Medical Informatics in a United and Healthy Europe
K.-P. Adlassnig et al. (Eds.)
IOS Press, 2009

doi:10.3233/978-1-60750-044-5-195

Ontology-Driven Health Information Systems Architectures

Bernd BLOBEL [a,1], Frank OEMIG [b]
[a] *eHealth Competence Center, University Hospital Regensburg, Germany*
[b] *Agfa Healthcare GmbH, Bonn, Germany*

Abstract. Following an architecture vision such as the Generic Component Model (GCM) architecture framework, health information systems for supporting personalized care have to be based on a component-oriented architecture. Representing concepts and their interrelations, the GCM perspectives system architecture, domains, and development process can be described by the domains' ontologies. The paper introduces ontology principles, ontology references to the GCM as well as some practical aspects of ontology-driven approaches to semantically interoperable and sustainable health information systems.

Keywords. health information system architecture, semantic interoperability, ontologies, knowledge representation

1. Introduction

As healthcare paradigms turn towards ubiquitous care, which is personalized, intelligent and independent of time, location and resources, future-proof and sustainable health information systems have to support this paradigm by enabling ubiquitous computing [1]. Necessary multi-disciplinary, intelligent solutions for the health business domain must be based on a comprehensive architecture framework formally expressing all components of the business system in structure, behavior and interrelationships throughout the entire development process, covering requirement analysis, design, implementation, evaluation, use, and maintenance. The business view (i.e., the computation independent model of the system) is the challenging part of the game. From an architectural perspective, architecture vision, business architecture, information system architecture, and technology architecture of the business system have to be provided consistently at the required level of comprehensiveness. The challenge in eHealth systems and their services is to achieve semantic interoperability and sustainability by avoiding misconception of the integrated business processes and solutions. Standards and methodologies must be selected or developed, and a unified process has to be established [1].

The paper is restricted to the architectural issues of eHealth systems and their appropriate representation, thereby especially reflecting their multi-disciplinary properties to approach the different domains involved.

[1] Corresponding Author: Bernd Blobel, PhD, Associate Professor, eHealth Competence Center, Regensburg University Hospital, Franz-Josef-Strauss-Allee 11, D-93053 Regensburg, Germany; Phone: +49-941-944-6769; Fax: +49-941-944-6766; E-mail: bernd.blobel@klinik.uni-regensburg.de; http://www.ehealth-cc.de.

2. Material and Methods

For analyzing, designing and deploying advanced eHealth systems architectures, their recursively composable/decomposable components, functions and internal as well as external relations have to be modeled properly. The structure and behavior of systems, i.e., the components, their functions and interrelationships, determine the achievable business objectives the systems have been designed for. Real world systems usually serve different domains, follow domain-specific constraints and are domain-specifically modeled. The challenge to be met is the appropriate description of the intended flexible, scalable, portable, service-oriented, multi-lingual, multi-disciplinary, multi-modal, knowledge-based, trustworthy, and sometimes very complex system. This description can be performed a) by representing the system in natural language or domain-specifically specialized language (nominalism); b) by representing the thoughts/concepts about the system expressed in algorithmic or logical way (conceptualism); c) by constraining the concepts according to the reality (restricting values, operations, interrelations, etc.) (realism), as discussed later on in this section. In any case, the first step is constraining (simplification, reduction of) the system model in question following an architecture framework such as the Generic Component Model (GCM) [1], which has been successfully applied in the context of system developments [2] or international standards and specifications, e.g., [3]. Thereby, its components, functions and relations will be classified and separated according to the domain served, the architecture perspective established by composing/decomposing the domain-specific components, or the views on the system according to the development process perspective. All those perspectives are represented through specific ontologies. Ontology engineering and management provide the intended component models and their aggregation thereafter during the design phase, using techniques such as creation, coordination and merging [4].

According to the granularity/complexity level the system and its environment represent, different ontologies apply ranging from philosophical ontology down to ICT ontology. The philosophical ontology is the science of being qua being, i.e., the study of the nature of being, existence or reality in general, as well as of the basic categories of being and their relations. It deploys formal methodologies such as logic for representing them. Independent of our knowledge and investigation, general ontologies serve as overarching and eternal umbrella for upper level ontologies, domain ontologies or even application ontologies up to the level of an ICT ontology.

Hence the GCM architecture framework supports an abstract and at desired level formalized description of any component-based system, it can be used to deal with ontology systems as well. Following, the ontology engineering and management challenge will be discussed in some detail.

Ontologies typically consist of: Names for important concepts in the domain; Relationships between concepts including a Hierarchy of concepts; Attributes, properties, constraints; Individuals. Why is it hard but inevitable to engineer and to manage ontologies for eHealth system architectures?

Dealing with the study of things, an ontology describes and explains the universe in its organization, designation and categorization. Transferred into the ICT world, an ontology enables an explicit and precise description and communication of recognized and conceived knowledge of a domain, provided in a logic-based language with well-defined semantics allowing for machine processing and logical deduction [5].

A formal ontology must be structurally and functionally rich enough for principally enabling the description of every scientific theory of the real world and beyond. For that purpose, the general framework of a formal ontology has to be extended by appropriate non-logical constants, axioms, and meaning postulates that represent the basic concepts and principles of that science. Predicate quantifiers of type theory help to bridge between the domain language and the constructural language of that ontology [4].

Beside the aforementioned different ways of predication for reflecting a business domain, also the variety of categories and types of quantifiable variables for expressions offered to represent the ontological category of being defines the system of formal ontology applied [4]. The formal description of such domain representation embraces first-order logics at the one end or modal logic at the other one. The limitation of the paper does not allow going into deeper details

Hence, domains are logical theories which can be expressed as pairs of ontological signatures for describing the domain's vocabulary and a set of ontological axioms for specifying its intended interpretation [4]. For specifying domains' content, ontologies consist of sets of following entities: class or concept (unary predicate) representing a set of objects with common properties; instance representing the individual object; relation (n-ary predicate) representing functions; and axiom representing a formalized statement or assertion to model non-deducible knowledge. These formally modeled elements are called the primitives of an ontology, which have to be provided or at least to be qualified by domain experts. Domain knowledge might be represented in diverse ways leading to different ontologies for the corresponding domain. As domain ontologies comply with the concept representation of system components, they follow the architectural perspective of an architecture framework like the GCM, which offers concepts, relation networks, aggregations and details. Different combinations (business integrations) generate different ontologies at the corresponding level.

3. Approaches to Ontology-Driven System Architectures

Following paradigms are ruling the advancement of eHealth: business integration; acceptance of multiple ontologies for properly representing the different domains by domain experts; evolutionary growth of the knowledge base representing the corresponding domain ontology; semantic interoperability, seamless information exchange beyond the deployment of syntax standards. For realizing this, the process to develop ontology-driven system architectures has to be dynamic, evolutionary and based on open virtual communities [4].

Ontology engineering and management in the healthcare domain have a long tradition without usually naming the process accordingly. Starting with controlled vocabularies with restricted lists of terms such as catalogs, unstructured glossaries, structured arrangements of words such as dictionaries, thesauri and taxonomies including structured glossaries, the most advanced ontology type are formal ontologies using mathematical procedures and logical statements. As knowledge representation has been engineered by logicians and artificial intelligence experts using formal logics, the advent of meta-languages opened the opportunity of formalizing knowledge to business domain experts. As a result, a bunch of ontology system representations above nominalism and below formal logics has been introduced. Figure 1 presents a hierarchy of ontology types regarding formalization and expressiveness without claiming the completeness of the ontology type instances provided.

Considering the GCM terminology domain using SNOMED CT, the practically introduced ontology presentation styles (nominalism and logical realism) consider being as genus. This results in first-order logical representation of concepts [6]. Considering other GCM domains, some constraint subdomains (e.g., the isolated consideration of special security services such as primitive RBAC) are meeting the paradigm of one ontological category represented by one type of quantifiable variable, so they can be formalized through first-order logic. Extending the approach to policy driven privilege management and access control, there are contextual conditions to be reflected within categories requiring higher level of predicate logic to be established.

Because an ontology is defined as explicit formal specification of concepts in a domain and their interrelationships, ontologies can be described by meta-languages such as the Unified Modeling Language (UML), expressing the concepts in classes with attributes and operations as well as the interrelations in associations. This is done in newest HL7 V3 specifications [7]. The knowledge representation regarding both structural and behavioral aspects can be enhanced using the Object Constraint Language (OCL) to include logical relations. This is exemplified, e.g., by openEHR Archetypes, for marketing reasons using the openEHR-specific constraint language Archetype Definition Language (ADL), however [8].

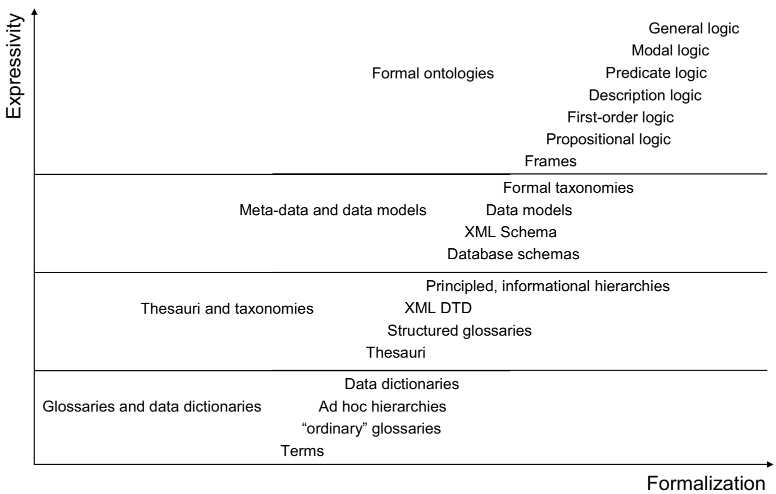

Figure 1. Types of ontologies (after [4], changed)

4. Discussion

As ontologies are intended to make the conceptualization of a domain explicit, knowledge representation as well as processing (reasoning) languages are practically suitable ontology languages. Besides expressing ontologies in logical terms, ontology languages such as OIL, DAML_OIL and OWL structuring the domain knowledge in terms of classes of objects and relationships between them are used. They represent a kind of description logic, which meanwhile covers a bunch of class-based representation formalisms used in artificial intelligence, software engineering as well as database design and management [9]. Such systems consist of the components description language, knowledge specification mechanism and reasoning procedures. Ontology engineering and management comprises the entire ontology lifecycle from creation (building, development, pruning, learning, modularization (extraction), evolution (versioning)), over coordination (matching, mapping, alignment (mediation)) through merging ontologies [4]. A health information system architecture based on the GCM framework has to reflect all those stages, as the GCM architecture perspective has to be represented by the aforementioned ontology creation mechanisms, while the domain perspective is represented by coordination and merging. For harmonizing ontologies to be considered in the case of complex systems or system integration, the reference ontology used has to be at least one level above the ontologies to be managed, regarding Figure 1. An instructive example for expressing specifications aiming at health information systems interoperability based at HL7 is given in [10]. An ontological approach to the GCM development process perspective (ICT ontology) is given in [5]. Following the IBM Standard for Architecture Description, an IT system development ontology has four major components: architecture assets, architecture decisions, stakeholder concerns, and architecture roadmap. Prior to the definition of the technical system with its components expressed by nodes and data, the components relations and interfaces to access them, the concerns of the stakeholders involved, covering business needs, risks, the required service quality, capabilities and changes have to be defined and properly expressed. The architecture decision is based on assumptions and possible alternatives to meet the aforementioned concerns. The roadmap finally initiates the development project considering all implications caused by the other basic classes [5]. This approach exemplifies the last step of an ontology-driven IT system architecture.

Acknowledgement. The authors are indebted to Stefan Schulz (University of Freiburg) and Sylvia Thun (DIMDI, Cologne) for kind cooperation and support.

References

[1] Blobel, B., Pharow, P. (2009) Analysis and evaluation of EHR approaches. *Methods of Information in Medicine* 48, 162–169.
[2] Lopez, D.M., Blobel, B. (2009) A development framework for semantically interoperable health information systems. *International Journal of Medical Informatics* 78(2):83–103.
[3] ISO TS 22600 Health informatics – Privilege management and access control.
[4] Rebstock, M., Fengel, J., Paulheim, H. (2008) *Ontologies-Based Business Integration*. Springer-Verlag, Berlin.
[5] Akerman, A., Tyree, J. (2006) Using ontology to support development of software architectures. *IBM Systems Journal* 45(4):813–825.
[6] Cocchiarella, N.B. (2007) *Formal Ontology and Conceptual Realism*. Springer, Dordrecht.
[7] Health Level 7, Inc, http://www.hl7.org.
[8] OpenEHR Foundation, http://www.openehr.org.
[9] World Wide Web Consortium, http://www.w3.org.
[10] Oemig, F., Blobel, B. (2009) Semantic interoperability between health communication standards through formal ontologies. *In this volume*.