-based mappings.

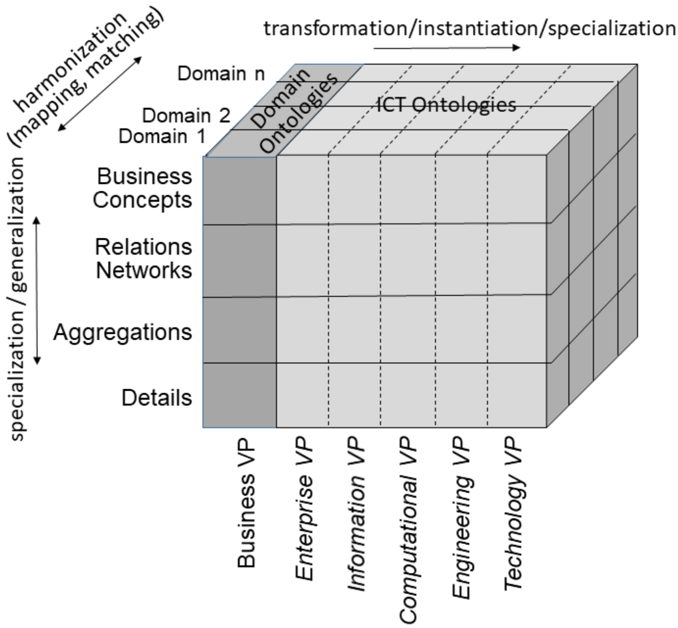

Figure 4. ISO DIS 23903 Mandatory Model and Framework [56].

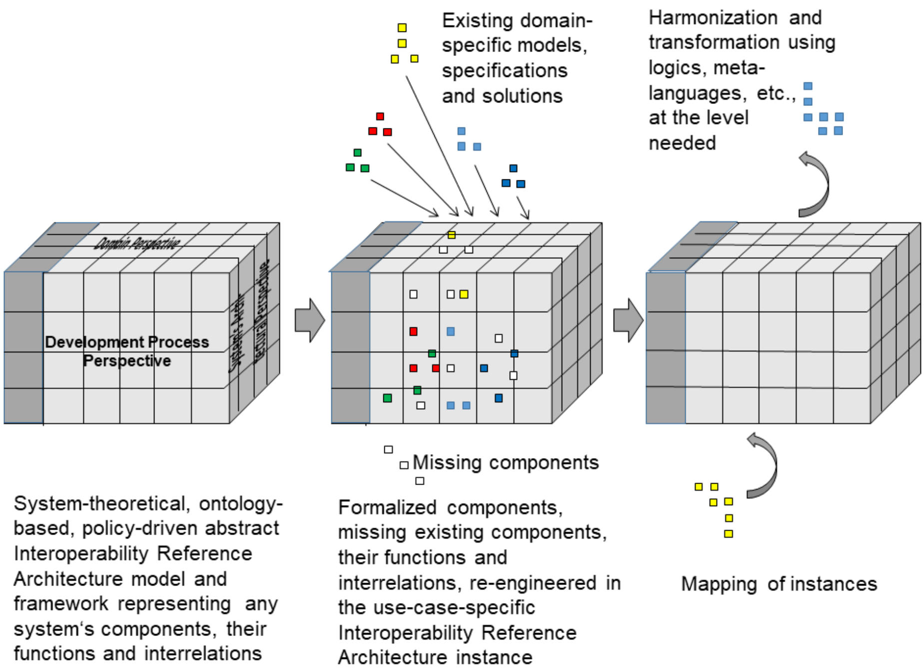

Figure 5. Integration of standards and specifications using the Reference Architecture model [56].

Before becoming an international standard, the presented approach developed by the first author and his teams has been successfully deployed in several cross-domain ISO specifications, such as ISO 22600 Health Informatics - Privilege Management and Access Control [64], ISO 21298 Health Informatics - Functional and Structural Roles [65]. Its feasibility has been practically demonstrated for automatically harmonizing HL7 v2.x and HL7 v3 specifications [66, 67] or for automatically designing inter-domain Web services to facilitate multi-disciplinary approaches to Type 2 Diabetes Care management [8, 9]. The approach also allows a comparative analysis and evaluation of ICT Enterprise Architectures [7].

9. Discussion

The presented data model levels [49, 50] roughly comply with the phases of OMG's Model Driven Architecture (MDA), defining the computation-independent modeling (CIM), platform-independent modeling (PIM) and platform-specific modeling (PSM). They are also comparable to the extended ISO/IEC 10746 Open Distributed Processing – Reference Model (RM-ODP) according to the ISO DIS 23903 Interoperability and Integration Reference Architecture model and framework introduced in the previous chapter. The Very High Level Data Model corresponds to the ISO DIS 23903 Business View, while the High Level Data Model corresponds to the RM-ODP Enterprise View, the Logical Data Model to the RM-ODP Information View and Computational View, and the Physical Data Model complies with the RM-ODP Engineering View and Technology View. The latter views represent implementation and deployment.

Table 5 compares the different Data Model Levels [46], Modeling Dimensions [48], related Modeling Actors and Model Scopes with ISO 23903 as well as other standards and specifications, demonstrating how ISO 23903 combines and advances all those approaches.

Table 5. Comparing Data Model Levels, Dimensions of Modeling, Data Model at Different Information Level and the ISO Interoperability and Integration Reference Architecture Model, applied to specification examples [19]

Data Model Level [46]	Dimension of Modeling [48]	Data Models at Different Information Levels [47]	Modeling Actors	Model Scope	ISO 23903 Interop. & Integration RA	Examples		
Very-high-level data model	Knowledge space	External	Business domains stakeholders	Scope, requirements and related basic concepts of business case	Business View			ISO 23903 Interoperability and Integration Reference Architecture
High-level data model	Knowledge	Conceptual	Business domains stakeholders	Relevant information and representation & relationships of basic concepts	Enterprise View	DCM, CSO	ISO 10746 ODP-RM	
Logical data model	Information	Logical	Data modelers and analysts	Layout & types of data and object relationships	Information View	HL7 V3 (CMETs), HL7 CIMI, openEHR Archetypes, FHIM		
					Computational View	HL7 FHIR		
Physical data model	Data	Physical	Data modelers and developers	Implementation-related and platform-specific aspects	Engineering View			

According to the basics of cognition theories and principles of knowledge representation and management, including appropriate representation tools (languages and grammars), KR and KM facilitate the representation of (parts of) reality to serve specific objectives and interests intended by the development teams. Available KR tools allow representing anything. Any representations are possible. However, a platform-specific model (view) wrongly representing the corresponding platform-independent model (view), or a platform-independent model (view) not complying with the enterprise view, etc., are useless and even dangerous. Therefore, the correctness, completeness and consistency of models (components, functions and relationships) as well as their integration and interoperability can only be justified at the systems levels they represent. The same holds also for the representation of a viewpoint by the more specialized/constrained view in the development process.

The presented system-theoretical, architecture-centered, ontology-based representation of business systems does not replace the other views and their models and specifications, but it facilitates their correct selection, constraining, completion and interrelation for systems integration and interoperability as demonstrated in Figure 5. That way, the correct deployment of existing work products such as HL7's Application Programming Interfaces (APIs) or Fast Healthcare Interoperability Resources (FHIR) [64] is enabled. Practitioners emphasize the importance of the approach derived in this paper as well. HIMSS Chief Technology and Innovation Officer Steve Wretling stated for example: "Interoperability needs an architecture in addition to APIs" [69].

10. Conclusions

All aspects or levels of knowledge representation and management from cognitive theories and modeling processes through notation up to processing, tooling and implementation undermine the need for a top down approach for any new system, new context or new use case at least in the first step. Only the business view allows the complete and consistent definition of concepts and constraints of components and relations. A solely data level focused development and implementation process cannot meet the challenges of highly dynamic and complex pHealth ecosystems. As models can be used to create new knowledge, aggregations in the sense of integration or interoperability can provide new insights at any viewpoint. However, the correctness of the outcome must be justified at reality. In other words, correct integration and interoperability cannot be guaranteed at data level or by just using terminologies, i.e. terms without the related concepts and relations. Standards defining platforms and frameworks must cover the entire continuum of knowledge representation and management presented in this paper. Just very few Standards Developing Organizations (SDOs) such as the Object Management Group (OMG) with its Model Driven Architecture (MDA) partially meet this challenge. Others like IEEE, ISO/TC215, CEN/TC251 or HL7 International have to extend their main focus from data models and implementable artifacts towards processes and related frameworks. With the recently approved ISO 23903, ISO/TC215 undertook an important step into that direction.

Acknowledgement

The authors are indebted to thank the reviewers of the paper, but especially Prof. Mathias Brochhausen, University of Florida, College of Medicine, for valuable comments and constructive recommendations. Furthermore, they feel obliged to their colleagues from HL7, ISO TC 215 and CEN TC 251 for their kind and constructive support and cooperation.

References

[1] Blobel B. Challenges and Solutions for Designing and Managing pHealth Ecosystems. Front. Med. 2019; 6: 83. doi: 10.3389/fmed.2019.00083.
[2] Blobel B, Ruotsalainen P, Lopez DM, Oemig F. Requirements and Solutions for Personalized Health Systems. Stud Health Technol Inform. 2017;237:3-21.
[3] Blobel B. Standards and Solutions for Architecture Based, Ontology Driven and Individualized Pervasive Health. Stud Health Technol Inform. 2012;177:147-157.

[4] Uribe GA, Lopez DM, Blobel B. Architectural Analysis of Clinical Ontologies for pHealth Interoperability. Stud Health Technol Inform. 2012;177:176-182.
[5] Blobel B. Translational Medicine Meets New Technologies for Enabling Personalized Care. Stud Health Technol Inform. 2013;189:8-23.
[6] Uribe G, Lopez D, Blobel B. Towards Automated Biomedical Ontology Harmonization. Stud Health Technol Inform. 2014;200:62-68.
[7] Blobel B, Oemig F. The Importance of Architectures for Interoperability. Stud Health Technol Inform. 2015;211:18-56.
[8] Uribe GA, Blobel B, López DM, Schulz S. A Generic Architecture for an Adaptive, Interoperable and Intelligent Type 2 Diabetes Mellitus Care System. Stud Health Technol Inform. 2015;211:121-131.
[9] Uribe GA, Blobel B, López DM, Ruiz AA. Specializing Architectures for the Type 2 Diabetes Mellitus Care Use Cases with a Focus on Process Management. Stud Health Technol Inform. 2015;211:132-142.
[10] Blobel B, Brochhausen M, Ruotsalainen P. Modeling the Personal Health Ecosystem. Stud Health Technol Inform. 2018;249:3-16.
[11] Brochhausen M, Bona J, Blobel B. The Role of Axiomatically-Rich Ontologies in Transforming Medical Data to Knowledge. Stud Health Technol Inform. 2018;249:38-49.
[12] Blobel B, Ruotsalainen P. Healthcare Transformation Towards Personalized Medicine – Chances and Challenges. Stud Health Technol Inform. 2019;261:3-21.
[13] Blobel B, Gonzalez C, Oemig F, Lopez DM, Nykänen P, Ruotsalainen P. The Role of Architecture and Ontology for Interoperability. Stud Health Technol Inform. 2010;155:33-39.
[14] Blobel B. Ontologies, Knowledge Representation, Artificial Intelligence – Hype or Prerequisite for International pHealth Interoperability? Stud Health Technol Inform. 2011;165:11-20.
[15] Blobel B, Oemig F, López DM. Why do We Need and How Can We Realize a Multi-Disciplinary Approach to Health Informatics? Stud Health Technol Inform. 2015;210:100-104.
[16] Blobel B. Knowledge Representation and Management Enabling Intelligent Interoperability – Principles and Standards. Stud Health Technol Inform. 2013; 186: 3-21.
[17] Blobel B. Reference Architecture Model Enabling Standards Interoperability. Stud Health Technol Inform. 2017; 235: 401-405.
[18] Blobel B. Architectural approach to eHealth for enabling paradigm changes in health. Methods Inf Med 2010; 49,2: 123-134.
[19] Blobel B, Oemig F (2018) Solving the Modeling Dilemma as a Foundation for Interoperability. European Journal for Biomedical Informatics, 2018; 14,3: 3-12.
[20] Oemig F, Blobel B. Natural Language Processing Supporting Interoperability in Healthcare. In: Biemann C, Mehler A (Edrs.) Text Mining – From Ontology Learning to Automated Text Processing Applications, 137-156. Series Theory and Applications of Natural Language Processing. Heidelberg: Springer-Verlag; 2014.
[21] Alter S. Information Systems – A Management Perspective. Reading: Addison-Wesley Publishing Company; 1992.
[22] https://www.merriam-webster.com/dictionary/knowledge
[23] Davenport TH, De Long DW, Beers MC. Successful Knowledge Management Projects. Sloan Management Review; Winter 1998; 39, 2; ABI/INFORM Global, pp. 43-57.
[24] Quine WV. From stimulus to science. Harvard University Press, Cambridge, Mass., 1995.
[25] Doerner H. Knowledge Representation. Ideas – Aspects – Formalisms. In: Grabowski J, Jantke KP, Thiele H (Edrs). Foundations of Artificial Intelligence. Berlin: Akademie-Verlag; 1989.
[26] Benjamins VR, Fensel D, Gómez-Pérez A. Knowledge Management through Ontologies. Conference: PAKM 98, Practical Aspects of Knowledge Management, Proceedings of the Second International Conference, Basel, Switzerland, October 29-30, 1998.
[27] Guarino: Formal Ontology, Conceptual Analysis and Knowledge Representation. Int. J. Hum. Comput. Stud. 1995;43:625-640. DOI:10.1006/ijhc.1995.1066.
[28] Makhfi P. Introduction to Knowledge Modeling. MAKHFI.com: www.makhfi.com/KCM_intro.htm
[29] Floridi L. Information. A very short introduction." Oxford: Oxford University Press, 2010.
[30] Schulz S, Jansen L. Formal Ontologies in Knowledge Representation. In: IMIA Yearbook of Medical Informatics 2013, 131-146.
[31] Davis R, Shrobe H, Szolovits P. What is a Knowledge Representation? AI Magazine, 14(1):17-3, 1993.
[32] Nonaka I, Takeuchi H. The Knowledge Creating Company. Oxford University Press; 1995.
[33] Blobel B. Knowledge Representation and Knowledge Management as Basis for Decision Support Systems. International Journal on Biomedicine and Healthcare 2017;5,1:13-20.
[34] Blobel B. Ontology driven health information systems architectures enable pHealth for empowered patients. International Journal of Medical Informatics 2011;80,2:e17-e25.
[35] Rebstock M, Fengel J, Paulheim H. Ontologies-Based Business Integration, Springer-Verlag, Berlin Heidelberg, 2008.

[36] Gruber TR. A Translation Approach to Portable Ontology Specifications. Knowledge Acquisition 1993; **5**(2):199-220.
[37] de Bruijn J, Lausen H. Active Ontologies for Data Source Queries. Lecture Notes in Computer Science May 2004;3053:107-120.
[38] Gruber T. Ontology Definition. In: Liu L and Özsu MT (Edrs.) The Encyclopedia of Database Systems. New York: Springer; 2009.
[39] https://en.wikipedia.org/wiki/Knowledge_representation_and_reasoning
[40] Barash M. Grammars for programming languages. Medium.com. https://medium.com/@mikhail.barash.mikbar/grammars-for-programming-languages-fae3a72a22c6
[41] Moeschler J. Formal and natural languages: what logic tells us about natural language? RoutIdge, January 2016. In: Barron, A., Gu, Y. & Steen, D. Routledge Handbook of Pragmatics. New York: Routledge. 2017, p. 241-256. Series Routledge Handbook in Applied Linguistics.
[42] Oviatt SL, Coulston R, Lunsford R. When do we interact multimodally? Cognitive load and multimodal communication patterns. Proceedings of the 6th International Conference on Multimodal Interfaces, ICMI 2004, State College, PA, USA, October 13-15, 2004. DOI: 10.1145/1027933.1027957.
[43] Catarci T, Leotta F, Marrella A, Mecella M, Sharf M. Process-Aware Enactment of Clinical Guidelines trough Multimodal Interfaces. Computers 2019;8(3):10.3390.
[44] Landgrebe J, Smith B. There is no Artificial General Intelligence. arXiv:1906.05833v2 2019. https://arxiv.org/abs/1906.05833v2
[45] Chomsky, N. 1956 Three models for the description of language. IRE Trans. Inform. Theory 2, 113–124. (doi:10.1109/TIT.1956.1056813).
[46] Jaeger G, Rogers J. Formal language theory: refining the Chomsky hierarchy. Phil. Trans. R. Soc. B 2012);367:1956–1970, doi:10.1098/rstb.2012.0077.
[47] Lankhorst M, et al., Enterprise Architecture at Work. The Enterprise Engineering Series. Berlin Heidelberg: Springer- Verlag; 2009.
[48] Wikipedia: Scientific modelling. https://en.wikipedia.org/wiki/Scientific_modelling
[49] 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.
[50] Ponniah P. Data Modeling Fundamentals. Hoboken, New Jersey: A John Wiley &Sons, INC.; 2007.
[51] Krogstie J. Business Information Systems Utilizing the Future Internet. LNBIP 2011;90:1-18.
[52] Sruthi Sirpur: Data Modeling for Systems Analysis, Course material from Information Systems Analysis 6840, https://www.umsl.edu/~sauterv/analysis/Fall2013Papers/Sirpur/index.html
[53] Nolle T. CIMI Corporation An enterprise architect's guide to the data modeling process. TechTarget, App Architecture, 10 Jun 2020.
[54] IEEE Standards Association. ANSI/IEEE 1471-2000 IEEE Recommended Practice for Architectural Description for Software-Intensive Systems. Piscataway: IEEE SA; 2000.
[55] International Organization for Standardization. ISO/IEC 42010:2011 Systems and software engineering - Architecture description. Geneva: ISO; 2011.
[56] International Organization for Standardization. ISO 23903 Health informatics – Interoperability Reference Architecture. Geneva: ISO; 2020.
[57] International Organization for Standardization. ISO/IEC 10746 Information technology – Reference Model – Open Distributed Processing. Geneva; 1996. www.iso.org.
[58] Zachman JA. The Zachman Framework: Primer for Enterprise Engineering and Manufacturing. Zachman International 2003. http://www.zachmaninternational.com.
[59] The Open Group. SOA Reference Architecture. 2016. http://www.opengroup.org/soa/source-book/soa_refarch/
[60] Arsanjani A, Gosh S, Allam A, Abdollah T, Ganapathy S, Holley K. SOMA: A method for developing service-oriented solutions. IBM Systems Journal 2008; 47, 3: 377-396.
[61] The Open Group. The Open Group Architecture Framework (TOGAF), Version 9.1. http://www.opengroup.org
[62] Object Management Group, Inc. http://www.omg.org
[63] International Organization for Standardization. ISO/IEC CD 21838 Information technology – Top-level ontologies. Geneva: ISO; 2020.
[64] International Organization for Standardization. ISO 22600:2014 Health informatics - Privilege management and access control. Geneva: ISO; 2014.
[65] International Organization for Standardization. ISO 21298:2017 Health informatics - Functional and structural roles. Geneva: ISO; 2017.
[66] Oemig F, Blobel B. A Communication Standards Ontology Using Basic Formal Ontologies. Stud Health Technol Inform. 2010;156:105-113.
[67] Oemig F, Blobel B. Establishing semantic interoperability between HL7 v2.x and V3: a Communication Standards Ontology (CSO). Journal of Health Informatics 2011;3(4):153-157.

[68] Health Level 7 International, Inc. https://www.hl7.org
[69] Sullivan T. Interoperability needs an architecture in addition to APIs. HealthcareITNews, September 19, 2018.