PAPER II

Marco-Ruiz L, Pedrinaci C, Maldonado JA, Panziera L, Chen R, Bellika JG. Publication, discovery and interoperability of Clinical Decision Support Systems: A Linked Data approach. Journal of Biomedical Informatics 2016;62:243–64. doi:10.1016/j.jbi.2016.07.011.

Journal of Biomedical Informatics 62 (2016) 243–264

Contents lists available at ScienceDirect

Journal of Biomedical Informatics

journal homepage: www.elsevier.com/locate/yjbin

Publication, discovery and interoperability of Clinical Decision Support Systems: A Linked Data approach

Luis Marco-Ruiz [a,b,*], Carlos Pedrinaci [c], J.A. Maldonado [d,e], Luca Panziera [c], Rong Chen [f], J. Gustav Bellika [a,b]

[a] Norwegian Centre for e-Health Research, University Hospital of North Norway, P.O. Box 35, N-9038 Tromsø, Norway
[b] Department of Clinical Medicine, Faculty of Health Sciences, UIT The Arctic University of Norway, 9037 Tromsø, Norway
[c] Knowledge Media Institute, The Open University Walton Hall, Milton Keynes MK7 6AA, UK
[d] Instituto de Aplicaciones de las Tecnologías de la Información y de las Comunicaciones Avanzadas, Universitat Politècnica de València, Camino de Vera s/n, 46022 Valencia, Spain
[e] Veratech for Health SL, Doctor Lluch 299, 46011 Valencia, Spain
[f] Department of Learning, Informatics, Management and Ethics, Karolinska Institutet, Tomtebodavägen 18, SE 17177 Stockholm, Sweden

ARTICLE INFO

Article history:
Received 7 March 2016
Revised 5 July 2016
Accepted 7 July 2016
Available online 8 July 2016

Keywords:
Clinical Decision Support
Linked Data
Semantic Web Service
Semantic interoperability
Service Oriented Architecture
Medical ontologies

ABSTRACT

Background: The high costs involved in the development of Clinical Decision Support Systems (CDSS) make it necessary to share their functionality across different systems and organizations. Service Oriented Architectures (SOA) have been proposed to allow reusing CDSS by encapsulating them in a Web service. However, strong barriers in sharing CDS functionality are still present as a consequence of lack of expressiveness of services' interfaces. Linked Services are the evolution of the Semantic Web Services paradigm to process Linked Data. They aim to provide semantic descriptions over SOA implementations to overcome the limitations derived from the syntactic nature of Web services technologies.
Objective: To facilitate the publication, discovery and interoperability of CDS services by evolving them into Linked Services that expose their interfaces as Linked Data.
Materials and methods: We developed methods and models to enhance CDS SOA as Linked Services that define a rich semantic layer based on machine interpretable ontologies that powers their interoperability and reuse. These ontologies provided unambiguous descriptions of CDS services properties to expose them to the Web of Data.
Results: We developed models compliant with Linked Data principles to create a semantic representation of the components that compose CDS services. To evaluate our approach we implemented a set of CDS Linked Services using a Web service definition ontology. The definitions of Web services were linked to the models developed in order to attach unambiguous semantics to the service components. All models were bound to SNOMED-CT and public ontologies (e.g. Dublin Core) in order to count on a lingua franca to explore them. Discovery and analysis of CDS services based on machine interpretable models was performed reasoning over the ontologies built.
Discussion: Linked Services can be used effectively to expose CDS services to the Web of Data by building on current CDS standards. This allows building shared Linked Knowledge Bases to provide machine interpretable semantics to the CDS service description alleviating the challenges on interoperability and reuse. Linked Services allow for building 'digital libraries' of distributed CDS services that can be hosted and maintained in different organizations.

1. Introduction

1.1. Clinical Decision Support Systems interoperability and reuse

The term Clinical Decision Support Systems (CDSS) encompasses a wide range of recommendation systems that vary in purpose and complexity ranging from small logic modules that implement simple lists of order sets, to complex decision algorithms that compile the knowledge contained in nationally recommended guidelines [1]. Nowadays, it is generally acknowledged that CDSS contribute to improve health care, reduce costs and support access to the latest evidence [2–4]. However, their development costs are high as a consequence of the highly skilled professionals needed for knowledge engineering and development tasks [5–7]. For example, Field et al. estimated a cost of circa 49,000 USD only for the initial development of a set of CDS artifacts for medication alerts [8]. When it comes to more complex CDSS such as Computer Interpretable Guidelines (CIGs), the

* Corresponding author at: Norwegian Centre for e-Health Research, University Hospital of North Norway, P.O. Box 35, N-9038 Tromsø, Norway.
E-mail address: Luis.Marco.Ruiz@telemed.no (L. Marco-Ruiz).

http://dx.doi.org/10.1016/j.jbi.2016.07.011
1532-0464/

development becomes even more complicated elapsing over longer periods and being more error-prone [9]. Furthermore, for large deployments, it is necessary to create dictionaries of terms and data templates which need significant resources to be maintained. Maviglia and Sordo [10] estimated the cost of including one concept definition in their CDS terms dictionary at 6 h, allocating around 300 h to cover the approximately 50 concepts that they process per month.

As a consequence of the high development and maintenance costs, several studies have pointed to the need of reusing CDS artifacts and dictionaries of terms across organizational boundaries to avoid replicating development efforts [11–14].

In order to reuse artifacts, early approaches such as the Arden syntax [15] or GLIF [16] focused on defining standards to specify reusable decision logic. However, the deployment of the artifacts in a new Electronic Health Record (EHR) still required re-implementation of the other components that compose the CDSS. This requires redefining data mappings to the EHR, mapping terminological concepts, and re-testing the CDSS behavior. In summary, building all the parts that are not logic from scratch so it complies with the data and execution restrictions of the new environment where the CDS artifact is deployed.

To alleviate this problem, CDS researchers turned their sight to the SOA paradigm aiming for the reutilization of the whole CDS system [12,17–19]. In a nutshell, a SOA implementation encapsulates the CDS system inside a Web service which is shared among several clients. This approach switches CDS reuse from a paradigm of sharing decision logic to a paradigm of sharing CDS functionality. Thus avoiding the need to re-implement the system when a new client requires its functionality.

The encapsulation of CDS artifacts as Web services allows delegating expensive tasks related to the implementation, maintenance and governance of the CDS system. However, this delegation comes at a price. When a client relies on a CDS service maintained by a third party, the client does not have precise information about the features of the system beyond a syntactic definition in an Interface Definition Language[1] (IDL). As a consequence, barriers to enable client-service semantic interoperability (SIOp) have been detected related to difficulties understanding the semantics of the CDS service interfaces. Dixon et al. [12] and Wright et al. [20] detected major challenges to enable client-service SIOp related to difficulties in understanding the semantics of the CDS service interfaces when sharing CDS services among 4 organizations. When it comes to large health networks, such as those in European public health systems, SIOp becomes much more complex, and yet reusing such artifacts becomes even more appealing. The systems in a health network usually employ different standards and terminologies. In fact, not even the representation of Web service messages as Clinical Information Models (CIMs) [21] annotated with standard terminologies has resolved this issue [12]. When the clinical models are annotated with a standard terminology, the terminology codes add a certain degree of semantics indirectly, but the structure is still a syntactic description with a standard code as identifier. This structure contains implicit knowledge in the labels and descriptions expressed as natural language, but lacks a proper ontological foundation [22]. The attributes and labels of the model specify information structures, but they are not defined as concepts interrelated by meaningful machine-understandable relationships. The relationships among concepts need to be not only human readable, but also machine computable for CDS systems to function effectively and safely across EHRs [23]. For example, it is not possible to unambiguously infer that a particular label refers to a semantic relationship between two concepts inside a CIM, or that an attribute is semantically equivalent to another one in another CIM. This represents a major issue in CDS functionality reuse since accurate understanding of the concepts and the relationships referenced in the CDS service interface is a necessary condition to understand how to invoke it. For example, let us consider that a CDS for drug dosing is available and its valid input is an anticoagulant drug. It is not possible to automatically infer that the system may be invoked with an instance of Xarelto® because it is the trade name of the active substance Rivaroxaban; which, in turn, is a subtype of anticoagulant. These limitations are not only related to clinical knowledge specification, but also to the properties needed to express metadata for the governance of the system.

Overcoming these challenges requires adequate support for capturing and sharing clear unambiguous definitions of every CDSS, covering among other aspects the information structures consumed and produced by the service, the version of the system, the institution hosting it etc. Such definitions cannot be provided by Web services alone due to the syntactic nature of their underlying technologies (e.g. SOAP, WSDL or UDDI) [24].

Several areas of software engineering and artificial intelligence have already studied these challenges. The research on software components reuse has provided powerful mechanisms to unambiguously specify the system interfaces and also allow to automate tasks traditionally performed by humans.

1.2. Software components reuse

One of the most prominent research efforts regarding software components reuse has been performed by the Semantic Web community. As a result, they defined the Semantic Web Services (SWS) paradigm as Web services that are extended with semantic annotations to define the system properties in a machine interpretable fashion [25,26]. Thus encapsulating the component in a Web service that describes the system interfaces using an Interface Definition Language (IDL), at a syntactic level (e.g. WSDL), and semantic annotations to reference ontologies, at a semantic level. Examples of ontologies to attach semantic descriptions to Web services are WSMO [27] and OWL-S [28].

The reuse of software components through SWS lies in the implementation of mechanisms that allow the **publication** of the component; the **discovery** of the component by third parties; and, once discovered, the analysis of the component interfaces by the clients to understand the meaning of the information exchanged; i.e. **interoperate at a semantic level** [25]. These mechanisms should allow consumers to automate discovery and analysis of the system using machine-interpretable descriptions. In the SWS domain, to express the various types of system properties and interfaces four different types of semantics have been defined [26]:

(1) Functional semantics – describe which task the system performs (e.g. the system provides support for the treatment of Atrial Fibrillation).
(2) Data semantics – describe the information model consumed by the service operations (e.g. the system processes as input a *stroke prevention review* and provides as output a *stroke risk alert*).
(3) Execution semantics – describe exceptional behaviors such as the correctness of the service execution, conditions to execute the system and runtime errors. These type of semantics appear at runtime and are not usually covered by CDS standards.
(4) Non-functional semantics – describe properties of the system deployed not included in the previous categories. Examples of these properties are the issuer of the service, the version, the date of publication etc. (e.g. the system was issued by Cambio Healthcare Systems).

[1] In this paper the term Interface Definition Language makes reference to the languages used to specify Web services interfaces. Examples are the Web Service Definition Language (WSDL) or the Web Application Description Language (WADL) used to describe the Web service operations, messages, data types etc.

In order to reach the objective of publishing, discovering and exploring how to interoperate with services based on their functionality, non-functional properties and data interfaces, it is necessary to specify functional, non-functional and data semantics respectively. The models used in their specification must be shared across systems since they are the common language in performing the discovery and analysis of services properties. Conveniently, Semantic Web research has led to initiatives such as Linked Data that have defined the principles to interlink the contents of the Web in a machine interpretable format [29]. Data published following such principles are known as Linked Data. Linked Data has opened the door to offer the knowledge implicit in Web documents as explicit machine interpretable conceptual models (RDFS descriptions and ontologies). Therefore it allows to share machine interpretable knowledge bases (KBs) across applications. Based on the possibilities offered by Linked Data, the SWS paradigm has evolved to provide the processing layer to compute Linked Data leading to the so called Linked Services [30].

1.3. Objective

This work aims to provide insights into how SWS can be used to enhance CDS services with the final goal of developing methods for CDS services discovery and reuse. The study is oriented from the perspective of the current paradigm of SWS, i.e. Linked Services. Particularly, we study how to: (1) develop methods to define the CDS service semantics; (2) discover services based on such methods; (3) analyze their metadata, functionality and data models based on their semantic description to learn how to interact with them.

The remainder of the paper is organized as follows. Section 2 presents the state of the art in CDS standards and Linked Data. Sections 3 and 4 present the Linked Services platform used, the models developed to express different types of semantics and the use case used to validate our approach. Section 5 presents the discussion about the challenges found in knowledge representation, the benefits of exposing KBs and services as Linked Data and a summary of the steps to follow in order to publish CDS services as Linked Data. It is important to note that, when the term KB is used, we refer to static knowledge (concepts and relationships) rather than dynamic knowledge (decision algorithms).

2. Background and significance

2.1. The CDS standards ecosystem diversity

Several semantic dimensions are implicit in a CDS specification: system management properties (non-functional semantics), functionality (functional semantics), and information models (data semantics). There is a wide diversity in the standards available to specify the different types of semantics linked to a CDS. Even within each standardization body, several standards share different approaches for the specification of the same semantic dimension. Examples of the standards available to specify each type of semantics in the CDS domain are:

Non-functional and functional semantics: these two types of semantics are usually specified together in CDSS. Examples of standards that devote a section to cover these semantics are the Arden Syntax (maintenance and library categories), HL7 Knowledge Artifacts (KA) standard for detailed knowledge management control [31], SAGE (Metadata class) [32], the openEHR (ResourceDescription class) [33] and the properties described by the HL7 Decision Support Service Implementation Guideline (HL7 DSS IG) [19]. All properties in those standards and guidelines overlap significantly.

Data-semantics: In the CDS domain, data semantics are typically modeled as CIMs. Several standards are available to specify the CIMs referenced from the CDS algorithms. Formalisms like PROforma [34], Arden, openEHR Guideline Definition Language (GDL) [35], SAGE etc. have completely different ways of treating the link with their information models. Some use standards that diverge significantly one from another; while others, in turn, do not even define any mechanism to link to data models. For example, openEHR GDL uses archetypes; the Arden Syntax links directly to the database encapsulating queries in its data section; SAGE uses a VMR based on HL7 RIM [6,36]; and recent developments such as the EU project Mobiguide [37] advocate for the use of HL7 vMR [38]. Both openEHR archetypes and HL7 templates (created from CDA or vMR) can be bound to terminologies to enrich the data structures with a certain level of clinical semantics.

When implementing a CDSS, not only do we need standards to define functional, non-functional or data semantics, but we also need architectures that enable the consumption of CDS artifacts that follow these standards and terminologies. Currently, SOA architectural principles are recommended for wide implementations of CDS systems [17,19]. Several works have covered the definition of SOA architectures to leverage the use of information standards and terminologies [17,18,20,37]. Recently, the HL7 CDS group published the HL7 Decision Support Service Implementation Guideline (HL7 DSS IG) [19], based on the experiences gathered in the last decade. It provides a detailed description for leveraging the HL7 vMR and terminologies with SOA architectures (both SOAP and RESTful). As a reference implementation it is worth mentioning openCDS [39].

2.2. Limitations of current CDS specification standards: Illustrative example

To facilitate understanding this paper, we introduce the challenges faced with an illustrative example. Let us examine the CDSS presented in Fig. 1. The Stroke Risk CDS, represented as a hexagon, is a CDS module specified with the openEHR GDL formalism that takes as input an instance of the archetype with the *CHA2DS2-VASc score* and returns an estimation of *stroke risk* in percentage. Its metadata for authorship, life cycle and functionality are defined using the class ResourceDescription as per the GDL guideline. Information models are specified as openEHR templates that contain the archetypes openEHR-EHR-OBSERVATION.chadsvas_score.V1 (input) and openEHR-EHR-OBSERVATION.stroke_risk.V1 (output). GDL clinical logic references those archetypes to infer the outcome of the stroke risk for a patient in the next year expressed as a percentage. The archetypes and ResourceDescription class provide a description of the different semantic dimensions of the CDS. However, this description does not allow for reasoning. The consequence is that, if the system is exposed to third parties, they would not be able to automate the discovery based on intelligent queries; or perform reasoning to analyze how the concepts that describe the CDS semantics are related. For example, if we attempt to discover which CDS services for the treatment of heart diseases are available, a CDS that aims for the treatment of Atrial Fibrillation should be retrieved. However the *atrial fibrillation - heart disease* subclass relationship would not be evaluated and the system will not be discovered. Furthermore, the client would need a deep understanding of the openEHR specifications to explore the system and determine how to interact with it. These obstacles explode in complexity when not only one but several standards are used to implement different CDS artifacts. In fact, as discussed before, there is a considerable amount of equivalent CDS standards to represent each type of semantics. Therefore, besides the limitations described, it is unrealistic to expect that CDS clients could deal with all of them when discovering and analyzing the systems. This

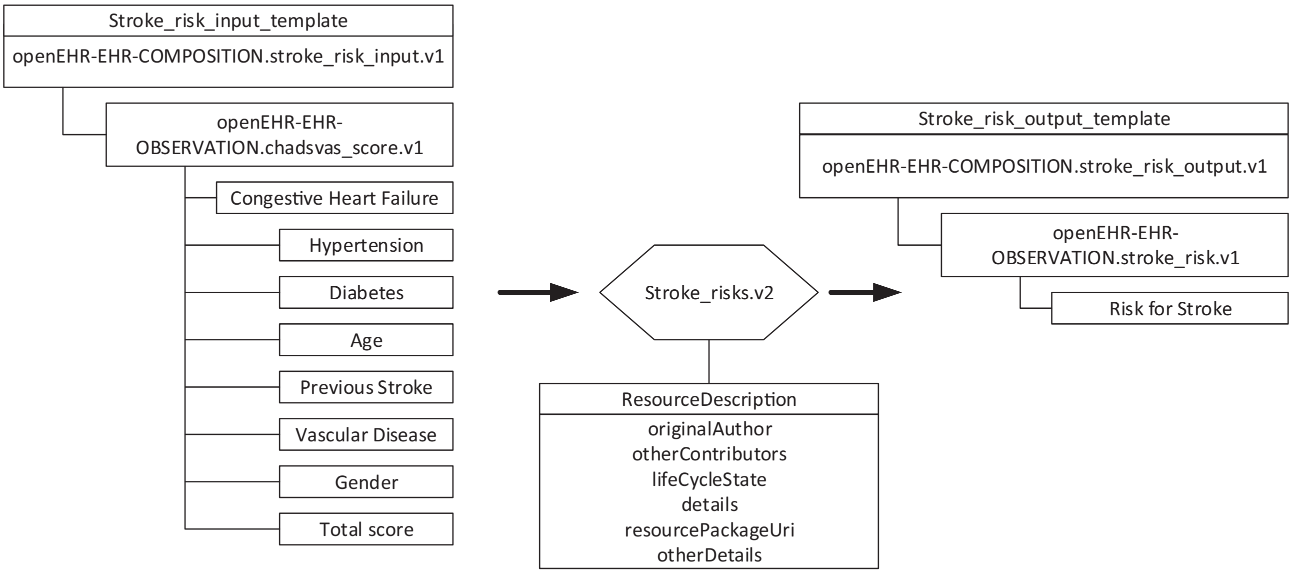

Fig. 1. Schema of the CDS module Stroke Risk.

would limit the level of integration and reuse of CDS across institutions.

CDS standards and terminologies have provided valuable mechanisms to decouple the CDS logic from the EHR. However, although they can implicitly express some degree of semantics in their attribute names and terminological annotations, they do not represent the full conceptual model of the concepts and relationships that describe the CDSS. This is simply because they have been designed prioritizing efficient information model representations rather than knowledge model representations. Thus, they effectively define information structures but do not use broadly accepted formats to define knowledge models in a machine interpretable manner (e.g. RDFS or OWL). This translates to: (a) limitations in the expressivity to define the CDS semantics; and (b) the lack of support for reasoning, for example, exploiting subsumption or equivalence relationships. This limited expressivity and reasoning support in turn has implications over the automated support that can be provided for advanced and flexible services discovery.

Besides, once a service has been discovered, every client needs to analyze it to determine what the appropriate way of invoking it is. For this analysis, it is not enough to explore the text annotations, data models and terminology codes of the system interface. Rather it is necessary to understand at a conceptual level how the entities that comprise the system are related among themselves and with other models. For example, in the example of the anticoagulant drug presented it is necessary to understand that a particular drug in the CDS knowledge base is semantically equivalent to another concept in the client conceptual model to invoke the service. Again, the ability to deal with more general or more specific concepts in this process is fundamental. The lack of ontological foundation in the CDS specification also hampers the reuse of models already expressed as Linked Data such as those available at Bioportal [40], forcing to replicate modeling efforts and re-map terminologies.

A constructive way of dealing with those limitations is to build upon the existing CDS standards, using them for what they were designed, i.e. the specification of information models in the case of archetypes, but extending them with mechanisms that allow to overcome the challenges discussed. This can be done without disrupting the CDSS already deployed applying the Linked Services paradigm. In fact, despite being different in the standard representation, when it comes to implementing the CDS system as a Web service, the basic technologies to specify data models and services at a syntactic level are mostly the same in all standards. For example, in data semantics specification, both HL7 vMR templates and openEHR archetypes are specified as XML Schemas at an implementation level to expose their data structures through a SOAP or RESTful Web service. Despite being structurally different, the specifications at this implementation level can be exploited enriching service descriptions and data schemas with a CDS standard neutral light-weight semantic layer. Such layer should be expressed in a machine interpretable format. As a result of this annotation, their CDS semantic descriptions can be published following Linked Data principles enabling the discovery and exploration of the artifacts regardless of the CDS specification standard as discussed later.

Additionally, all the infrastructure developed for Linked Data could be used on top of the CDS artifact specification regardless of the original specification formalism. This may enable not only reasoning but linking to other knowledge bases available in the Web of Data about drugs, genomics, proteins etc. Thus bringing further contextual information to facilitate CDS reuse. Moreover, all the tooling developed to navigate across RDF graphs can be used to perform such tasks alleviating the complexity of browsing Linked Data [41]. This facilitates the management of published CDS modules and the understanding of how to interact with the CDS at an abstract unambiguous level regardless of the standards used in their implementation.

2.3. Linked Data and The Web of Data

Linked Data is a set of principles derived from Semantic Web research to enable the publication of data on the Web in machine computable standard formats accepted by the World Wide Web Consortium (W3C) [42,29]. Besides, data published following Linked Data principles is also identified with the term Linked Data [43]. Linked Data is based on four principles [44]: (1) every resource exposed should be identified by a URI; (2) HTTP URIs should be used so people can look up resources; (3) the resource, when accessed, should offer machine computable information using standards such as RDF; (4) links to other URIs to discover related information should be offered [26]. The gradual incorporation of these principles and techniques is exposing the information contained in documents as interconnected computable data that can be navigated, discovered and reused using universal standard languages. This has driven the transformation of the Web of Documents into the so called Web of Data [45]. The Web of Data can be envisioned as a global growing repository in the form of navigable graphs that contains machine interpretable semantic descriptions of each object [45]. The most prominent developments in extending the Web with the Web of Data have been carried out by the Linked Open Data Project [46] and its central dataset DBpedia [47]. The collection of Linked Data published on the Web is known as the Linked Open Data cloud (LOD cloud). It contains interlinked ontologies about persons, places, drugs, genetics etc. Especially relevant are the ontologies for life sciences (e.g. Bio2RDF). The developments of the LOD project have led to the creation of an extensive global knowledge base. An example of how knowledge is expressed in the LOD cloud is shown by performing a search in DBpedia of Rivaroxaban. The search reveals that the RDF resource Rivaroxaban holds a meaningful relationship in RDF indicating that it is a type of Anticoagulant drug and that Xarelto® is both a synonym and a tradename of it.

The Web of Data has opened the door to produce applications that use its massive body of knowledge to navigate across services providing a processing layer to it. Based on that, Pedrinaci and Domingue proposed to evolve the paradigm of SWS into Linked Services [26]. Linked Services are based on principles for publishing service annotations (RDFS vocabularies) in the Web of Data to develop services that process Linked Data. The paradigm of Linked Services relies on models published as Linked Data to ensure an appropriate level of automation in discovery, analysis and invocation [30]. These models need to leverage expressivity and computation power; therefore light-weight ontologies must be prioritized to simplify the annotation of services. Complex models should be avoided since they are often downplayed as a consequence of the difficulties in their adoption. Additionally, when a new data set is created to annotate a service, it should be linked to existing ones to facilitate scalability and analysis of knowledge bases.

If applied to CDS services, the Linked Services paradigm can enable the specification and publication of services properties using the set of common ontologies in the LOD cloud; the integration of heterogeneous CDS services using a common semantic layer; and reasoning over semantic descriptions to perform discovery and analysis tasks.

3. Materials and methods

3.1. Overview

Aiming to facilitate the reuse of CDS functionalities, we analyzed how Linked Data techniques can be used to expose CDSS as Linked Services to enable their publication, discovery and SIOp. The definition of Linked Services for CDS involves two types of tasks: (a) the description of the service with a Web service modeling ontology; (b) the development of ontologies to attach non-functional, functional and clinical data semantics to services descriptions.

Regarding the first type (a), the services description ontology chosen was the Minimal Service Model (MSM) (described later) for its simplicity and the software available to support it.

Regarding the second type (b), the development of ontologies to express each type of semantics was done in three stages. First, we analyzed how to link the CDSS non-functional properties to existing ontologies for non-functional semantics specification. Second, we developed a taxonomy of possible functionalities of CDS systems to model their functional semantics. Third, we examined how to create ontologies for data semantics specification taking as starting point the implicit semantics contained in CIMs (archetypes) annotated with SNOMED-CT. Appendix A contains the description of the archetypes used in the examples presented.

Table 1 presents a summary of the technological challenges covered. The table compares the features provided by exposing CDS services using only existing standards vs. if the CDS implementations were evolved into Linked Services. The first row of the table presents some of the competency questions that each type of semantics helps to answer. To validate our developments we implemented a set of CDSS deployed for medical practice as Linked Services and performed discovery and analysis tasks over them.

3.2. Technological framework

As the technological platform we employed iServe to publish, discover and enable analysis of Linked Services [48,49]. iServe is a service warehouse developed in the EU project SOA4All [50] that runs on top of a triple store or reasoner. It enables Linked Services to be published, discovered and analyzed. It is based on light-

Table 1
Technological challenges and comparison of the approach presented with existing CDS standards.

	Non-functional semantics		Functional semantics		Data semantics	
Competency questions →	Which CDSS are issued by Cambio Healthcare Systems?		Which CDSS are available for stroke prevention? Which CDSS are available for stroke medication?		Which CDSS process a stroke prevention review and provide an alert for the stroke risk?	
Technological challenge ↓	Current CDS Standards	Linked Services/MSM	Current CDS Standards	Linked Services/MSM	Current CDS Standards	Linked Services/MSM
Publication	Limited to proprietary portals and text definitions	Exposed to the Web of Data as machine interpretable models	Limited to proprietary portals and text definitions	Exposed to the Web of Data as machine interpretable models	Based on specific standards like openEHR or HL7 vMR	Exposed to the Web of Data as machine interpretable models
Discovery based on semantics	Not supported(text descriptions)	Based on intelligent queries	Not supported(text descriptions)	Based on intelligent queries	Ad-hoc developments inside one repository (openEHR CKM, CIMI etc.)	Based on intelligent queries
Semantic Interoperability	Not supported(text descriptions)	Based on published ontologies in the Web of Data	Not supported(text descriptions)	Based on published ontologies in the Web of Data	Based on information standards and terminologies	Based on interlinked ontologies in the Web of Data

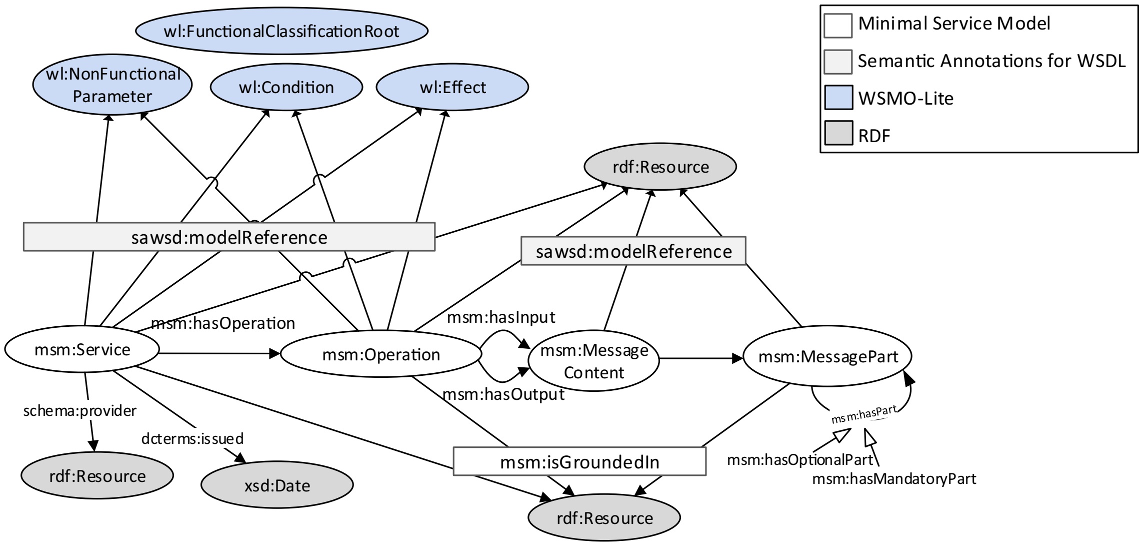

Fig. 2. Excerpt of the Minimal Service Model.

weight semantics [48] and supports services descriptions expressed in several formalisms. iServe allows annotating both SOAP and RESTful Web services [30]. Its underlying model is MSM which realizes the principles of Linked Services. MSM (depicted in Fig. 2) is an RDFS based on the principle of minimal ontological commitment that distills the common core between pre-existent conceptual models for SWS specification [30,51]. Therefore, it summarizes the core concepts of models such as WSMO or OWL-S overcoming the problem of the heterogeneity of SWS ontologies.

MSM acts as a proxy that connects the syntactic specification of the service (e.g. expressed as WSDL files and XML Schemas) with semantic models specified as domain ontologies (in RDFS or OWL) that contain the precise meaning of the system components. Several properties of MSM are available to attach the different types of semantics. Fig. 3 shows how MSM is used to connect the Atrial Fibrillation treatment CDS service (represented by a hexagon) to different ontologies (represented by clouds) each describing a type of semantics. The following describes how MSM is used to implement different types of semantics and the link to the syntactic implementation.

Non-functional semantics: Properties from models such as the Dublin core (*dc:* and *dcterms:* prefixes) or schema.org (*schema:* prefix) can be used to define non-functional semantics such as the date when the system was issued (dcterms:issued), the author (dc:creator), the organization responsible for it (schema:provider) etc. MSM can be extended with other properties using the class *NonFunctionalParameter* from WSMO-Lite (*wl:* prefix) if necessary. During discovery this type of semantics answers competency questions such as: Which CDS services were provided by Cambio Healthcare Systems? What is the last version of a CDS service? Which bibliographic citations support the algorithm contained by a CDS service?

Functional semantics: the relationship *modelReference* from the model Semantic Annotations for WSDL (*sawsdl:* prefix) [52] is used to reference concepts from a domain ontology that define the functionality of the service. That ontology should extend the concept *FunctionalClasificationRoot* from the model WSMO-Lite (*wl:* prefix) [53]. This indicates where the root of the hierarchy of concepts representing the different categories of CDS functionalities is located. During discovery this type of semantics answers competency questions such as: Which CDS services provide support for anticoagulant drug dosing? Which CDS services provide recommendations for laboratory test ordering? Which CDS services are available for recommendations on the treatment of COPD?

Data semantics: For data semantics specification, operations and messages that are consumed or produced by the system are defined. The example shown in Fig. 3 defines that the Atrial Fibrillation CDS has one operation (*msm:Operation*) called *RECOMMEND_AF_TEATMENT_PROCESS* which has one input message called *PROCESS_input* and one output message called *PROCESS_output*. The representation in MSM of the messages used by the CDS artifact can be itemized using *msm:MessageContent*, to identify the whole template structure, or *msm:MessagePart*, to describe the parts of it (e.g. LVEF). Parts can be nested without a limit and they can be defined as mandatory if needed. Equivalently to what is done in functional semantics, each part of the message can be linked to a domain ontology using *sawsdl:modelReference*. For a CDS service, these domain ontologies should define the essence of the implicit conceptual model contained in CIMs regardless of their implementation standard (openEHR archetypes/templates or HL7 CDA/vMR templates). During discovery this type of semantics allows answering competency questions such as: Which systems process an echocardiography? Which systems provide as output an estimation of stroke risk? Which systems accept as input a set of drugs and provide as output an alert when a possible drug-drug interaction exists?

Link to syntactic layer: When a service is discovered and analyzed, the syntactic layer where its implementation resides needs to be found to invoke it. This is done following the *msm:isGroundedIn* relationship. The grounding of the system and its components references a resource with the service URL in which an IDL description of the service should be available.

3.3. Ontology models

MSM provides the service description and the link between the syntactic level and the domain ontologies that provide the precise description of each component (clouds in Fig. 3). Without these ontologies the MSM description would only provide a structural definition. These ontologies have been developed as follows.

Non-functional semantics: properties from existing standards (Arden, SAGE and the HL7 DSS IG) have been revised by extracting their maximum common denominator. The set of properties

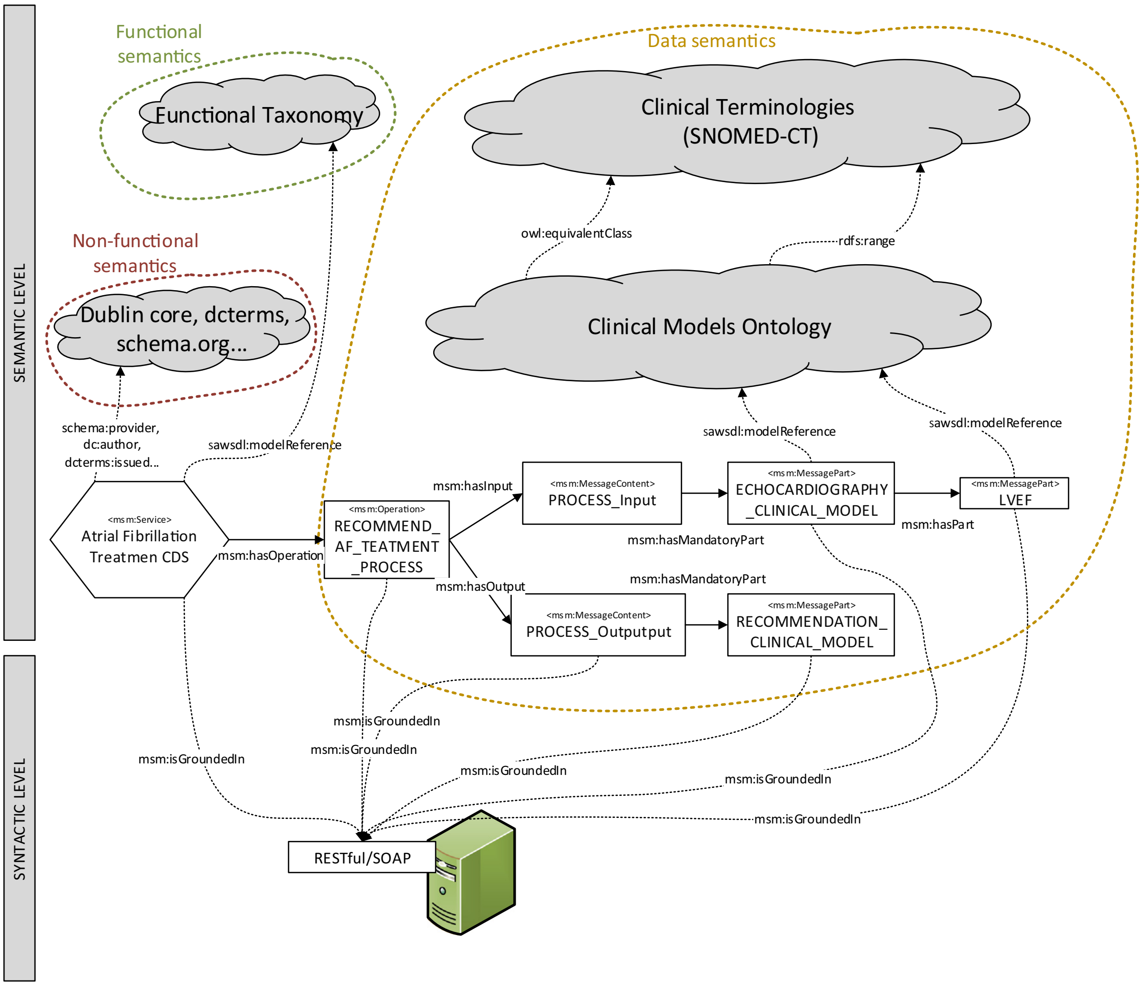

Fig. 3. Overview of the CDS Linked Service conceptual models involved in the CDS semantic description.

selected have been mapped to standard ontologies such as the Dublin Core [54].

Functional semantics: To define a complete taxonomy of functionalities, we reviewed the studies available [1,55–60] merging them in a RDFS taxonomy focused on their clinical target task. In order to define the clinical domain linked to the functionality, we studied several alternatives based on SNOMED-CT. We needed to find a compromise between the expressivity allowed, on the one hand, by OWL and SNOMED-CT post-coordination; and, on the other hand, the performance and viability of models for Linked Services annotation. We defined a RDFS poly-hierarchy of functional types linked to Clinical Findings and Procedures defining a subset of the post-coordinated expressions that can be defined composing[2] *Decision Making Support* with the attribute *hasFocus* (e.g. *Decision Making Support hasFocus Malignant tumor of colon).*

Data-semantics: To specify data semantics, we found it appropriate to follow a separation in three layers. Namely, the *service message model layer*, the *clinical models layer* and the *domain ontologies layer* (Fig. 9). Specifically, the *service message model layer* must be specified using MSM to represent the structures of the input and output messages of the WS. The different sections of the message need to be linked to the ontology that represents the *clinical models layer*. It was determined that the *clinical models layer* needs to define a Clinical Models Ontology (CMO) that distills only the clinical conceptual model implicit in the archetypes. This leaves data constraints divided between the *service message model layer* and the syntactical level (XML Schema types in the service specification) while sharing a common CMO among all services. Besides, the *clinical models layer* must be linked to the *domain ontologies layer* to enable semantic descriptions and discovery based on SNOMED-CT and other domain ontologies. We determined that the reference ontology of the *domain ontologies layer* used to attach clinical semantics should be an RDFS of SNOMED-CT concepts that, when needed, can be used to define mappings to other terminologies.

4. Results

This section presents the ontologies developed to attach each type of semantics to the services components and their linkage to the MSM service description. It is also illustrated how the grounding of the different components is performed. At the end of the section, we present the use case implemented to validate the models and methods developed.

4.1. Non-functional semantics

The set of properties representing non-functional semantics have been selected considering the properties usually attached to the CDS artifacts in existing standards. In particular, the properties provided by the Arden Syntax [61], SAGE meta data class [32] and

[2] Composition is the ontology term for coordination in the SNOMED-CT domain. Here it is used as an equivalent to post-coordination.

the HL7 DSS IG [19] metadata were considered. In a second stage, following the principle of ontology reuse and pursuing the maximum standardization level across different CDS specifications, the selected set of properties were mapped to standard ontologies such as the Dublin core [54] rather than define another model for CDS non-functional properties. Only properties common to several specifications were considered. However, if very accurate management of non-functional semantics is desired, other ontologies can be imported for that purpose. For example, non-functional properties available in PROV-O [62] and WSMO [27] have a good correspondence with those specified in standards for knowledge management (KM) such as HL7 KA. Table 2 shows non-functional properties in Arden, SAGE and HL7 DSS IG in the first three columns, and the properties in existing ontologies selected to annotate CDS services in the fourth column. Fig. 4 illustrates how the MSM service can be annotated with non-functional properties to

Table 2
Non-functional semantics in existing standards and in our proposal (last column).

Arden syntax	SAGE	HL7 DSS IG	Standard ontology equivalents used in our approach
Title	Description	Explanation	rdfs:comment
MLM Name	Label		rdfs:label
Arden syntax version			–
Version	Revision plan Release Version		dcterms:hasversion
Institution	Issuing organization	Steward Organization	schema:provider
Author		Author list	dc:creator
Specialist			–
Date		Creation date Last Review date	dcterms:datesubmitted, dcterms:dateaccepted
Validation			–
Purpose		Purpose	*(implemented as functional semantics)*
Explanation			dc:description
Key words		FreeTextKeywordList CodedValueKeywordList	dcterms:subject
Citations			dcterms:bibliographiccitation
Links	Endorsements		rdfs:seealso
Type	Category		dcterms:type
Data			–
Priority			–
Evoke	Usage context Enrolment criteria		wl:condition
Logic			dcterms:conformsto
Action			–
Urgency			–
	Knowledge development		–
	External review		–
	Recommendation		–

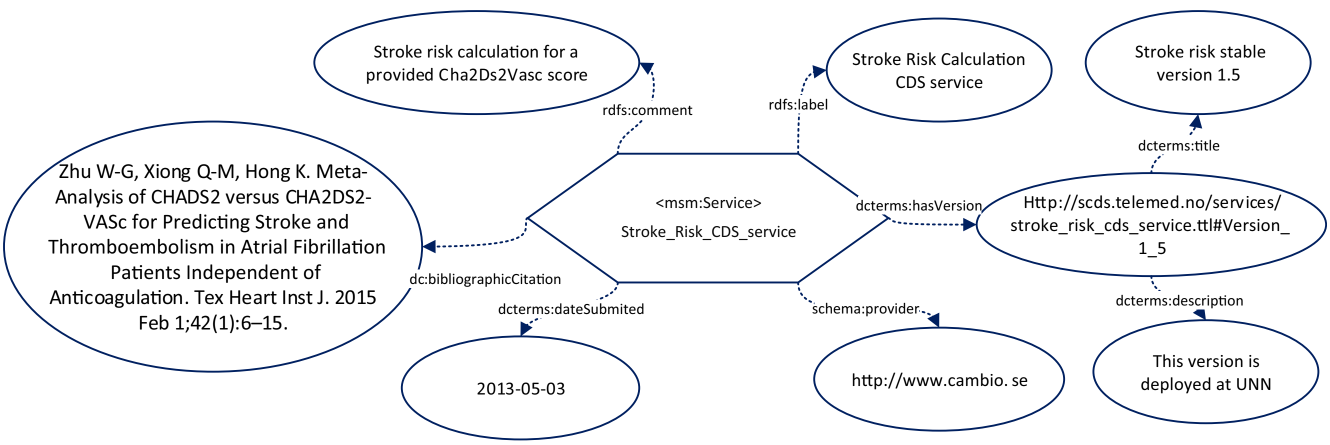

Fig. 4. Service annotated with non-functional properties.

define its version, provider etc. Only properties common to at least two standards, or considered of major importance have been mapped, the remainder have been left for not belonging to the common core of properties among the standards considered (hyphen symbol).

4.2. Functional semantics

Functional semantics were modeled in two phases. First, the literature was reviewed to define a taxonomy of possible functionalities. Second, the defined taxonomy was extended to link the clinical domain of application using SNOMED-CT.

4.2.1. Functional taxonomy

Fig. 5 shows the taxonomy coded in RDFS as a result of merging pre-existing taxonomies. Functional types are modeled defining a parent class *CDS functionality* as subclass of *wl:FunctionalClassificationRoot*. The class *CDS functionality* is specialized in the different functional types of CDSS extracted from the taxonomies available in the literature [1,55–60]. Seven studies were reviewed [1,55–60], five of them were complete taxonomies [55–59] whereas two were lists of CDSS types [1,60]. In the reviews we found that the term 'functional' was used with different meaning between CDS taxonomy studies and SWS models. The reviewed taxonomies considered as 'functional' some concepts that in SWS models are considered data and non-functional semantics. For example, in Wright's study [55] the category *Input data* would be considered data semantics in the domain of Linked Services. Another example can be found in Sim & Berlin and Berlin et al. [58,59] studies where the concept *knowledge and data sources* relates to non-functional semantics in the domain of Linked Services (e.g. 'Data source' is treated as dcterms:bibliographicCitation or rdfs:seeAlso in our non-functional properties set).

As a consequence, we determined that the aim of the functional taxonomy was to identify the pure functional aspect of the CDS artifact referring to a particular clinical focus; i.e. the CDS clinical task target.

Therefore, we did not select any section from Wright's taxonomy since it is focused in business process types (e.g. trigger that causes invocation, write an order, log etc.) and input data (e.g. family history). Regarding the other reviewed taxonomies we left aside those sections not specifying CDS clinical target tasks (functional semantics) and considered the others. From Sim & Berlin's taxonomy we selected the *Clinical task* subtaxonomy inside the *Context Axes* section; from Wang's we partially considered the *Benefit* level; and from Osheroff, Berner and HIMSS taxonomies top level concepts were all considered. For brevity, the rationale behind its organization is available as supplementary material in Appendix B.

4.2.2. Clinical domain specification

The defined functionality classes are very broad. In order to answer the type of competency questions presented (e.g. Which CDS systems are available for stroke prevention?), the clinical focus also needs to be specified.

For this purpose, the functional classification taxonomy in Fig. 5 is extended with the possible clinical focus concepts that can be post-coordinated to the concept *Decision making support in SNOMED-CT*, i.e. *clinical finding* and *procedure* hierarchies. The concepts in the extension are identified with an id whose first part is the task target type, and second part is the SNOMED-CT concept code. For example, the concept *Chronic Disease Management* with clinical focus on *Atrial Fibrillation* has the id *Chronic_Disease_management_focused_on_49436004*. The approach followed for specifying functionalities as a clinical target task linked to a clinical focus resembles the one proposed by Fox et al. for specifying clinical goals in PROforma (*Goal = Verb: Object*) [63]. Nevertheless, the goals ontology proposed by Fox et al. is designed to define clinical goals in a general way; whereas the presented taxonomy is designed to enable the discovery and specification of CDS services by functionality.

Fig. 6 shows an excerpt of the extended taxonomy. The concepts belonging to the pure functional part are represented as colored ellipses; while white ellipses represent extension concepts. The dotted rdfs:subClassOf relationships between white and gray colored concepts represent the extension point of the functional classification taxonomy with SNOMED-CT terms. Reasoning over subsumptive relationships in the extension is granted by the relationships colored in red. To facilitate the understanding of Fig. 6, the SNOMED-CT hierarchy is displayed in Fig. 7.

The generated taxonomy provides the vocabulary for all possible post-coordinated terms that can be attached to a functional concept. It enables the reasoning capabilities needed for discovery and matching of Linked Services. An example of the inferences that

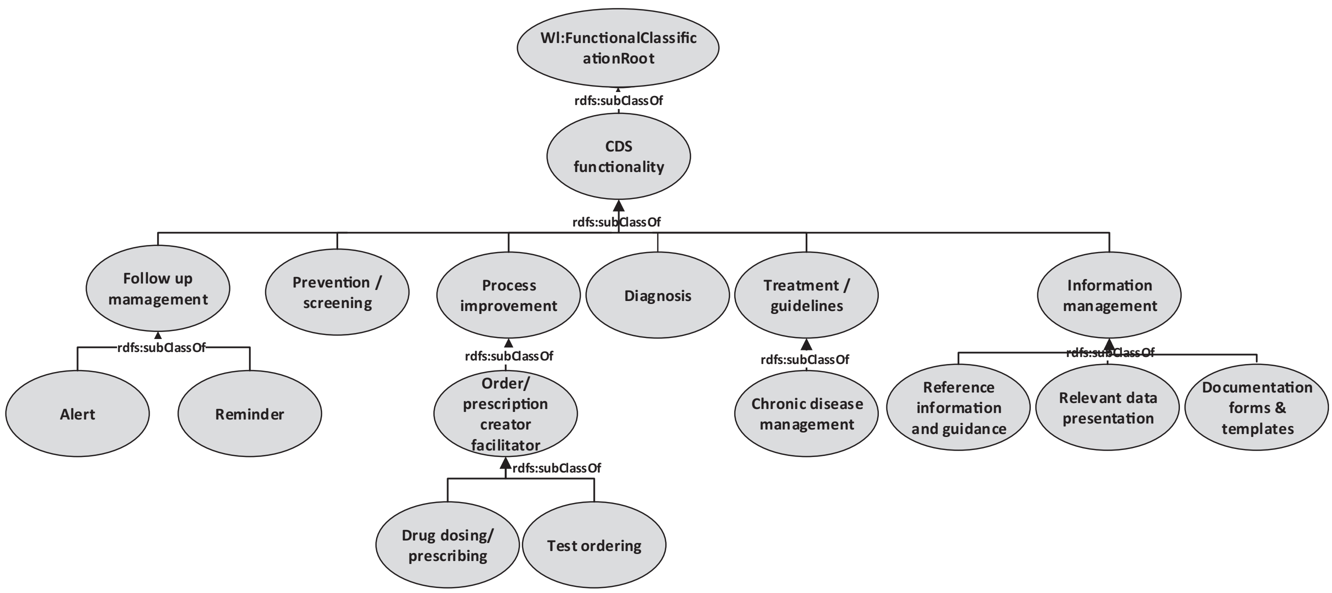

Fig. 5. Functional classification taxonomy.

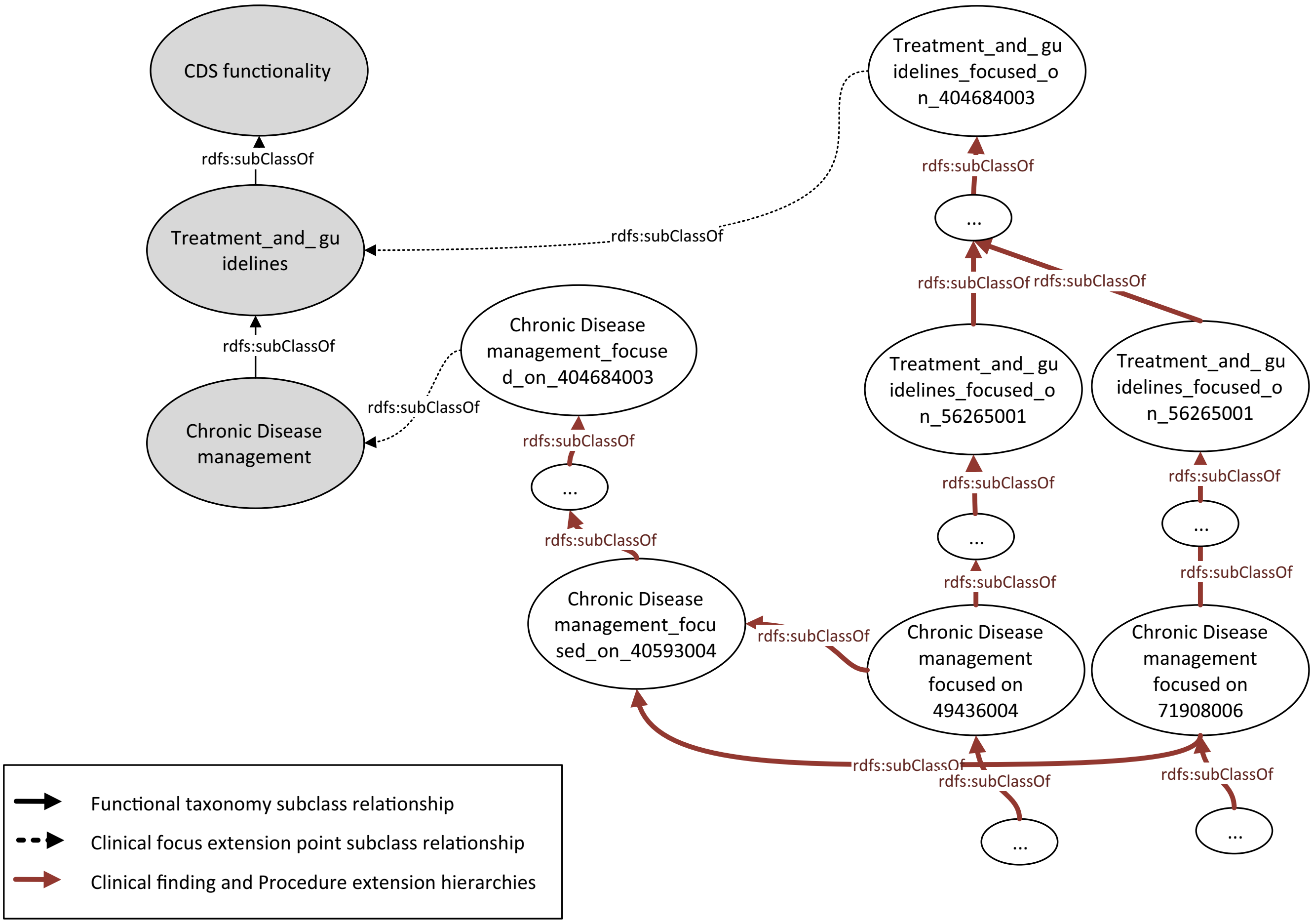

Fig. 6. Excerpt of the taxonomy generated to link the clinical focus to each functionality.

can be performed is to search for the CDS artifact annotated with the functional concept *Treatment_and_guidelines_focused_on_56265001* (56265001|Heart_Disease (disorder)) and retrieve by subsumption the CDS artifact annotated with *'Chronic_disease_management_focused_on_49436004'* (Atrial Fibrillation|49436004). Fig. 8 shows the functional annotation of the service to estimate the Stroke Risk.

4.3. Data semantics

The specification of data semantics for CDS services involves special challenges since terminologies, information models and medical knowledge need to be assembled to provide the expressivity and reasoning features needed. It is appropriate to implement the data semantics in the three separate layers shown in Fig. 9: the *service message model layer*, the *clinical models layer* and the *domain ontologies layer*.

The *service message model layer* represents the structure of the data contained in the input and output messages of the CDS service. The *clinical models layer* represents the clinical knowledge embedded in the archetypes [21]; i.e. how different concepts from the clinical domain are organized to define a more complex concept. The *domain ontologies layer* represents ontologies used to bind reference knowledge models to the *clinical models layer*.

The separation of models allows to: (a) define only one common CMO that is reused among service message models, this avoids repeating ontology binding tasks which are the most complex as discussed later; (b) use the CMO to refine the semantics of concepts from reference domain ontologies when the binding to only one ontology does not suffice to reach the desired level of expressivity; (c) separate maintenance and reasoning over the models [64]. An example of separate maintenance happens when mappings to other terminologies need to be performed from the SNOMED-CT ontology without disrupting the CMO or the service message models.

4.3.1. The service message model layer

The *service message model layer* is defined by instantiating MSM to define the service operations and data structures managed by each CDS service. *Service message models* do not define any clinical semantics beyond the content of message labels and they are defined for each service. Fig. 10 shows the MSM instantiation performed to describe the input and output message models of the CDS service *Cha2ds2Vasc score calculation*. From left to right, the hexagon represents the *msm:Service* class instance that contains the *Cha2ds2Vasc score* CDS. The service has one operation that performs the calculation of the Cha2Ds2Vasc score. This type of operation has one input and one output message, identified in MSM as instances of *msm:MessageContent*. Input and output messages have both mandatory parts called _BASIC_DEMOGRAPHIC_CLINICAL_MODEL, _CHADSVAS_DIAGNOSIS_REVIEW_CLINICAL_MODEL and _CHADS2Score_CLINICAL_MODEL. Those parts specify how clinical models and their attributes are organized in the input and output messages of the CDS artifact. In particular, the archetypes referenced by the parts are openEHR-EHR-OBSERVATION.basic_demographic.v1, openEHR-EHR-EVALUATION.chadsvas_diagnosis_review.v1 and openEHR-EHR-OBSERVATION.chadsvas_score.v1. The next level allows to break-down the message into its subparts. This

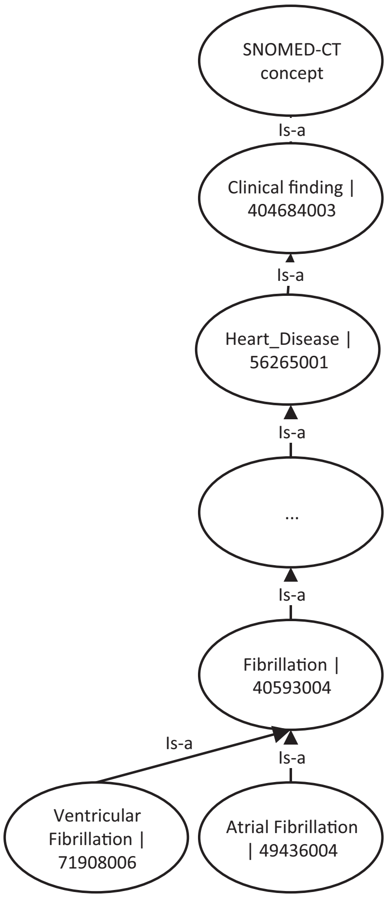

Fig. 7. SNOMED-CT concepts hierarchy excerpt.

process can be applied iteratively allowing to itemize input and output messages at any granularity level.

4.3.2. Clinical models layer

The *clinical models layer* is designed to attach clinical semantics to service message models. Clinical models (e.g. archetypes), in addition to data constraints, define a consensus of domain experts about how concepts from the clinical domain are organized to define more complex conceptual entities, thus defining the conceptual model of clinical content [65]. Provided that the elements contained in clinical models often correspond to concepts in biomedical ontologies, they convey knowledge about how concepts from reference ontologies (i.e. SNOMED-CT) may be combined to define clinical content models. For Linked CDS services this knowledge can be used to specify a common CMO that *service message models* can reference to attach clinical semantics.

The implicit ontology and terminology that archetypes convey [22,65] is taken as starting point to define the CMO that acts as a machine-interpretable projection of their underlying conceptual models. The CMO is limited to clinical concepts and properties that belong to the domain of clinical knowledge. Thus, it is free of the constraints that belong to the information model.

In order to truly represent unambiguous semantics and enable discovery based on standard vocabularies, the concepts and properties of the CMO must reference domain ontologies and terminologies. On the one hand, a general reference clinical ontology (SNOMED-CT) is needed to both link clinical semantics to the concepts of the CMO and provide a standard vocabulary to perform discovery. On the other hand, in many cases other ontologies from the LOD cloud need to be leveraged to refine the semantics of concepts which are not fully represented by SNOMED-CT.

Fig. 11 shows how concepts of clinical models defined as classes can be directly linked to SNOMED-CT defining them as an *equivalentClass* of the SNOMED-CT concept. With regards to the attributes of the clinical models that are modeled as properties, their domain is used to restrict the clinical models that may use them. Additionally, the range of the property is set, when possible, to the SNOMED-CT candidate that best identifies the attribute. For example, the attribute *hasSelectedDrug* for the clinical model *Stroke Prevention Review*, is modeled as a property with range 410942007 (*Drug or medicament*).

In cases where the term pointed by the clinical model attribute has no SNOMED-CT candidate or requires further refinement to preserve its semantics, it is possible to use standard ontologies from other domains available in the LOD cloud (striped ellipses in Fig. 11). For example, in addition to the range specified by the SNOMED-CT code, the duration of the delay is restricted by defining the concept *DurationInDays* as subclass of the W3C time ontology class *DurationDescription* that constraints units to days.

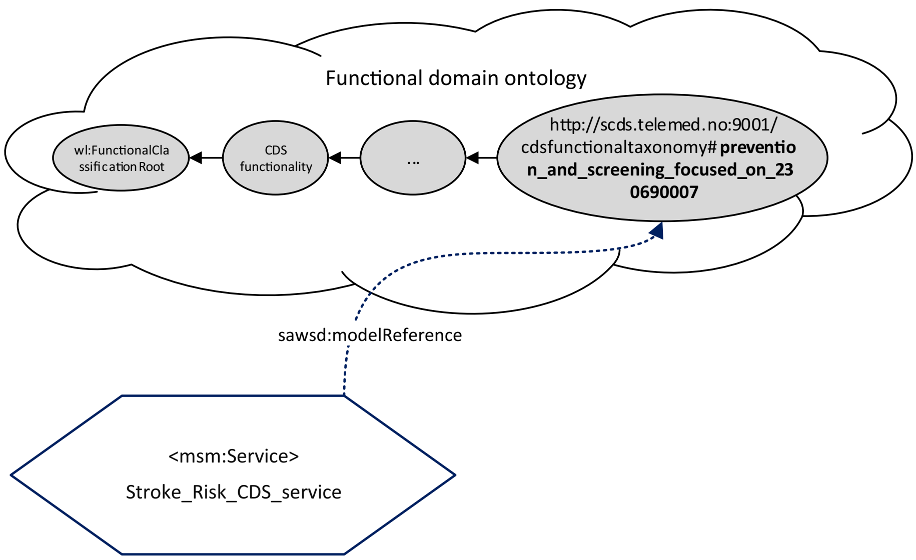

Fig. 8. Annotation of the MSM description of Stroke Risk CDS with its functionality.

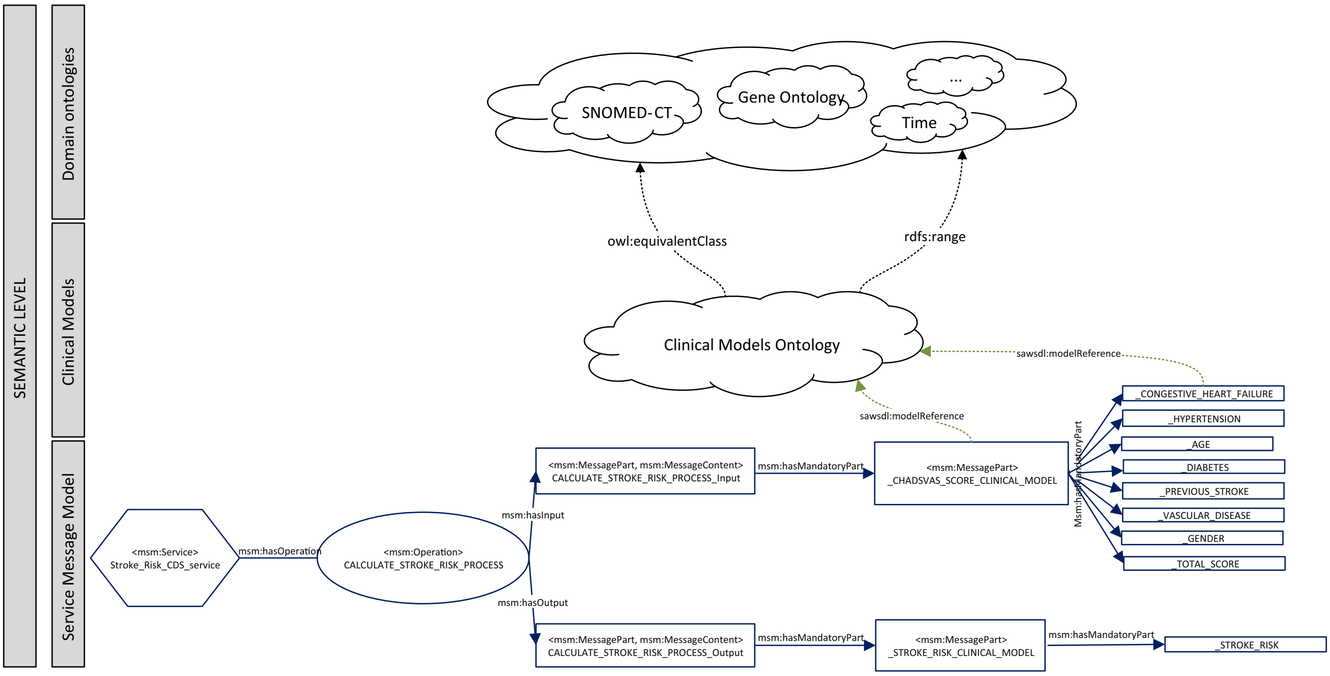

Fig. 9. Data semantics models and relationships.

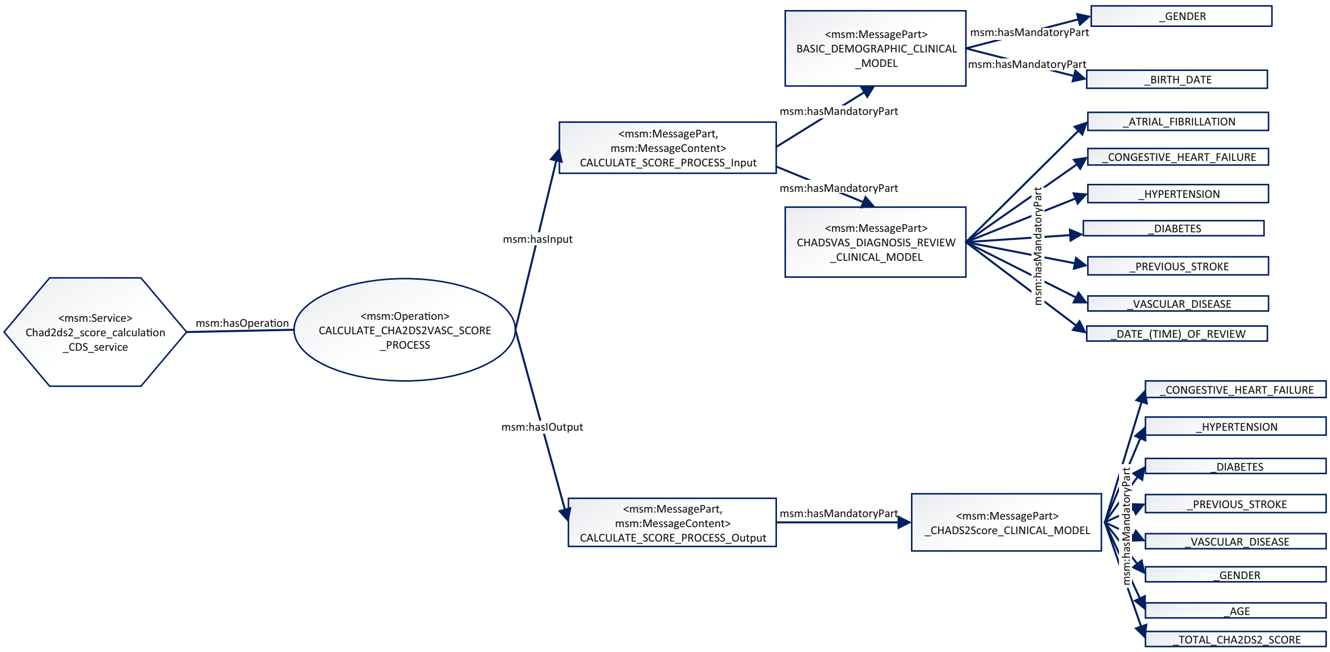

Fig. 10. MSM service message model specification for the Chad2ds2_score_calculation CDS service.

In some situations, not even the refinement of SNOMED-CT concepts with other ontologies suffice to fully describe a concept in the CMO. For example, the concept *Cha2Ds2Vasc score* is linked to the concept *CHADS2 score* (438367009) to allow discovery based on SNOMED-CT. However, it should be noted that these two concepts are not fully semantically equivalent. Although both are scoring systems to estimate the stroke risk, they are slightly different: the first is a refinement of the latter [66], but only *CHADS2 score* is available in the current release of SNOMED-CT. This needs to be clarified adding a comment as shown in Fig. 11 to disambiguate the meaning when a service is discovered using that concept.

4.3.3. Domain ontologies layer

The *domain ontologies layer* represents all those public ontologies available in the LOD cloud referenced by the CMO. Those ontologies represent validated knowledge models available in the WWW that anyone can access.

As shown in the previous section, it is appropriate to choose one main clinical domain ontology to serve as the main model to drive the binding of concepts to the CMO. Using SNOMED-CT as the main reference ontology allows to narrow the ontology binding task and, when necessary, refine its concepts semantics using other ontologies.

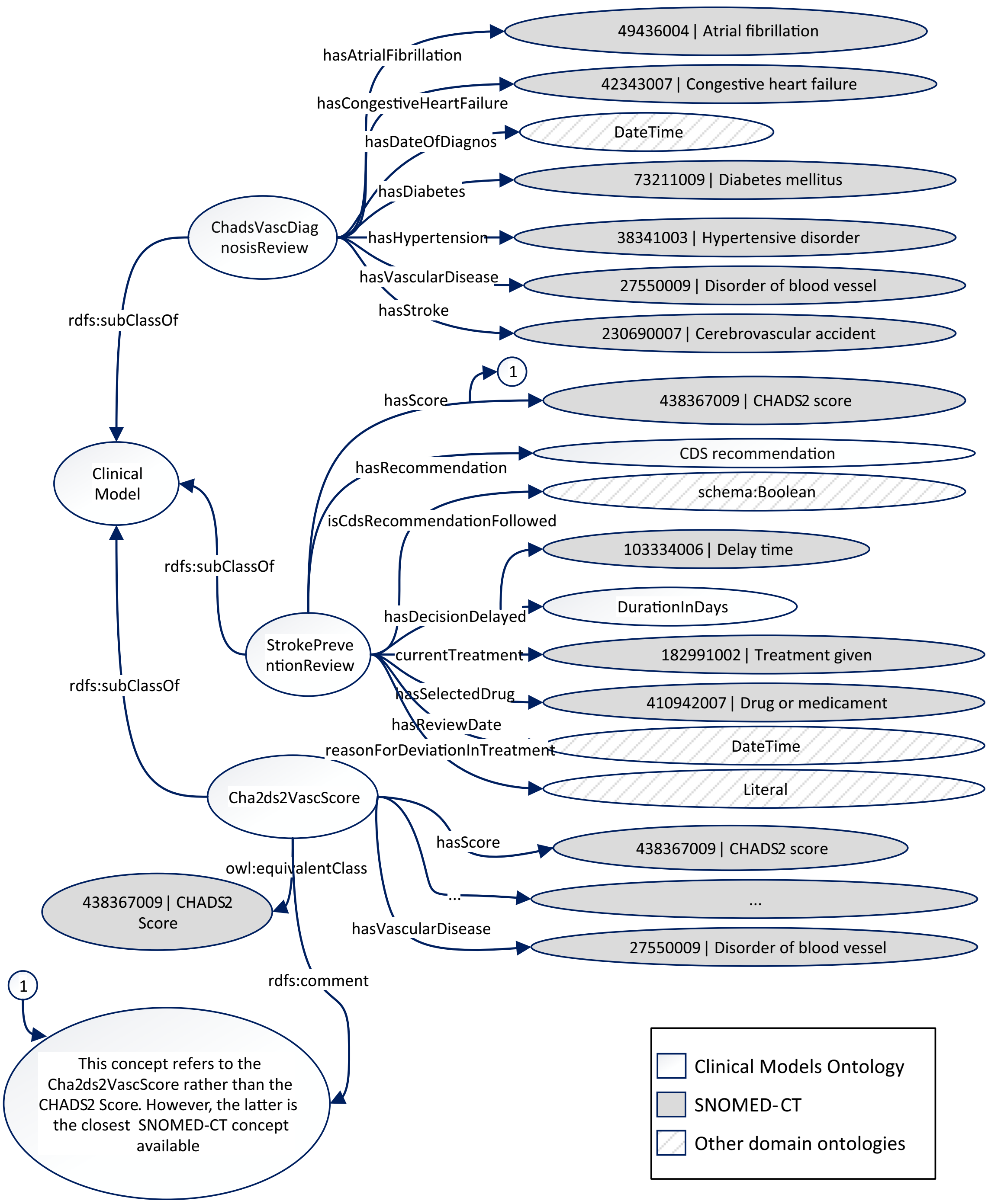

Fig. 11. Clinical Models Ontology (CMO) binding with SNOMED-CT and other domain ontologies.

There are some requirements that the main reference ontology implementation needs to fulfill. The first is to be broad enough to serve as standard vocabulary for discovery. For that purpose the ontology deployment needs to also guarantee fast reasoning in order to efficiently answer discovery queries. The second one is established by previous research in Linked Services that points to the need of prioritizing light-weight models with subsumption as a main requirement over complex models [30]. This ensures that the majority of reasoners are able to process the ontology. The third one is to serve as a place to perform mappings to other terminologies both standard and proprietary.

These requirements are not fulfilled by the OWL SNOMED-CT implementation generated from the standard release. The semantics of the full OWL version are too heavy to enable fast reasoning and cannot be loaded in many reasoners without information loss. Additionally, from a design point of view, it would not be appropriate to perform ontology mapping inside it. Based on that, the reference clinical ontology implementation proposed is an RDFS of SNOMED-CT containing all parent-child relationships rather than a complete SNOMED-CT OWL ontology. Such representation provides a light-weight and fast reasoning replica of the terminology. In addition, it enables the use of RDFS classes as placeholders of the SNOMED-CT codes that can eventually be used to map to other terminologies. Although most standard terminologies are already mapped, this is particularly useful when an organization uses proprietary codes. In scenarios when the full

SNOMED-CT knowledge needs to be exploited, Linked Data principles can be applied to define a link from the RDFS classes to the concepts of a full SNOMED-CT version placed in a reasoner capable to process it. This way the aforementioned requirements are not compromised.

4.4. Link to the syntactic service implementation

Once the MSM *service message model* is defined, it is necessary to link the MSM components to the syntactic implementation. For example, Fig. 12 shows how the Stroke Risk service is exposed at a syntactic level in the WSDL file located at the URL referenced by *msm:isGroundedIn.* Also, the service input and output messages (numbers 1 and 2 in the figure) are grounded to the WSDL messages. The clinical models and their attributes are grounded referencing the XML Schema data types that represent them (numbers 3 and 4 in the figure).

4.5. Use case and validation

4.5.1. Implementation

The set of CDS artifacts shown in Table 3 have been analyzed and exposed as Linked Services with iServe. All CDS modules are systems provided by our industrial partner Cambio Healthcare Systems [67] currently deployed to support medical practice. All the CDS artifacts were expressed in openEHR archetypes and GDL rules, including terminology bindings. The CDS artifacts functionality varies among prevention, screening, treatment and diagnosis of stroke. The set of archetypes that compose the information model referenced from the decision logic of the CDS modules is shown in Appendix A.

The set of GDL files encompassing decision logic, terminology bindings and data models (archetypes) have been analyzed in depth to understand how these components fit together to produce the CDS outcome. Once the behavior of the CDS artifacts was clear to us, we coded a description of each CDS service instantiating the MSM to describe CDS non-functional properties and its service message model. From the MSM instance we also defined links to the CDS functional taxonomy and the CMO to attach semantic descriptions to the service functionality, the message content and its parts. Table 4 shows the functionality assigned to each CDS service. Grounding to the syntactic level of the WS implementation was also defined. After defining the specification for each CDS we uploaded them into iServe and tested the services publication, analysis and discovery capabilities.

Among the lessons learned during the implementation of the use case we found that, once the behavior of a CDS service is clearly understood, the tasks that can be performed with less effort are the specification in MSM and the definition of non-functional and functional semantics. On the other hand, the tasks that consumed most of the effort were the definition of the clinical models in the CMO and their binding with SNOMED-CT and the domain ontolo-

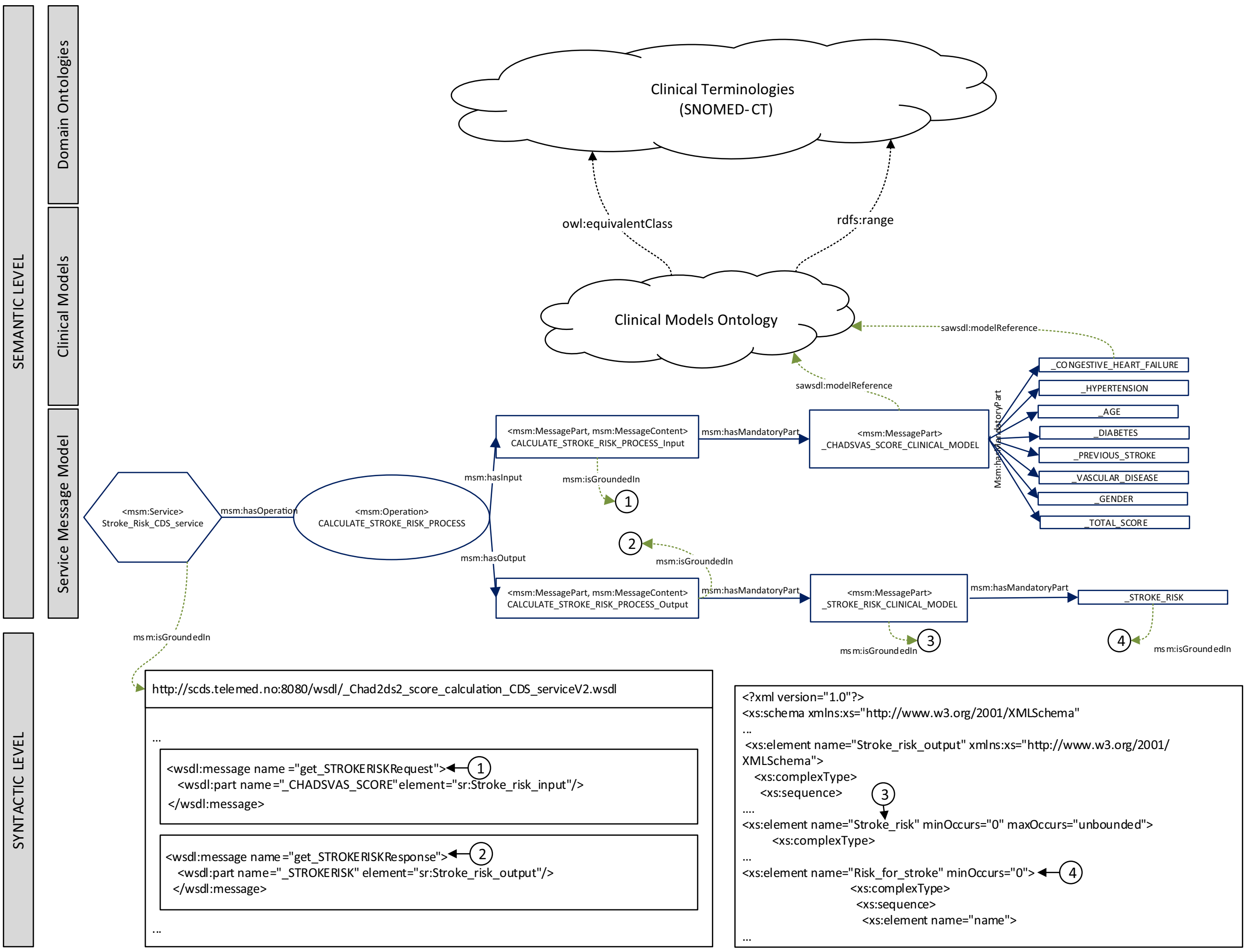

Fig. 12. Grounding of the Stroke Risk CDS module.

Table 3
CDS modules implemented as Linked Services.

Input template archetypes	CDS MODULE	Output template archetypes
openEHR-EHR-EVALUATION.problem-diagnosis.v1 openEHR-EHR-EVALUATION.exclusion-problem_diagnosis.v1 openEHR-EHR-OBSERVATION.imaging_exam-Echocardiography.v1	Atrial_Fibrillation_Treatment	openEHR-EHR-EVALUATION.recommendation.v1
openEHR-EHR-OBSERVATION.problem_diagnosis.v1 openEHR-EHR-EVALUATION.chadsvas_diagnosis_review.v1	CHA2DS2VASc_diagnosis_review	openEHR-EHR-EVALUATION. chadsvas_diagnosis_review.v1
openEHR-EHR-OBSERVATION.basic_demographic.v1 openEHR-EHR-EVALUATION.chadsvas_diagnosis_review.v1	CHA2DS2VASc_Score_calculation	openEHR-EHR-OBSERVATION.chadsvas_score.v1
openEHR-EHR-EVALUATION.problem-diagnosis.v1 openEHR-EHR-EVALUATION.chadsvas_diagnosis_review.v1 openEHR-EHR-EVALUATION.stroke_prevention_review.v1 openEHR-EHR-EVALUATION.cha2ds2vasc_compliance.v1 openEHR-EHR-OBSERVATION.chadsvas_score.v1	Stroke_prevention_alert	openEHR-EHR-EVALUATION.alert.v1
openEHR-EHR-EVALUATION.cha2ds2vasc_compliance.v1 openEHR-EHR-OBSERVATION.chadsvas_score.v1 openEHR-EHR-INSTRUCTION.medication.v1 openEHR-EHR-OBSERVATION.basic_demographic.v1 openEHR-EHR-EVALUATION.chadsvas_diagnosis_review.v1	Stroke_prevention_compliance_ checking_in_AF	openEHR-EHR-EVALUATION. cha2ds2vasc_compliance.v1
openEHR-EHR-INSTRUCTION.medication.v1 openEHR-EHR-EVALUATION.alert.v1	Stroke_prevention_medication_ recommendation	openEHR-EHR-INSTRUCTION.medication.v1
openEHR-EHR-OBSERVATION.chadsvas_score.v1	Stroke_risks	openEHR-EHR-OBSERVATION.stroke_risk.v1

Table 4
Functionality associated to each of the services.

CDS MODULE	Functional annotation
Atrial_Fibrillation_Treatment	chronic_disease_management_focused_on_49436004
CHA2DS2VASc_diagnosis_review	prevention_and_screening_focused_on_438367009
CHA2DS2VASc_Score_calculation	prevention_and_screening_focused_on_438367009
Stroke_prevention_alert	alert_focused_on_135875009
Stroke_prevention_compliance_checking_in_AF	chronic_disease_management_focused_on_49436004
Stroke_prevention_medication_recommendation	drug_dosing_prescribing_focused_on_135875009
Stroke_risk	prevention_and_screening_focused_on_230690007

gies available in the LOD cloud. This binding must be manually performed since it is crucial for the appropriate semantic specification of clinical models. The quality of ontology binding determines how successful the discovery and analysis of the CDS services by third parties will be.

4.5.2. Discovery of CDS services

The implementation was tested by querying on each type of semantics. We expressed the competency questions as queries and checked the results consistency. Table 5 presents the implemented queries, the types of semantics used by the system for their resolution and the CDS services retrieved. The queries are shown as URLs sent to iServe representing the concepts to search and the parameters of the query. iServe translates them to SPARQL queries that are executed against a triple store.

4.5.3. Analysis of services

To test how Linked Data allows to analyze services, we checked that the system displayed the CDS module non-functional properties, message structures and links to the different conceptual models providing semantic descriptions. We compared the data displayed by iServe GUI with the data in the GDL files and the archetypes referenced. The appropriate nesting structure and semantic descriptions based in the CMO and SNOMED-CT were checked for inconsistences. Fig. 13 shows, from left to right, the iServe GUI displaying the service Atrial Fibrillation Treatment after being discovered, the input for its operation and one of the clinical models in the input message (*Problem Diagnosis*). Since the system is described following Linked Data principles it is possible to check each of the components in detail by simply using the iServe web GUI to navigate across the URLs links (black arrows). For example, from the service screen, we can check the input of the operation. Then, from the input screen, we can check the *Problem Diagnosis Clinical Model* which is part of that input. Additionally, to see the details of the clinical model we can check its attributes following their URL links. All the components are in turn linked by a URL to the ontologies defined to attach semantics to them (green arrows). In turn, these ontologies can be explored in the same way using Linked Data browsers (e.g. tabulator) checking their concepts and discovering their relationship with other ontologies and terminologies. Since those ontologies are available in the Web of Data this provides implementers with a clear semantic description of each CDS service based on a shared knowledge model common to all them.

5. Discussion

5.1. What steps need to be followed to expose a CDSS as a Linked Service?

We have analyzed how the different semantic dimensions of a CDS system can be modeled to expose the system as a Linked Service to the Web of Data. An overview of the steps needed to publish a CDS service as Linked Service applying the methods developed follows.

1. Analysis of the system: Identification of the CDS artifact semantic dimensions and how they fit together to produce outcomes.

Table 5
Competency questions represented as URL queries and the results provided.

ID	Competency question	Type of semantics used for discovery	Query	Result
1	*Which CDSS are issued by Cambio Healthcare Systems?*	Non-functional	NF = schema:provider = http://www.cambio.se	• Atrial_Fibrillation_Treatment • CHA2DS2VASc_diagnosis_review • CHA2DS2VASc_Score_calculation • Stroke_prevention_alert • Stroke_prevention_compliance_checking_in_AF • Stroke_prevention_medication_recommendation • Stroke_risks
2	*Which CDSS are available for Stroke medication?*	Functional	F = http://scds.telemed.no:9001/cdsfunctionaltaxonomy#drug_dossing_prescribing_focused_on_135875009	• Stroke prevention medication recommendation
3	*Which CDSS are available for the treatment of heart diseases?*	Functional	F = http://scds.telemed.no:9001/cdsfunctionaltaxonomy#treatment_and_guideline_focused_on_56265001	• Atrial_Fibrillation_Treatment • Stroke_prevention_compliance_checking_in_AF
4	*Which CDSS are available to treat heart diseases that process as input an echocardiography report?*	Advanced (functional + data)	F: http://scds.telemed.no:9001/cdsfunctionaltaxonomy#treatment_and_guideline_focused_on_56265001 && Input:http://scds.telemed.no:9001/snomedfull/id/sn40701008\“, \”type\“: \”svc\“}}]}}	• Atrial_Fibrillation_Treatment
5	*Which CDSS process an echocardiography?*	Data	Input = http://scds.telemed.no:9001/snomedfull/id/sn40701008	• Atrial_Fibrillation_Treatment
6	*Which CDSS processes a Chadsvasc_score and provide as output an estimation of stroke risk?*	Data	Input = http://scds.telemed.no:9001/snomedfull/id/438367009 && Output = http://scds.telemed.no:9001/snomedfull/id/135877001	• Stroke_risk
7	*Which CDSS provide as output some type of stroke risk estimation?*	Data	Output = http://scds.telemed.no:9001/snomedfull/id/135877001	• Stroke risk

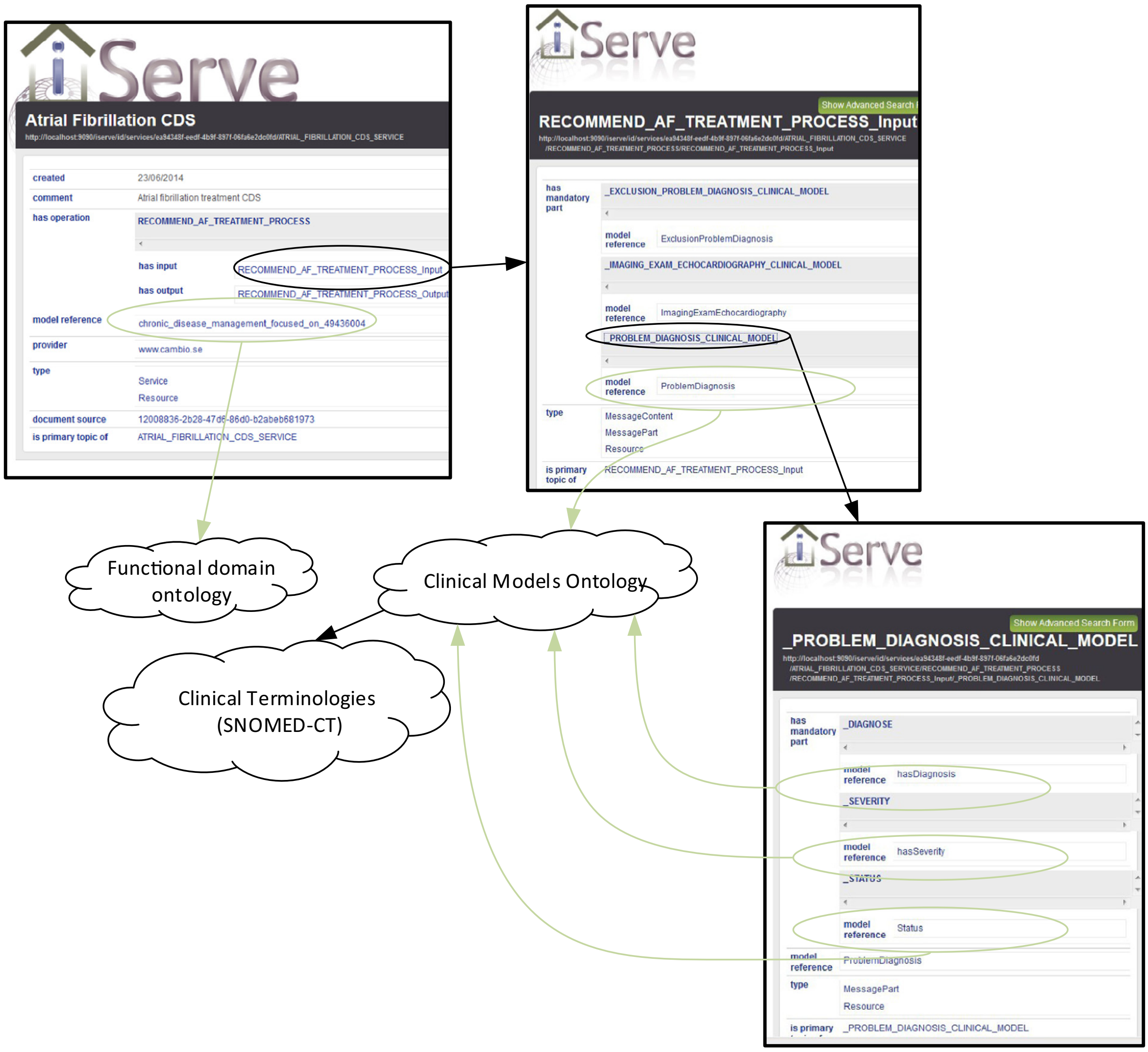

Fig. 13. Excerpt of the Atrial Fibrillation CDS service analysis after its discovery.

2. Semantic description specification: using MSM the service should be described at a semantic level. The implementer should annotate non-functional semantics with the standard ontologies discussed (see Table 2) and specify the service message model with MSM (e.g. Fig. 10).
3. Grounding: the implementer should link the service definition and the service message model parts to the IDL syntactic specification (WSDL for SOAP or Swagger/html for REST). Ground the information model to the data types used in the service interface (e.g. Fig. 12).
4. Ontology binding: functional semantics should be specified referencing the functional taxonomy presented. The implementer will need to decide which functional concept is the best to specify the clinical target task (see Fig. 5), and which term in the hierarchy of Clinical Finding or Procedure best describes the clinical focus to produce the annotation. Regarding the binding of the service message model, it is paramount to attempt ontology reuse by checking the models available in the LOD cloud; for example, Bioportal [40] and schema.org. If existing models fit the semantic description desired they should be reused; otherwise define a CMO with the patterns specified trying to reuse parts of other ontologies (e.g. Fig. 11). When reusing models link their classes to standard ontologies such as SNOMED-CT. A placeholder that extends the reused term with a relationship linking to the ontology can be useful for this. Relationships owl:sameAs or owl:equivalentClass can be used to link the code placeholders to terminologies. However, it is important to be aware of the reasoning implications of each of them [68].
5. Depending on the organization strategy, the implementer may decide to keep the models in a private health network or publish them in the LOD cloud. For the latter, it is important to create a VoID file, use sitemaps, register the data set in semantic indexes (e.g. Sindice, Data Hub) and define links from DBpedia. A good description of the details on how to publish data sets in the LOD cloud can be found in [43].

5.2. Knowledge Base modeling

5.2.1. Reasoning over SNOMED-CT

It may be argued that implementing the main clinical domain ontology by reducing SNOMED-CT to a mere RDFS hierarchy neglects much of the knowledge contained in it. We considered the possibility of using the SNOMED-CT OWL version. However, we found that the OWL representation of the complete terminology was too heavy to be fully processed in our OWL Horst triple store with the hardware resources we had available. After 48 h of processing, the computer (Intel Core i5 2.6 GHz with 8 GB RAM, 5 GB devoted to the triple store) ran out of memory. It is possible to load the full OWL DL representation in much less time (15 min) in a RDF triple store such as Apache Fuseki [69]. However, since this kind of repository is designed for RDFS, many properties and relationships expressed in OWL are not processed and therefore not exploited. This caused children to lose some of their parent concepts. Then, provided that our reasoning need is only subsumption, it was decided to aim for efficiency and limit the main clinical domain ontology to a SNOMED-CT RDFS representation. Besides, as discussed before, counting on a RDFS SNOMED-CT replica allows for the use of its concepts as placeholders for mapping heterogeneous terminologies.

Regarding the specification of the functional semantics, alternative approaches can be followed. We explored the possibility of using OWL axioms to define links between a CDS functionality and a clinical focus. However, if any post-coordinated expression length is allowed, the number of possible post-coordinated alternatives drives to a combinatorial explosion of possible expressions. For the reasoner this is unaffordable in terms of memory since all expressions should be loaded a priori for their classification in order to be used in a discovery query.

Nowadays some reasoners optimized for SNOMED-CT have proven to be very efficient. Karlsson et al. [70] documented the classification of 3,000,000 post-coordinated expressions in 3.5 min before the system ran out of memory with ELK [71] . Our functional taxonomy represents around 2.3 million post-coordinations. Therefore it may be feasible to use ELK if post-coordination is restricted to the combination with only one concept from *Clinical Finding* or *Procedure* hierarchies. However, we aim to create models that can be shared in the LOD cloud to offer them as shared knowledge bases. Nowadays, this kind of reasoners are not widely deployed in the Web of Data since they do not enable the performance and scalability that less expressive languages (RDFS) provide. Therefore forcing the users to adopt a particular set of reasoners to use some KB is not convenient.

With the light-weight models designed, the times for queries in our deployment environment (Intel Xeon 2.0 GHz with 20 GB RAM, 5 GB for the triple store/10 GB for iServe) are acceptable (less than 2 s) for all queries, except number 3, before being cached. After caching, all queries took less than half a second except number 3 that took 1.1 s. Although light-weight models fulfilled our requirements, it can be expected that some scenarios in the biomedical domain will need further reasoning capabilities. In such cases linked data principles allow to link light-weight models (e.g. the SNOMED-CT RDFS) to richer models hosted in advanced reasoners with relationships such as owl:sameAs and a proper URL.

5.2.2. Data semantics modeling

Similar difficulties to those described in the previous section are found when deciding how to model data semantics. Actually, when designing the CDS system KB it would be questionable whether direct annotation of the MSM information model (MessageContent and MessageParts) with SNOMED-CT would be a more effective and much simpler approach than creating the CMO. However, several limitations were identified when the *service message model* is directly annotated with SNOMED-CT.

The first limitation is that the lack of a CMO would force to replicate the ontology binding task for all service message models that specify the same CIM (archetype) structure in different messages. This is a major issue since ontology binding must be performed manually to ensure the proper semantic definition of clinical models and it is the most time consuming task.

The second limitation was the existence of concepts in CIMs that did not have a SNOMED-CT candidate. An example is the clinical model *CHADVAS_DIAGNOSIS_REVIEW*, whose closest expression in SNOMED-CT is *425268008|Review of care plan|:363702006|Has focus|=438367009|CHADS2 score*. Although both of them are used to estimate the stroke risk, the *CHA2DS2-VASc score* is not equivalent to the concept *CHA2DS2 score* available in SNOMED-CT. Rather, it is an extension that adds extra features and different scoring values to assess the risk of stroke [66]. Expressing the concept in the CMO allows us to attach a textual definition or annotate it with other terminologies to clarify this. This loss of meaning when no SNOMED-CT candidate is available appears as well in the semantics expressed in the labels of the clinical model nodes that cannot be fully expressed in SNOMED-CT without post-coordination. For example, in many cases the CIM attribute refers to a relationship (Property) rather than a concept (Class) as is the case of the *currentTreatment* attribute in the *StrokePreventionReview* archetype. A drawback of prioritizing light-weight semantics is that we cannot rely on post-coordination to express this term as *182991002|Treatment given|:408731000|Temporal context|=15240007|Current|*. Therefore using only the concept *Treatment Given |182991002* the semantics of the treatment being the one administered at the current time would be lost. Defining *currentTreatment* as a property in the CMO with range *Treatment Given|182991002* allows to deal with these situations preserving the meaning in the relationship but at the price of sacrificing the formality and precision that the post-coordinated expression would provide. Additionally, the CMO makes it easier to express concepts that require several ontologies to be expressed. An example of this was explained using the W3C time ontology [72] to refine the meaning from *delay time of decision* into *days of delay in decision*.

The third limitation is the inability to preserve the semantics implicit in the hierarchy and labels of the clinical models. An example of this situation is the openEHR archetype openEHR-EHR-EVALUATION.family_history.v1 containing the problem/diagnosis Diabetes. The hierarchy and the combination of the semantics implicit in the term diabetes included in a Family History tree allows to understand that the term diabetes is a diagnosis that does not refer to the subject but to a family member. In this case, the concept can be expressed using post-coordination (*57177007|Family history of|:246090004|Associated finding|=73211009|Diabetes mellitus|*) but again this would force the use of reasoners optimized for SNOMED-CT for the reasons described in previous sections. This case is even more challenging when the clinical model or some of its parts have no candidate in SNOMED-CT to be annotated with, and could not be approached without the definition of the CMO.

5.3. Linked knowledge bases

Having shared KBs is one of the main requirements for sharing CDS capabilities across organizations [12]. Following Linked Data principles all the concepts in the KB are identified by a valid URL that points to a computable resource. When exposed to the Web of Data, this allows for direct exploration of the semantic dimensions of the CDS services. This way, the models that specify the CDS service have the potential to be shared and reused by other developers that may, in turn, enhance them by adding new relationships, concepts or additional mappings to existing ontologies and terminologies. KB modeling based on Linked Data principles drives, in the end, the development of Linked Knowledge Bases (LKBs). When developing new CDS services, parts of such Linked Knowledge Bases may be used to define service semantics with a shared model making them semantically interoperable regardless the CDS standard used in their development. This would eventually provide linkage among the most common existing ontologies and terminologies that is currently another important barrier to share CDS capabilities across organizations [12].

It is interesting that most of the modeling work needed to develop a LKB has already been performed but, to our knowledge, not exposed as Linked Data. For example, the openEHR CKM [73] counts on a complete set of CIMs expressed as archetypes. It has recently accelerated the publication of stable versions of clinical models as a result of joint ventures with national CKMs such as the Norwegian one [74]. Another extensive modeling initiative is openCIMI [75]. Its objective is the development of a library of iso-semantic models that are made available in several formats (ADL, CDISC, SOA Payload etc.). One of the formats planned for the publication of models is OWL, which may open the door to expose clinical models as Linked Data. Some studies have used OWL to represent CIMs [76,77]. However, they represented the whole standard reference model entities as an ontology. This leads to the previously discussed problem of being limited by one particular standard. We believe that LKBs should be general enough to work on any standard representation and be free of data constraints. Although the modeling of such structures with OWL in a generic way presents some challenges, the work in the EU project SemanticHealthNet has provided patterns to facilitate such tasks [22]. The application of such patterns to the CMO could improve its design in many ways. For example, it could provide linkage of the CMO to upper level ontologies. Nevertheless, a thorough study on semantic patterns application using light-weight semantics to design LKBs remains as future work.

If those models are eventually exposed to the Web of Data, the years of modeling efforts invested by organizations such as openEHR, Intermountain, the Mayo Clinic etc. could be reused. This would lead to a machine interpretable shared core of clinical models available to specify data semantics of CDS services.

5.4. Reuse of CDS artifacts

Some studies [11,12] have proposed the implementation of repositories located at academic institutions such as medical schools to gather CDS artifacts that providers can access in the cloud. Projects like openclinical.org or DeGel have been precursors of such an approach [13]. Although this is a sensible strategy, it would require the allocation of significant resources to maintain such KM centers. Also, the standards chosen in the CDS artifact specifications may suit some providers but penalize others. We propose to orientate this approach in a distributed manner where those institutions that already maintain CDS artifacts can expose them as linked services without the need of a centralized KM body. The intrinsic nature of the Web of Data can provide the linkage among all those services creating a distributed 'digital library' of interlinked CDS services. These services would be located at the organization that uses them, and they would therefore be in charge of their maintenance.

Finally, it is important to remark that exposing services as Linked Data does not necessarily mean to open their interfaces to the global WWW. Linked Data techniques are a set of best practices that can be used on either side of the enterprise firewalls [43]. We envision the use of this approach inside medium-large health networks that decide to decouple the CDS functionalities from the EHR. CDS services would be offered to any CIS in the network that requires its functionality based on a shared LKB. Eventually, that health network may decide to make its KB and/or catalog of services available to other networks or the LOD cloud. Third parties could therefore discover them and, if interested, apply for access to the service. When logic needs to be modified, third parties may decide to apply local adaptions of the algorithm to fit their needs and eventually build upon the LKB defining their implementation as a subtype of the original one.

5.5. Limitations

This study is limited by the number of CDS modules implemented in this case study and their homogeneity since they are all are specified in openEHR GDL. However, we believe that our approach can be easily generalized to the represented modules by using the HL7 vMR/RIM information model. This would require grounding the concepts of the MSM in the vMR/RIM template XML Schema data elements rather than the XML Schema representation of the openEHR template. In fact, MSM and iServe define a generic approach to Linked Services that has already been broadly used in heterogeneous data schemas from other domains [30].

Another limitation is the need of a broad qualitative evaluation with CDS implementers and researchers on how representative of clinical target tasks our functional taxonomy is. It would be positive to study its combination with generic clinical goals ontologies such as that proposed by Fox et al. [63] to augment its expressivity. However it would be necessary to count on an open implementation based on W3C standards of such ontologies.

Finally, local adaption of CDS systems is one of the main barriers to share CDS capabilities. However, approaching it from a Linked Data perspective would require the study of techniques to explore the internal CDS logic [78] which are beyond this paper's scope.

6. Conclusion

The reuse of CDS capabilities is a task that involves major political, legal, organizational and technical challenges which is far from being resolved [11,12]. There is no silver bullet to overcome these challenges and the work described here is no exception. However, Linked Data, and in particular Linked Services, provide a paradigm that allows to enhance the valuable work performed in CDS standards development with the features of the Semantic Web, thus providing a paradigm that exploits the common body of knowledge available as Linked Data to specify CDS services properties. This enables the integration of heterogeneous CDS services under a common Linked Knowledge Base, the use of intelligent queries to discover them in a health network and the analysis of services interfaces based in rich unambiguous semantics facilitating their interoperability and reuse.

Contributors

All authors contributed in the inception of the study. LMR, CP and JAM have been the main contributors to the manuscript. LMR led the developments and drafting of the manuscript. CP supervised the study from the SWS perspective and contributed to the developments. JAM supervised the methods developed from the medical informatics perspective. LP contributed with technical developments and advised the implementation of the use case. RC and JGB have critically revised the manuscript and advised the study from the industry and medical informatics perspective.

Conflict of interest

We wish to confirm that there are no known conflicts of interest associated with this publication and there has been no significant financial support for this work that could have influenced its outcome.

Acknowledgments

This work was supported by Helse Nord [grant HST1121-13]; the EU project COMPOSE [FP7-317862]; the Spanish Ministry of

Economy and Competitiveness and the FEDER program [grants TIN2014-53749-C2-1-R and PTQ-12-05620]. We would like to thank Dr. Catalina Martínez-Costa for providing insights into Semantic Patterns and the outcomes of the SemanticHealthNet project. We are grateful to Cambio Healthcare Systems for providing the CDS modules used in the study. We would like to acknowledge Dr. Ian McNicoll, Dr. Gro Berntsen and Dr. Nils Kolstrup for providing information about how clinicians discover and analyze CDS guidelines and systems. Finally, we thank IHTSDO, Indizen and Ontotext for the technologies and licenses provided.

Appendix A. Set of archetypes used in the use case CDS modules

Table 6 presents the set of archetypes used as VMR by the set of GDL modules of the case study. Only the attributes of the archetypes referenced by the decision logic are displayed for brevity.

Appendix B. Rationale behind the CDS functional taxonomy

Table 7 shows how many clinical target task concepts are common to most taxonomies with a different level of specialization. We have selected those common concepts of the reviewed studies and aimed for a minimum level of specialization to model our functional taxonomy. The result of merging the reviewed taxonomies in our model is depicted in Fig. 5. The first level of the functional classification taxonomy is formed by the concept *follow-up management* borrowed from Berner's report. *Prevention and screening* was selected from Sim & Berlin's study and considered equivalent to Berner's *Preventive care*. The concept *process improvement* was borrowed from Wang's study and considered to be general enough to comprehend *cost reduction and improved patient convenience* from Berner's report. The reason for not specializing *cost reduction and improved patient convenience* in a new concept was that the beneficiary of the improvement is not in the domain modeled in this taxonomy. *Diagnosis* is common in Sim & Berlin, Berner and HIMMS taxonomies; it references functionality for diagnosis support. The concept *Treatment/guidelines* is a generalization to encompass the concepts *treatment* (Sim & Berlin), *protocol/pathway* (Osheroff), *planning or implementing treatment* (Berner) and *clinical guideline* (HIMMS). Similarly, *information management* is a concept created to include Osheroff's concepts for information management. 'Hospital efficacy' (Berner) was

Table 6
Archetypes defining the information model referenced from the logic of the CDS modules in the case study.

Archetypes defining the information model referenced from the logic of the CDS modules in the case study	
Archetype	Attribute
openEHR-EHR-EVALUATION.chadsvas_diagnosis_review.v1	Atrial fibrillation Congestive heart failure Hypertension Diabetes Stroke Vascular disease Atrial fibrillation Date (time) of diagnosis review
openEHR-EHR-EVALUATION.exclusion-problem_diagnosis.v1	Exclusion statement
openEHR-EHR-EVALUATION.problem-diagnosis.v1	Diagnose Severity Status
openEHR-EHR-EVALUATION.recommendation.v1	Narrative description of the recommendation
openEHR-EHR-OBSERVATION.imaging_exam-Echocardiography.v1	Left Ventricular Ejection Fraction (LVEF)
openEHR-EHR-EVALUATION.stroke_prevention_review.v1	Follow CDS recommendation CHA2DS2-VASc Score Reason for deviation Actual treatment Review date Decision delayed Reason for delayed decision
openEHR-EHR-EVALUATION.cha2ds2vasc_compliance.v1	Stroke prevention treatment compliance
openEHR-EHR-OBSERVATION.chadsvas_score.v1	Total score Congestive heart failure Vascular disease Diabetes Age Previous stroke Hypertension Gender Total score
openEHR-EHR-EVALUATION.alert.v1	Category Alert Status
openEHR-EHR-OBSERVATION.basic_demographic.v1	Gender Birth date
openEHR-EHR-INSTRUCTION.medication.v1	Name of medication Generic name Date of last administration
openEHR-EHR-OBSERVATION.stroke_risk.v1	Risk for stroke

Table 7
Common concepts in the selected taxonomies to identify clinical target tasks.

Sim & Berlin + Berlin & Sorani & Sim	Wang (restricted to CPOE)	Osheroff	Berner	HIMSS
Prevention & Screening			Preventive care	
Drug Dossing Test Ordering	Process improvement Error prevention	Order prescription facilitator	Hospital, provider efficacy Cost reductions and improved patient convenience	Order sets
Treatment Chronic disease management		Protocol/pathway support	Planning or implementing treatment	Clinical guideline
Health related behavior				
Diagnosis			Diagnosis	Diagnostic support
		Alert & Reminder	Follow-up management	Alerts and reminders
		Documentation forms and templates		Documentation templates
		Relevant data presentation		Patient data reports, dashboards
		Reference information and guidance		Reference information

considered to be included in *Process improvement* and not specialized since it does not identify a new CDS clinical task target but rather the organization that benefits from it.

In the second level of specialization, from left to right, appear *Alert* and *Reminder* from Osheroffs's and HIMMS taxonomies. Both concepts are encompassed by *follow-up management* parent concept. *Alert* and *Reminder* were modeled as two separate concepts as a result of conversations with clinicians and other CDS researchers. The concept from Osheroff's taxonomy *Order/prescription creator facilitator* was modeled as a subtype of *Process improvement* given that orders are found inside *process improvement* in Wang's study. *Order sets* (HIMSS) was considered included in *Order/prescription creator facilitator*. *Chronic disease management* (Sim & Berlin) is modeled as specialization of *treatment/guidelines*. Finally, the three types of information management concepts extracted from Osheroff's taxonomy *Reference information and guidance*, *Relevant data presentation* and *Documentation forms and templates* were defined as specializations of *Information management*. The HIMSS concept *Patient data reports, dashboards* was considered to be well represented by *Relevant data representation* and not included. HIMSS's *documentation templates* is represented by Osheroff's *Documentation forms and templates*. HIMSS *Diagnostic support* is represented by *Diagnosis*; and *Reference information* by *Reference information and guidance*.

The third level of specialization contains only the concepts *drug dosing/prescribing* and *test ordering* selected from Sim & Berlin's taxonomy specializing *Order/prescription creator facilitator*.

Wang's *error prevention* has been considered represented by *process improvement*, while the other types in Wang's levels *Benefits* (*policy implementation* and *decision support*), *domain* and *class* have not been included as they are not considered to specify any category of CDS clinical task target or are considered to be too specific.

References

[1] E.S. Berner, Clinical Decision Support Systems: State of the Art, AHRQ Publication No. 09-0069-EF, 2009.
[2] K. Kawamoto, C.A. Houlihan, E.A. Balas, D.F. Lobach, Improving clinical practice using clinical decision support systems: a systematic review of trials to identify features critical to success, BMJ 330 (2005) 765, http://dx.doi.org/10.1136/bmj.38398.500764.8F.
[3] A.X. Garg, N.K.J. Adhikari, H. McDonald, M.P. Rosas-Arellano, P.J. Devereaux, J. Beyene, et al., Effects of computerized clinical decision support systems on practitioner performance and patient outcomes: a systematic review, JAMA 293 (2005) 1223–1238, http://dx.doi.org/10.1001/jama.293.10.1223.
[4] D.W. Bates, G.J. Kuperman, S. Wang, T. Gandhi, A. Kittler, L. Volk, et al., Ten commandments for effective clinical decision support: making the practice of evidence-based medicine a reality, J. Am. Med. Inform. Assoc. 10 (2003) 523–530, http://dx.doi.org/10.1197/jamia.M1370.
[5] Managing clinical knowledge for health care improvement, n.d., <http://www.ihi.org/resources/Pages/Publications/Managingclinicalknowledgeforhealthcareimprovement.aspx> (accessed October 9, 2015).
[6] S.W. Tu, J.R. Campbell, J. Glasgow, M.A. Nyman, R. McClure, J. McClay, et al., The SAGE guideline model: achievements and overview, J. Am. Med. Inform. Assoc. 14 (2007) 589–598, http://dx.doi.org/10.1197/jamia.M2399.
[7] E.S. Berner (Ed.), Clinical Decision Support Systems: Theory and Practice, second ed., Springer, New York, NY, 2007.
[8] T.S. Field, P. Rochon, M. Lee, L. Gavendo, S. Subramanian, S. Hoover, et al., Costs associated with developing and implementing a computerized clinical decision support system for medication dosing for patients with renal insufficiency in the long-term care setting, J. Am. Med. Inform. Assoc. 15 (2008) 466–472, http://dx.doi.org/10.1197/jamia.M2589.
[9] M. Peleg, L.A. Gutnik, V. Snow, V.L. Patel, Interpreting procedures from descriptive guidelines, J. Biomed. Inform. 39 (2006) 184–195, http://dx.doi.org/10.1016/j.jbi.2005.06.002.
[10] S. Maviglia, M. Sordo, Practical Approaches to Knowledge Management: Focus on Clinical Decision Support, MedInfo 2015, Sao Paulo, 2015.
[11] K. Kawamoto, T. Hongsermeier, A. Wright, J. Lewis, D.S. Bell, B. Middleton, Key principles for a national clinical decision support knowledge sharing framework: synthesis of insights from leading subject matter experts, J. Am. Med. Inform. Assoc. 20 (2013) 199–207, http://dx.doi.org/10.1136/amiajnl-2012-000887.
[12] B.E. Dixon, L. Simonaitis, H.S. Goldberg, M.D. Paterno, M. Schaeffer, T. Hongsermeier, et al., A pilot study of distributed knowledge management and clinical decision support in the cloud, Artif. Intell. Med. 59 (2013), http://dx.doi.org/10.1016/j.artmed.2013.03.004.
[13] M. Peleg, Computer-interpretable clinical guidelines: a methodological review, J. Biomed. Inform. 46 (2013) 744–763, http://dx.doi.org/10.1016/j.jbi.2013.06.009.
[14] V.L. Patel, E.H. Shortliffe, Chapter 10 – Human-Intensive Techniques, in: R.A. Greenes (Ed.), Clinical Decision Support, second ed., Academic Press, Oxford, 2014, pp. 285–308.
[15] M. Samwald, K. Fehre, J. de Bruin, K.-P. Adlassnig, The Arden Syntax standard for clinical decision support: experiences and directions, J. Biomed. Inform. 45 (2012) 711–718, http://dx.doi.org/10.1016/j.jbi.2012.02.001.
[16] A.A. Boxwala, M. Peleg, S. Tu, O. Ogunyemi, Q.T. Zeng, D. Wang, et al., GLIF3: a representation format for sharable computer-interpretable clinical practice guidelines, J. Biomed. Inform. 37 (2004) 147–161, http://dx.doi.org/10.1016/j.jbi.2004.04.002.
[17] K. Kawamoto, D.F. Lobach, Proposal for fulfilling strategic objectives of the U.S. roadmap for national action on decision support through a service-oriented architecture leveraging HL7 services, J. Am. Med. Inform. Assoc. 14 (2007) 146–155, http://dx.doi.org/10.1197/jamia.M2298.
[18] K. Kawamoto, D.F. Lobach, Design, implementation, use, and preliminary evaluation of SEBASTIAN, a standards-based Web service for clinical decision support. AMIA Annu. Symp. Proc., 2005, pp. 380–384.
[19] HL7 Standards Product Brief – HL7 Implementation Guide: Decision Support Service, Release 1, n.d., <http://www.hl7.org/implement/standards/product_brief.cfm?product_id=334> (accessed October 8, 2015).
[20] A. Wright, D.F. Sittig, J.S. Ash, J.L. Erickson, T.T. Hickman, M. Paterno, et al., Lessons learned from implementing service-oriented clinical decision support at four sites: a qualitative study, Int. J. Med. Inform. 84 (2015) 901–911, http://dx.doi.org/10.1016/j.ijmedinf.2015.08.008.

[21] A. Moreno-Conde, D. Moner, W.D. da Cruz, M.R. Santos, J.A. Maldonado, M. Robles, et al., Clinical information modeling processes for semantic interoperability of electronic health records: systematic review and inductive analysis, J. Am. Med. Inform. Assoc. 22 (2015) 925–934, http://dx.doi.org/10.1093/jamia/ocv008.

[22] C. Martínez-Costa, R. Cornet, D. Karlsson, S. Schulz, D. Kalra, Semantic enrichment of clinical models towards semantic interoperability. The heart failure summary use case, J. Am. Med. Inform. Assoc. 22 (2015) 565–576, http://dx.doi.org/10.1093/jamia/ocu013.

[23] V. Stroetmann, D. Kalra, P. Lewalle, A. Rector, J.M. Rodrigues, K.A. Stroetmann, et al. European Commission, Semantic Interoperability for Better Health and Safer Healthcare, Deployment and Research Roadmap for Europe, 2009.

[24] P.D. Fensel, D.F.M. Facca, D.E. Simperl, I. Toma, Web service modeling ontology, Semantic Web Services, Springer, Berlin, Heidelberg, 2011, pp. 107–129.

[25] S.A. McIlraith, T.C. Son, H. Zeng, Semantic web services, IEEE Intell. Syst. 16 (2001) 46–53, http://dx.doi.org/10.1109/5254.920599.

[26] C. Pedrinaci, J. Domingue, A.P. Sheth, Semantic web services, in: J. Domingue, D. Fensel, J.A. Hendler (Eds.), Handbook of Semantic Web Technologies, Springer, Berlin, Heidelberg, 2011, pp. 977–1035.

[27] Web Service Modeling Ontology (WSMO), n.d., <http://www.w3.org/Submission/WSMO/> (accessed October 16, 2015).

[28] OWL-S: Semantic Markup for Web Services, n.d., <http://www.w3.org/Submission/OWL-S/> (accessed November 22, 2015).

[29] C. Bizer, T. Heath, T. Berners-Lee, Linked Data - The Story So Far, IJSWIS 5 (2009) 1–22, http://dx.doi.org/10.4018/jswis.2009081901.

[30] C. Pedrinaci, J. Domingue, Toward the next wave of services: linked services for the web of data, J. Univ. Comput. Sci. 16 (2010) 1694–1719, http://dx.doi.org/10.3217/jucs-016-13-1694.

[31] HL7 Standards Product Brief - HL7 Version 3 Standard: Clinical Decision Support Knowledge Artifact Specification, Release 1.2, n.d., <http://www.hl7.org/implement/standards/product_brief.cfm?product_id=337> (accessed December 17, 2014).

[32] Microsoft Word - SAGEGuidelineModelSpec1.65.doc - SAGEGuidelineModelSpec1.65.pdf, n.d., <http://sage.wherever.org/references/docs/SAGEGuidelineModelSpec1.65.pdf> (accessed October 15, 2015).

[33] OpenEHR Common Information Model, n.d., <http://www.openehr.org/releases/RM/latest/docs/common/common.html#_common_information_model> (accessed January 26, 2016).

[34] D.R. Sutton, J. Fox, The syntax and semantics of the PROforma guideline modeling language, J. Am. Med. Inform. Assoc. 10 (2003) 433–443, http://dx.doi.org/10.1197/jamia.M1264.

[35] Guideline Definition Language (GDL), n.d., <http://www.openehr.org/releases/CDS/latest/docs/GDL/GDL.html> (accessed February 25, 2016).

[36] P. Ram, D. Berg, S. Tu, G. Mansfield, Q. Ye, R. Abarbanel, et al., Executing clinical practice guidelines using the SAGE execution engine, Stud. Health Technol. Inform. 107 (2004) 251–255.

[37] MobiGuide, n.d., <http://www.mobiguide-project.eu/> (accessed November 14, 2014).

[38] Virtual Medical Record (vMR), n.d., <http://wiki.hl7.org/index.php?title=Virtual_Medical_Record_%28vMR%29>.

[39] OpenCDS, n.d., <http://www.opencds.org/> (accessed October 13, 2015).

[40] Welcome to the NCBO BioPortal|NCBO BioPortal, n.d., <http://bioportal.bioontology.org/> (accessed November 24, 2015).

[41] T. Berners-lee, J. Hollenbach, K. Lu, J. Presbrey, E.P. D'ommeaux, M. Schraefel, Tabulator Redux: Writing Into the Semantic Web, 2007.

[42] Data - W3C, n.d., <http://www.w3.org/standards/semanticweb/data> (accessed November 20, 2015).

[43] Linked Data, first ed., Shelter Island, NY: Manning Publications, 2014.

[44] T. Berners-Lee, Linked Data - Design Issues, n.d., <https://www.w3.org/DesignIssues/LinkedData.html> (accessed March 7, 2016).

[45] P.D. Fensel, D.F.M. Facca, D.E. Simperl, I. Toma, Semantic web, Semantic Web Services, Springer, Berlin, Heidelberg, 2011, pp. 87–104.

[46] Linked Data|Linked Data - Connect Distributed Data across the Web, n.d., <http://linkeddata.org/> (accessed November 11, 2015).

[47] About|DBpedia, n.d., <about:reader?url=http%3A%2F%2Fdbpedia.org%2Fabout> (accessed November 11, 2015).

[48] C. Pedrinaci, D. Liu, M. Maleshkova, D. Lambert, J. Kopecky, J. Domingue, iServe: a linked services publishing platform, in: Proceedings of Ontology Repositories and Editors for the Semantic Web at 7th ESWC, 2010.

[49] iserve technology full details|Knowledge Media Institute|The Open University, n.d., <http://kmi.open.ac.uk/technologies/name/iserve> (accessed February 22, 2015).

[50] SOA4All|Knowledge Media Institute|The Open University, n.d., <http://projects.kmi.open.ac.uk/soa4all/index.html> (accessed October 12, 2015).

[51] C. Pedrinaci, iServe - Data Model. iServe - Data Model, n.d., <http://kmi.github.io/iserve/latest/data-model.html> (accessed September 17, 2015).

[52] Semantic Annotations for WSDL and XML Schema, n.d., <http://www.w3.org/TR/sawsdl/> (accessed November 22, 2015).

[53] WSMO-Lite: Lightweight Semantic Descriptions for Services on the Web, n.d., <http://www.w3.org/Submission/2010/SUBM-WSMO-Lite-20100823/> (accessed November 22, 2015).

[54] DCMI Home: Dublin Core® Metadata Initiative (DCMI), n.d., <http://dublincore.org/> (accessed October 12, 2015).

[55] A. Wright, H. Goldberg, T. Hongsermeier, B. Middleton, A description and functional taxonomy of rule-based decision support content at a large integrated delivery network, J. Am. Med. Inform. Assoc. 14 (2007) 489–496, http://dx.doi.org/10.1197/jamia.M2364.

[56] J.A. Osheroff, Healthcare Information and Management Systems Society, Improving Medication use and Outcomes with Clinical Decision Support: A Step-by-step Guide, Healthcare Information and Management Systems Society Mission, Chicago, IL, 2009.

[57] J.K. Wang, M.M. Shabot, R.G. Duncan, J.X. Polaschek, D.T. Jones, A clinical rules taxonomy for the implementation of a computerized physician order entry (CPOE) system, Proc. AMIA Symp. (2002) 860–863.

[58] A. Berlin, M. Sorani, I. Sim, A taxonomic description of computer-based clinical decision support systems, J. Biomed. Inform. 39 (2006) 656–667, http://dx.doi.org/10.1016/j.jbi.2005.12.003.

[59] I. Sim, A. Berlin, A Framework for Classifying Decision Support Systems, in: AMIA Annu. Symp. Proc. 2003, 2003, pp. 599–603.

[60] Healthcare, Information and Management, The Healthcare Information and Management Society (HIMMS), Healthcare. Types of Clinical Decision Support, 2011.

[61] Arden Syntax, n.d., <http://www.hl7.org/Special/Committees/arden/index.cfm> (accessed November 13, 2014).

[62] PROV-O: The PROV Ontology, n.d., <http://www.w3.org/TR/2013/REC-prov-o-20130430/> (accessed March 16, 2015).

[63] J. Fox, A. Alabassi, V. Patkar, T. Rose, E. Black, An ontological approach to modelling tasks and goals, Comput. Biol. Med. 36 (2006) 837–856, http://dx.doi.org/10.1016/j.compbiomed.2005.04.011.

[64] A.L. Rector, R. Qamar, T. Marley, Binding ontologies and coding systems to electronic health records and messages, Appl. Ontol. 4 (2009) 51–69, http://dx.doi.org/10.3233/AO-2009-0063.

[65] W. Goossen, A. Goossen-Baremans, M. van der Zel, Detailed CLINICAL MODELS: A REview, Healthcare Inform. Res. 16 (2010) 201, http://dx.doi.org/10.4258/hir.2010.16.4.201.

[66] W.-G. Zhu, Q.-M. Xiong, K. Hong, Meta-analysis of CHADS2 versus CHA2DS2-VASc for predicting stroke and thromboembolism in atrial fibrillation patients independent of anticoagulation, Tex. Heart Inst. J. 42 (2015) 6–15, http://dx.doi.org/10.14503/THIJ-14-4353.

[67] Cambio – Start, n.d., <http://www.cambio.se/> (accessed October 12, 2015).

[68] W3C. OWL 2 Web Ontology Language Document Overview (second ed.), n.d., <http://www.w3.org/TR/owl2-overview/>.

[69] Apache Jena - Fuseki: serving RDF data over HTTP, n.d., <https://jena.apache.org/documentation/serving_data/> (accessed November 2, 2015).

[70] D. Karlsson, M. Nyström, R. Cornet, Does SNOMED CT post-coordination scale?, Stud Health Technol. Inform. 205 (2014) 1048–1052.

[71] Y. Kazakov, M. Krötzsch, F. Simancík, Concurrent Classification of EL Ontologies, in: Proceedings of the 10th International Conference on The Semantic Web – Volume Part I, Springer-Verlag, Berlin, Heidelberg, 2011, pp. 305–320.

[72] Time Ontology in OWL, n.d., <http://www.w3.org/TR/owl-time/> (accessed October 27, 2015).

[73] Ocean Informatics. Clinical Knowledge Manager. OpenEHR Clinical Knowledge Manager, n.d., <http://www.openehr.org/ckm/> (accessed October 14, 2013).

[74] Nasjonal IKT. Clinical Knowledge Manager, n.d., <http://arketyper.no/ckm/> (accessed October 20, 2014).

[75] Mission and Goals|www.opencimi.org, n.d., <http://www.opencimi.org/> (accessed October 5, 2015).

[76] C.M. Costa, M. Menárguez-Tortosa, J.T. Fernández-Breis, Clinical data interoperability based on archetype transformation, J. Biomed. Inform. 44 (2011) 869–880, http://dx.doi.org/10.1016/j.jbi.2011.05.006.

[77] L. Lezcano, M.-A. Sicilia, C. Rodríguez-Solano, Integrating reasoning and clinical archetypes using OWL ontologies and SWRL rules, J. Biomed. Inform. 44 (2011) 343–353, http://dx.doi.org/10.1016/j.jbi.2010.11.005.

[78] E. Sanchez, C. Toro, M. Graña, C. Sanin, E. Szczerbicki, Extended Reflexive Ontologies for the generation of clinical recommendations, Cybernet. Syst. 46 (2015) 4–18.